HAL
open science

Stochastic analysis of rumor spreading with multiple pull operations

Frédérique Robin, Bruno Sericola, Emmanuelle Anceaume, Yves Mocquard

To cite this version:

Frédérique Robin, Bruno Sericola, Emmanuelle Anceaume, Yves Mocquard. Stochastic analysis of rumor spreading with multiple pull operations. Methodology and Computing in Applied Probability, 2022, 24, pp.2195–2211. hal-03128118v2

HAL Id: hal-03128118

https://hal.science/hal-03128118v2

Submitted on 22 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Noname manuscript No.
(will be inserted by the editor)

Stochastic analysis of rumor spreading with multiple pull operations

Frdrique Robin · Bruno Sericola · Emmanuelle Anceaume · Yves Mocquard

Received: date / Accepted: date

Abstract We propose and analyze a new asynchronous rumor spreading protocol to deliver a rumor to all the nodes of a large-scale distributed network. This spreading protocol relies on what we call a k-pull operation, with $k \geq 2$. Specifically a k-pull operation consists, for an uninformed node s, in contacting $k-1$ other nodes at random in the network, and if at least one of them knows the rumor, then node s learns it. We perform a thorough study of the total number $T_{k,n}$ of k-pull operations needed for all the n nodes to learn the rumor. We compute the expected value and the variance of $T_{k,n}$, together with their limiting values when n tends to infinity. We also analyze the limiting distribution of $(T_{k,n} - \mathbb{E}(T_{k,n}))/n$ and prove that it has a double exponential distribution when n tends to infinity. Finally, we show that when $k > 2$, our new protocol requires less operations than the traditional 2-push-pull and 2-push protocols by using stochastic dominance arguments. All these results generalize the standard case $k = 2$.

Keywords Rumor spreading · Pull protocol · Markov chain · Asymptotic analysis

F. Robin
Inria, Univ. Rennes, CNRS, IRISA
Inria, Campus de Beaulieu, 35042 Rennes Cedex, France E-mail: frederique.robin@inria.fr

B. Sericola
Inria, Univ. Rennes, CNRS, IRISA
Inria, Campus de Beaulieu, 35042 Rennes Cedex, France E-mail: bruno.sericola@inria.fr

E. Anceaume
CNRS, Univ. Rennes, Inria, IRISA
IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France E-mail: emmanuelle.anceaume@irisa.fr

Y. Mocquard
Inria, Univ. Rennes, CNRS, IRISA
Inria, Campus de Beaulieu, 35042 Rennes Cedex, France E-mail: yves.mocquard@irisa.fr

1 Introduction

This paper focuses on the dissemination of information from users to users in a decentralized manner. Peer-to-peer (P2P) networks allow users or more generally nodes to exchange information by relying on gossip protocols, also called randomized rumor spreading protocols. Such protocols describe the rules required for one or more pieces of information known to an arbitrary node in the network to be spread to all the nodes of the network [5,8]. Among the different variants of randomized rumor spreading protocols, the push protocol provides a single operation, called the push operation, that allows an informed node to contact some randomly chosen node and sends it the rumor. The pull protocol, on the other hand allows, through the pull operation, an uninformed node to contact some random node to ask for the rumor. The same node can perform both operations according to whether it knows or not the rumor, which corresponds to the push-pull protocol.

One of the important questions raised by these randomized rumor spreading protocols is the spreading time, that is the time needed for all the nodes of the network to know the rumor.

Several models have been considered to answer this question. The most studied one is the synchronous model. This model assumes that all the nodes of the network act in synchrony, which allows the algorithms designed in this model to divide time in synchronized rounds. During each synchronized round, each node i of the network selects at random one of its neighbors j and either sends to j the rumor if i knows it (push operation) or gets the rumor from j if j knows the rumor (pull operation). In the synchronous model, the spreading time of a rumor is defined as the number of synchronous rounds necessary for all the nodes to know the rumor. When the underlying graph is complete, it has been shown by Frieze [12] that the ratio of the number of rounds over $\log_2(n)$ converges in probability to $1 + \ln(2)$ when the number n of nodes in the graph tends to infinity. Further results have been established (see for example [25,19] and the references therein), the most recent ones resulting from the observation that the rumor spreading time is closely related to the conductance of the graph of the network, see [14]. Investigations have also been done in different topologies of the network as in [2,4,11,23], in the presence of link or nodes failures as in [10], in dynamic graphs as in [3] and spreading with node expansion as in [15]. Another alternative consists for the nodes to make more than one call during the push or pull operations [24]. This alternative is of particular interest since it does not require any particular network structure. The synchronous case has been tackled in [24] where the authors show that the push-pull protocol takes $O(\log n/\log\log n)$ rounds in expectation assuming that nodes can connect to a random number of neighbors following a specific power law during each single round.

In large scale and open networks, assuming that all nodes act in synchrony is a very strong assumption since it requires that all the nodes have access to some global synchronization mechanism and that message delays are upper bounded. Several authors, including [16,20,1,7,26], suppose that nodes asyn-

chronously trigger operations with randomly chosen nodes. In [26], the authors model a multiple call by tuning the clock rate of each node with a given probability distribution. Some authors have focused on the message complexity by optimizing the network structure [21,6,8,16]. For instance, in [6], the authors show that in a preferential attachment graph the spreading time of an asynchronous push-pull protocol is $O\left(\sqrt{\log n}\right)$. Another way of limiting the number of interactions is by finely tuning the push and pull operations to take advantage of both of them as achieved for example in [9,5].

The pull algorithm attracted very little attention because this operation was long considered inefficient to spread a rumor within a large scale network [27]. It is actually very useful in systems fighting against message saturation (see for instance [30]). The ineffectiveness of the pull protocol stems from the fact that it takes some time before the rumour reaches a phase of exponential growth. Conversely, the push protocol initiates the rumor very quickly but then struggles to reach the last few uninformed nodes.

The objective of this paper is to push further this line of inquiry by presenting and studying a new pull protocol, called the k-pull protocol. This protocol is characterized by successive operations during which an initiator node asks for the rumor to a fixed number $(k-1)$ of other nodes in parallel. We propose a Markov model and we introduce the random variable $T_{k,n}$ representing the total number of such an operation, also called pull operation, needed for all the n nodes of the system to learn the rumor.

