Gossip vs. Markov Chains, and Randomness-Efficient Rumor Spreading

Zeyu Guo* He Sun†

Abstract

We study gossip algorithms for the rumor spreading problem which asks one node to deliver a rumor to all nodes in an unknown network, and every node is only allowed to call one neighbor in each round. In this work we introduce two fundamentally new techniques in studying the rumor spreading problem:

First, we establish a new connection between the rumor spreading process in an *arbitrary* graph and certain Markov chains. While most previous work analyzed the rumor spreading time in general graphs by studying the rate of the number of (un-)informed nodes after every round, we show that the mixing time of a certain Markov chain suffices to bound the rumor spreading time in an arbitrary graph.

Second, we construct a reduction from rumor spreading processes to branching programs. This reduction gives us a general framework to derandomize the rumor spreading and other gossip processes. In particular, we show that, for *any* n-vertex expander graph, there is a protocol which informs every node in $O(\log n)$ rounds with high probability, and uses $O(\log n \cdot \log\log n)$ random bits in total. The runtime of our protocol is tight, and the randomness requirement of $O(\log n \cdot \log\log n)$ random bits almost matches the lower bound of $\Omega(\log n)$ random bits. We further show that, for many graph families (defined with respect to the expansion and the degree), $O(\text{poly}\log n)$ random bits in total suffice for fast rumor spreading. These results give us an almost complete understanding of the role of randomness in the rumor spreading process, which was extensively studied over the past years.

1 Introduction

Gossip algorithms are one of the most important communication primitives in large networks, and have been studied under different names such as rumor spreading, information dissemination, and broadcasting. Efficient gossip algorithms for information spreading have wide applications in failure detection [37], resource discovery [27], replicated database systems [10], data aggregation [3], etc.

*California Institute of Technology, Pasadena, USA. This work is supported by NSF CCF-1423544, CCF-111611, and BSF grant 2010120. Part of this work was done while visiting Max Planck Institute for Informatics.

†Max Planck Institute for Informatics, Saarbrücken, Germany, and Cluster of Excellence, "Multimodal Computing and Interaction", Universität des Saarlandes, Saarbrüken, Germany. This work has partially been funded by the Cluster of Excellence "Multimodal Computing and Interaction" within the Excellence Initiative of the German Federal Government. Part of this work was done while visiting California Institute of Technology.

The simplest and widely studied form of gossip algorithms is the so-called *push model* of rumor spreading [18, 32]. Initially, a message, called *a rumor*, is placed on an arbitrary node of an unknown network with n nodes. In subsequent synchronous rounds, every node that knows the rumor picks a neighbor uniformly at random and sends the rumor to the chosen neighbor. This process continues until every node gets the rumor. It was shown that this simple protocol is very efficient on several network topologies [12, 13, 14, 16, 19]. In addition, the protocol is local, simple, and can tolerate link failures.

Over the past decades extensive studies have focused on the *rumor spreading time*, i.e., the number of rounds required before every node gets the rumor with high probability. While usually good expansion of the underlying graph implies fast rumor spreading [3, 5, 6, 19, 21, 34], it was far from clear if graph expansion is the *only* reason for fast rumor spreading and more general gossip processes. For instance, most of these gossip algorithms are inherently randomized, and all of the analysis of these algorithms crucially rely on choosing neighbors *independently and uniformly at random* in each round. However, from a theoretical point of view, it is not clear if randomization is essential for efficiently spreading a rumor. From a practical point of view, choosing neighbors independently and uniformly at random in each round requires every node of the graph to have access to a random source of *unbiased and independent* coins, whose physical realization is unknown. Hence, both of theoretical considerations and practical difficulties in obtaining truly random sources lead to the following fundamental question: What is the role of randomness in fast rumor spreading and other gossip algorithms? More specifically, how many random bits are sufficient for efficiently spreading a rumor? While for any graph with n nodes, the above-mentioned *fully-random* push protocol requires $O(T \cdot n \log n)$ random bits for spreading a rumor within T rounds, it is not difficult to show that, for any graph with n nodes, there is a protocol which uses $O(\log n)$ random bits in total, and whose rumor spreading time is as fast as the standard fully-random protocol. However, explicit

constructions of such protocols are widely open, and a long line of research has been devoted to constructing randomness-efficient and deterministic protocols for rumor spreading and similar problems, e.g. [11, 22, 23, 26].

1.1 Our Results In this work we present several randomness-efficient protocols for which both the rumor spreading time and the randomness requirement are *almost-optimal*. One fundamentally new technique introduced in our work is to establish a connection between rumor spreading processes and branching programs, a well-studied computation model in complexity theory. Based on this connection, we present a novel reduction from the problem of designing rumor spreading protocols to the problem of constructing pseudorandom generators (PRGs) for branching programs. This reduction gives the first application of the model of branching programs in the area of parallel and distributed computing, and provides a powerful tool for designing and analyzing gossip algorithms.

To informally discuss the reduction, we notice one natural connection between gossip processes and random walks in a branching program: First, random walks over a branching program resemble the rumor spreading process where nodes send rumors to random neighbors. Second, in the rumor spreading process, each node has access to only its own list of neighbors, and is oblivious to the structure of the network. This is an analogue of the *oblivious derandomization* achieved by PRGs.

However, rumor spreading appears much more complicated than small-space computation due to the following facts: First, in the rumor spreading process, rumors are "duplicated" every round, although every "existing" rumor viewed individually performs a random walk. Hence, instead of considering every single random walk performed by any fixed rumor, we need to study the dynamics of the whole rumor spreading process. Second, the state of the process at some round essentially depends on the past behavior of all nodes and is by no means computable in small-space. Indeed, even knowing if a single node u gets the rumor at some round requires knowing the set of its neighbors having the rumor in the previous rounds, and may require $\deg(u) = \Theta(n)$ bits for dense graphs. Hence, this connection between rumor spreading and small-space computation is delicate and not obvious.

Surprisingly, we show that such a reduction from designing rumor spreading protocols to constructing PRGs for branching programs exists. Based on this reduction, numerous techniques developed in the constructions of PRGs for small-space computation can be applied in designing rumor spreading and other gossip processes. In particular, we prove that certain PRGs with optimal parameters imply rumor spreading protocols whose randomness complexity matches the lower bound and the best known upper bound of the existential results from probabilistic methods. Our main result is summarized as follows:

Theorem 1.1 (**Main Result**). *Let G be a graph with n nodes, spectral gap $\alpha \in (0,1)$, and irregularity $\beta \triangleq \Delta/\delta$. Then there is an explicit protocol using $O((\log(1/\alpha) + \log\beta + \log\log n)\cdot\log n)$ random bits, such that with high probability all nodes get the rumor in $T = O(C\log n)$ rounds, where $C = (1/\alpha)\cdot\beta^2\cdot\max\{1, 1/(\alpha\cdot\Delta^{0.499})\}$.*

Theorem 1.1 implies that, for *any* expander graph with n nodes, $\alpha = \Theta(1)$ and $\beta = O(1)$, the protocol informs all nodes in $O(\log n)$ rounds with high probability and uses $O(\log n\cdot\log\log n)$ random bits in total. Note that any protocol needs at least $\Omega(\log n)$ rounds to spread a rumor to all nodes, hence the rumor spreading time of Theorem 1.1 for expander graphs is tight. For the randomness complexity, our result improves the previous best bound of $O(\log^2 n)$ random bits [22]. We also proved that, for any expander graph with minimum degree $\delta = n^{\Theta(1)}$, any protocol that finishes in $O(\log n)$ rounds with high probability needs at least $\Omega(\log n)$ random bits, therefore the result in Theorem 1.1 is almost tight.

We further introduce a different technique, and address the rumor spreading problem in a slightly restricted setting, where every node u knows the IDs of its neighbors and its index for each of its neighbors[1]. We prove that, under this condition, the randomness complexity in Theorem 1.1 can be further improved. Comparing with Theorem 1.1, we analyze a more general gossip process. Our result is summarized as follows:

Theorem 1.2. *Let G be a graph with n nodes, spectral gap $\alpha \in (0,1)$, and irregularity $\beta \triangleq \Delta/\delta$. Let $\mathsf{List}(u)$ be the adjacency list of node u, and $N(u)$ be the set of neighbors of u. We assume that each node u knows the ID of its neighbors $v \in N(u)$, and its index in $\mathsf{List}(v)$ for any neighbor $v \in N(u)$. Then there is an explicit protocol using $O((\log(1/\alpha) + \log\beta + \log\log n)\cdot\log n)$ random bits, such that with high probability all nodes get the rumor in $T = O((1/\alpha)\cdot\beta^2\cdot\log n)$ rounds.*

We further present protocols for graphs with small Δ. In contrast to Theorem 1.1 and Theorem 1.2 that are based on branching programs, the following result relies on the observation that the rumor spreading process enjoys nice locality when Δ is small.

[1] We remark that similar assumptions are also made in other references, e.g. [26], and one can deterministically use $O(\Delta)$ preprocessing time to guarantee this assumption.

Theorem 1.3. *Let G be a graph with n nodes, conductance ϕ and irregularity β. Then there is an explicit protocol using $O\big((1/\phi)\cdot\beta\cdot\log n\cdot(\log\log n+\log\Delta)\big)$ random bits in total, such that with high probability all nodes get the rumor in $O((1/\phi)\cdot\beta\cdot\log n)$ rounds.*

The rumor spreading time in Theorem 1.3 matches the upper bound known in the truly random protocol, and is tight, in the sense that there are graphs with diameter $\Omega((1/\phi)\cdot\beta\cdot\log n)$ [5]. Our result improves the previous best result [22], which needs $O((1/\phi)\cdot\log^2 n)$ random bits in total and only holds for graphs with $\beta=O(1)$.

Our protocol takes advantage of the locality by using a "two-level hashing" construction: We use a family of objects called *unbalanced expanders* to hash the node IDs into a smaller space, and then apply the classical pairwise independent generators. This construction yields much smaller seed length than using pairwise independent generators alone [22]. The protocol has the advantage of being very simple. Furthermore, a variant of this protocol using PRGs for combinatorial rectangles achieves the best possible rumor spreading time for strong expander graphs:

Theorem 1.4. *Let G be a graph with n nodes, $\alpha=1-o(1)$ and irregularity $\beta=1+o(1)$. Then there is an explicit protocol using $O(\log n\cdot(\log\log n+\log\Delta))$ random bits in total, such that with high probability all nodes get the rumor in $\log n+\ln n+o(\log n)$ rounds.*

In comparison to the protocol in Theorem 1.1 which requires $O(\log n)$ rounds for expander graphs, the protocol in Theorem 1.4 is applied for *strong expander graphs* (i.e., graphs with $\alpha=1-o(1)$), and the rumor spreading time of $\log n+\ln n+o(\log n)$ rounds matches the *precise rumor spreading time* for the fully-random protocol [13, 14, 16], which is known to be tight [14]. Moreover, our protocol uses $O(\log n\cdot(\log\log n+\log\Delta))$ random bits in total, in contrast to $\Omega(\log^3 n)$ random bits used in [22]. These four results (Theorem 1.1–Theorem 1.4), together with the existential result and the lower bound analysis, give us an almost complete understanding of the randomness requirement of this fundamental gossip process on expander graphs, and greatly extend our understanding of the randomness requirement for more general graph families.

Remark 1.5. *The analysis of our protocols can be easily adapted to the real-world settings, where communication failures are present. We assume that each node v fails to send the rumor with probability $\gamma_{v,i}$ in round i, where $\gamma_{v,i}$ is upper bounded by some parameter $\gamma<1$. Moreover, we assume that these link failures are independent for different nodes and/or different rounds. Under these two assumptions, our results still hold with the runtime bound multiplied by a factor of $O((1-\gamma)^{-1})$.*

1.2 Techniques In the current work we introduce several new techniques in studying gossip processes. In addition to the reduction from rumor spreading processes to branching programs, other techniques include: approximating rumor spreading time via random walks, analyzing the rumor spreading process via Doeblin coupling, and simulating pull operations by push operations in designing randomness-efficient protocols. Let us briefly discuss these techniques in this subsection.

Approximation via Random Walks. The usual analyses for fast rumor spreading proceed by showing that the number of informed/uninformed nodes increases/decreases over time, e.g. [5, 6, 19, 21, 29, 34]. Our approach is fundamentally different from previous work. Roughly speaking, we approximate the rumor spreading process by a collection of random walks, and use the rapid mixing of these random walks to prove the property of fast rumor spreading. It turns out that the pieces of local information provided by these random walks give a surprisingly good control of the global behavior of rumor spreading, despite that the walks are complicatedly and highly correlated.

Formally, we approximate the rumor spreading process by various random walks, distinguished by whether the walks are lazy or non-lazy in each round. Each walk is associated with a non-negative number called its *weight*. A node u is informed within T rounds if the total weight of random walks of length T reaching u is positive. By Cauchy-Schwarz inequality, we lower bound the probability of this event in terms of the expectation of the total weights of the random walks reaching u as well as its second moment.

Analysis of Markov Chains. We further show that, by choosing the weights intelligently, the expectation and the second moment of the total weights reaching a node are computed by certain Markov chains. The expected total weights are computed by the chain $\mathbf{M}$ representing a lazy random walk in the graph. It follows from the rapid mixing of $\mathbf{M}$ that it can be well estimated using the stationary distribution of $\mathbf{M}$. However, the case for the second moments is more complicated as they correspond to a *non-reversible* Markov chain $\mathbf{M}'$. One key result we manage to show is that $\mathbf{M}'$ and $\mathbf{M}\otimes\mathbf{M}$ have very close stationary distributions and comparable mixing time. This result is also very interesting on its own, since $\mathbf{M}'$ is a very natural Markov chain and closely related to *Doeblin coupling* [30].

Simulating Pull by Push. While a randomness-efficient protocol using a global seed can be easily implemented in the push model, the "dual" protocol

in the pull model is not physically realizable, as it is impossible for a node to perform a random pull operation before getting the random seed. Using the technique called *simulating pull by push*, we are able to use the analysis for pull operations while actually performing push operations. This is crucial in our analysis, since when most nodes already have the rumor, the random walks defined via push operations become too congested and correlated, whereas the "reversed" random walks using pull operations work well.

1.3 Related Work Gossip and rumor spreading have been extensively studied under various settings and aspects over the past decades. The first line of research focuses on the rumor spreading time and graph expansion. Chierichetti et al. [5, 6] studied the rumor spreading time and the conductance of the underlying graph, and proved that the push-pull protocol informs every node in $\widetilde{O}((1/\phi)\cdot\log n)$ rounds with high probability, where ϕ is the conductance of the graph and $\widetilde{O}(\cdot)$ hides a $\operatorname{poly}\log(1/\phi)$ factor. Their result was improved to $O((1/\phi)\cdot\log n)$ by Giakkoupis [19], and this bound is tight. Sauerwald and Stauffer [34] proved that, for any regular graph with vertex expansion α, the push protocol informs every node in $O((1/\alpha)\cdot\log^5 n)$ rounds with high probability. Subsequent work on rumor spreading versus vertex expansion includes [20, 21]. Giakkoupis [20] showed that the push-pull protocol informs every node in $O((1/\alpha)\cdot\log^2 n)$ rounds with high probability, and this bound is tight.

While all these studies indicate that good expansion properties imply fast rumor spreading, it was not clear whether fast rumor spreading is also due to extensive use of randomness in the whole process. Hence, the second line of the research is to understand and determine the amount of randomness required in fast rumor spreading and other gossip algorithms.

Doerr et al. [11] proposed a quasirandom version of the rumor spreading push protocol: Every node has a list of its neighbors, and chooses a random position on the list when it gets the rumor the first time. From then on, it informs its neighbors starting from that position and continues in the order of the list. Their protocol uses $O(n\log n)$ random bits in total, and informs every node in polylogarithmic number of rounds with high probability in several network topologies. Giakkoupis and Woelfel [23] further presented a modification of the quasirandom protocol which uses $O(n\log\log n)$ random bits in total and informs every node in $O(\log n)$ rounds with high probability. Their result and analysis only hold for complete graphs.

Recently, Giakkoupis et al. [22] introduced two low-randomness rumor spreading protocols. Their protocols are based on pairwise independent hash functions, and Nisan's pseudorandom generators. Comparing with the fully-random protocol that requires $O(n\log n)$ random bits in each round, the protocols in [22] only require polylogarithmic number of random bits per round. However, their analysis requires that the random choices in different rounds are independent. Since $\Omega(\log n)$ random bits per round are needed in their analysis, these two protocols need $\Omega(\log^2 n)$ random bits in total.

Randomness requirement of other gossip processes has also been studied. Haeupler [26] studied the k-local broadcast problem and the global broadcast problem, and presented a deterministic protocol. These two problems in [26] assume that every node has one rumor and the protocol requires $\Omega(\log^2 n)$ rounds, hence the techniques of designing deterministic protocols developed there seem difficult to be applied in our setting.

