Noisy rumor spreading and plurality consensus

Pierre Fraigniaud, Emanuele Natale

Noisy Rumor Spreading and Plurality Consensus

Pierre Fraigniaud

IRIF, CNRS and University Paris Diderot
pierref@irif.fr

Emanuele Natale

Max Planck Institute for Informatics
emanuele.natale@mpi-inf.mpg.de

Abstract

Error-correcting codes are efficient methods for handling *noisy* communication channels in the context of technological networks. However, such elaborate methods differ a lot from the unsophisticated way biological entities are supposed to communicate. Yet, it has been recently shown by Feinerman, Haeupler, and Korman [PODC 2014] that complex coordination tasks such as *rumor spreading* and *majority consensus* can plausibly be achieved in biological systems subject to noisy communication channels, where every message transferred through a channel remains intact with small probability $\frac{1}{2}+\epsilon$, without using coding techniques. This result is a considerable step towards a better understanding of the way biological entities may cooperate. It has nevertheless been established only in the case of 2-valued *opinions*: rumor spreading aims at broadcasting a single-bit opinion to all nodes, and majority consensus aims at leading all nodes to adopt the single-bit opinion that was initially present in the system with (relative) majority. In this paper, we extend this previous work to k-valued opinions, for any constant $k \geq 2$.

Our extension requires to address a series of important issues, some conceptual, others technical. We had to entirely revisit the notion of noise, for handling channels carrying *k-valued* messages. In fact, we precisely characterize the type of noise patterns for which plurality consensus is solvable. Also, a key result employed in the bivalued case by Feinerman et al. is an estimate of the probability of observing the most frequent opinion from observing the mode of a small sample. We generalize this result to the multivalued case by providing a new analytical proof for the bivalued case that is amenable to be extended, by induction, and that is of independent interest.

1 Introduction

1.1 Context and Objective

To guarantee reliable communication over a network in the presence of noise is the main goal of Network Information Theory [24]. Thanks to the achievements of this theory, the impact of noise can often be drastically reduced to almost zero by employing *error-correcting codes*, which are practical methods whenever dealing with artificial entities. However, as observed in [27], the situation is radically different for scenarios in which the computational entities are biological. Indeed, from a biological perspective, a computational process can be considered "simple" only if it consists of very basic primitive operations, and is extremely lightweight. As a consequence, it is unlikely that biological entities are employing techniques like error-correcting codes to reduce the impact of noise in communications between them. Yet, biological signals are subject to noise, when generated, transmitted, and received. This rises the intriguing question of how entities in biological ensembles can cooperate in presence of noisy communications, but in absence of mechanisms such as error-correcting codes.

An important step toward understanding communications in biological ensembles has been achieved recently in [27], which showed how it is possible to cope with noisy communications in absence of coding mechanisms for solving complex tasks such as *rumor-spreading* and *majority consensus*. Such a result provides highly valuable hints on how complex tasks can be achieved in frameworks such as the immune system, bacteria populations, or super-organisms of social insects, despite the presence of noisy communications.

In the case of rumor-spreading, [27] assumes that a source-node initially handles a bit, set to some binary value, called the *correct opinion*. This opinion has to be transmitted to all nodes, in a noisy environment, modeled as a complete network with unreliable links. More precisely, messages are transmitted in the network according to the classical *uniform push model* [19, 32, 38] where, at each round, every node can send one binary opinion to a neighbor chosen uniformly and independently at random and, before reaching the receiver, that opinion is flipped with probability at most $\frac{1}{2} - \epsilon$ with $\epsilon > 0$. It is proved that, even in this very noisy setting, the rumor-spreading problem can be solved quite efficiently. Specifically, [27] provides an algorithm that solves the noisy rumor-spreading problem in $O(\frac{1}{\epsilon^2} \log n)$ communication rounds, with high probability[1] (w.h.p.) in n-node networks, using $O(\log \log n + \log(1/\epsilon))$ bits of memory per node[2]. Again, this algorithm exchanges solely opinions between nodes.