The remainder of the paper is organized as follows. In Section 2, we present the asynchronous k-pull protocol. We prove in Section 3 that the mean number of k-pull operations needed to inform all the n nodes of the system, assuming that a single node initially knows the rumor, that is $\mathbb{E}(T_{k,n})$, is equivalent to $n\ln(n)/(k-1)$. We also show that the variance of $T_{k,n}$ is equivalent to $n^2\pi^2/(6(k-1)^2)$. The distribution of $T_{k,n}$ is analyzed in Section 4. A recurrence relation is proposed to compute it and we provide bounds of this distribution. We also prove in this section that the limiting distribution of $(T_{k,n}-\mathbb{E}(T_{k,n})/n$, when n tends to infinity, is a double exponential distribution. Finally, we prove, in Section 5 by using stochastic dominance argument, that when $k \geq 3$, our new protocol requires less interactions than the standard push-pull protocol and also, as expected, less interactions than the standard push and pull protocols. Moreover, this efficiency increases strictly with k. Section 6 concludes the paper.

2 The k-pull protocol and its model

We consider a complete network of size n in which each node may be asked for a piece of information (pull operation). Nodes ask for a piece of information, i.e. the rumor, in an asynchronous way. The algorithm starts with a single node informed of the rumor. At each discrete time t, a single uninformed node s contacts $k-1$ distinct nodes, chosen at random uniformly among the $n-1$ other nodes. If at least one of these $k-1$ contacted nodes knows the rumor

then node s learns it. Otherwise nothing happens. This is a k-pull operation. We analyze the distribution of the number of k-pull operations needed for all the nodes to be informed of the rumor, and we compare it to the standard (i.e. $k = 2$) asynchronous push, pull and push-pull protocols. In order to avoid mistakes, these standard protocols will be denoted by 2-push, 2-pull and 2-push-pull. Note that it is the first time, to the best of our knowledge, that such a protocol is analyzed in the case $k > 2$.

To analyze the k-pull protocol, we introduce the discrete-time stochastic process $Y = \{Y_t,\ t \geq 0\}$ where Y_t represents the number of informed nodes at time t. Stochastic process Y is a discrete-time homogeneous Markov chain with n states where states $1, \ldots, n-1$ are transient and state n is absorbing. From the description of the protocol, we deduce that when the Markov chain Y is in state i at time t, then at time $t+1$, either it remains in state i if none of the $k-1$ chosen nodes were informed of the rumor or it transits to state $i+1$ if at least one of the $k-1$ chosen nodes were informed of the rumor. We denote by P the transition probability matrix of Markov chain Y. The non zero entries of matrix P are thus $P_{i,i}$ and $P_{i,i+1}$, for any $i = 1, \ldots, n-1$.

We then have $P_{i,i} = 1 - P_{i,i+1}$, which is given, for any $i = 1, \ldots, n-1$, by

$$P_{i,i} = \begin{cases} \dfrac{\binom{n-1-i}{k-1}}{\binom{n-1}{k-1}} & \text{if } i \leq n-k \\ \\ 0 & \text{otherwise.} \end{cases}$$

$\frac{\binom{n-1-i}{k-1}}{\binom{n-1}{k-1}}$ Indeed, given that $Y_t = i$, i.e. when i nodes are informed of the rumor at time t, we have $Y_{t+1} = i$ if and only if, at time $t+1$, the set of $k-1$ chosen nodes (i.e. $k-1$ among $n-1$) must be chosen among the $n-1-i$ non informed nodes.

We denote by $T_{k,n}$ the random variable defined by

$$T_{k,n} := \inf\{t \geq 0 \mid Y_t = n\}$$

which represents the spreading time, that is the total number of k-pull operations needed for all the nodes in the network to know the rumor.

The spreading time distribution can thus be expressed as a sum of independent random variables $S_{k,n}(i)$, where $S_{k,n}(i)$ is the sojourn time of Markov chain Y in state i. For all $i = 1, \ldots, n-k$, $S_{k,n}(i)$ follows a geometric distribution with parameter $p_{k,n}(i)$, where

$$p_{k,n}(i) = 1 - P_{i,i} = 1 - \prod_{h=1}^{k-1} \left(1 - \frac{i}{n-h}\right), \tag{1}$$

and $S_{k,n}(i) = 1$, for $i = n-k+1, \ldots, n-1$. Thus $T_{k,n}$ verifies

$$T_{k,n} = \sum_{i=1}^{n-1} S_{k,n}(i) = k - 1 + \sum_{i=1}^{n-k} S_{k,n}(i). \tag{2}$$

In the next section, we analyze the mean and the variance of the spreading time $T_{k,n}$ and their asymptotic behavior when n goes to infinity.

3 Mean and variance of the spreading time

To analyze the asymptotic behavior of the mean and variance of the spreading time $T_{k,n}$ when n goes to infinity, we provide upper and lower bounds of the probabilities $p_{k,n}(i)$, $i = 1, \ldots, n-k$. We first need the following technical lemma for which we introduce the functions $P_{k,n}(x)$ defined for all $x \in \mathbb{R}$, for every $n \geq 3$ and $k = 1, \ldots, n-1$, by

$$P_{k,n}(x) = 1 - \prod_{h=1}^{k} \left(1 - \frac{x}{n-h}\right).$$

Note that

$$p_{k,n}(i) = P_{k-1,n}(i). \tag{3}$$

Lemma 1 *For all $x \in \mathbb{R}$, we have*

$$\frac{d}{dx} P_{k,n}(x) = \frac{(-1)^{k+1} k}{\prod_{h=1}^{k}(n-h)} \prod_{h=1}^{k-1} (x - \mu_h),$$

where, for all $h = 1, \ldots, k-1$, μ_h are positive real numbers such that $n-(h+1) < \mu_h < n-h$.

Proof Note that for all $h = 1, \ldots, k$, we have $P_{k,n}(n-h) = 1$. Since $P_{k,n}(x)$ is a continuous function, it follows that there exists necessarily at least one local extremum point, denoted by μ_h, in each interval $(n-h-1, n-h)$, for $h = 1, \ldots, k-1$. The point μ_h is therefore a root of the polynomial $dP_{k,n}(x)/dx$. Note also that since the polynomial $1 - P_{k,n}(x)$ has only simple roots, we necessarily have $\mu_h \neq n-h-1$ and $\mu_h \neq n-h$. Using the fact that $P_{k,n}(x)$ is a k-degree polynomial, we deduce that $dP_{k,n}(x)/dx$ is a $(k-1)$-degree polynomial. The number of extremum μ_h being at least equal to $k-1$, this implies that the μ_h are unique. We thus first deduce that

$$\frac{d}{dx} P_{k,n}(x) = K \prod_{h=1}^{k-1} (x - \mu_h),$$

where K is a constant. We conclude using the fact that the factor of term x^k in polynomial $P_{k,n}(x)$ is equal to $(-1)^{k+1}/\prod_{h=1}^{k}(n-h)$. ■

We are now able to get, in the following two lemmas, lower and upper bounds of polynomial $P_{k,n}(x)$.