Finally, we note that our work is closely related to multiple random walks [1, 15], derandomizing and deterministic random walks [8, 9], as well as derandomizing other distributed processes (e.g., averaging [3], and load balancing [17]), for which the role of randomness has been studied.

2 Preliminaries

Let $G=(V,E)$ be a connected, undirected, and simple graph with n nodes. For any node u, the degree of u is represented by $\deg(u)$. Let Δ,δ and d be the maximum, minimum and average degree of G, and call $\beta \triangleq \Delta/\delta$ the *irregularity* of G. The set of neighbors of a node u is represented by $N(u)$. Moreover, for any set $S\subseteq V$, let $N(S)\triangleq\bigcup_{u\in S}N(u)$, and $\mathrm{vol}(S)\triangleq\sum_{u\in S}\deg(u)$. For any set $S,T\subseteq V$, we define

$$E(S,T)\triangleq\{\{u,v\}:u\in S \text{ and } v\in T\},$$

and $e(S,T)\triangleq|E(S,T)|$.

Let $\mathbf{A}_G$ be the adjacency matrix of G, and $\mathbf{N}_G\triangleq\mathbf{D}^{-1/2}\mathbf{A}_G\mathbf{D}^{-1/2}$, where $\mathbf{D}$ is the diagonal matrix defined by $\mathbf{D}_{uu}=\deg(u)$ for $u\in V[G]$. Define the eigenvalues of $\mathbf{N}_G$ by $1=\lambda_1\geqslant\cdots\geqslant\lambda_n\geqslant-1$, and let $\lambda_{\max}\triangleq\max\{\lambda_2,|\lambda_n|\}$. The spectral gap α is defined by $\alpha\triangleq1-\lambda_2$, and let the absolute spectral gap be $1-\lambda_{\max}$. For simplicity, we also use α to express the spectral expansion of a reversible Markov chain if the chain is clear from the context.

For $m\in\mathbb{N}$, vector $\mathbf{u}\in\mathbb{R}^m$ and real number $p\geqslant1$, define the ℓ_p-norm by $\|\mathbf{u}\|_p=\left(\sum_{i=1}^m|\mathbf{u}_i|^p\right)^{1/p}$. In addition, we define $\|\mathbf{u}\|_\infty=\max_{1\leqslant i\leqslant m}|\mathbf{u}_i|$. The inner product of two vectors $\mathbf{u},\mathbf{v}\in\mathbb{R}^m$ is $\langle\mathbf{u},\mathbf{v}\rangle=\sum_{i=1}^m\mathbf{u}_i\mathbf{v}_i$. Let $\mathbf{e}_i$ be the vector that has an one in the ith entry and zero elsewhere. Write $\mathbf{I}_m$ or $\mathbf{I}$ for the $m\times m$ identity matrix. For a matrix $\mathbf{M}\in\mathbb{R}^{m\times m'}$, we use $\mathbf{M}_{i,j}$ to denote the entry on $\mathbf{M}$'s ith row and jth column. For

$p \in [1,\infty) \cup \{\infty\}$, define

$$\|\mathbf{M}\|_p = \sup_{\mathbf{u}\in\mathbb{R}^m\setminus\{\mathbf{0}\}} \frac{\|\mathbf{uM}\|_p}{\|\mathbf{u}\|_p},$$

where $\mathbf{0}$ is the zero vector. It is easy to show that $\|\mathbf{M}\|_1$ equals the maximum of the ℓ_1-norms of the rows of $\mathbf{M}$. Moreover, $\|\mathbf{M}\|_\infty$ equals the maximum of the ℓ_1-norms of the columns of $\mathbf{M}$, or equivalently $\|\mathbf{M}^\intercal\|_1$. We say a square matrix $\mathbf{M}$ is stochastic if all of its entries are non-negative and all of its rows have ℓ_1-norm 1. Clearly, if $\mathbf{M}$ is stochastic, then $\|\mathbf{M}\|_1 = 1$. We say $\mathbf{M}$ is doubly-stochastic if both $\mathbf{M}$ and $\mathbf{M}^\intercal$ are stochastic.

By $\log x$ we denote the binary logarithm of x. For any integer m, define $[m] \triangleq \{0,\dots,m-1\}$. The disjoint union of a family of sets $\{A_i : i \in I\}$ indexed by I is denoted by $\bigsqcup_{i\in I} A_i \triangleq \bigcup_{i\in I}\{(x,i) : x \in A_i\}$.

2.1 Pairwise Independent Generators We say $X_0,\dots,X_{d-1}$ with X_i distributed over $[m_i]$ are ε-pairwise independent if

- $\left|\mathbf{Pr}\left[X_i = x\right] - \frac{1}{m_i}\right| \leqslant \varepsilon$ for all $i \in [d]$ and $x \in [m_i]$, and
- $\left|\mathbf{Pr}\left[X_i = x \wedge X_j = x'\right] - \frac{1}{m_i\cdot m_j}\right| \leqslant \varepsilon$ for all distinct $i,j \in [d]$ and all $x \in [m_i]$, $x' \in [m_j]$.

We say they are pairwise independent if $\varepsilon = 0$. We say $\mathcal{G} : \{0,1\}^\ell \mapsto [m_0]\times\cdots\times[m_{d-1}]$ is an (ε-)pairwise independent generator if its outputs are (ε-)pairwise independent given a uniformly distributed seed.

Theorem 2.1 ([4]). *There exists an explicit pairwise independent generator $\mathcal{G} : \{0,1\}^\ell \mapsto [m]^d$ with seed length $\ell = O(\log m + \log d)$.*

Lemma 2.2. *Suppose $\mathcal{G} = (\mathcal{G}_0,\dots,\mathcal{G}_{d-1})$ is a pairwise independent generator where $\mathcal{G}_i : \{0,1\}^\ell \mapsto [m]$. Define $\mathcal{G}' = (\mathcal{G}'_0,\dots,\mathcal{G}'_{d-1})$, where $\mathcal{G}'_i(x) = \mathcal{G}_i(x) \bmod m_i$ for $i \in [d]$. Then $\mathcal{G}' : \{0,1\}^\ell \mapsto [m_0]\times\cdots\times[m_{d-1}]$ is an ε-pairwise independent generator, where $\varepsilon = 2/m$.*

2.2 PRGs for Combinatorial Rectangles Given $d \in \mathbb{N}$ and a finite set $S = \prod_{i\in[d]} S_i$, define $\mathsf{CR}_S \triangleq \left\{\prod_{i\in[d]} A_i : A_i \subseteq S_i\right\}$. The members of CR_S are called *combinatorial rectangles in S* and d the *dimension*. For $\varepsilon > 0$, $d \in \mathbb{N}$, and $S = \prod_{i\in[d]} S_i$, we call $\mathcal{G} : \{0,1\}^\ell \mapsto S$ an ε-PRG for CR_S with seed length ℓ if it holds for any $A \in \mathsf{CR}_S$ that

$$\left|\mathbf{Pr}_{x\in\{0,1\}^\ell}\left[\mathcal{G}(x) \in A\right] - |A|/|S|\right| \leqslant \varepsilon.$$

Theorem 2.3 ([24]). *Let $S = [m]^d$. There exists an explicit ε-PRG for CR_S with seed length $O(\log m + \log d) + \widetilde{O}(\log(1/\varepsilon))$.* [2]

Lemma 2.4. *Suppose $\mathcal{G} = (\mathcal{G}_0,\dots,\mathcal{G}_{d-1})$ is an ε-PRG for CR_S where $S = [m]^d$. Define $\mathcal{G}' = (\mathcal{G}'_0,\dots,\mathcal{G}'_{d'-1})$, where $\mathcal{G}'_j(x) = \mathcal{G}_{i_j}(x) \bmod m_j$ for $j \in [d']$ and $i_0,\dots,i_{d'-1} \in [d]$. Then $\mathcal{G}'$ is an $(\varepsilon + \sum_{i\in[d']} m_i/m)$-PRG for $\mathsf{CR}_{S'}$, where $S' = \prod_{i\in[d']}[m_i]$.*

Lemma 2.5. *There exists an explicit function $\mathcal{G} : \{0,1\}^\ell \mapsto [m]^d$ that is both a pairwise independent generator and an ε-PRG for $\mathsf{CR}_{[m]^d}$ with seed length $O(\log m + \log d) + \widetilde{O}(\log(1/\varepsilon))$.*

2.3 PRGs for Branching Programs A branching program of length L, width W and degree D, or an (L,W,D)-branching program, is a directed (multi)-graph with node set $[W]\times[L+1]$. We say the nodes in $[W]\times\{i\}$ are on the ith layer for $0 \leqslant i \leqslant L$. Each node (u,i) except those on the last layer has D outgoing edges to nodes on the next layer, and these D edges are associated with D distinct labels from $[D]$.

Let $\mathcal{B}$ be a branching program of length L, width W and degree D. For $x = (x_1,\dots,x_L) \in [D]^L$ and a node $(s,0)$ on the first layer, define $\mathcal{B}(s,x) \in [W]$ such that the walk that starts from $(s,0)$ and takes the edge with label x_i at the ith step for $1 \leqslant i \leqslant L$ finally arrives at $(\mathcal{B}(s,x),L)$. We call a function $\mathcal{G} : \{0,1\}^\ell \mapsto [D]^L$ an ε-PRG for (L,W,D)-branching programs if for any (L,W,D)-branching program, and any node $(s,0)$ on the first layer, it holds that

$$\sum_{u\in[W]} \Big|\mathbf{Pr}_{x\in\{0,1\}^\ell}\left[\mathcal{B}(s,\mathcal{G}(x)) = u\right] - \mathbf{Pr}_{x\in[D]^L}\left[\mathcal{B}(s,x) = u\right]\Big| \leqslant \varepsilon.$$

Theorem 2.6 ([28]). *There exists an explicit ε-PRG for (L,W,D)-branching programs with seed length $O(\log L(\log W + \log L + \log(1/\varepsilon)) + \log D)$.*

2.4 Unbalanced Expanders Another family of pseudorandom objects used in our paper is called unbalanced expanders.

Definition 2.7. *Let $\Gamma : [N]\times[D] \mapsto \bigsqcup_{i\in[D]}[M_i]$ be a function where $\Gamma(x,y) \in [M_y]$ for any $x \in [N]$, $y \in [D]$. Function Γ specifies a left-degree D bipartite graph with left vertex set $[N]$ and right vertex set $\bigsqcup_{i\in[D]}[M_i]$ in the*

[2] In [24] the seed length is presented as $O((\log\log m)(\log m + \log d + \log(1/\varepsilon))) + \widetilde{O}(\log(1/\varepsilon))$. But there are standard techniques of reducing m and d to $m' = (1/\varepsilon)^{O(1)}, d' = (1/\varepsilon)^{O(1)}$ using $O(\log m + \log d)$ randomness, cf. [2, 31].

following way: for $x \in [N]$ and $y \in [D]$, the yth neighbor of x is given by $\Gamma(x, y)$.

We are interested in graphs Γ exhibiting excellent expansion properties. This leads to the notion of unbalanced expanders [25, 36].

Definition 2.8 (Unbalanced expanders, [25, 36])**.** *Let $\Gamma : [N] \times [D] \mapsto \bigsqcup_{i\in[D]}[M_i]$ be as in Definition 2.7. We call Γ a (K, A)-expander if for any $S \subseteq [N]$ of size K, it holds that $|N(S)| \geqslant AK$. We call Γ a $(\leqslant K, A)$-expander if it is a (K', A)-expander for all $K' \leqslant K$.* [3]

In particular, we are interested in (K, A)-expanders, where the parameter $A = (1 - \varepsilon)D$ for small ε, i.e., for any subset S of size K from the left set $[N]$, there is almost no collision among the neighbors of nodes in S. Explicit constructions of such unbalanced expanders with near-optimal expansion are known.

Theorem 2.9 ([25])**.** *For any $N \in \mathbb{N}$, $K \leqslant N$, and $\varepsilon > 0$, there is an explicit $(K, (1 - \varepsilon)D)$-expander $\Gamma : [N] \times [D] \mapsto \bigsqcup_{i\in[D]}[M_i]$ with $D = \left(\frac{\log N}{\varepsilon}\right)^{O(1)}$ and $M_0 = \cdots = M_{D-1} \leqslant \max\left\{D, K^{O(1)}\right\}$.*

Assume that $\Gamma : [N] \times [D] \mapsto \bigsqcup_{i\in[D]}[M_i]$ is a $(K, (1 - \varepsilon)D)$-expander. We consider the map $\Gamma(\cdot, U)$ applied on any K elements of $[N]$, where U is uniformly distributed over $[D]$. The following lemma states that with high probability these K elements are mapped into $\bigsqcup_{i\in[D]}[M_i]$ with almost no collision.

Lemma 2.10. *Let $\Gamma : [N] \times [D] \mapsto \bigsqcup_{i\in[D]}[M_i]$ be a $(K, (1-\varepsilon)D)$-expander. Let S be a subset of $[N]$ of size K. Then for at least $(1 - \sqrt{\varepsilon})$-fraction of $y \in [D]$, it holds that*

$$|\{\Gamma(x, y) : x \in S\}| \geqslant (1 - \sqrt{\varepsilon})K.$$

3 Gossip vs. Markov Chains

Let G be a graph with $V[G] = [n]$. We consider only T'-round protocols for G, in which nodes send rumors in the first T' rounds, and assume that $T' = O(n^c)$ for a constant c. Throughout this section, we assume that each node has a unique ID, and each node initially only knows its own ID $\in [n^c]$. Let s be the initial node having the rumor. For simplicity, we assume that the adjacency list of each node u has length Δ, and the last $\Delta - \deg(u)$ neighbors are u itself, i.e., we add $\Delta - \deg(u)$ self-loops for every node u. We call the resulting regular graph $\mathrm{Reg}(G)$. However, we use $\deg(u)$ and $N(u)$ to represent the degree and the set of neighbors of u in the underlying simple graph.

[3] The definition here is slightly different from [25, 36] as we require $\Gamma(x, y) \in [M_y]$. This is analogous to the difference between standard and strong condensers.

3.1 Analysis of the Prototype Protocol

We present our new techniques of analyzing rumor spreading processes via Markov chains, and show how the mixing time of certain Markov chains relates to the rumor spreading time. We first analyze the following prototype of rumor spreading protocols.

Protocol 1 (Prototype of Rumor Spreading)**.** *Let $\mathcal{D}$ be a distribution over the set of functions $f : [T] \times V[G] \mapsto [\Delta]$. Sample f according to $\mathcal{D}$. In round i, an informed node u sends the rumor to its $f(i, u)$th neighbor in its adjacency list.*

We analyze Protocol 1 where $\mathcal{D} = \mathcal{U}$ is the uniform distribution, i.e., the fully-random protocol where every informed node chooses its neighbor uniformly at random, and derive the rumor spreading time for general graphs in the standard push model. One new approach that we introduce in studying Protocol 1 is to approximate the rumor spreading process via random walks, i.e., we compare the process of rumor spreading with a random walk on a branching program. For random walks, a walk always stays at a single node during the process, although this node keeps changing. On the other hand, in the process of rumor spreading, each informed node u randomly sends the rumor to one of its neighbors v in each round, and then u, v are both informed subsequently. So we may think of rumor spreading as many random walks in parallel: When node u sends the rumor to v, one random walk makes a *non-lazy step* and moves from u to v whereas another one makes a *lazy step* and stays at u. In order to characterize this behavior, we introduce the notion of forward and reversed random walks. For any round $i \in [T]$ and node $u \in V[G]$, denote by $\widetilde{f}(i, u)$ the ID of the $f(i, u)$th neighbor of u in its adjacency list.