In the case of majority consensus, [27] assumes that some nodes are supporting opinion 0, some nodes are supporting opinion 1, and some other nodes

[1] A series of events $\mathcal{E}_n$, $n \geq 1$, hold w.h.p. if $\Pr(\mathcal{E}_n) \geq 1 - O(1/n^c)$ for some $c > 0$.

[2] We remark that, while it would be more appropriate to measure the space complexity by the *number of states* here (in accordance with other work which is concerned with minimizing it, such as population protocols [4] or cellular automata [35]), we make use of the *memory bits* for consistency with the main related work [27].

are supporting no opinion. The objective is that all nodes eventually support the initially most frequent opinion (0 or 1). More precisely, let A be the set of nodes with opinion, and let $b \in \{0,1\}$ be the majority opinion in A. The *majority bias* of A is defined as $\frac{1}{2}(|A_b| - |A_{\bar{b}}|)/|A|$ where A_i is the set of nodes with opinion $i \in \{0,1\}$. In the very same noisy communication model as above, [27] provides an algorithm that solves the noisy majority consensus problem for $|A| = \Omega(\frac{1}{\epsilon^2}\log n)$ with majority-bias $\Omega(\sqrt{\log n/|A|})$. The algorithm runs in $O(\frac{1}{\epsilon^2}\log n)$ rounds, w.h.p., in n-node networks, using $O(\log\log n + \log(1/\epsilon))$ bits of memory per node. As for the case of rumor spreading, the algorithm exchanges solely opinions between nodes. In fact, the latter algorithm for majority consensus is used as a subroutine for solving the rumor-spreading problem. Note that the majority consensus algorithm of [27] requires that the nodes are initially aware of the size of A.

According to [27], both algorithms are optimal, since both rumor-spreading and majority consensus require $\Omega(\frac{1}{\epsilon^2}\log n)$ rounds w.h.p. in n-node networks.

Our objective is to extend the work of [27] to the natural case of an arbitrary number of opinions, to go beyond a proof of concept. The problem that results from this extension is an instance of the *plurality consensus* problem in the presence of noise, i.e., the problem of making the system converging to the initially most frequent opinion (i.e., the *plurality* opinion). Indeed, the plurality consensus problem naturally arises in several biological settings, typically for choosing between different directions for a flock of birds [11], different speeds for a school of fish [42], or different nesting sites for ants [30]. The computation of the most frequent value has also been observed in biological cells [15].

1.2 Our Contribution

1.2.1 Our Results

We generalize the results in [27] to the setting in which an arbitrary large number k of opinions is present in the system. In the context of rumor spreading, the correct opinion is a value $i \in \{1,\ldots,k\}$, for any constant $k \geq 2$. Initially, one node supports this opinion i, and the other nodes have no opinions. The nodes must exchange opinions so that, eventually, all nodes support the correct opinion i. In the context of (relative) majority consensus, also known as *plurality consensus*, each node u initially supports one opinion $i_u \in \{1,\ldots,k\}$, or has no opinion. The objective is that all nodes eventually adopt the *plurality opinion* (i.e., the opinion initially held by more nodes than any other, but not necessarily by an overall majority of nodes). As in [27], we restrict ourselves to "natural" algorithms [16], that is, algorithms in which nodes only exchange opinions in a straightforward manner (i.e., they do not use the opinions to encode, e.g., part of their internal state). For both problems, the difficulty comes from the fact that every opinion can be modified during its traversal of any link, and switched at random to any other opinion. In short, we prove that there are algorithms solving the noisy rumor spreading problem and the noisy plurality consensus problem for multiple opinions, with the same performances and probabilistic

guarantees as the algorithms for binary opinions in [27].

1.2.2 The Technical Challenges

Generalizing noisy rumor spreading and noisy majority consensus to more than just two opinions requires to address a series of issues, some conceptual, others technical.