Lemma 2 *For all $x \in [1, n-k]$, we have*

$$P_{k,n}(x) \leq \frac{kx}{n-k}.$$

Proof From Lemma 1 and using the fact that $P_{k,n}(0) = 0$, we deduce that, for all $x \geq 0$,

$$P_{k,n}(x) = \frac{(-1)^{k+1}k}{\prod_{h=1}^{k}(n-h)} \int_0^x \prod_{h=1}^{k-1}(s-\mu_h)ds = \frac{k}{\prod_{h=1}^{k}(n-h)} \int_0^x \prod_{h=1}^{k-1}(\mu_h - s)ds.$$

Since $n-k < \mu_{k-1} < \cdots < \mu_1$, we get for all $x \in [1, n-k]$,

$$P_{k,n}(x) \leq \frac{kx}{\prod_{h=1}^{k}(n-h)} \max_{s\in[0,x]} \prod_{h=1}^{k-1}(\mu_h - s) = \frac{kx}{\prod_{h=1}^{k}(n-h)} \prod_{h=1}^{k-1}\mu_h.$$

Since $\mu_h < n-h$, for all $h = 1, \ldots, k-1$, we conclude that, for all $x \in [1, n-k]$,

$$P_{k,n}(x) \leq \frac{kx}{n-k},$$

which completes the proof. ■

We now turn to the lower bound of polynomial $P_{k,n}(x)$.

Lemma 3 *For all $x \in [1, n-k]$, we have*

$$P_{k,n}(x) \geq \frac{kx}{n+kx}.$$

Proof We first prove by recurrence that, for all integers $k \geq 1$ and for all $x \in [1, n-k]$, we have

$$(n+kx)\prod_{h=1}^{k}(n-h-x) \leq \prod_{h=0}^{k}(n-h). \tag{4}$$

Relation (4) is true for $k = 1$, since for all $x \in [1, n-1]$, we have $(n+x)(n-1-x) = n(n-1) - x - x^2 \leq n(n-1)$. Suppose now that Relation (4) is true

at rank k. At rank $k+1$, using (4), we get, for all $x \in [1, n-k-1]$,

$$\begin{aligned}
(n+(k+1)x)\prod_{h=1}^{k+1}(n-h-x) &= (n-k-1-x)(n+(k+1)x)\prod_{h=1}^{k}(n-h-x) \\
&\leq (n-k-1-x)\left[1+\frac{x}{n+kx}\right]\prod_{h=0}^{k}(n-h) \\
&= \left[n-k-1-(k+1)\frac{x^2+x}{n+kx}\right]\prod_{h=0}^{k}(n-h) \\
&\leq (n-k-1)\prod_{h=0}^{k}(n-h) \\
&= \prod_{h=0}^{k+1}(n-h),
\end{aligned}$$

which proves Relation (4). Using this relation, we obtain

$$1-P_{k,n}(x) = \prod_{h=1}^{k}\left(1-\frac{x}{n-h}\right) = \frac{n\prod_{h=1}^{k}(n-h-x)}{\prod_{h=0}^{k}(n-h)} \leq \frac{n}{n+kx}.$$

This implies that $P_{k,n}(x) \geq kx/(n+kx)$, which completes the proof. ■

The following theorem provides an equivalent of the mean spreading time when n tends to infinity. Note that the mean spreading time $\mathbb{E}(T_{k,n})$ is given by

$$\mathbb{E}(T_{k,n}) = k-1+\sum_{i=1}^{n-k}\frac{1}{1-P_{i,i}} = k-1+\sum_{i=1}^{n-k}\frac{1}{p_{k,n}(i)}. \tag{5}$$

Theorem 4 (Asymptotic mean spreading time) *For every $k \geq 2$, we have*

$$\mathbb{E}(T_{k,n}) \underset{n\longrightarrow\infty}{\sim} \frac{n\ln(n)}{k-1}.$$

Proof Combining Relations (3) and (5), we get

$$\mathbb{E}(T_{k,n}) = k-1+\sum_{i=1}^{n-k}\frac{1}{P_{k-1,n}(i)}.$$

Applying Lemmas 2 and 3, we obtain

$$k-1+\frac{n-k+1}{k-1}\sum_{i=1}^{n-k}\frac{1}{i} \leq \mathbb{E}(T_{k,n}) \leq n-1+\frac{n}{k-1}\sum_{i=1}^{n-k}\frac{1}{i}. \tag{6}$$

The fact that, for every $k \geq 0$, we have

$$\sum_{i=1}^{n-k} \frac{1}{i} \underset{n \longrightarrow \infty}{\sim} \ln(n)$$

completes the proof. ■

Concerning the variance of $T_{k,n}$, which is given by

$$\begin{aligned} \mathbb{V}ar(T_{k,n}) &= \sum_{i=1}^{n-k} \mathbb{V}ar(S_{k,n}(i)) = \sum_{i=1}^{n-k} \frac{1 - p_{k,n}(i)}{\left(p_{k,n}(i)\right)^2} \\ &= \sum_{i=1}^{n-k} \frac{1}{\left(p_{k,n}(i)\right)^2} - \left(\mathbb{E}(T_{k,n}) - k + 1\right), \end{aligned} \tag{7}$$

we have the following asymptotic result.

Theorem 5 (Asymptotic spreading time variance)

$$\mathbb{V}ar(T_{k,n}) \underset{n \longrightarrow \infty}{\sim} \frac{n^2}{(k-1)^2} \frac{\pi^2}{6}.$$

Proof Applying Lemma 3, we get, from Relation (7),

$$\begin{aligned} \mathbb{V}ar(T_{k,n}) &\leq \sum_{i=1}^{n-k} \frac{1}{\left(p_{k,n}(i)\right)^2} \leq \sum_{i=1}^{n-k} \frac{\left(n + (k-1)i\right)^2}{(k-1)^2 i^2} \\ &= \frac{n^2}{(k-1)^2} \sum_{i=1}^{n-k} \frac{1}{i^2} + \frac{2n}{k-1} \sum_{i=1}^{n-k} \frac{1}{i} + n - k \\ &\underset{n \longrightarrow \infty}{\sim} \frac{n^2}{(k-1)^2} \frac{\pi^2}{6} \end{aligned}$$