Definition 3.1 (Forward random walks)**.** *Consider the rumor spreading process in T rounds on a graph G using Protocol 1 determined by $f \sim \mathcal{D} = \mathcal{U}$. A forward random walk of length $k \in [T]$ with pattern $S = (s_0, \ldots, s_{k-1}) \in \mathcal{C}_k \triangleq \{\textit{lazy}, \textit{non-lazy}\}^k$ is a sequence of $k + 1$ nodes $(p_0, \ldots, p_k)$ of G, such that for all $i \in [k]$: (1) if $s_i = \textit{lazy}$, then $p_{i+1} = p_i$; (2) if $s_i = \textit{non-lazy}$, then $p_{i+1} = \widetilde{f}(i, p_i)$.*

We also define reversed random walks, tailored to the idea of simulating pull operations using push operations. The basic idea of the reversed random walk is to view a push operation (or one step of a forward walk) as a pull operation (or one step of a reversed walk). However, there are several complications: (1) we let v "pull from u" only if u is the unique node pushing to v, since v is not allowed to pull from multiple nodes at the same time; (2) we need to use

auxiliary randomness $r_{i,u}$ to equalize the probabilities of successful pull operations made by different nodes[4]; (3) we want the pull operations to be pairwise independent. In particular, two nodes u and v pull from their common neighbor w at the same time with probability $1/\Delta^2$. To realize this, we combine two rounds into one round so that w can send two rumors, say to a and b at the same time. Also, note that there are two cases when w pushes to both u and v, or equivalently u and v both pull from w: $(a,b)=(u,v)$ or $(a,b)=(v,u)$. We admit only one of them, so that the event occurs with probability $1/\Delta^2$ rather than $2/\Delta^2$. The formal definition of the reversed random walks is as follows:

Definition 3.2 (Reversed random walks)**.** *Consider a random rumor spreading process in T rounds on a graph G using Protocol 1 determined by its own randomness $f\sim\mathcal{D}=\mathcal{U}$. Pick real numbers $r_{i,u}$ independently and uniformly from $[0,1]$ for all $i\in[T/2]$ and $u\in V[G]$.*

Fix an arbitrary total order $\preceq$ on $V[G]$. For $i\in[T/2]$ and $u\in V[G]$, define set $N_{i,u}$ by

$$N_{i,u}=\left\{\widetilde{f}(T-1-2i,u),\widetilde{f}(T-2-2i,u)\right\}$$

if $\widetilde{f}(T-1-2i,u)\preceq\widetilde{f}(T-2-2i,u)$, and $N_{i,u}=\emptyset$ otherwise. Moreover, let

$$N^{\vee}_{i,u}=\{v\in V[G]:v\neq u \text{ and } u\in N_{i,v}\}.$$

A reversed random walk of length $k\in[T/2]$ with pattern $S=(s_0,\ldots,s_{k-1})\in\mathcal{C}_k$ is a sequence of $k+1$ nodes $(p_0,\ldots,p_k)$ of G, such that for all $i\in[k]$ the following properties hold: (1) if $s_i=$ lazy, then $p_{i+1}=p_i$; (2) if $s_i=$ non-lazy, then $p_{i+1}=u$ if $N^{\vee}_{i,p_i}=\{u\}$ is a singleton and $r_{i,p_i}\leqslant(1-1/\Delta)^{\Delta-\deg(u)}$, and otherwise $p_{i+1}=p_i$.

Now we use forward and reversed random walks to analyze the rumor spreading process. For $k\in[T/2]$, $u,v\in V[G]$ and $S\in\mathcal{C}_k=\{\text{lazy},\text{non-lazy}\}^k$, let $X^S_{u,v}$ (resp. $Y^S_{u,v}$) be the indicator random variable of the event that the unique forward (resp. reversed) walk with pattern S and initial node u is at node v at the kth step. We use the following distributions in our analysis:

- Let $\mathcal{D}_{\gamma,k}$ be the distribution over $\mathcal{C}_k$ where entries are independently chosen to be lazy with probability $1-\gamma$.
- Let $r=(f,\{r_{i,u}\})$ be the whole randomness used in forward random walks and reversed random walks. Let $\widetilde{\mathcal{D}}$ be the distribution of r, which is the product of $\mathcal{D}$ with copies of the uniform distributions over $[0,1]$.

[4]The auxiliary randomness only appears in the analysis.

We fix an arbitrary node $w\in V[G]$, and study the probability that node w is informed in T rounds. Clearly, if there exist a forward random walk p from s to some node u and a reversed random walk p' from w to u, then the rumor is sent from s to u following p and then from u to w following the reversal of p'. Also, these two walks exist if and only if $X^S_{s,u}Y^{S'}_{w,u}>0$ for some S, S' and u. Therefore it holds for any $k\in[T/4]$ that [5]

$$\begin{aligned}&\mathbf{Pr}_{f\sim\mathcal{D}}\left[\,w\text{ receives the rumor in }T\text{ rounds}\,\right]\\ &\geqslant\mathbf{Pr}_{r\sim\widetilde{\mathcal{D}}}\left[\sum_{S,S'\in\mathcal{C}_k,u\in V[G]}X^S_{s,u}Y^{S'}_{w,u}>0\right].\end{aligned}\tag{3.1}$$

Now we reduce the global event $\sum_{S,S'\in\mathcal{C}_k,u\in V[G]}X^S_{s,u}Y^{S'}_{w,u}>0$ to local events $X^S_{s,u}$ and $Y^{S'}_{w,u}$. By Cauchy-Schwarz inequality and linearity of expectation, we prove that (3.1) is lower bounded by

$$\frac{\sum_{u,v}\mathbf{E}_{r,S}\left[X^S_{s,u}\right]\mathbf{E}_{r,S}\left[X^S_{s,v}\right]\mathbf{E}_{r,S}\left[Y^S_{w,u}\right]\mathbf{E}_{r,S}\left[Y^S_{w,v}\right]}{\sum_{u,v}\mathbf{E}_{r,S,S'}\left[X^S_{s,u}X^{S'}_{s,v}\right]\mathbf{E}_{r,S,S'}\left[Y^S_{w,u}Y^{S'}_{w,v}\right]}\tag{3.2}$$

where the subscripts r, S and S' are independent with distributions $\widetilde{\mathcal{D}}$, $\mathcal{D}_{\gamma,k}$ and $\mathcal{D}_{\gamma,k}$ respectively. Hence the rumor spreading time of Protocol 1 can be derived by analyzing multiple random walks individually or pairwisely. This presents a new approach to analyze the rumor spreading time, where informed nodes can choose their neighbors according to a general distribution.

Now we study the expectations in (3.2) in terms of finite-state Markov chains. For simplicity, we represent these Markov chains by stochastic matrices. Recall that a stochastic matrix $\mathbf{M}''\in\mathbb{R}^{n\times n}\otimes\mathbb{R}^{n\times n}$ is a *coupling* of $\mathbf{M},\mathbf{M}'\in\mathbb{R}^{n\times n}$ if (1) $\sum_{x\in[n]}\mathbf{M}''_{(u,w)(v,x)}=\mathbf{M}_{u,v}$ for any $u,w,v\in[n]$, and (2) $\sum_{v\in[n]}\mathbf{M}''_{(u,w)(v,x)}=\mathbf{M}'_{w,x}$ for any $u,w,x\in[n]$.

We define the "bi-lazy" analogue of lazy Markov chains with respect to a coupling, where the two chains choose to be lazy or non-lazy independently.

Definition 3.3. *For $\gamma\in[0,1]$, let*

$$\mathcal{L}_\gamma(\mathbf{M})\triangleq(1-\gamma)\mathbf{I}+\gamma\mathbf{M}$$

be the lazy Markov chain.

Definition 3.4 (Lazy coupling)**.** *Let $\mathbf{M}''$ be a coupling of $\mathbf{M},\mathbf{M}'\in\mathbb{R}^{n\times n}$. For $\gamma,\gamma'\in[0,1]$, define*

$$\begin{aligned}\mathcal{L}_{\gamma,\gamma'}(\mathbf{M}'')\triangleq&(1-\gamma)(1-\gamma')(\mathbf{I}\otimes\mathbf{I})+(1-\gamma)\gamma'(\mathbf{I}\otimes\mathbf{M}')\\&+\gamma(1-\gamma')(\mathbf{M}\otimes\mathbf{I})+\gamma\gamma'\mathbf{M}''.\end{aligned}$$

[5] We let $k\in[T/4]$ rather than $[T/2]$ as for technical reasons, since we have to define reversed walks in the way that each step takes two rounds instead of one.

Note that $\mathcal{L}_{\gamma,\gamma'}(\mathbf{M}'')$ *is a coupling of* $\mathcal{L}_\gamma(\mathbf{M})$ *and* $\mathcal{L}_{\gamma'}(\mathbf{M}')$.

Definition 3.5 (Doeblin coupling, [30])**.** *Let* $\mathbf{M} \in \mathbb{R}^{n\times n}$ *be a stochastic matrix. The Doeblin coupling* $\mathcal{Q}(\mathbf{M})$ *of two copies of* $\mathbf{M}$ *is defined as*

$$\mathcal{Q}(\mathbf{M})_{(u,w)(v,x)} \triangleq \begin{cases} (\mathbf{M}\otimes\mathbf{M})_{(u,w)(v,x)} & u \neq w, \\ \mathbf{M}_{u,v} & u = w, v = x, \\ 0 & u = w, v \neq x. \end{cases}$$

Using the above definitions, we are able to characterize the expectations in (3.2) in terms of Markov chains. For instance, the first and second moments $\mathbf{E}_{r,S}\left[X^S_{u,v}\right]$ and $\mathbf{E}_{r,S,S'}\left[X^S_{u,v}X^{S'}_{w,x}\right]$ about forward random walks are characterized by the chains $\mathcal{L}_\gamma\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$ and $\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$ respectively, and similar results hold for reversed random walks.

Lemma 3.6. *Let* r, S *and* S' *be independent with distributions* $\widetilde{\mathcal{D}}$ *(induced by* $\mathcal{D} = \mathcal{U}$*),* $\mathcal{D}_{\gamma,k}$ *and* $\mathcal{D}_{\gamma,k}$ *respectively. Then for stochastic matrices* $\mathbf{M}_1 = \mathcal{L}_\gamma\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$, $\mathbf{M}_2 = \mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$, $\mathbf{M}_3 = \mathcal{L}_\gamma\circ\mathcal{L}_{\gamma'}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$, $\mathbf{M}_4 = \mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}\circ\mathcal{L}_{\gamma'}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$, $\gamma' \triangleq (1-1/\Delta)^{\Delta-1}$, *and any* $u,v,w,x \in V[G]$, *the following statements hold:*

1. $\mathbf{E}_{r,S}\left[X^S_{u,v}\right] = \left\langle \mathbf{e}_u\mathbf{M}_1^k, \mathbf{e}_v\right\rangle$,
2. $\mathbf{E}_{r,S,S'}\left[X^S_{u,v}X^{S'}_{w,x}\right] = \left\langle \mathbf{e}_{(u,w)}\mathbf{M}_2^k, \mathbf{e}_{(v,x)}\right\rangle$,
3. $\mathbf{E}_{r,S}\left[Y^S_{u,v}\right] = \left\langle \mathbf{e}_u\mathbf{M}_3^k, \mathbf{e}_v\right\rangle$, *and*
4. $\mathbf{E}_{r,S,S'}\left[Y^S_{u,v}Y^{S'}_{w,x}\right] = \left\langle \mathbf{e}_{(u,w)}\mathbf{M}_4^k, \mathbf{e}_{(v,x)}\right\rangle$.

Notice that matrix $\mathcal{Q}(\mathbf{M})$ agrees with $\mathbf{M}\otimes\mathbf{M}$ except on the rows indexed by (u,u), $u \in V[G]$. This is a manifestation of the fact that the "non-lazy" steps from the same node made by two different forward/reversed random walks are not independent, i.e., every informed node can only send the rumor to one neighbor in each round. Despite this complication, we show that $\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})$ is actually quite close to $\mathcal{L}_{\gamma,\gamma}(\mathbf{M}\otimes\mathbf{M}) = \mathcal{L}_\gamma(\mathbf{M})\otimes\mathcal{L}_\gamma(\mathbf{M})$:

Lemma 3.7. *Suppose that* $\mathbf{M} \in \mathbb{R}^{n\times n}$ *is a doubly-stochastic matrix with spectral gap* $\alpha > 0$*, and suppose* $\mathbf{M}_{u,v} \leqslant \eta$ *for any distinct* $u,v \in V[G]$. *Then for any distribution* $\mathbf{u}$ *over* $V[G]\times V[G]$, $k \in \mathbb{N}$, *and* $0 \leqslant \gamma \leqslant \min\left\{1/3, \alpha\eta^{-1/2}/9\right\}$, *we have*

$$\left\|\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})\right)^k - \boldsymbol{\pi}\otimes\boldsymbol{\pi}\right\|_2 \leqslant (1-\gamma\alpha/2)^k + 2\sqrt{2}\gamma\alpha^{-1}n^{-3/2},$$

where $\boldsymbol{\pi}$ *denotes the uniform distribution over* $V[G]$.

Corollary 3.8. *Let* $\mathbf{M}$, γ *and* α *be as in Lemma 3.7. Let* $\boldsymbol{\pi}'$ *be the stationary distribution*[6] *of* $\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})$. *Then*

$$\|\boldsymbol{\pi}' - \boldsymbol{\pi}\otimes\boldsymbol{\pi}\|_2 \leqslant (1-\gamma\alpha/2)^k + 2\sqrt{2}\gamma\alpha^{-1}n^{-3/2}.$$

Define the ℓ_1*-mixing time*

$$\bar{\tau}(\varepsilon) \triangleq \max_{\mathbf{u}} \min\{k : \|\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})\right)^k - \boldsymbol{\pi}\otimes\boldsymbol{\pi}\|_1 \leqslant \varepsilon\},$$

where $\mathbf{u}$ *ranges over all distributions over* $V[G]\times V[G]$. *Assuming* $\gamma\alpha^{-1} = O(n^{1/2-c})$ *for some constant* $c > 0$, *we have* $\bar{\tau}(\varepsilon) = O(\gamma^{-1}\alpha^{-1}(\log n + \log\varepsilon^{-1}))$.

We know that the stationary distribution of $\mathcal{Q}(\mathbf{M})$ is the uniform distribution over the set of diagonal entries $\{(u,u) : u \in V[G]\}$. So is the stationary distribution of the lazy chain $\mathcal{L}_\gamma\circ\mathcal{Q}(\mathbf{M})$ for any $\gamma \in (0,1]$. Interestingly, Corollary 3.8 tells us that the "bi-lazy" chain $\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})$ behaves *very differently*, as its stationary distribution is close to $\boldsymbol{\pi}\otimes\boldsymbol{\pi}$ instead.

Using the rapid mixing of $\mathcal{L}_\gamma\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$ and $\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})$ (and similar chains for reversed random walks), we have the following upper bound of the rumor spreading time of Protocol 1:

Theorem 3.9. *Let* G *be a graph with* n *nodes, spectral expansion* $\alpha \in (0,1)$, *and irregularity* β. *Protocol 1 with the distribution* $\mathcal{D} = \mathcal{U}$ *informs all nodes in* $T = O(C\log n)$ *rounds with high probability, where* $C = (1/\alpha)\cdot\beta^2\cdot\max\{1, 1/(\alpha\cdot\Delta^{0.499})\}$.

We remark that the analysis above provides a fundamentally new approach to analyze the rumor spreading process, and the rumor spreading time is tight for certain graph families. For instance, for any expander graph with n nodes and $\beta = O(1)$, Protocol 1 informs all nodes in $O(\log n)$ rounds with high probability, which is known to be tight.

3.2 Proof of Lemma 3.7

In this subsection, we will show that $\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})$ behaves similarly as $\mathcal{L}_\gamma(\mathbf{M})\otimes\mathcal{L}_\gamma(\mathbf{M})$, in the sense that it almost preserves the vector $\boldsymbol{\pi}\otimes\boldsymbol{\pi}$ and shrinks vectors orthogonal to $\boldsymbol{\pi}\otimes\boldsymbol{\pi}$. For a distribution $\mathbf{u}$ over $V[G]\times V[G]$, we have the decomposition $\mathbf{u} = \boldsymbol{\pi}\otimes\boldsymbol{\pi} + \mathbf{u}^\perp$, where $\mathbf{u}^\perp \triangleq \mathbf{u} - \boldsymbol{\pi}\otimes\boldsymbol{\pi}$ is orthogonal to $\boldsymbol{\pi}\otimes\boldsymbol{\pi}$. The following two results will be used in our proof.

Lemma 3.10. *Let* $\mathbf{M}$, $\boldsymbol{\pi}$ *and* γ *be as in Lemma 3.7. Then*

$$\left\|((\boldsymbol{\pi}\otimes\boldsymbol{\pi})\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M}))^\perp\right\|_2 \leqslant \sqrt{2}\gamma^2 n^{-3/2}.$$

[6] The laziness and $\alpha > 0$ guarantee that $\mathcal{L}_{\gamma,\gamma}\circ\mathcal{Q}(\mathbf{M})$ is ergodic and has a unique stationary distribution.