Using Lemma 2 and applying Theorem 4, we obtain

$$\mathbb{V}ar(T_{k,n}) \geq \frac{(n-k+1)^2}{(k-1)^2} \sum_{i=1}^{n-k} \frac{1}{i^2} - \mathbb{E}(T_{k,n}) + k - 1 \underset{n \longrightarrow \infty}{\sim} \frac{n^2}{(k-1)^2} \frac{\pi^2}{6},$$

which completes the proof. ■

4 Distribution of $T_{k,n}$, bounds and asymptotic analysis

It is well-known, see for instance [28], that the distribution of $T_{k,n}$ is given, for every integer $t \geq 0$, by

$$\mathbb{P}\{T_{k,n} > t\} = \alpha Q^t \mathbb{1}, \tag{8}$$

where α is the row vector containing the initial probabilities of states $1, \ldots, n-1$, that is $\alpha_i = \mathbb{P}\{Y_0 = i\} = 1_{\{i=1\}}$, Q is the matrix obtained from the transition matrix P containing the transition probabilities between transient states and $\mathbb{1}$ is the column vector of dimension $n-1$ with all its entries equal to 1. Note that the submatrix Q of the transition probability matrix P, is upper triangular with a single non zero upper-diagonal, that is $Q_{i,j} = 0$ for all $i = 1, \ldots, n-1$ and $j \neq i, i+1$. The computation of the distribution of $T_{k,n}$ can be easily done using the following recurrence relations. Let $V(t) = (V_1(t), \ldots, V_{n-1}(t))$ be the column vector defined by $V_i(t) = \mathbb{P}\{T_{k,n} > t \mid Y_0 = i\}$. From Relation (8), we have $V(t) = Q^t \mathbb{1}$. Since $V(0) = \mathbb{1}$, writing $V(t) = QV(t-1)$ for $t \geq 1$, we get for any $t \geq 1$:

$$\begin{cases} V_i(t) = P_{i,i}V_i(t-1) + (1 - P_{i,i})V_{i+1}(t-1),\ i = 1, \ldots, n-2 \\ V_{n-1}(t) = P_{n-1,n-1}V_{n-1}(t-1), \end{cases}$$

where the $P_{i,i}$ have been obtained in Section 2. This backward recursion leads to the computation of $V_1(t) = \mathbb{P}\{T_{k,n} > t\}$. It has been used for the drawing of Figure 1.

In order to get bounds of this distribution, we follow the approach used in [20]. We apply the bounds of the distribution of a sum of independent, but not necessarily identically distributed, geometric random variables provided in [18] to deduce bounds for the distribution $T_{k,n}$, both for all n (Theorem 6) and when n tends to infinity (Theorem 7).

We denote by H_n the Harmonic series defined, for every $n \geq 1$, by $H_n = \sum_{i=1}^{n} 1/i$ and we introduce the function g defined, for all $x > 0$, by $g(x) = x - 1 - \ln(x)$.

Theorem 6 *For any $c \geq 1$, we have*

$$\mathbb{P}\left\{T_{k,n} \geq c\mathbb{E}\left(T_{k,n}\right)\right\} \leq \exp\left(-\frac{(k-1)^2 + (n-k+1)H_{n-k}}{n-1} g(c)\right).$$

For any $c \leq 1$, we have

$$\mathbb{P}\left\{T_{k,n} > c\mathbb{E}\left(T_{k,n}\right)\right\} \geq 1 - \exp\left(-\frac{(k-1)^2 + (n-k+1)H_{n-k}}{n-1} g(c)\right).$$

Proof It is easily checked from Relation (1) that, for every $i = 1, \ldots, n-k$, we have

$$p_{k,n}(i) \geq p_{k,n}(1) = 1 - \prod_{h=1}^{k-1}\left(1 - \frac{1}{n-h}\right) = \frac{k-1}{n-1}.$$

We can now apply Theorem 13 (see Appendix) and deduce that for any $c \geq 1$,

$$\mathbb{P}\left\{T_{k,n} \geq c\mathbb{E}\left(T_{k,n}\right)\right\} \leq \exp\left(-\frac{k-1}{n-1}\mathbb{E}\left(T_{k,n}\right) g(c)\right).$$

Note that $g(c) = c - 1 - \ln(c) \geq 0$ for any $c > 0$. Using Relation (6), that is $\mathbb{E}(T_{k,n}) \geq k - 1 + (n-k+1)H_{n-k}/(k-1)$, we obtain, for any $c \geq 1$,

$$\mathbb{P}\{T_{k,n} \geq c\mathbb{E}(T_{k,n})\} \leq \exp\left(-\frac{(k-1)^2 + (n-k+1)H_{n-k}}{n-1} g(c)\right),$$

which concludes the first part of the proof.
From Theorem 14 (see Appendix), we deduce that for any $c \leq 1$,

$$\mathbb{P}\{T_{k,n} \leq c\mathbb{E}(T_{k,n})\} \leq \exp\left(-\frac{k-1}{n-1}\mathbb{E}(T_{k,n})\, g(c)\right).$$

Again, since $g(c) = c - 1 - \ln(c) \geq 0$ for $c > 0$, the same lower bound of $\mathbb{E}(T_{k,n})$ used for the case $c \geq 1$, yields the second result. ■

When n tends to infinity, we obtain the following result.

Corollary 7 *For every $k \geq 2$, we have*

$$\lim_{n \longrightarrow +\infty} \mathbb{P}\{T_{k,n} > c\mathbb{E}(T_{k,n})\} = \begin{cases} 0 & \text{if } c > 1 \\ 1 & \text{if } c < 1. \end{cases}$$

Proof First, note that $g(c) = c - 1 - \ln(c) > 0$ for any $c \in (0,1) \cup (1,\infty)$ and that $\lim_{n \longrightarrow +\infty} H_{n-k} = +\infty$. Hence,

$$\lim_{n \longrightarrow \infty} \exp\left(-\frac{(k-1)^2 + (n-k)H_{n-k}}{n-1} g(c)\right) = 0.$$

Applying Theorem 6 completes the proof. ■

One can observe that for any $c \neq 1$,

$$\exp\left(-\frac{(k-1)^2 + (n-k)H_{n-k+1}}{n-1}(c - 1 - \ln(c))\right) \underset{n \longrightarrow \infty}{\sim} 1/n^{c-1-\ln(c)}$$

which indicates that, for large values of n, the distribution of $T_{k,n}$ becomes closer to its mean at a speed of $1/n^{c-1-\ln(c)}$.
When $c = 1$, Corollary 7 does not allow us to figure out neither the existence of $\lim_{n \to +\infty} \mathbb{P}\{T_{k,n} \geq \mathbb{E}(T_{k,n})\}$ nor its value. We deal with this case in a more general situation which consists in analyzing the limiting distribution of $(T_{k,n} - \mathbb{E}(T_{k,n}))/n$. We first need the following lemma.