Lemma 3.11. *Let $\mathbf{M}$, $\boldsymbol{\pi}$ and γ be as in Lemma 3.7. For any vector $\mathbf{u} \in \mathbb{R}^n \otimes \mathbb{R}^n$ orthogonal to $\boldsymbol{\pi} \otimes \boldsymbol{\pi}$, we have $\mathbf{u}\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M}) \perp \boldsymbol{\pi} \otimes \boldsymbol{\pi}$ and*

$$\|\mathbf{u}\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\|_2 \leqslant \left(1 - (1-\gamma)\gamma\alpha + \gamma^2\sqrt{2\eta}\right)\|\mathbf{u}\|_2.$$

Proof of Lemma 3.7. Note that we are bounding the ℓ_2-norm of

$$\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^k - \boldsymbol{\pi} \otimes \boldsymbol{\pi} = \left(\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^k\right)^{\perp}.$$

The proof is based on the induction on k. When $k = 0$, we have

$$\left\|\left(\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^k\right)^{\perp}\right\|_2 \leqslant \left\|\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^k\right\|_2 \leqslant 1,$$

and hence the claim holds. For $k > 0$, assume the claim holds for $k' < k$. Let $\mathbf{v} = \mathbf{u}\left(\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{k-1}$. We have

$$\begin{aligned}
&\left(\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^k\right)^{\perp} \\
&= \left(\mathbf{v}\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{\perp} \\
&= \left((\boldsymbol{\pi} \otimes \boldsymbol{\pi})\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{\perp} + \left(\mathbf{v}^{\perp}\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{\perp}.
\end{aligned}$$

By Lemma 3.10, we have

$$\left\|\left((\boldsymbol{\pi} \otimes \boldsymbol{\pi})\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{\perp}\right\|_2 \leqslant \sqrt{2}\gamma^2 n^{-3/2}.$$

By Lemma 3.11, we have

$$\left(\mathbf{v}^{\perp}\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{\perp} = \mathbf{v}^{\perp}\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M}),$$

whose ℓ_2-norm is at most

$$\left(1 - (1-\gamma)\gamma\alpha + \gamma^2\sqrt{2\eta}\right)\|\mathbf{v}\|_2 \leqslant (1 - \gamma\alpha/2)\|\mathbf{v}\|_2,$$

where we use the condition $\gamma \leqslant \{1/3, \alpha\eta^{-1/2}/9\}$. This is bounded by

$$\begin{aligned}
&(1-\gamma\alpha/2)\left((1-\gamma\alpha/2)^{k-1} + 2\sqrt{2}\gamma\alpha^{-1}n^{-3/2}\right) \\
&= (1-\gamma\alpha/2)^k + 2\sqrt{2}\gamma\alpha^{-1}n^{-3/2}(1-\gamma\alpha/2),
\end{aligned}$$

by the induction hypothesis. Then,

$$\begin{aligned}
&\left\|\left(\mathbf{u}\left(\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^k\right)^{\perp}\right\|_2 \\
&\leqslant \left\|\left((\boldsymbol{\pi} \otimes \boldsymbol{\pi})\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{\perp}\right\|_2 + \left\|\left(\mathbf{v}^{\perp}\mathcal{L}_{\gamma,\gamma} \circ \mathcal{Q}(\mathbf{M})\right)^{\perp}\right\|_2 \\
&\leqslant \sqrt{2}\gamma^2 n^{-3/2} + (1-\gamma\alpha/2)^k + \\
&\qquad 2\sqrt{2}\gamma\alpha^{-1}n^{-3/2}(1-\gamma\alpha/2) \\
&= (1-\gamma\alpha/2)^k + 2\sqrt{2}\gamma\alpha^{-1}n^{-3/2},
\end{aligned}$$

as desired. □

3.3 Proof of Theorem 3.9 We are now ready to derive a bound on the runtime of Protocol 1. It suffices to prove the following lemma:

Lemma 3.12. *Let G be a graph with n nodes, spectral gap $\alpha \in (0,1)$, and irregularity β. Using Protocol 1 with distribution $\mathcal{D} = \mathcal{U}$, any node gets the rumor in $T = O(C\log n)$ rounds with probability at least $1 - O(n^{-2c})$, where $C = (1/\alpha)\cdot\beta^2\max\{1, 1/(\alpha\cdot\Delta^{0.5-c})\}$ and $c > 0$ is an arbitrary small constant.*

We first define the ℓ_2-mixing time $\tau_{\mathbf{M}}(\varepsilon)$, and present an upper bound of $\tau_{\mathbf{M}}(\varepsilon)$, which will be used in our proof. For an ergodic Markov chain represented by the stochastic matrix $\mathbf{M}$ and $\varepsilon > 0$, define its ℓ_2-mixing time as

$$\tau_{\mathbf{M}}(\varepsilon) = \max_{\mathbf{u}} \min\{k : \|\mathbf{u}\mathbf{M}^k - \boldsymbol{\pi}\|_2 \leqslant \varepsilon\},$$

where $\boldsymbol{\pi}$ is the stationary distribution of $\mathbf{M}$ and $\mathbf{u}$ ranges over all distributions over the state set of the chain.

Lemma 3.13 ([33])**.** *Suppose that $\mathbf{M} \in \mathbb{R}^{n\times n}$ represents a reversible Markov chain with absolute spectral gap $\alpha > 0$. Then it holds that $\tau_{\mathbf{M}}(\varepsilon) < \log_{1-\alpha}\varepsilon + 1$.*

Proof of Lemma 3.12. Let $s \in V[G]$ be the initial node, and fix a target node $w \in V[G]$. Let $c > 0$ be any constant. Choose

$$\gamma = \min\left\{1/3, \Delta^{0.5-c}\alpha/9\right\} \leqslant n^{0.5-c}\alpha/9.$$

Choose $k = (\gamma\gamma'\alpha)^{-1}\beta^2\log n + 1$, and let $T = 4k$. So $T = O(C\log n)$. Define the distributions $\mathbf{u} = \mathbf{e}_s\mathbf{M}_1^k$, $\mathbf{v} = \mathbf{e}_{(s,s)}\mathbf{M}_2^k$, $\mathbf{u}' = \mathbf{e}_w\mathbf{M}_3^k$, and $\mathbf{v}' = \mathbf{e}_{(w,w)}\mathbf{M}_4^k$, where $\mathbf{M}_1, \ldots, \mathbf{M}_4$ are as in Lemma 3.6. Let $\boldsymbol{\pi}$ be the uniform distribution over $V[G]$. As before, let $\mathbf{u}^{\perp} = \mathbf{u} - \boldsymbol{\pi}$ and $\mathbf{v}^{\perp} = \mathbf{v} - \boldsymbol{\pi}\otimes\boldsymbol{\pi}$, and similarly for $\mathbf{u}'$ and $\mathbf{v}'$. By (3.2) and Lemma 3.6, the probability that w gets the rumor in k rounds is lower bounded by

$$\begin{aligned}
&\frac{\sum_{u,v\in V[G]}\langle\mathbf{u},\mathbf{e}_u\rangle\langle\mathbf{u},\mathbf{e}_v\rangle\langle\mathbf{u}',\mathbf{e}_u\rangle\langle\mathbf{u}',\mathbf{e}_v\rangle}{\sum_{u,v\in V[G]}\langle\mathbf{v},\mathbf{e}_{(u,v)}\rangle\langle\mathbf{v}',\mathbf{e}_{(u,v)}\rangle} \\
&= \frac{\langle\mathbf{u},\mathbf{u}'\rangle^2}{\langle\mathbf{v},\mathbf{v}'\rangle} = \frac{\left(1/n + \langle\mathbf{u}^{\perp},\mathbf{u}'^{\perp}\rangle\right)^2}{1/n^2 + \langle\mathbf{v}^{\perp},\mathbf{v}'^{\perp}\rangle}.
\end{aligned} \tag{3.3}$$

Note that $\mathbf{M}_1 = \mathcal{L}_{\gamma}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$ and $\mathbf{M}_3 = \left(\mathcal{L}_{\gamma}\circ\mathcal{L}_{\gamma'}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)\right)$ have absolute spectral gaps $\gamma\alpha\beta^{-2}$ and $\gamma\gamma'\alpha\beta^{-2}$, respectively. This follows from the definition of lazy Markov chains (Also, the laziness guarantees that the eigenvalues are all non-negative, and hence the bounds are about absolute spectral gaps, not just spectral gaps). By Lemma 3.13 and the fact that

$$k \geqslant (\gamma\gamma'\alpha)^{-1}\beta^2\log n + 1 \geqslant \log_{1-\gamma\gamma'\alpha\beta^{-2}}(1/n) + 1,$$

we have $|\langle \mathbf{u}^{\perp}, \mathbf{u}'^{\perp}\rangle| \leqslant \left\|\mathbf{u}^{\perp}\right\|_2 \left\|\mathbf{u}'^{\perp}\right\|_2 \leqslant 1/n^2$. By Lemma 3.7 (with $\eta = 1/\Delta$), we have

$$|\langle \mathbf{v}^{\perp}, \mathbf{v}'^{\perp}\rangle| \leqslant \left\|\mathbf{v}^{\perp}\right\|_2 \left\|\mathbf{v}'^{\perp}\right\|_2 \leqslant \left((1-\gamma\alpha/2)^k + 2\sqrt{2}\gamma\alpha^{-1}n^{-3/2}\right)^2 \leqslant 1/n^{2+2c}.$$

So (3.3) is lower bounded by

$$\frac{(1/n - 1/n^2)^2}{1/n^2 - 1/n^{2+2c}} = 1 - O(n^{-2c}). \qquad \square$$

Theorem 3.9 is obtained by repeating the protocol $O(1)$ times and applying the union bound.

3.4 A Randomness-Efficient Protocol The discussion above relates the rumor spreading process to multiple random walks. The transitions of these random walks from a fixed node only depend on local information and are characterized by combinatorial rectangles. This memoryless property of random walks/Markov chains allows us to compute them in logspace, or branching programs with polynomial width. Using PRGs for combinatorial rectangles and those for branching programs, we obtain the distribution which is samplable with a short random seed and has almost the same performance as $\mathcal{D} = \mathcal{U}$ in Protocol 1. This gives Protocol 2, which corresponds to Theorem 1.1.

Protocol 2. *Pick the following objects: (1) an explicit ε-PRG $\mathcal{G} = (\mathcal{G}_0, \ldots, \mathcal{G}_{n-1}) : \{0,1\}^{\ell} \mapsto [m]^n$ for $\mathsf{CR}_{[m]^n}$ with seed length ℓ, and (2) an explicit ε'-PRG $\mathcal{G}' = (\mathcal{G}'_0, \ldots, \mathcal{G}'_{T/2-1}) : \{0,1\}^{\ell'} \mapsto \left(\{0,1\}^{\ell}\right)^{T/2}$ for $(T/2, n^2, 2^{\ell})$-branching programs with seed length ℓ', where $\varepsilon^{-1}, \varepsilon'^{-1}, m = n^{\Theta(1)}$ are sufficiently large.*

The initial node having the rumor independently chooses random strings $x, y \in \{0,1\}^{\ell'}$. These random strings are appended with the rumor and sent to other nodes. (1) In the ith round for $0 \leqslant i < T/2$, an informed node u sends the rumor to the neighbor with index $\mathcal{G}_u(\mathcal{G}'_i(x)) \bmod \Delta$ in its adjacency list. (2) In the ith round for $T/2 \leqslant i < T$, let $j = \lfloor \frac{T-i-1}{2} \rfloor$. For $u \in V[G]$, let $(r_0, r_1) = \mathcal{G}_u(\mathcal{G}'_j(y)) \bmod \Delta^2 \in [\Delta]^2$. Then u sends the rumor to the r_0th neighbor if $i = T-1-2j$, or to the r_1th neighbor if $i = T-2-2j$.

Setting

$$C = (1/\alpha) \cdot \beta^2 \cdot \max\left\{1, 1/(\alpha \cdot \Delta^{0.499})\right\},$$

Protocol 2 uses $2\ell'$ random bits, and with high probability informs all nodes in $T = O(C \log n)$ rounds. The main technique behind designing Protocol 2 is a reduction from PRGs for branching programs to rumor spreading protocols. This reduction implies the following results: First, combining the reduction with known explicit constructions of PRGs (Theorem 2.3, Theorem 2.6), we obtain Theorem 1.1. Second, we show that the existence of an explicit ε-PRG for $\left(\max\{T/2, n\}, n^2, n^{\Theta(1)}\right)$-branching programs with seed length $O(\log n)$ and $\varepsilon^{-1} = n^{\Theta(1)}$ implies an explicit rumor spreading protocol which uses $O(\log n)$ random bits, and with high probability informs all nodes in $T = O(C \log n)$ rounds (Theorem 3.17). That is, explicit constructions of PRGs for branching programs with the *optimal parameters* imply explicit constructions of *optimal* rumor spreading protocols for expander graphs with respect to both the rumor spreading time and the randomness complexity. This suggests that there may exist some further and interesting connections between gossip processes and branching programs.

3.5 Analysis of Protocol 2 In this subsection, we analyze Protocol 2, and prove Theorem 1.1. Let $\mathcal{P}$ be the distribution over the set of functions $f : [T] \times V[G] \mapsto [\Delta]$ associated with Protocol 2. The values $f(i,u)$ in the ith round are generated using the PRG $\mathcal{G}$, and the seeds of $\mathcal{G}$ in different rounds are generated by the PRG $\mathcal{G}'$. We will show that Protocol 1 with distribution $\mathcal{D} = \mathcal{P}$ has almost the same performance as the one with $\mathcal{D} = \mathcal{U}$. As an intermediate step, we consider the distribution $\mathcal{P}'$ defined as follows: the values of f in each round are determined by the PRG $\mathcal{G}$ in the same way as for $\mathcal{P}$ but the seeds of $\mathcal{G}$ in different rounds are now independent and random, instead of being generated by $\mathcal{G}'$. With $\mathcal{D} = \mathcal{P}'$, Definition 3.1 and Definition 3.2 about random walks are still valid, and the lower bound of (3.2) still holds by exactly the same proof. Moreover, Lemma 3.6 "almost holds" in the following sense.

Lemma 3.14. *Let r, S and S' be independent with distributions $\widetilde{\mathcal{D}}$ (induced by $\mathcal{D} = \mathcal{P}'$), $\mathcal{D}_{\gamma,k}$ and $\mathcal{D}_{\gamma,k}$, respectively. Then there exist stochastic matrices $\mathbf{M}'_1, \mathbf{M}'_3 \in \mathbb{R}^{n\times n}$, $\mathbf{M}'_2, \mathbf{M}'_4 \in \mathbb{R}^{n\times n} \otimes \mathbb{R}^{n\times n}$ such that $\|\mathbf{M}'_i - \mathbf{M}_i\|_1 \leqslant 12\gamma\Delta^2(\varepsilon + 2\Delta^3/m)$ for $1 \leqslant i \leqslant 4$, where $\mathbf{M}_i$ are as in Lemma 3.6 and ε, m are as in Protocol 2. Moreover, for any $u, v, w, x \in V[G]$, the following statements hold:*

1. $\mathbf{E}_{r,S}\left[X^S_{u,v}\right] = \left\langle \mathbf{e}_u \mathbf{M}'^k_1, \mathbf{e}_v \right\rangle$,
2. $\mathbf{E}_{r,S,S'}\left[X^S_{u,v} X^{S'}_{w,x}\right] = \left\langle \mathbf{e}_{(u,w)} \mathbf{M}'^k_2, \mathbf{e}_{(v,x)} \right\rangle$,
3. $\mathbf{E}_{r,S}\left[Y^S_{u,v}\right] = \left\langle \mathbf{e}_u \mathbf{M}'^k_3, \mathbf{e}_v \right\rangle$,
4. $\mathbf{E}_{r,S,S'}\left[Y^S_{u,v} Y^{S'}_{w,x}\right] = \left\langle \mathbf{e}_{(u,w)} \mathbf{M}'^k_4, \mathbf{e}_{(v,x)} \right\rangle$.

Next we consider the case $\mathcal{D} = \mathcal{P}$. By the same proof, we can show that (3.1) is lowered bounded by (3.2). Furthermore we show that the expectations are almost the same as in $\mathcal{D} = \mathcal{P}'$, since they can be computed by small-width branching programs:

Lemma 3.15. *For any $u, w \in V[G]$, the quantities*

$$\sum_{v \in V[G]} \left| \mathbf{E}_{r\sim\widetilde{\mathcal{P}}',S}\left[X^S_{u,v}\right] - \mathbf{E}_{r\sim\widetilde{\mathcal{P}},S}\left[X^S_{u,v}\right]\right| \tag{3.4}$$

and

$$\sum_{v,x \in V[G]} \left| \mathbf{E}_{r\sim\widetilde{\mathcal{P}}',S,S'}\left[X^S_{u,v}X^{S'}_{w,x}\right] - \mathbf{E}_{r\sim\widetilde{\mathcal{P}},S,S'}\left[X^S_{u,v}X^{S'}_{w,x}\right]\right| \tag{3.5}$$

are bounded by ε', where $\widetilde{\mathcal{P}}$ (resp. $\widetilde{\mathcal{P}}'$) is the distribution of r induced by $\mathcal{P}$ (resp. $\mathcal{P}'$), S, S' in the subscripts are independent and have distribution $\mathcal{D}_{\gamma,k}$, and ε' is as in Protocol 2. The same statement holds with $X^S_{u,v}$ and $X^S_{w,x}$ replaced by $Y^S_{u,v}$ and $Y^S_{w,x}$, respectively.

Now we are ready to prove a derandomized version of Lemma 3.12.

Theorem 3.16. *Let G be a graph with spectral gap α and irregularity β. Using Protocol 1 with distribution $\mathcal{D} = \mathcal{P}$, any node gets the rumor in $T = O(C\log n)$ rounds with probability at least $1 - n^{-2c}$, where $C = (1/\alpha)\cdot\beta^2\max\{1, 1/(\alpha\cdot\Delta^{0.5-c})\}$ and $c > 0$ is an arbitrary small constant.*

Proof. Let $s \in V[G]$ be the initial node, and fix a target node $w \in V[G]$. Let $c, \gamma, k, T, \boldsymbol{\pi}, \mathbf{u}, \mathbf{u}', \mathbf{v}, \mathbf{v}'$ be as in the proof of Lemma 3.12 and $T = O(C\log n)$. Define $\bar{\mathbf{u}} = \mathbf{e}_s\mathbf{M}_1'^k$, $\bar{\mathbf{v}} = \mathbf{e}_{(s,s)}\mathbf{M}_2'^k$, $\bar{\mathbf{u}}' = \mathbf{e}_w\mathbf{M}_3'^k$, and $\bar{\mathbf{v}}' = \mathbf{e}_{(w,w)}\mathbf{M}_4'^k$, where $\mathbf{M}_1', \ldots, \mathbf{M}_4'$ are as in Lemma 3.14. Then

$$\begin{aligned}\|\bar{\mathbf{u}} - \mathbf{u}\|_1 = \left\|\mathbf{e}_s\left(\mathbf{M}_1'^k - \mathbf{M}_1^k\right)\right\|_1 &\leqslant \left\|\mathbf{M}_1'^k - \mathbf{M}_1^k\right\|_1 \\ &\leqslant k\left\|\mathbf{M}_1' - \mathbf{M}_1\right\|_1 \leqslant k\varepsilon_0,\end{aligned}$$

where $\varepsilon_0 = 12\gamma\Delta^2(\varepsilon + 2\Delta^3/m)$ (cf. Lemma 3.14). Here the second inequality holds by a simple induction on k. Similarly $\|\bar{\mathbf{u}}' - \mathbf{u}'\|_1, \|\bar{\mathbf{v}} - \mathbf{v}\|_1, \|\bar{\mathbf{v}}' - \mathbf{v}'\|_1 \leqslant k\varepsilon_0$. Define $\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}' \in \mathbb{R}^n$ and $\widetilde{\mathbf{v}}, \widetilde{\mathbf{v}}' \in \mathbb{R}^n \otimes \mathbb{R}^n$ such that $\widetilde{\mathbf{u}}_u = \mathbf{E}_{r,S}\left[X^S_{s,u}\right]$, $\widetilde{\mathbf{u}}'_u = \mathbf{E}_{r,S}\left[Y^S_{w,u}\right]$, $\widetilde{\mathbf{v}}_{u,v} = \mathbf{E}_{r,S,S'}\left[X^S_{s,u}X^S_{s,v}\right]$ and $\widetilde{\mathbf{v}}'_{u,v} = \mathbf{E}_{r,S,S'}\left[Y^S_{w,u}Y^S_{w,v}\right]$, where r, S and S' are independent with distributions $\widetilde{\mathcal{P}}$ (induced by $\mathcal{P}$), $\mathcal{D}_{\gamma,k}$ and $\mathcal{D}_{\gamma,k}$, respectively. Then Lemma 3.14 and Lemma 3.15 altogether imply that $\|\widetilde{\mathbf{u}} - \bar{\mathbf{u}}\|_1 \leqslant \varepsilon'$ and hence $\|\widetilde{\mathbf{u}} - \mathbf{u}\|_1 \leqslant k\varepsilon_0 + \varepsilon'$. Obviously we have $\|\widetilde{\mathbf{u}} - \mathbf{u}\|_\infty \leqslant 1$. By Hölder's inequality, we have $\|\widetilde{\mathbf{u}} - \mathbf{u}\|_2 \leqslant \sqrt{k\varepsilon_0 + \varepsilon'}$. Similarly,

$$\|\widetilde{\mathbf{u}}' - \mathbf{u}'\|_2, \|\widetilde{\mathbf{v}} - \mathbf{v}\|_2, \|\widetilde{\mathbf{v}}' - \mathbf{v}'\|_2 \leqslant \sqrt{k\varepsilon_0 + \varepsilon'},$$

As shown in the proof of Lemma 3.12, we have $\left\|\mathbf{u}^\perp\right\|_2, \left\|\mathbf{u}'^\perp\right\|_2 \leqslant 1/n$, and $\left\|\mathbf{v}^\perp\right\|_2, \left\|\mathbf{v}'^\perp\right\|_2 \leqslant n^{-(1+c)}$. Note that

$$\widetilde{\mathbf{u}}^\perp = \widetilde{\mathbf{u}} - \boldsymbol{\pi} = (\widetilde{\mathbf{u}} - \mathbf{u}) + (\mathbf{u} - \boldsymbol{\pi}) = (\widetilde{\mathbf{u}} - \mathbf{u}) + \mathbf{u}^\perp.$$

So we have $\left\|\widetilde{\mathbf{u}}^\perp\right\|_2 \leqslant \sqrt{k\varepsilon_0 + \varepsilon'} + 1/n$ and similarly $\left\|\widetilde{\mathbf{u}}'^\perp\right\|_2 \leqslant \sqrt{k\varepsilon_0 + \varepsilon'} + 1/n$, and $\left\|\widetilde{\mathbf{v}}^\perp\right\|_2, \left\|\widetilde{\mathbf{v}}'^\perp\right\|_2 \leqslant \sqrt{k\varepsilon_0 + \varepsilon'} + n^{-(1+c)}$.

By (3.1) and (3.2), the probability that node w gets the rumor in k rounds is lower bounded by

$$\begin{aligned}&\frac{\sum_{u,v\in V[G]}\langle\widetilde{\mathbf{u}}, \mathbf{e}_u\rangle\langle\widetilde{\mathbf{u}}, \mathbf{e}_v\rangle\langle\widetilde{\mathbf{u}}', \mathbf{e}_u\rangle\langle\widetilde{\mathbf{u}}', \mathbf{e}_v\rangle}{\sum_{u,v\in V[G]}\left\langle\widetilde{\mathbf{v}}, \mathbf{e}_{(u,v)}\right\rangle\left\langle\widetilde{\mathbf{v}}', \mathbf{e}_{(u,v)}\right\rangle} \\ &= \frac{\langle\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}'\rangle^2}{\langle\widetilde{\mathbf{v}}, \widetilde{\mathbf{v}}'\rangle} = \frac{\left(\langle\boldsymbol{\pi}, \boldsymbol{\pi}\rangle + \left\langle\widetilde{\mathbf{u}}^\perp, \widetilde{\mathbf{u}}'^\perp\right\rangle\right)^2}{\langle\boldsymbol{\pi}\otimes\boldsymbol{\pi}, \boldsymbol{\pi}\otimes\boldsymbol{\pi}\rangle + \left\langle\widetilde{\mathbf{v}}^\perp, \widetilde{\mathbf{v}}'^\perp\right\rangle} \\ &= \frac{\left(1/n + \left\langle\widetilde{\mathbf{u}}^\perp, \widetilde{\mathbf{u}}'^\perp\right\rangle\right)^2}{1/n^2 + \left\langle\widetilde{\mathbf{v}}^\perp, \widetilde{\mathbf{v}}'^\perp\right\rangle}.\end{aligned} \tag{3.6}$$

We have

$$\left|\left\langle\widetilde{\mathbf{u}}^\perp, \widetilde{\mathbf{u}}'^\perp\right\rangle\right| \leqslant \left\|\widetilde{\mathbf{u}}^\perp\right\|_2\left\|\widetilde{\mathbf{u}}'^\perp\right\|_2 = O\left(k\varepsilon_0 + \varepsilon' + n^{-2}\right),$$
$$\left|\left\langle\widetilde{\mathbf{v}}^\perp, \widetilde{\mathbf{v}}'^\perp\right\rangle\right| \leqslant \left\|\widetilde{\mathbf{v}}^\perp\right\|_2\left\|\widetilde{\mathbf{v}}'^\perp\right\|_2 = O\left(k\varepsilon_0 + \varepsilon' + n^{-(2+2c)}\right).$$

So (3.6) is lower bounded by $1 - O(n^2(k\varepsilon_0 + \varepsilon') + n^{-2c})$, where $\varepsilon_0 = 12\gamma\Delta^2(\varepsilon + 2\Delta^3/m)$. The claim follows since we pick $\varepsilon^{-1}, \varepsilon'^{-1}, m = n^{\Theta(1)}$ sufficiently large in Protocol 2. □

By repeating the protocol $O(1)$ times and applying the union bound, we obtain the following result, which implies Theorem 1.1.

Theorem 3.17. *Given an explicit ε-PRG for $\mathsf{CR}_{[m]^n}$ with seed length ℓ and an explicit ε'-PRG for $(T/2, n^2, 2^\ell)$-branching programs with seed length ℓ', where $\varepsilon^{-1}, \varepsilon'^{-1}, m = n^{\Theta(1)}$ are sufficiently large, there exists an explicit protocol using $2\ell'$ random bits such that, with high probability all nodes get the rumor in $T = O(C\log n)$ rounds. In particular, given an explicit ε-PRG for (L, W, D)-branching programs with seed length $O(\log n)$ where $L = \max\{T/2, n\}$, $W = n^2$, and $D, \varepsilon^{-1} = n^{\Theta(1)}$ are sufficiently large, there exists an explicit protocol which uses $O(\log n)$ random bits, and with high probability informs all nodes in $T = O(C\log n)$ rounds.*[7]

[7] This follows from the simple observation that combinatorial rectangles in $[m]^n$ can be computed by $(n, 2, m)$-branching programs.

4 The Averaging Protocol

In this section, we assume that every node u knows the IDs of its neighbors and its index for each of its neighbors, and present one protocol with improved rumor spreading time. We remark that such condition can be guaranteed with $O(\Delta)$ preprocessing time. In comparison to Protocol 2, the result in this section is obtained by analyzing a more general gossip process, called the *averaging protocol*, which is closely related to other gossip processes, e.g. load balancing [35].

4.1 The Protocol

Protocol 3 (The Averaging Protocol)**.** *Let m be a prime power. Pick the following objects:*

- *an explicit pairwise independent generator $\mathcal{G} = (\mathcal{G}_0, \ldots, \mathcal{G}_{n-1}) : \{0,1\}^\ell \mapsto [m]^n$ with seed length ℓ, and*
- *an explicit ε-PRG $\mathcal{G}' = (\mathcal{G}'_0, \ldots, \mathcal{G}'_{T-1}) : \{0,1\}^{\ell'} \mapsto \left(\{0,1\}^\ell\right)^T$ for $(T, n^2, 2^\ell)$-branching programs with seed length ℓ'*

where $\varepsilon^{-1}, m = n^{\Theta(1)}$ are sufficiently large.

The initial node having the rumor independently chooses a random string $x \in \{0,1\}^{\ell'}$, which is appended with the rumor and sent to other nodes. Once one node gets the rumor, it gets the ID u. Let $y = (y_0, \ldots, y_{T-1})$ be the sequence of seeds generated by $\mathcal{G}'$, i.e., $y_i = \mathcal{G}'_i(x)$. For $i \in [T]$ and $u \in V[G]$, define $(w_{u,i}, z_{u,i}) = \mathcal{G}_u(y_i) \bmod 4\Delta \in [2\Delta] \times \{\text{active}, \text{inactive}\}$. We say u is active in the ith round if $z_{u,i}$ is active, and otherwise inactive. We say u selects *v if v is the $w_{u,i}$th neighbor of u. In the ith round, an informed node u sends the rumor to the unique neighbor v (if exist) if $\{u, v\}$ is a good pair, where we call $\{u, v\}$ is a good pair if either of the following two conditions is met: (1) u is active, v is inactive, and u is the unique node selecting v, or (2) the same holds with u and v swapped.*

In addition, each node maintains a value that is initially specified by the distribution $\mathbf{e}_s$, i.e., the value of node u is one if u is the initial node, and zero otherwise. When node u sends the rumor to node v, set both the values of u and v as the average of their original values.

Checking the conditions in Protocol 3 requires that u and v know their index in the lists of their neighbors as well as the IDs of the neighbors. One can deterministically use $O(\Delta)$ preprocessing time to guarantee this assumption. Then Condition (2) can be checked directly by u. For Condition (1), note that an active node u can send the rumor and the seed to its unique inactive neighbor v specified by $w_{i,u}$ and then v can check if the condition is satisfied, i.e., if u is the unique node selecting v [8].

In the following, we represent the values of nodes after k rounds by a vector $\mathbf{v}(k) \in \mathbb{R}^n$, and initially the values of nodes are represented by $\mathbf{v}(0)$. Define the *averaging time* $\tau_{\mathsf{avg}}(\kappa)$ of the protocol as the smallest $t \in \mathbb{N}$ such that $\mathbf{Pr}\left[\|\mathbf{v}(t)^\perp\|_2 < \kappa\right] > 1 - \kappa$ for any distribution $\mathbf{v}$, or ∞ if there is no such t.

The main result in this section is as follows:

Theorem 4.1. *For $\delta > 0$, assume $2\varepsilon < \kappa^2$ where ε is as in Protocol 3. Then Protocol 3 uses $2\ell'$ random bits, and $\tau_{\mathsf{avg}}(\kappa) = O((1/\alpha) \cdot \beta^2 \log(1/\kappa))$.*

Corollary 4.2. *Let G be a graph with spectral gap α and irregularity β. Then Protocol 3 uses 2ℓ random bits, and with high probability informs all nodes of G in $T = O((1/\alpha) \cdot \beta^2 \cdot \log n)$ rounds.*

Proof. Set $\kappa = 1/n$. By Theorem 4.1, we know $\|\mathbf{v}(T)^\perp\|_2 < 1/n$ for sufficiently large $T = O((1/\alpha) \cdot \beta^2 \cdot \log n)$. Then $\mathbf{v}(k)_u$ must be nonzero for all $u \in V[G]$, i.e., all nodes are informed in T rounds. □

As a consequence, we have the following reduction from rumor spreading protocols to PRGs for branching programs:

Corollary 4.3. *Assume each node knows its index in the lists of its neighbors as well as the IDs of its neighbors. Then the following statements hold:*

1. *Given an explicit ε-PRG for $(T/2, n^2, 2^\ell)$-branching programs with seed length ℓ', where $\varepsilon^{-1} = n^{\Theta(1)}$ and $\ell = O(\log n)$ are sufficiently large, there exists an explicit protocol using $2\ell'$ random bits, such that with high probability all nodes get the rumor in $T = O((1/\alpha) \cdot \beta^2 \log n)$ rounds.*
2. *In particular, given an explicit ε-PRG for $(T/2, n^2, \varepsilon)$-branching programs with seed length $O(\log n)$ where $\varepsilon^{-1} = n^{\Theta(1)}$ is sufficiently large, there exists an explicit protocol using $O(\log n)$ random bits, such that with high probability all nodes get the rumor in $T = O((1/\alpha) \cdot \beta^2 \log n)$ rounds.*

Combining the reduction above with known explicit constructions of PRGs (Theorem 2.6), we obtain Theorem 1.2.

In Theorem 4.1 we only consider initial values specified by $\mathbf{v}(0) = \mathbf{e}_s$. Before giving a formal proof of Theorem 4.1, we first remark that the runtime bound and the randomness complexity hold for more

[8] The uniqueness requirement in Condition (1) is necessary only for analyzing the associated averaging algorithm. For the sake of rumor spreading, dropping the requirement only makes the rumor spread faster.

general initial distributions. Assuming $\varepsilon/\kappa^2 = n^{-\Theta(1)}$ is sufficiently small, it is easy to establish an upper bound $O((1/\alpha)\cdot\beta^2(\log n+\log(1/\kappa)))$ on the averaging time regarding a general distribution $\mathbf{v}(0)$: first use $T = O((1/\alpha)\cdot\beta^2\log(1/\kappa))$ rounds to inform all the nodes with high probability. Then set the new initial values $\mathbf{v}'(0) = \mathbf{v}(T)$, and run the averaging protocol for another $O((1/\alpha)\cdot\beta^2(\log n+\log(1/\kappa)))$ rounds. The process with the initial value distribution $\mathbf{v}'(0)$ can be viewed as a convex combination of those with initial value distribution $\mathbf{e}_u$, $u\in V[G]$ (note that each node u is already informed). With high probability, for all initial value distributions $\mathbf{e}_u$, the values converge to the average up to ℓ_2-distance κ. So the same is true for $\mathbf{v}'(0)$.

4.2 Analysis of the Protocol Now we analyze Protocol 3 and prove Theorem 4.1. For $x\in\{0,1\}^\ell$, define the matrix $\mathbf{M}(x)$ such that

$$\mathbf{M}(x)_{u,v} = \begin{cases} 1/2 & u\neq v \text{ and } \{u,v\} \text{ is a good pair,}\\ 1/2 & u=v \text{ and } \{u,v'\} \text{ is a good pair for some } v',\\ 1 & u=v \text{ and } \{u,v'\} \text{ is not a good pair for any } v',\\ 0 & u\neq v \text{ and } \{u,v\} \text{ is not a good pair,}\end{cases}$$

where the set of good pairs is determined by the seed $y_i = x$ (c.f. Protocol 3, where the definition of good pairs is independent from the round number i). It is easy to check that $\mathbf{M}(x)$ is doubly stochastic, symmetric, and $\mathbf{M}(x)^2 = \mathbf{M}(x)$ for all $x\in\{0,1\}^\ell$. Moreover, it characterizes the averaging operations using the seed $y_i = x$.