Lemma 8 *For every $m \geq 1$ and $n \geq m + k$, we have*

$$\frac{1}{n^2} \sum_{i=m}^{n-k} \mathbb{V}ar(S_{k,n}(i)) \leq \sum_{i=m}^{n-k} \frac{1}{i^2} \leq \frac{\pi^2}{6}.$$

Proof Since $1 \leq i \leq n-k$, all the terms $1 - i/(n-h)$ in Relation (1) are less than one. We thus have $p_{k,n}(i) \geq 1 - (1 - i/(n-1)) = i/(n-1)$. Using Relation (7), we get

$$\frac{1}{n^2}\sum_{i=m}^{n-k} \mathbb{V}ar(S_{k,n}(i)) = \frac{1}{n^2}\sum_{i=m}^{n-k} \frac{1-p_{k,n}(i)}{(p_{k,n}(i))^2} \leq \frac{1}{n^2}\sum_{i=m}^{n-k} \frac{1}{(p_{k,n}(i))^2}$$
$$\leq \frac{(n-1)^2}{n^2}\sum_{i=m}^{n-k} \frac{1}{i^2} \leq \frac{\pi^2}{6},$$

which completes the proof. ■

Theorem 9 *Let $(Z_i)_{i\geq 1}$ be a sequence of i.i.d. random variables exponentially distributed with rate 1 and let W be defined by*

$$W = \frac{1}{k-1}\sum_{i=1}^{\infty} \frac{Z_i - 1}{i}.$$

We then have

$$\frac{T_{k,n} - \mathbb{E}(T_{k,n})}{n} \xrightarrow{\mathcal{L}} W \text{ as } n \longrightarrow \infty$$

and, for all $x \in \mathbb{R}$,

$$\lim_{n\longrightarrow\infty} \mathbb{P}\left\{\frac{T_{k,n} - \mathbb{E}(T_{k,n})}{n} \leq x\right\} = e^{-e^{-(k-1)x-\gamma}},$$

where γ is the Euler-Mascheroni constant given by $\gamma \approx 0.5772156649$.

Proof For each fixed i, we have $\lim_{n\longrightarrow\infty} p_{k,n}(i) = 0$. It follows that for every $x \geq 0$, we have

$$\mathbb{P}\{p_{k,n}(i)S_{k,n}(i) > x\} = \mathbb{P}\{S_{k,n}(i) > x/p_{k,n}(i)\}$$
$$= n(1 - p_{k,n}(i))^{\lfloor x/p_{k,n}(i)\rfloor},$$

which tends to e^{-x} when n tends to infinity. If Z_i is a random variable exponentially distributed with rate 1, we have shown that

$$p_{k,n}(i)S_{k,n}(i) \xrightarrow{\mathcal{L}} Z_i \text{ as } n \longrightarrow \infty.$$

Moreover since the $(S_{k,n}(i))_{i=1,\ldots,n-k}$ are independent, the $(Z_i)_{i\geq 1}$ are also independent.

Observing now that for each fixed i, we have, from Relation (1),

$$p_{k,n}(i) = i\sum_{h=1}^{k-1} \frac{1}{n-h} + o(1/n),$$

we obtain $\lim_{n\longrightarrow\infty} np_{k,n}(i) = (k-1)i$. Defining $R_{k,n}(i) = S_{k,n}(i) - \mathbb{E}(S_{k,n}(i))$ we obtain, since $\mathbb{E}(S_{k,n}(i)) = 1/p_{k,n}(i)$,

$$\begin{aligned}\frac{R_{k,n}(i)}{n} &= \frac{S_{k,n}(i) - \mathbb{E}(S_{k,n}(i))}{n} = \frac{p_{k,n}(i)S_{k,n}(i) - 1}{np_{k,n}(i)} \\ &\xrightarrow{\mathcal{L}} \frac{Z_i - 1}{(k-1)i} \text{ as } n \longrightarrow \infty. \end{aligned} \tag{9}$$

The rest of the proof consists in checking the hypothesis of the principle of accompanying laws of Theorem 3.1.14 of [29]. Introducing the random variables V_n and $V_{m,n}$ defined by

$$V_n = \frac{T_{k,n} - \mathbb{E}(T_{k,n})}{n} = \frac{1}{n}\sum_{i=1}^{n-k} R_{k,n}(i) \text{ and } V_{m,n} = \frac{1}{n}\sum_{i=1}^{m-1} R_{k,n}(i), \tag{10}$$

we obtain, using the fact that $\mathbb{E}(R_{k,n}(i)) = 0$ and that the $R_{k,n}(i)$ are independent,

$$\begin{aligned}\mathbb{E}((V_n - V_{m,n})^2) &= \mathbb{E}\left(\left[\frac{1}{n}\sum_{i=m}^{n-k} R_{k,n}(i)\right]^2\right) = \mathbb{V}ar\left(\frac{1}{n}\sum_{i=m}^{n-k} R_{k,n}(i)\right) \\ &= \frac{1}{n^2}\sum_{i=m}^{n-k} \mathbb{V}ar(R_{k,n}(i)) = \frac{1}{n^2}\sum_{i=m}^{n-k} \mathbb{V}ar(S_{k,n}(i)).\end{aligned}$$

Using now Lemma 8, we get

$$\mathbb{E}((V_n - V_{m,n})^2) \leq \sum_{i=m}^{n-k} \frac{1}{i^2},$$

that is

$$\lim_{m\longrightarrow\infty} \limsup_{n\longrightarrow\infty} \mathbb{E}((V_n - V_{m,n})^2) \leq \lim_{m\longrightarrow\infty} \sum_{i=m}^{\infty} \frac{1}{i^2} = 0.$$

Using now the Markov inequality, we obtain, for all $\varepsilon > 0$,

$$\mathbb{P}\{|V_n - V_{m,n}| \geq \varepsilon\} = \mathbb{P}\{(V_n - V_{m,n})^2 \geq \varepsilon^2\} \leq \frac{\mathbb{E}((V_n - V_{m,n})^2)}{\varepsilon^2}.$$

Putting together these results, we have shown that for all $\varepsilon > 0$, we have

$$\lim_{m\longrightarrow\infty} \limsup_{n\longrightarrow\infty} \mathbb{P}\{|V_n - V_{m,n}| \geq \varepsilon\} = 0. \tag{11}$$

Let us introduce the notation

$$W_m = \frac{1}{k-1}\sum_{i=1}^{m-1} \frac{Z_i - 1}{i}.$$

Using (9) and the fact that the $R_{k,n}(i)$ are independent, we have

$$V_{m,n} \xrightarrow{\mathcal{L}} W_m \text{ as } n \longrightarrow \infty. \tag{12}$$