Lemma 4.4. *For any $i\in[T]$, it holds that $\mathbf{v}(i+1) = \mathbf{v}(i)\mathbf{M}(y_i)$.*

Let $\mathbf{M} = \mathbf{E}_{x\in\{0,1\}^\ell}[\,\mathbf{M}(x)\,]$. Then $\mathbf{M}$ is doubly-stochastic. We have the following lemma:

Lemma 4.5. *It holds that $\mathbf{M}_{u,v}\geqslant c\mathcal{L}_{1/2}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)_{u,v}$, for some constant $c\in(0,1)$.*

Proof. Each edge $\{u,v\}$ with $u\neq v$ is a good pair if either of the two mutually exclusive conditions (cf. Protocol 3) is satisfied. The first one holds with probability at least

$$\mathbf{Pr}_{x\in\{0,1\}^\ell}[\,u \text{ is active and selects } v\,] - \sum_{\substack{u'\in N(v)\\ u'\neq u}} \mathbf{Pr}_{x\in\{0,1\}^\ell}[\,u \text{ is active and both } u,u' \text{ select } v\,]$$

taken over the seed $y_i = x$. As $\mathcal{G}$ is a pairwise independent generator, by Lemma 2.2 this probability is lower bounded by

$$\left(\frac{1}{4\Delta}-\frac{2}{m}\right) - \Delta\cdot\left(\frac{1}{4\Delta}\cdot\frac{1}{2\Delta}+\frac{2}{m}\right) \geqslant \frac{c}{2\Delta}$$

for some $c>0$ and $m=\Omega(\Delta^2)$. The case for the second condition is the same. So $\{u,v\}$ is a good pair with probability at least $\frac{c}{\Delta}$. Note that $\mathbf{M}(x)_{u,v} = 1/2$ whenever $\{u,v\}$ is a good pair. Therefore

$$\mathbf{M}_{uv} = \mathbf{E}_{x\in\{0,1\}^\ell}[\,\mathbf{M}(x)_{uv}\,] \geqslant \frac{c}{2\Delta} = c\mathcal{L}_{1/2}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)_{u,v}.$$

For $u=v$, note that $\mathbf{M}_{u,v}\geqslant 1/2$ by definition, and $\mathcal{L}_{1/2}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)_{u,v}\leqslant 1$. □

Lemma 4.6. *For any $\mathbf{v}\in\mathbb{R}^n$ orthogonal to $\boldsymbol{\pi}$, it holds that*

$$0\leqslant \mathbf{E}_{x\in\{0,1\}^\ell}\left[\,\|\mathbf{v}\mathbf{M}(x)\|_2^2\,\right] \leqslant (1-c\beta^{-2}\alpha)\|\mathbf{v}\|_2^2$$

for some constant $c\in(0,1)$.

Proof. The first inequality is obvious. For the upper bound, we have

$$\begin{aligned}\mathbf{E}_{x\in\{0,1\}^\ell}\left[\,\|\mathbf{v}\mathbf{M}(x)\|_2^2\,\right] &= \mathbf{E}_{x\in\{0,1\}^\ell}[\,\mathbf{v}\mathbf{M}(x)\mathbf{M}(x)^\intercal\mathbf{v}^\intercal\,]\\ &= \mathbf{v}\mathbf{E}_{x\in\{0,1\}^\ell}[\,\mathbf{M}(x)\mathbf{M}(x)^\intercal\,]\,\mathbf{v}^\intercal\\ &= \mathbf{v}\mathbf{E}_{x\in\{0,1\}^\ell}[\,\mathbf{M}(x)\,]\,\mathbf{v}^\intercal\\ &= \mathbf{v}\mathbf{M}\mathbf{v}^\intercal.\end{aligned}$$

Let $\mathbf{M}' = \mathbf{M} - c\cdot\mathcal{L}_{1/2}(\mathbf{M}_{\mathrm{Reg}(G)})$, where c is as in Lemma 4.5. Then $\mathbf{M}'$ is a non-negative matrix by Lemma 4.5. As both $\mathbf{M}$ and $\mathcal{L}_{1/2}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)$ are doubly-stochastic, so is $\mathbf{M}'/(1-c)$. Then $\lambda_{\max}(\mathbf{M}')\leqslant\|\mathbf{M}'\|_2\leqslant 1-c$. Note that $\lambda_{\max}\left(\mathcal{L}_{1/2}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)\right)\leqslant 1-\beta^{-2}\alpha/2$. Therefore

$$\begin{aligned}\lambda_{\max}(\mathbf{M}) &\leqslant \lambda_{\max}(\mathbf{M}') + c\cdot\lambda_{\max}\left(\mathcal{L}_{1/2}\left(\mathbf{M}_{\mathrm{Reg}(G)}\right)\right)\\ &\leqslant 1-(c/2)\beta^{-2}\alpha,\end{aligned}$$

and the claim follows. □

Recall that $\boldsymbol{\pi}$ denotes the uniform distribution over $V[G]$. We need the following lemma stating that a random matrix $\mathbf{M}(x)$ shrinks vectors orthogonal to $\boldsymbol{\pi}$ greatly:

Lemma 4.7. *For any $\mathbf{v}\in\mathbb{R}^n$ orthogonal to $\boldsymbol{\pi}$ and $k\in[T]$, it holds that*

$$\mathbf{E}_{y_0,\ldots,y_{k-1}\in\{0,1\}^\ell}\left[\left\|\mathbf{v}\prod_{i=0}^{k-1}\mathbf{M}(y_i)\right\|_2^2\right] \leqslant (1-c\beta^{-2}\alpha)^k\|\mathbf{v}\|_2^2$$

for some constant $c\in(0,1)$.

Proof. The proof is by induction k. The claim is trivial for $k = 0$. For $k > 0$, assume the claim holds for $k' < k$. Let $\mathbf{v} \in \mathbb{R}^n$ be a vector orthogonal to $\boldsymbol{\pi}$, and define $\mathbf{v}' = \mathbf{v}\prod_{i=0}^{k-2}\mathbf{M}(y_i)$. Then $\mathbf{v}'$ is also orthogonal to $\boldsymbol{\pi}$. Hence,

$$\begin{aligned}
&\mathbf{E}_{y_0,\ldots,y_{k-1}\in\{0,1\}^\ell}\left[\left\|\mathbf{v}\prod_{i=0}^{k-1}\mathbf{M}(y_i)\right\|_2^2\right]\\
&=\mathbf{E}_{y_0,\ldots,y_{k-2}\in\{0,1\}^\ell}\left[\mathbf{E}_{y_{k-1}\in\{0,1\}^\ell}\left[\|\mathbf{v}'\mathbf{M}(y_{k-1})\|_2^2\right]\right]\\
&\leqslant\mathbf{E}_{y_0,\ldots,y_{k-2}\in\{0,1\}^\ell}\left[(1-c\beta^{-2}\alpha)\|\mathbf{v}'\|_2^2\right]\\
&\leqslant(1-c\beta^{-2}\alpha)^k\|\mathbf{v}\|_2^2,
\end{aligned}$$

where the first inequality uses Lemma 4.6 and the second one uses the induction hypothesis. □

Let $\mathcal{P}$ be the distribution of $y = (y_0, \ldots, y_{T-1})$ in Protocol 3. The following lemma bounds the difference of the expected ℓ_2-norms of the resulting value vectors in T rounds by using the distribution $\mathcal{P}$ and the uniform distribution.

Lemma 4.8. *For any $u \in V[G]$, it holds that*

$$\begin{aligned}
\left|\mathbf{E}_{y\sim\mathcal{P}}\left[\left\|\mathbf{e}_u\prod_{i=0}^{T-1}\mathbf{M}(y_i)\right\|_2^2\right]\right.\\
\left.-\mathbf{E}_{y\in(\{0,1\}^\ell)^T}\left[\left\|\mathbf{e}_u\prod_{i=0}^{T-1}\mathbf{M}(y_i)\right\|_2^2\right]\right|\leqslant\varepsilon,
\end{aligned}$$

where ε is as in Protocol 3.

Proof of Theorem 4.1. By Lemma 4.7, we have

$$\mathbf{E}_{y\in(\{0,1\}^\ell)^T}\left[\left\|\mathbf{e}_s^\perp\prod_{i=0}^{T-1}\mathbf{M}(y_i)\right\|_2^2\right]\leqslant(1-c\beta^{-2}\alpha)^T.$$

Combining this with Lemma 4.8 and using the fact that $\|\mathbf{v}\|_2^2 = \|\mathbf{v}^\perp\|_2^2 + \|\boldsymbol{\pi}\|_2^2$ for any distribution $\mathbf{v}$, we obtain

$$\begin{aligned}
&\mathbf{E}_{y\sim\mathcal{P}}\left[\left\|\left(\mathbf{e}_s\prod_{i=0}^{T-1}\mathbf{M}(y_i)\right)^\perp\right\|_2^2\right]\\
&=\mathbf{E}_{y\sim\mathcal{P}}\left[\left\|\mathbf{e}_s^\perp\prod_{i=0}^{T-1}\mathbf{M}(y_i)\right\|_2^2\right]\leqslant(1-c\beta^{-2}\alpha)^T+\varepsilon<\kappa^2
\end{aligned}$$

for sufficiently large $T = O(\beta^2\alpha^{-1}\log\kappa^{-1})$. The claim follows from Lemma 4.4 and Markov's inequality. □

5 Two-Level Hashing Protocols

In this section we further present two protocols. These two protocols are based on pairwise independent generators and unbalanced expanders with near-optimal expansion. In contrast to the protocols in Section 3, the protocols in this section do not need to assume that nodes have initial IDs, and we can combine the protocols with an ID distribution mechanism [22] so that every node gets a unique ID once it gets the rumor. This mechanism [22] further guarantees that all informed nodes have different IDs, and the IDs are in $\left[2^T\right]$ if the protocol finishes in T rounds.

5.1 Protocol For Graphs with Certain Conductance

We first present one protocol for general graphs with conductance ϕ. Formally, for a graph G of n nodes, the *conductance* $\phi(G)$ of G is defined by

$$\phi(G)\triangleq\min_{S\subseteq V,0<|S|<n}\frac{e(S,V\setminus S)}{\min\{\mathrm{vol}(S),\mathrm{vol}(V\setminus S)\}}.$$

At a high level, our protocol for general graphs is based on a nice "two-level hashing" framework: The first level is based on a pairwise independent generator $\mathcal{G}$. While the PRG-based protocol in [22] needs to generate $O(n)$ blocks and different nodes need to use different blocks, our protocol only needs $M = (\Delta\log n)^{O(1)}$ blocks, and hence $O(\log\log n + \log\Delta)$ random bits suffice for this purpose. The second level uses unbalanced expanders to map the node with ID $u \in [n^c]$ to $r \in \left[\Delta^{O(1)}\right]$ by using $O(\log\log n + \log\Delta)$ random bits. After these, node u uses the value of the rth block of $\mathcal{G}$ to choose the neighbors. It is easy to see that every informed node u only needs $O(\mathrm{poly}\log n)$ arithmetic operations per round in order to determine its neighbor. Moreover, the protocol finishing in T rounds only needs $O(T\cdot(\log\log n+\log\Delta))$ random bits in total, see Figure 1 for an illustration and Protocol 4 below for the formal description.

Protocol 4 (Protocol for General Graphs)**.** *Let $\varepsilon = \Delta^{-\Theta(1)}$ be sufficiently small, and $m = 2^{\lceil\log(4/\varepsilon)\rceil}$. Pick the following objects:*

- *An explicit $(K, (1-\varepsilon^2/4)D)$-expander $\Gamma : [n^c]\times[D] \mapsto \bigsqcup_{i\in[D]}[M_i]$, where $K = 2$, $D = ((\log n)/\varepsilon)^{O(1)}$ and $M_0 = \cdots = M_{D-1} = M \leqslant D$.*
- *An explicit pairwise independent generator $\mathcal{G} = (\mathcal{G}_1, \ldots, \mathcal{G}_M) : \{0,1\}^\ell \mapsto [m]^M$, where $\ell = O(\log m + \log M) = O(\log\log n + \log\Delta)$. These two objects $\mathcal{G}$ and Γ can be uniquely constructed from n^c and $\Delta^{\Theta(1)}$, and hence are known to every informed node.*

The initial node having the rumor chooses a random string $(s_1, \ldots, s_T)$, where every s_i is of the form $(x_i, y_i) \in [D] \times \{0,1\}^\ell$. This random string is appended with the rumor and sent to other nodes. Once one node gets the rumor, it gets the ID u. In the ith round, node u computes $r = \Gamma(u, x_i)$ that is in $[M_u]$, the uth copy of $[M]$. Node u computes $y \triangleq \mathcal{G}_r(y_i) \bmod \Delta$, and chooses the neighbor with index y in its adjacency list to send the rumor if $y \leqslant \deg(u)$.

Using the explicit constructions of unbalanced expanders in [25] and pairwise independent generators in [4], our protocol is very simple and can be described as follows: Assign each node with ID $u \in [n^c]$ with a distinct polynomial p_u of degree at most $\lceil c \log_q n \rceil$ over a finite field $\mathbb{F}_q$ of size $q = (\Delta \log n)^{\Theta(1)}$. The protocol then uses the random string $(s_1, \ldots, s_T)$, where every $s_i \triangleq (x_i, a_i, b_i) \in \mathbb{F}_q^3$. In the ith round, an informed node u computes $z = a_i \cdot p_u(x_i) + b_i$ (over $\mathbb{F}_q$), and chooses the neighbor with index $(z \bmod \deg(u))$ in its adjacency list to send the rumor.

Now we analyze the protocol above, and prove Theorem 1.3. We start by analyzing a single round t and see the properties of our protocol. Let I_t be the set of informed nodes after round t, and U_t the set of uninformed nodes after round t. Remember that all the random choices in round t are determined by (x_t, y_t), and in Protocol 4 different rounds use independently chosen random seeds. We need the following lemma:

Lemma 5.1. *Fix any round $0 \leqslant t < T$. For any $u \in U_t$, $v \in I_t$, let $X_{v\to u}$ be the boolean random variable whose value is 1 if and only if v informs u in round $t+1$. Then the following two statements hold:*

1. $|\mathbf{E}[X_{v\to u}] - 1/\Delta| \leqslant \varepsilon$ *for any* $u \in U_t$, $v \in I_t$;
2. $\mathbf{Cov}[X_{v\to u}, X_{v'\to u'}] \leqslant \varepsilon$ *for any* $u, u' \in U_t$, $v, v' \in I_t$ *satisfying* $(u, v) \neq (u', v')$.

Proof. For any $u \in U_t$ and $v \in I_t$, suppose the index of u in the adjacency list of v is z. By construction, $X_{v\to u}$ equals 1 if and only if $\mathcal{G}_{\Gamma(v,x_t)}(y_t) \bmod \Delta = z$. Fix x_t. The fact that $\mathcal{G}$ is a pairwise independent generator together with Lemma 2.2 shows that

$$|\mathbf{E}[X_{v\to u}] - 1/\Delta| \leqslant 2/m \leqslant \varepsilon.$$

For any $u, u' \in U_t$ and $v, v' \in I_t$, first assume $v \neq v'$. Suppose the index of u (resp. u') in the adjacency list of v (resp. v') is z (resp z'). By construction, $X_{v\to u}$ equals 1 if and only if $\mathcal{G}_{\Gamma(v,x_t)}(y_t) \bmod \Delta = z$, and similarly for $X_{v'\to u'}$. By Lemma 2.10 and the fact that Γ is a $(K, (1-\varepsilon^2/4)D)$-expander, the event $|\{\Gamma(v, x_t), \Gamma(v', x_t)\}| \geqslant (1-\varepsilon/2) \cdot 2 > 1$ occurs with probability at least $1 - \varepsilon/2$ over the choices of x_t. Condition on any x_t such that this event occurs. We have $\Gamma(v, x_t) \neq \Gamma(v', x_t)$. Using the fact that $\mathcal{G}$ is pairwise independent together with Lemma 2.2, we have $\mathbf{Cov}[X_{v\to u}, X_{v'\to u'}] \leqslant 2/m$. For the other choices of x_t, we have $\mathbf{Cov}[X_{v\to u}, X_{v'\to u'}] \leqslant 1$ since $X_{v\to u}, X_{v'\to u'}$ are boolean random variables. Therefore

$$\mathbf{Cov}[X_{v\to u}, X_{v'\to u'}] \leqslant (1-\varepsilon/2)(2/m) + (\varepsilon/2) \leqslant \varepsilon$$

for random x_t.

Now assume $v = v'$ and hence $u \neq u'$. We have

$$\begin{aligned}&\mathbf{Cov}[X_{v\to u}, X_{v\to u'}]\\ &= \mathbf{E}[X_{v\to u} \cdot X_{v\to u'}] - \mathbf{E}[X_{v\to u}] \cdot \mathbf{E}[X_{v\to u'}]\\ &= 0 - \mathbf{E}[X_{v\to u}] \cdot \mathbf{E}[X_{v\to u'}] \leqslant 0.\end{aligned}$$ □

Lemma 5.2. *Fix a round $0 \leqslant t < T$, and the set I_t of informed nodes after round t. Fix also an arbitrary set of edges $F \subseteq E(I_t, U_t)$. Let J be the set of nodes that become informed in round $t+1$ if we consider only transmissions of the rumor along the edges in F. Then the following two statements hold:*

1. $\mathbf{Pr}[J \neq \emptyset] \geqslant c_1 \min\{|F|/\Delta, 1\}$ *for some constant* $c_1 > 0$.
2. *If* $|F| = \Omega(\Delta)$, *then* $\mathbf{Pr}[|J| \geqslant c_2|F|/\Delta] \geqslant c_3$ *for some constant* $c_2, c_3 > 0$.

Now we prove Theorem 1.3. We first define a matrix $\mathcal{M} \in \mathbb{R}^{n\times n}$ that is associated with graph G. For any $u, v \in V[G]$, let $\mathcal{M}_{u,v} = 1/\Delta$ if $\{u, v\} \in E[G]$, $\mathcal{M}_{u,v} = 1 - \deg(u)/\Delta$ if $u = v$, and $\mathcal{M}_{u,v} = 0$ otherwise. Notice that matrix $\mathcal{M}$ is doubly-stochastic. We further define the conductance of matrix $\mathcal{M}$ by

$$\Phi(\mathcal{M}) \triangleq \min_{\substack{A\subset V\\ |A|\leqslant n/2}} \frac{e(A, \overline{A})}{\Delta \cdot |A|}.$$

Notice that $\Phi(\mathcal{M}) \leqslant \phi(G) \leqslant \Phi(\mathcal{M}) \cdot \beta$. Hence it suffices to work with $\Phi(\mathcal{M})$ in the following.

Proof of Theorem 1.3. By the construction of Protocol 4, the randomness requirement in Theorem 1.3 is obvious, and we only need to analyze the runtime of Protocol 4. The proof is divided into four phases, depending on the number of informed nodes $|I_t|$ after round t.

Phase 1: $1 \leqslant |I_t| \leqslant 1/\Phi$. This phase is divided into several subphases. For every $1 \leqslant i \leqslant \log(1/\phi)$, subphase i begins when the number of informed nodes is at least 2^{i-1} and ends when this number is at least 2^i. Assume that we are at the beginning of the ith subphase. Fix an arbitrary round t of the ith subphase

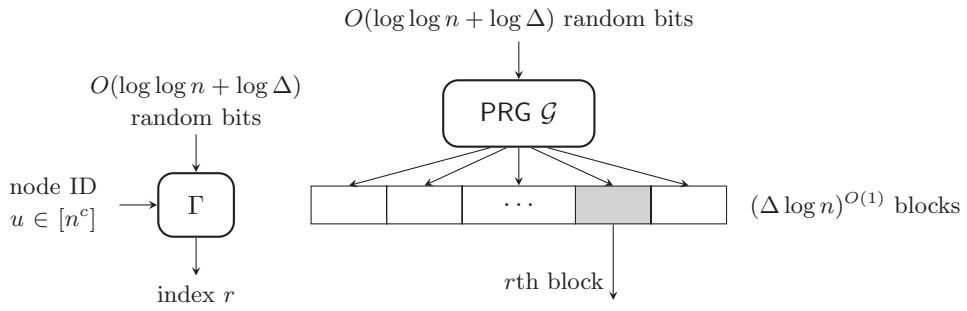

Figure 1: Illustration of the protocol for general graphs. Every node u uses an unbalanced expander Γ to generate an index r, and uses the rth block of a pairwise independent generator $\mathcal{G}$ to choose a neighbor to send the rumor.

and the set of informed nodes I_t; thus, $2^{i-1} \leqslant |I_t| < 2^i$. We consider the number of nodes that become informed in round $t+1$. Applying Lemma 5.2(1) with $F = E(I_t, U_t)$ gives

$$\begin{aligned}\mathbf{Pr}\left[\,|I_{t+1} \setminus I_t| \geqslant 1\,\right] &\geqslant c_1 \min\{e(I_t,U_t)/\Delta, 1\} \\ &\geqslant c_1 \min\{\Phi \cdot |I_t|/\beta, 1\}.\end{aligned}$$

Let $p \triangleq c_1 \min\{\Phi \cdot |I_t|/\beta, 1\}$, and hence $p = O(\Phi \cdot |I_t|)$, since $|I_t| \leqslant 1/\Phi$ and $\beta \geqslant 1$. Therefore, the expected time to increase $|I_t|$ from 2^{i-1} to 2^i is at most $2^{i-1}/p = O(1/\Phi)$. By Markov's inequality,

$$\mathbf{Pr}\left[\,|I_{t+\tau}| \leqslant 2^i \;\middle|\; |I_t| \geqslant 2^{i-1}\,\right] \leqslant 1/2$$

for some $\tau = O(\Phi^{-1})$. Hence the time to complete Phase 1 can be upper bounded by $\tau = O((1/\Phi))$ multiplied with the sum of $\log(1/\Phi) = O(\log n)$ independent geometric random variables each with parameter $1/2$. Applying a Chernoff bound for the sum of independent geometric random variables yields that the number of rounds required for Phase 1 is at most $O((1/\Phi)\cdot\log n) = O((1/\phi)\cdot\beta\cdot\log n)$ with high probability.

Phase 2: $1/\Phi \leqslant |I_t| \leqslant n/2$**.** Fix a round t and the set of informed nodes I_t. We apply Lemma 5.2(2), with $F = E(I_t, U_t)$. Note that the precondition $|F| = \Omega(\Delta)$ is satisfied, as $|F| = e(I_t,U_t) \geqslant \Phi\cdot\Delta\cdot|I_t| \geqslant \Phi\cdot\Delta\cdot(1/\Phi) = \Omega(\Delta)$. Hence we conclude from Lemma 5.2(2) that

$$\mathbf{Pr}\left[\,|I_{t+1} \setminus I_t| \geqslant c_2\cdot\phi\cdot\delta\cdot|I_t|/\Delta\,\right] \geqslant c_3,$$

for some constant $c_2, c_3 > 0$. When this event occurs, we have $|I_{t+1}| \geqslant (1 + c_2\cdot\phi/\beta)|I_t|$. So, the number of rounds until we have $|I_t| \leqslant n/2$ can be upper bounded by the sum of $\log_{1+c_2\cdot\phi/\beta}(n/2) = O((1/\phi)\cdot\beta\cdot\log n)$ independent geometric random variables with parameter c_3. Applying a Chernoff bound again we obtain that Phase 2 is completed within at most $O((1/\phi)\cdot\beta\cdot\log n)$ rounds with high probability.

Phase 3: $n/2 \leqslant |I_t| \leqslant n - 1/\Phi$**.** The analysis is the same as in Phase 2 with the roles of I_t and U_t switched.

Phase 4: $n - 1/\Phi \leqslant |I_t| \leqslant n$**.** Again, the analysis is the same as in Phase 1 with the roles of I_t and U_t switched.

Since each of these four phases requires only $O((1/\phi)\cdot\beta\cdot\log n)$ rounds with high probability, the result follows by applying the union bound. □

5.2 Protocol For Strong Expander Graphs We further present a protocol for strong expander graphs. Let $\mathcal{G} = \{G_i\}_{i\geqslant 0}$ be a family of graphs. We call $\mathcal{G}$ a family of *strong expander graphs*, if every G_i in $\mathcal{G}$ has spectral gap $\alpha = 1 - o(1)$, and irregularity $\beta = 1 + o(1)$, where the term $o(1)$ tends to 0 as n goes to the infinity. This graph family includes several interesting graphs, e.g. Ramanujan graphs, complete graphs, random graphs $G(n,p)$ with $p = \omega(\log n/n)$, and random d-regular graphs.

It is known that, for any strong expander graph with n nodes, the fully-random push protocol informs all nodes in $\log n + \ln n + o(\log n)$ rounds with high probability [13, 14, 16], i.e., we know the *precise* rumor spreading time for strong expander graphs, and this bound is tight [14]. Our next protocol shows that, this tight runtime can be achieved by only using $O(\log n \cdot (\log\log n + \log\Delta))$ random bits in total.

Protocol 5 (Protocol for Strong Expanders)**.** *Let* $\varepsilon = \Delta^{-\Theta(1)}$ *be sufficiently small,* $\varepsilon' = 2^{-\sqrt{\log\log n}}$*, and* $m = \Theta((\log n)/\varepsilon)$ *a power of 2. Pick the following objects: (1) An explicit* $(\leqslant K, (1-\varepsilon^2/4)D)$*-expander* $\Gamma : [n^c] \times [D] \mapsto \bigsqcup_{i\in[D]}[M_i]$*, where* $K = \Delta$*,* $D = ((\log n)/\varepsilon)^{O(1)}$ *and*

$$M_0 = \cdots = M_{D-1} = M \leqslant \max\left\{D, \Delta^{O(1)}\right\}.$$

(2) An explicit function $\mathcal{G} = (\mathcal{G}_1, \ldots, \mathcal{G}_M) : \{0,1\}^\ell \mapsto$

$[m]^M$ *that is both a pairwise independent generator and an* ε'*-PRG for* $\mathsf{CR}_{[m]^M}$*, where* $\ell = O(\log m + \log M) + \widetilde{O}(\log(1/\varepsilon')) = O(\log\log n + \log\Delta)$[9]*. These two objects* $\mathcal{G}$ *and* Γ *can be uniquely constructed from* n^c *and* $\Delta^{\Theta(1)}$*, and hence are known to every informed node.*

The initial node having the rumor chooses a random string $(s_1, \ldots, s_T)$*, where every* s_i *is of the form* $(x_i, y_i) \in [D] \times \{0,1\}^\ell$*. This random string is appended with the rumor and sent to other nodes. Once one node gets the rumor, it gets the ID whose value is* u*. In the* i*th round, node* u *computes* $r = \Gamma(u, x_i)$ *that is in* $[M_u]$*, the* u*th copy of* $[M]$*. It then chooses the neighbor with index* $\mathcal{G}_r(y_i) \bmod \deg(u)$ *in its adjacency list to send the rumor.*

Proposition 5.3. *Assume that Protocol 5 finishes in* T *rounds. Then it uses* $O(T \cdot (\log\log n + \log\Delta))$ *random bits in total.*

In this subsection we analyze Protocol 5, which corresponds to Theorem 1.4. We first remark that the condition $\alpha = 1 - o(1)$ is equivalent to $\lambda \triangleq \lambda_2 = o(1)$, which will be used in the following. The following lemma will be used in our proof.

Lemma 5.4 (Expander Mixing Lemma, [7])**.** *Let* $G = (V, E)$ *be a graph. Then for any subset* X *and* Y *of* V *it holds that*

$$\left| e(X,Y) - \frac{\mathrm{vol}(X)\cdot\mathrm{vol}(Y)}{\mathrm{vol}(G)} \right| \leqslant \lambda \cdot \frac{\sqrt{\mathrm{vol}(X)\cdot\mathrm{vol}(Y)\cdot\mathrm{vol}(\overline{X})\cdot\mathrm{vol}(\overline{Y})}}{\mathrm{vol}(G)}.$$

Proof of Theorem 1.4. By the construction of Protocol 5, the randomness requirement is obvious, and we only need to analyze the runtime of Protocol 5.

For any round t and $u \in U_t$, $v \in I_t$, let $X_{v\to u}$ be the indicator random variable whose value is 1 if and only if v informs u in round $t+1$. Note that Γ is a $(\leqslant K, (1-\varepsilon^2/4)D)$-expander, and hence a $(2, (1-\varepsilon^2/4)D)$-expander. And $\mathcal{G}$ is a pairwise independent generator. Then we observe that the statements in Lemma 5.1 hold here as well by the same proof. Notice that it holds by Lemma 5.4 that

$$e(I_t, U_t) \geqslant (1-\lambda)\cdot\frac{\mathrm{vol}(I_t)\cdot(\mathrm{vol}(G)-\mathrm{vol}(I_t))}{\mathrm{vol}(G)}. \tag{5.7}$$

The proof is divided into three phases, depending on the number of informed nodes $|I_t|$ after round t.

[9]The order $\widetilde{O}(\cdot)$ here hides poly $\log\log(1/\varepsilon')$.

Phase 1: $1 \leqslant |I_t| \leqslant n/\log n$. We will show that there is $\tau = \log n + o(\log n)$ such that $|I_{t+\tau}| > n/\log n$. By (5.7) we have

$$e(I_t, U_t) \geqslant (1-\lambda)\cdot\delta\cdot|I_t|\left(1 - \frac{\Delta\cdot|I_t|}{nd}\right),$$

where d is the average degree. Since $\lambda = o(1)$ and $|I_t| \leqslant n/\log n$, we have

$$e(I_t, U_t) \geqslant \left(1 - \frac{1}{\log n} - o(1)\right)\cdot\Delta\cdot|I_t|. \tag{5.8}$$

Hence

$$|N(I_t)\setminus I_t| \geqslant \frac{e(I_t,U_t)}{\Delta} \geqslant \left(1 - \frac{1}{\log n} - o(1)\right)\cdot|I_t|.$$

Define $\gamma \triangleq \lambda + \frac{1}{\log n}$, and $A \triangleq \{u \in N(I_t)\setminus I_t : |N(u)\cap I_t| \geqslant 2d\sqrt{\gamma}\}$. Then $e(A, I_t) \geqslant |A|\cdot 2d\cdot\sqrt{\gamma}$. On the other hand, by Lemma 5.4 it holds that

$$\begin{aligned} e(A,I_t) &\leqslant \frac{\mathrm{vol}(A)\cdot\mathrm{vol}(I_t)}{\mathrm{vol}(G)} + \lambda\sqrt{\mathrm{vol}(A)\cdot\mathrm{vol}(I_t)} \\ &\leqslant \frac{\Delta^2\cdot|A|\cdot|I_t|}{nd} + \gamma\Delta\cdot\sqrt{|A|\cdot|I_t|}. \end{aligned}$$

We know $e(A, I_t) \geqslant 2d\sqrt{\gamma}\cdot|A|$ by the definition of set A , and hence

$$\begin{aligned} |A|\cdot 2d\cdot\sqrt{\gamma} &\leqslant \frac{\Delta^2\cdot|A|\cdot|I_t|}{nd} + \gamma\Delta\cdot\sqrt{|A|\cdot|I_t|} \\ &\leqslant (1+o(1))\cdot\frac{\Delta\cdot|A|}{\log n} + \gamma\Delta\cdot\sqrt{|A|\cdot|I_t|}, \end{aligned}$$

which implies $|A| \leqslant \gamma\cdot|I_t|$.

Now define $B \triangleq N(I_t)\setminus I_t\setminus A$. We have

$$\begin{aligned} e(B,I_t) &= e(N(I_t),I_t) - e(A,I_t) \\ &\geqslant \left(1 - \frac{1}{\log n} - o(1) - \gamma\right)\Delta\cdot|I_t|. \end{aligned}$$

With the above estimate at hand, we compute the expected value of $|N(I_t)\cap B|$. Note that for any $u \in B$, the chance that it gets informed in round $t+1$ is

$$p_{t+1}(u) \triangleq \mathbf{Pr}\left[\bigvee_{v\in N(u)\cap I_t}(X_{v\to u} = 1)\right],$$

which is lower bounded by

$$\sum_{v\in N(u)\cap I_t}\mathbf{Pr}\left[X_{v\to u} = 1\right] - \sum_{\substack{v_1,v_2\in N(u)\cap I_t \\ v_1<v_2}}\mathbf{Pr}\left[\bigwedge_{i=1,2}(X_{v_i\to u} = 1)\right],$$

by the Bonferroni inequalities. Hence

$$\begin{aligned} &p_{t+1}(u) \\ &\geqslant |N(u)\cap I_t|\left(\frac{1}{\Delta}-\varepsilon\right)-\binom{|N(u)\cap I_t|}{2}\left(\frac{1}{\delta^2}+\varepsilon\right) \\ &\geqslant (1-o(1))\cdot\frac{|N(u)\cap I_t|}{\Delta}, \end{aligned} \tag{5.9}$$

where the first inequality follows from Lemma 5.1 and the fact that $\varepsilon = (1/\Delta)^{\Theta(1)}$ is sufficiently small, and the last step uses the condition that $|N(u)\cap I_t| \leqslant 2d\sqrt{\gamma} = o(\Delta)$. Hence, we have

$$\begin{aligned} \mathbf{E}\left[\,|I_{t+1}\setminus I_t|\,\right] &\geqslant \mathbf{E}\left[\,|I_{t+1}\cap B|\,\right] = \sum_{u\in B} p_{t+1}(u) \\ &\geqslant (1-o(1))\cdot\frac{e(B,I_t)}{\Delta} \geqslant (1-o(1))\cdot|I_t|. \end{aligned}$$

Since $|I_{t+1}\setminus I_t| \leqslant |I_t|$, it follows by using Markov's inequality (applied to $|I_t| - |I_{t+1}\setminus I_t|$) that

$$\mathbf{Pr}\left[\,|I_{t+1}| \geqslant (2-f(n))|I_t|\,\right] \geqslant 1-g(n),$$

where $f(n)$ and $g(n)$ are both functions that tend to zero. Hence, the time to reach $|I_t| \geqslant n/\log n$ can be upper bounded by the sum of $\log_{2-f(n)} n$ independent, identically distributed geometric random variables with expectation at most $1-o(1)$ each. Using a Chernoff bound yields that

$$\mathbf{Pr}\left[\,|I_{t+\tau}| > n/\log n\,\right] = 1-o(1).$$

Phase 2: $|I_t| \in [n/\log n, n-n/\log n]$. We will show that there is $\tau = o(\log n)$ such that $|I_{t+\tau}| > n - n/\log n$. We start with the first case $|I_t| \in [n/\log n, n/2]$.