The hypothesis of the principle of accompanying laws of Theorem 3.1.14 of [29] are properties (9) and (12). We can thus conclude that

$$V_n \xrightarrow{\mathcal{L}} W \text{ as } n \longrightarrow \infty.$$

To obtain the distribution of W, we recall that Louis Gordon has proved in [17] that

$$-\gamma + \sum_{i=1}^{+\infty} \frac{1 - Z_i}{i} \overset{\mathcal{L}}{=} \ln(Z_1),$$

where γ is the Euler-Mascheroni constant. Thus, by definition of W, we have

$$W \overset{\mathcal{L}}{=} -\frac{\gamma + \ln(Z_1)}{k-1}.$$

It follows that, for all $x \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{P}\{W \leq x\} &= \mathbb{P}\{\ln(Z_1) > -(k-1)x - \gamma\} \\ &= \mathbb{P}\left\{Z_1 > e^{-(k-1)x-\gamma}\right\} = e^{-e^{-(k-1)x-\gamma}}, \end{aligned}$$

which completes the proof. ■

In particular, by taking $x = 0$ in Theorem 9, we obtain

$$\lim_{n \longrightarrow \infty} \mathbb{P}\{T_{k,n} \leq \mathbb{E}(T_{k,n})\} = e^{-e^{-\gamma}} \approx 0.5703760017,$$

which solves the case $c = 1$ of Corollary 7.

5 Comparison of the pull, push and push-pull protocols

In this section, we compare the spreading time of the k-pull protocol with the standard 2-push-pull, 2-pull and 2-push ones. We summarize in Table 1 some characteristics of the spreading time distribution of each protocol. The notation $\sum_{i=1}^{n-1} \mathcal{G}(p_i)$ represents the sum of independent random variables geometrically distributed with parameter p_i for the i-th variable. Note that the spreading time distributions of the 2-push and 2-pull protocols are the same. In addition, the mean spreading times of all the standard protocols (2-push, 2-pull and 2-push-pull) are the same.

In this section, we denote by $T_{2,n}^{push}$, $T_{2,n}^{push-pull}$ and $T_{k,n}^{pull}$, for $k \geq 2$, the spreading time associated with respectively the 2-push, 2-push-pull and k-pull protocols. As recalled in the introduction, in the 2-push protocol, an informed node contacts some random node and sends it the rumor (i.e. triggers a push operation), while in the 2-pull protocol, an uninformed node contacts some

	$T_{k,n}$ distribution	$\mathbb{E}\left(T_{k,n}\right)$
2-push-pull	$\sum_{i=1}^{n-1} \mathcal{G}\left(\frac{2i(n-i)}{n(n-1)}\right)$	$(n-1)H_{n-1} \underset{n\longrightarrow\infty}{\sim} n\ln n$
2-push	$\sum_{i=1}^{n-1} \mathcal{G}\left(\frac{n-i}{n-1}\right)$	$(n-1)H_{n-1} \underset{n\longrightarrow\infty}{\sim} n\ln n$
2-pull	$\sum_{i=1}^{n-1} \mathcal{G}\left(\frac{i}{n-1}\right)$	$(n-1)H_{n-1} \underset{n\longrightarrow\infty}{\sim} n\ln n$
k-pull	$k-1+\sum_{i=1}^{n-k} \mathcal{G}\left(p_{k,n}(i)\right)$	(5) $\underset{n\longrightarrow\infty}{\sim} n\ln(n)/(k-1)$

Table 1 Spreading time distributions of the 2-push-pull, 2-push and k-pull protocols. A detailed analysis of the 2-push and 2-push-pull protocols are respectively provided in [22] and [20].

random node and asks for the rumor (i.e. triggers a pull operation). Finally, in the 2-push-pull protocol, the same node can perform both push and pull operations according to whether it knows or not the rumor.

To compare the spreading time distributions of each protocol, we use stochastic dominance tools. We recall the following definition (see for instance [13]).

Definition 10 (Stochastic dominance definition) *Let X and Y two independent real random variables.*

a) X strictly stochastically dominates Y if for all x, we have $\mathbb{P}\{X > x\} > \mathbb{P}\{Y > x\}$. We then write $X \overset{s.t.}{\succ} Y$.

b) X stochastically dominates Y if for all x we have $\mathbb{P}\{X > x\} \geq \mathbb{P}\{Y > x\}$. We then write $X \overset{s.t.}{\succeq} Y$.

Comparing the spreading time distributions of each protocol amounts in comparing sums of geometric distributions. We thus first start by the following lemma which is used to prove the next theorem.

Lemma 11 (Stochastic dominance for geometric distributions) *Let G_1 and G_2 be two independent geometric random variables with parameters g_1 and g_2 respectively. If $g_1 > g_2$ then G_1 is strictly stochastically dominated by G_2.*

Proof Since $g_1 > g_2$, we have $(1-g_1)^\ell < (1-g_2)^\ell$ for every integer $\ell \geq 1$. Hence,

$$\mathbb{P}\left\{G_1 > \ell\right\} = (1-g_1)^\ell < (1-g_2)^\ell = \mathbb{P}\left\{G_2 > \ell\right\},$$

which implies that G_1 is strictly stochastically dominated by G_2. ■

Theorem 12 *For all $n \geq 4$ and $2 \leq k \leq n$, we have*

$$T_{k,n}^{pull} \overset{s.t.}{\prec} T_{k-1,n}^{pull} \overset{s.t.}{\prec} \ldots \overset{s.t.}{\prec} T_{3,n}^{pull} \overset{s.t.}{\prec} T_{2,n}^{pull} \overset{\mathcal{L}}{=} T_{2,n}^{push}$$

and

$$T_{3,n}^{pull} \overset{s.t.}{\prec} T_{2,n}^{push-pull}.$$

Proof We first show that $p_{k,n}(i) > p_{k-1,n}(i)$ for all i. For all $i = 1, \ldots, n-k$, we have

$$1 - p_{k,n}(i) = \prod_{h=1}^{k-1} \left(1 - \frac{i}{n-h}\right) < \prod_{h=1}^{k-2} \left(1 - \frac{i}{n-h}\right) = 1 - p_{k-1,n}(i)$$

which implies that for all $i = 1, \ldots, n-k$, $p_{k,n}(i) > p_{k-1,n}(i)$. If $i = n-k+1$, $p_{k-1,n}(n-k+1) < 1$ and $p_{k,n}(n-k+1) = 1$.