For any $u \in N(I_t)\setminus I_t$, the probability $p_{t+1}(u)$ that u gets informed in round $t+1$ is lowered bounded by

$$(1-o(1))\cdot\frac{|N(u)\cap I_t|}{\Delta}\left(1-\frac{(1+o(1))\cdot|N(u)\cap I_t|}{2\Delta}\right),$$

by the same argument as in (5.9). This is then lower bounded by

$$(1-o(1))\cdot\frac{|N(u)\cap I_t|}{2\Delta},$$

since we have $|N(u)\cap I_t| \leqslant \Delta$.

By (5.7), we have $e(I_t,U_t) = (1-o(1))\cdot\delta/2\cdot|I_t|$. Similar to the analysis of Phase 1, we can lower bound the expected number of nodes that become informed in round $t+1$:

$$\mathbf{E}\left[\,|I_{t+1}\setminus I_t|\,\right] \geqslant \sum_{u\in N(I_t)\setminus I_t} p_{t+1}(u) \geqslant \frac{\delta}{8\Delta}|I_t|.$$

Since $|I_{t+1}| \leqslant 2|I_t|$, we obtain that, as long as $|I_t| \leqslant n/2$, there are constants $\alpha,\beta > 0$ so that $\mathbf{Pr}\left[\,|I_{t+1}| \geqslant (1+\alpha)|I_t|\,\right] \geqslant \beta$. Hence the time to reach $|I_t| \geqslant n/2$ can be upper bounded by the sum of $\log_{1+\alpha}(\log n)$ independent, identically distributed geometric random variables with expectation at most $1/\beta$ each. Using a Chernoff bound for the sum of geometric random variables yields that with probability $1-o(1)$, we reach $|I_t| \geqslant n/2$ within at most $o(\log n)$ additional rounds.

Consider now the case $|I_t| \in [n/2, n-n/\log n]$. To analyze this case, we examine the shrinking of $U_t = V\setminus I_t$. Note that for any $u \in U_t$, the probability $p_{t+1}(u)$ that u gets informed in round $t+1$ is lowered bounded by

$$(1-o(1))\cdot\frac{|N(u)\cap I_t|}{\Delta}\left(1-\frac{(1+o(1))\cdot|N(u)\cap I_t|}{2\Delta}\right),$$

by the same argument as in (5.9). This is then lower bounded by

$$(1-o(1))\cdot\frac{|N(u)\cap I_t|}{2\Delta},$$

since we have $|N(u)\cap I_t| \leqslant \Delta$.

Again, as $|U_t| \leqslant n/2$, by (5.7) we have

$$e(I_t,U_t) \geqslant (1-o(1))\cdot\frac{\delta}{2}|U_t|.$$

Let us now compute the expected number of uninformed nodes after one additional round:

$$\mathbf{E}\left[\,|U_{t+1}|\,\right] = \sum_{u\in U_t}(1-p_{t+1}(u)) \leqslant \left(1-\frac{\delta}{8\Delta}\right)|U_t|.$$

A simple inductive argument yields for any integer τ that

$$\mathbf{E}\left[\,|U_{t+\tau}|\,\right] \leqslant \left(1-\frac{\delta}{8\Delta}\right)^{\tau}|U_t|,$$

so for

$$\tau \triangleq \log\log n/\log(1/(1-\frac{\delta}{8\Delta})) + \omega(1),$$

where $\omega(1)$ is an arbitrarily slow growing function, we have $\mathbf{E}\left[\,|U_{t+\tau}|\,\right] = o(n/\log n)$. Hence it holds by Markov's inequality that

$$\mathbf{Pr}\left[\,|U_{t+\tau}| \geqslant n/\log n\,\right] = o(1).$$

Phase 3: $|I_t| \in [n-n/\log n, n]$. We will show that there is $\tau = \ln n + o(\log n)$ such that $|I_{t+\tau}| = n$.

Again, we analyze the shrinking of the set U_t. By Lemma 2.10, for at least $(1-\varepsilon/2)$-fraction of the

choices of x_t, it holds for any $u \in U_t$ that the size of $\{\Gamma(v, x_t) : v \in N(u) \cap I_t\}$ is at least $(1-\varepsilon/2)|N(u) \cap I_t|$. From now on, we fix x_t such that this event occurs.

For any $u \in U_t$, we have

$$\mathbf{Pr}\left[\, u \notin I_{t+1} \,\right] = \mathbf{Pr}\left[\bigwedge_{v \in N(u) \cap I_t} (X_{v \to u} = 0)\right].$$

Let F be a subset of $N(u) \cap I_t$ of size $(1-\varepsilon/2)|N(u) \cap I_t|$, such that the map $\Gamma(\cdot, x_t)$ is injective when restricted to F. By Lemma 2.4, the function $y \mapsto \big(\mathcal{G}_{\Gamma(v,x_t)}(y) \bmod \deg(v)\big)_{v \in F}$ is an $(\varepsilon' + |F|\Delta/m)$-PRG for CR_S where $S = \prod_{v \in F}[\deg(v)]$. Therefore, we have

$$\begin{aligned}
&\mathbf{Pr}\left[\, u \notin I_{t+1} \,\right] \\
&\leqslant \mathbf{Pr}\left[\bigwedge_{v \in F} (X_{v \to u} = 0)\right] \\
&\leqslant \left(1 - \frac{1}{\Delta} + \varepsilon\right)^{(1-\varepsilon/2)|N(u) \cap I_t|} + \varepsilon' + \Delta^2/m,
\end{aligned}$$

where the second inequality follows from the properties of PRGs for combinatorial rectangles, and the third inequality follows from using pairwise independent generators. Since $\varepsilon \leqslant \frac{1}{\Delta}$, a simple induction shows that

$$\left(1 - \frac{1}{\Delta} + \varepsilon\right)^k \leqslant \left(1 - \frac{1}{\Delta}\right)^k + k\varepsilon$$

for any $k \geqslant 0$. So we have

$$\begin{aligned}
&\mathbf{Pr}\left[\, u \notin I_{t+1} \,\right] \\
&\leqslant \left(1 - \frac{1}{\Delta}\right)^{(1-\varepsilon/2)|N(u) \cap I_t|} + (1-\varepsilon/2)\Delta\varepsilon + \varepsilon' + \frac{\Delta^2}{m}.
\end{aligned}$$

The bound above applies for any choice of x_t such that the size of $\{\Gamma(v, x_t) : v \in N(u) \cap I_t\}$ is at least $(1-\varepsilon/2)|N(u) \cap I_t|$. The probability of choosing such x_t is at least $1 - \varepsilon/2$. So for random x_t, we have

$$\mathbf{Pr}\left[\, u \notin I_{t+1} \,\right] \leqslant \left(1 - \frac{1}{\Delta}\right)^{(1-\varepsilon/2) \cdot |N(u) \cap I_t|} + o(1),$$

where we use the fact that $\varepsilon = (1/\Delta)^{\Theta(1)}$ is sufficiently small, and $m = \Theta((\log n)/\varepsilon)$.

By (5.8) it holds that $e(I_t, U_t) \geqslant (1 - \frac{1}{\log n} - o(1)) \cdot \Delta|U_t|$. Let $A \subseteq U_t$ be the set of nodes v for which $|N(v) \cap I_t| \leqslant (1 - \sqrt{\gamma}/2) \cdot \Delta$, where $\gamma \triangleq \frac{1}{\log n} + o(1)$. We assume for a contradiction that $|A| > 2\sqrt{\gamma} \cdot |U_t|$. Hence,

$$\begin{aligned}
e(I_t, U_t) &= \sum_{v \in A} |N(v) \cap I_t| + \sum_{v \in U_t \setminus A} |N(v) \cap I_t| \\
&< \left(1 - \frac{1}{\log n} - o(1)\right) \cdot \Delta|U_t|,
\end{aligned}$$

which yields the desired contradiction. Hence $|A| \leqslant 2\sqrt{\gamma}|U_t|$. Now define $B \triangleq U_t \setminus A$ so that for each $u \in B$, $|N(v) \cap I_t| > (1 - \sqrt{\gamma}/2)\Delta$ and $|B| \geqslant (1 - 2\sqrt{\gamma})|U_t|$. Using linearity of expectation,

$$\begin{aligned}
\mathbf{E}\left[\, |U_{t+1}| \,\right] &\leqslant \sum_{u \in B} \mathbf{Pr}\left[\, u \notin I_{t+1} \,\right] + \sum_{u \in A} \mathbf{Pr}\left[\, u \notin I_{t+1} \,\right] \\
&\leqslant \sum_{u \in B} \left(1 - \frac{1}{\Delta}\right)^{(1-\varepsilon/2)|N(u) \cap I_t|} + o(|U_t|).
\end{aligned}$$

Using the inequalities that $(1 - 1/k) \leqslant \mathrm{e}^{-1/k}$ for $k \geqslant 1$, $\mathrm{e}^x \leqslant 1 + 2x$ for sufficiently small enough constant $x > 0$, and the condition that $|N(u) \cap I_t| \geqslant (1 - \sqrt{\gamma}/2) \cdot \Delta$ for $u \in B$, we get $\mathbf{E}\left[\, |U_{t+1}| \,\right] \leqslant (1 + o(1)) \cdot \mathrm{e}^{-1} \cdot |U_t|$.

By induction, it follows that for any step $\tau > 0$,

$$\mathbf{E}\left[\, |U_{t+\tau}| \,\right] \leqslant ((1 + o(1)) \cdot \mathrm{e}^{-1})^\tau \cdot |U_t|.$$

We choose $\tau \triangleq -\log_{(1+o(1)) \cdot \mathrm{e}^{-1}}(n) = \ln n + o(\log n)$ and obtain that $\mathbf{E}\left[\, |U_{t+\tau}| \,\right] \leqslant (1/\log n)$. So

$$\mathbf{Pr}\left[\, |U_{t+\tau}| \geqslant 1 \,\right] \leqslant \mathbf{E}\left[\, |U_{t+\tau}| \,\right] \leqslant 1/\log n.$$

Combing these three phase together and applying the union bound, we obtain the desired statement. □

References

[1] N. Alon, C. Avin, M. Koucký, G. Kozma, Z. Lotker, and M. R. Tuttle. Many random walks are faster than one. In *20th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA'08)*, pages 119–128, 2008.

[2] R. Armoni, M. Saks, A. Wigderson, and S. Zhou. Discrepancy sets and pseudorandom generators for combinatorial rectangles. In *37th Annual IEEE Symposium on Foundations of Computer Science (FOCS'96)*, pages 412–421, 1996.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory and IEEE/ACM Transactions on Networking*, 52(6):2508–2530, 2006.

[4] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18(2):143 – 154, 1979.

[5] F. Chierichetti, S. Lattanzi, and A. Panconesi. Almost tight bounds on rumour spreading by conductance. In *42nd Annual ACM Symposium on Theory of Computing (STOC'10)*, pages 399–408, 2010.

[6] F. Chierichetti, S. Lattanzi, and A. Panconesi. Rumour spreading and graph conductance. In *21th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'10)*, pages 1657–1663, 2010.

[7] F. R. K. Chung. Spectral graph theory. *Regional Conference Series in Mathematics, American Mathematical Society*, 92:1–212, 1997.

[8] C. Cooper. Random walks, interacting particles, dynamic networks: Randomness can be helpful. In *18th International Colloquium on Structural Information and Communication Complexity (SIROCCO'11)*, pages 1–14, 2011.

[9] C. Cooper, D. Ilcinkas, R. Klasing, and A. Kosowski. Derandomizing random walks in undirected graphs using locally fair exploration strategies. *Distributed Computing*, 24(2):91–99, 2011.

[10] 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 *6th Annual ACM-SIGOPT Principles of Distributed Computing (PODC'87)*, pages 1–12, 1987.

[11] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading. In *19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'08)*, pages 773–781, 2008.

[12] R. Elsässer, U. Lorenz, and T. Sauerwald. On randomized broadcasting in star graphs. *Discrete Applied Mathematics*, 157(1):126–139, 2009.

[13] R. Elsässer and T. Sauerwald. Broadcasting vs. mixing and information dissemination on cayley graphs. In *24th International Symposium on Theoretical Aspects of Computer Science (STACS'07)*, pages 163–174, 2007.

[14] R. Elsässer and T. Sauerwald. On the runtime and robustness of randomized broadcasting. *Theoretical Computer Science*, 410(36):3414–3427, 2009.

[15] R. Elsässer and T. Sauerwald. Tight bounds for the cover time of multiple random walks. *Theoretical Computer Science*, 412(24):2623–2641, 2011.

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

[17] T. Friedrich, M. Gairing, and T. Sauerwald. Quasirandom load balancing. *SIAM Journal on Computing*, 41(4):747–771, 2012.

[18] A. M. Frieze and G. R. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10(1):57–77, 1985.

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

[20] G. Giakkoupis. Tight bounds for rumor spreading with vertex expansion. In *25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'14)*, pages 801–815, 2014.

[21] G. Giakkoupis and T. Sauerwald. Rumor spreading and vertex expansion. In *23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'12)*, pages 1623–1641, 2012.

[22] G. Giakkoupis, T. Sauerwald, H. Sun, and P. Woelfel. Low randomness rumor spreading via hashing. In *29th International Symposium on Theoretical Aspects of Computer Science (STACS'12)*, pages 314–325, 2012.

[23] G. Giakkoupis and P. Woelfel. On the randomness requirements of rumor spreading. In *22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'11)*, pages 449–461, 2011.

[24] P. Gopalan, R. Meka, O. Reingold, L. Trevisan, and S. Vadhan. Better pseudorandom generators from milder pseudorandom restrictions. In *53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*, pages 120–129, 2012.

[25] V. Guruswami, C. Umans, and S. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *Joural of ACM*, 56(4):20:1–20:34, 2009.

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

[27] M. Harchol-Balter, F. T. Leighton, and D. Lewin. Resource discovery in distributed networks. In *18th Annual ACM-SIGOPT Principles of Distributed Computing (PODC'99)*, pages 229–237, 1999.

[28] R. Impagliazzo, N. Nisan, and A. Wigderson. Pseudorandomness for network algorithms. In *26th Annual ACM Symposium on Theory of Computing (STOC'94)*, pages 356–364, 1994.

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

[30] T. Lindvall. *Lectures on the Coupling Method*. John Wiley & Sons Inc., New York, 2002.

[31] C.-J. Lu. Improved pseudorandom generators for combinatorial rectangles. *Combinatorica*, 22(3):417–434, 2002.

[32] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1):213–223, 1987.

[33] L. Saloff-Coste. Lectures on finite markov chains. In Pierre Bernard, editor, *Lectures on Probability Theory and Statistics*, volume 1665 of *Lecture Notes in Mathematics*, pages 301–413. Springer, 1997.

[34] T. Sauerwald and A. Stauffer. Rumor spreading and vertex expansion on regular graphs. In *22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'11)*, pages 462–475, 2011.

[35] T. Sauerwald and H. Sun. Tight bounds for randomized load balancing on arbitrary network topologies. In *53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*, pages 341–350, 2012.

[36] A. Ta-Shma, C. Umans, and D. Zuckerman. Lossless condensers, unbalanced expanders, and extractors. *Combinatorica*, 27(2):213–240, 2007.

[37] R. van Renesse, Y. Minsky, and M. Hayden. A gossip-style failure detection service. In *15th IFIP Intl. Conf. on Distributed Systems Platforms (Middleware)*, pages 55–70, 1998.