Applying Proposition 11, we deduce that $\mathcal{G}(p_{k,n}(i)) \overset{s.t.}{\prec} \mathcal{G}(p_{k-1,n}(i))$ for all $i = 1, \ldots, n-k$. Summing for $i = 1$ to $n-k+1$, we have

$$\sum_{i=1}^{n-k+1} \mathcal{G}\left(p_{k,n}(i)\right) \overset{s.t.}{\prec} \sum_{i=1}^{n-k+1} \mathcal{G}\left(p_{k-1,n}(i)\right).$$

Adding the quantity $k-2$ to both terms and since $p_{k,n}(n-k+1) = 1$, we obtain

$$T_{k,n}^{pull} \overset{s.t.}{\prec} T_{k-1,n}^{pull} \overset{s.t.}{\prec} \ldots \overset{s.t.}{\prec} T_{2,n}^{pull}.$$

Note that from Table 1, the random variables $T_{2,n}^{pull}$ and $T_{2,n}^{push}$ have the same distribution.

We turn now to the second part of the proof by comparing of the parameters of both distributions $T_{3,n}^{pull}$ and $T_{2,n}^{push-pull}$, that is $p_{3,n}(i)$, which is the parameter of the i-th geometric distribution of $T_{3,n}^{pull}$, with $2i(n-i)/(n(n-1))$, which is the parameter of the i-th geometric distribution of $T_{2,n}^{push-pull}$. For all $i = 1, \ldots, n-1$, we have

$$\begin{aligned} p_{3,n}(i) - \frac{2i(n-i)}{n(n-1)} &= 1 - \prod_{h=1}^{2} \left(1 - \frac{i}{n-h}\right) \\ &= 1 - \left(1 - \frac{i}{n-1}\right)\left(1 - \frac{i}{n-2}\right) - \frac{2i(n-i)}{n(n-1)} \\ &= \frac{i(n + i(n-4))}{n(n-1)(n-2)}. \end{aligned}$$

It is obvious that $n + i(n-4) \geq 0$ for $n \geq 4$. Hence, we deduce from Lemma 11 that for all $i = 1, \ldots, n-1$,

$$\mathcal{G}\left(p_{3,n}(i)\right) \overset{s.t.}{\prec} \mathcal{G}\left(\frac{2i(n-i)}{n(n-1)}\right).$$

Summing for $i = 1, \ldots, n-1$, we conclude that $T_{3,n}^{pull} \overset{s.t.}{\prec} T_{2,n}^{push-pull}$, which completes the proof. ■

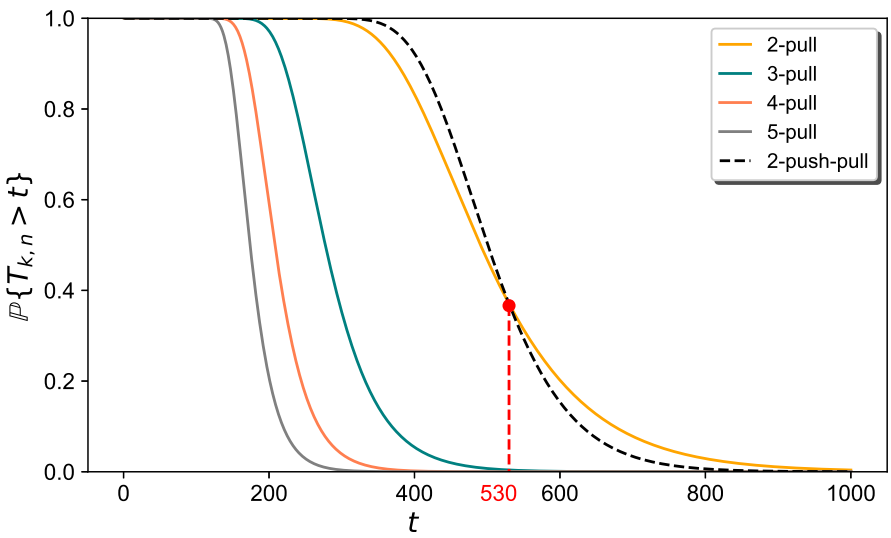

Fig. 1 Stochastic dominance illustration. Applying the recursion detailed in Relation 8, we compute the distribution $\mathbb{P}\left\{T_{k,n} > t\right\}$ for the 2-push, 2-push-pull and k-pull protocols. The total number of nodes has been set to $n = 100$.

Theorem 12 shows that the k-pull protocol requires significantly less operations than the other standard protocols.

Figure 1 illustrates the fact that $T_{2,n}^{pull}$ (or equivalently $T_{2,n}^{push}$) and $T_{2,n}^{push-pull}$ cannot be stochastically ordered. Indeed, this figure shows that there is a threshold instant $t(n)$ such that

$$\mathbb{P}\{T_{2,n}^{pull} > t\} < \mathbb{P}\{T_{2,n}^{push-pull} > t\} \text{ for } \ t < t(n)$$
$$\mathbb{P}\{T_{2,n}^{pull} > t\} > \mathbb{P}\{T_{2,n}^{push-pull} > t\} \text{ for } \ t > t(n).$$

In this figure, for which we have $n = 100$, the threshold instant $t(100)$ is equal to 530. This phenomenon can be explained as follows.

On the one hand, for $i = 1, \ldots, \lfloor n/2 \rfloor - 1$ we have $p_{2,n}(i) = i/(n-1) < 2i(n-i)/(n(n-1))$ and for $i = \lfloor n/2 \rfloor$ we have $p_{2,n}(i) = 2i(n-i)/(n(n-1))$. It follows from Theorem 12 that

$$\sum_{i=1}^{\lfloor n/2 \rfloor} \mathcal{G}(p_{2,n}(i)) \overset{s.t.}{\succ} \sum_{i=1}^{\lfloor n/2 \rfloor} \mathcal{G}\left(2\frac{i(n-i)}{n(n-1)}\right).$$

This means that the time needed for $\lfloor n/2 \rfloor$ nodes to learn the rumor is stochastically smaller when using the 2-push-pull protocol instead of the 2-pull protocol.

On the other hand, for $i = \lfloor n/2 \rfloor + 1, \ldots, n-1$ we have $p_{2,n}(i) > 2i(n-i)/(n(n-1))$. It follows from Theorem 12 that

$$\sum_{i=\lfloor n/2 \rfloor+1}^{n-1} \mathcal{G}(p_{2,n}(i)) \overset{s.t.}{\prec} \sum_{i=\lfloor n/2 \rfloor+1}^{n-1} \mathcal{G}\left(2\frac{i(n-i)}{n(n-1)}\right),$$

which means that the time needed for all the n nodes to learn the rumor, starting initially with $\lfloor n/2 \rfloor + 1$ nodes knowing the rumor, is stochastically smaller when using the 2-pull protocol instead of the 2-push-pull protocol.

6 Conclusion

In this paper, we have proposed a new rumor spreading protocol that allows each node to asynchronously interact with $k-1$ other nodes during each operation. We have analyzed its limiting behavior when the number of nodes goes to infinity and we have shown that it generalizes the standard (i.e. $k = 2$) pull protocol and improves it when $k > 2$. Further research would allow us to manage competing rumours more finely. For instance, the initiator of the k-pull operation might take advantage of this interaction scheme to decide which rumors to learn when different rumors are allowed to compete. Such a problem has a great impact in the context of blockchain protocols. Another interesting research direction would be to generalize the 2-push and 2-push-pull protocols to the corresponding k-push and k-push-pull protocols.

7 Appendix

Let $X_1, \ldots, X_n$ be n independent geometric random variables with possibly distinct parameters, i.e. such that $X_i \sim \mathcal{G}(p_i)$ with $p_i \in (0,1]$. Let $X = X_1 + \cdots + X_n$, $\mu = \mathbb{E}(X)$ and $p_* = \min_{i=1,\ldots,n} p_i$. We then have the following results which have been proved in [18].

Theorem 13 *For any $p_1, \ldots, p_n \in (0,1]$ and any $\lambda \geq 1$,*

$$\mathbb{P}\{X \geq \lambda\mu\} \leq e^{-p_*\mu(\lambda-1-\ln(\lambda))}.$$

Theorem 14 *For any $p_1, \ldots, p_n \in (0,1]$ and any $\lambda \leq 1$,*

$$\mathbb{P}\{X \leq \lambda\mu\} \leq e^{-p_*\mu(\lambda-1-\ln(\lambda))}.$$

References

1. Acan, H., Collevecchio, A., Mehrabian, A., Wormald, N.: On the push & pull protocol for rumour spreading. Trends in Mathematics **6**, 3–10 (2017)
2. Chierichetti, F., Lattanzi, S., Panconesi, A.: Rumor spreading in social networks. Theoretical Computer Science **412**(24), 2602–2610 (2011)

3. Clementi, A., Crescenzi, P., Doerr, C., Fraigniaud, P., Pasquale, F., Silvestri, R.: Rumor spreading in random evolving graphs. Random structures and Algorithms **48**(2), 290–312 (2015)
4. Daum, S., Kuhn, F., Maus, Y.: Rumor spreading with bounded indegree. In: Proceedings of the International Colloquium on Structural Information and Communication Complexity (SIROCCO) (2016)
5. Demers, A.J., Greene, D.H., Hauser, C.H., Irish, W., Larson, J., Shenker, S.J., Sturgis, H.E., Swinehart, D.C., Terry, D.B.: Epidemic algorithms for replicated database maintenance. PODC 87 pp. 1–12 (1987)
6. Doerr, B., Fouz, M., Friedrich, T.: Asynchronous rumor spreading in preferential attachment graphs. SWAT 2012 pp. 307–315 (2012)
7. Doerr, B., Fouz, M., Friedrich, T.: Experimental analysis of rumor spreading in social networks. MedAlg 2012 pp. 159–173 (2012)
8. Doerr, B., Kostrygin, A.: Randomized Rumor Spreading Revisited. In: I. Chatzigiannakis, P. Indyk, F. Kuhn, A. Muscholl (eds.) ICALP 2017, vol. 80, pp. 138:1–138:14. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik (2017). DOI 10.4230/LIPIcs.ICALP.2017.138
9. Duan, Z., Gopalan, K., Dong, Y.: Probabilistic reliable dissemination in large-scale systems. SRUTI (2005)
10. Feige, F., Peleg, D., Raghavan, P., Upfal, E.: Randomized broadcast in networks. Random Structures and Algorithms **1**(4), 447–460 (1990)
11. Fountoulakis, N., Panagiotou, K.: Rumor spreading on random regular graphs and expanders. Random Structures and Algorithms **43**(2), 201–220 (2013)
12. Frieze, A., Grimmet, G.: The shortest-path problem for graphs with random arc-lengths. Discrete Applied Mathematics **10**(1), 57–77 (1985)
13. Ganesh, A.J.: Rumour spreading on graphs. Technical report:https://people.maths.bris.ac.uk/ maajg/teaching/complexnets/rumours.pdf (2015)
14. Giakkoupis, G.: Tight bounds for rumor spreading in graphs of a given conductance. In: Proceedings of the International Symposium on Theoretical Aspects of Computer Science (STACS) (2011)
15. Giakkoupis, G.: Tight bounds for rumor spreading with vertex expansion. In: Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) (2014)
16. Giakkoupis, G., Nazari, Y., Woelfel, P.: How asynchrony affects rumor spreading time. PODC '16 pp. 185–194 (2016)
17. Gordon, L.: Bounds for the distribution of the generalized variance. The Annals of Statistics **17**(4), 1684–1692 (1989)
18. Janson, S.: Tail bounds for sums of geometric and exponential variables. Statistics & Probability Letters **135**, 1–6 (2018)
19. Karp, R., Schindelhauer, C., Shenker, S., Vocking, B.: Randomized rumor spreading. In: Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS) (2000)
20. Mocquard, Y., Robert, S., Sericola, B., Anceaume, E.: Analysis of the propagation time of a rumour in large-scale distributed systems. NCA 2016 (2016)
21. Moreno, Y., Nekovee, M., Pacheco, A.F.: Dynamics of rumor spreading in complex networks. Phys. Rev. **69**, 248–258 (2004)
22. Nazari, Y.: Analysis of asynchronous and synchronous rumor spreading protocols. Master's thesis, University of Calgary (2016). http://hdl.handle.net/11023/3136
23. Panagiotou, K., Perez-Gimenez, X., Sauerwald, T., Sun, H.: Randomized rumor spreading: the effect of the network topology. Combinatorics, Probability and Computing **24**(2), 457–479 (2015)
24. Panagiotou, K., Pourmiri, A., Sauerwald, T.: Faster rumor spreading with multiple calls. The electronic journal of combinatorics **22** (2015)
25. Pittel, B.: On spreading a rumor. SIAM Journal on Applied Mathematics **47**(1), 213–223 (1987)
26. Pourmiri, A., Ramezani, F.: Brief announcement: Ultra-fast asynchronous randomized rumor spreading. SPAA 2019 (2019)
27. Sanghavi, S., Hajek, B., Massouli, L.: Gossiping with multiple messages. IEEE Transactions on Information Theory **53**(123) (2007)

28. Sericola, B.: Markov Chains. Theory, Algorithms and Applications. John Wiley & Sons (2013)
29. Stroock, D.W.: Probability Theory: An Analytic View, second edn. Cambridge University Press (2010)
30. Yao, G., Bi, J., Wang, S., Zhang, Y., Li, Y.: A pull model IPv6 duplicate address detection. LCN 2010 (2010)