Conceptually, one needs first to redefine the notion of noise. In the case of binary opinions, the noise can just flip an opinion to its complement. In the case of multiple opinions, an opinion i subject to a modification is switched to another opinion i', but there are many ways of picking i'. For instance, i' can be picked uniformly at random (u.a.r.) among all opinions. Or, i' could be picked as one of the "close opinions", say, either $i+1$ or $i-1$ modulo k. Or, i' could be "reset" to, say, $i = 1$. In fact, there are very many alternatives, and not all enable rumor spreading and plurality consensus to be solved. One of our contributions is to characterize *noise matrices* $P = (p_{i,j})$, where $p_{i,j}$ is the probability that opinion i is switched to opinion j, for which these two problems are efficiently solvable. Similar issues arise for, e.g., redefining the majority bias into a *plurality bias*.

The technical difficulties are manifold. A key ingredient of the analysis in [27] is a fine estimate of how nodes can mitigate the impact of noise by observing the opinions of *many* other nodes, and then considering the mode of such sample. Their proof relies on the fact that for the binary opinion case, given a sample of size γ, the number of 1s and 0s in the sample sum up to γ. Even for the ternary opinion case, the additional *degree of freedom* in the sample radically changes the nature of the problem, and the impact of noise is statistically far more difficult to handle.

Also, to address the multivalued case, we had to cope with the fact that, in the uniform push model, the messages received by nodes at every round are correlated. To see why, consider an instance of the system in which a certain opinion b is held by one node only, and there is no noise at all. In one round, only one other node can receive b. It follows that if a certain node u has received b, no other nodes have received it. Thus, the messages each node receives are not independent. In [26] (conference version of [27]), probability concentration results are claimed for random variables (r.v.) that depend on such messages, using Chernoff bounds. However, Chernoff bounds have been proved to hold only for random variables whose stochastic dependence belongs to a very limited class (see for example [22]). In [27], it is pointed out that the binary random variables on which the Chernoff bound is applied satisfy the property of being negatively 1-correlated (see Section 1.7 in [27] for a formal definition). In our analysis, we show instead how to obtain concentration of probability in this dependent setting by leveraging Poisson approximation techniques. Our approach has the following advantage: instead of showing that the Chernoff bound can be directly applied to the specific involved random variables, we show that the execution of the given protocol, on the uniform push model, can be tightly approximated with the execution of the same protocol on a another suitable communication

model, that is not affected by the stochastic correlation that affects the uniform push model.

1.3 Other Related Work

By extending the work of [27], we contribute to the theoretical understanding of how communications and interactions function in biological distributed systems, from an algorithmic perspective [5, 7, 3, 2, 4, 34, 13, 14]. We refer the reader to [27] for a discussion on the computational aspects of biological distributed systems, an overview of the rumor spreading problem in distributed computing, as well as its biological significance in the presence of noise. In this section, we mainly discuss the previous technical contributions from the literature related to the novelty of our work, that is the extension to the case of several different opinions.

We remark that, in the following, we say that a protocol solves a problem within a given time if a correct solution is achieved with high probability within said time. In the context of population protocol, the problem of achieving majority consensus in the binary case has been solved by employing a simple protocol called *undecided-state dynamic* [4]. In the uniform push model, the binary majority consensus problem can be solved very efficiently as a consequence of a more general result about computing the median of the initial opinions [20]. Still in the uniform push model, the undecided state dynamic has been analyzed in the case of an arbitrarily large number of opinions, which may even be a function of the number of agents in the system [9]. A similar result has been obtained for another elementary protocol, so-called *3-majority dynamics*, in which, at each round, each node samples the opinion of three random nodes, and adopts the most frequent opinion among these three [10]. The 3-majority dynamics has also been shown to be fault-tolerant against an adversary that can change up to $O(\sqrt{n})$ agents at each round [20, 10]. Other work has analyzed the undecided-state dynamics in asynchronous models with a constant number of opinions [21, 31, 37], and the h-majority dynamics (or slight variations of it) on different graph classes in the uniform push model [17, 1]. The analysis of the undecided-state dynamics in [9] has been followed by a series of work which have used it to design optimal plurality consensus algorithms in the uniform pull model [28, 29].

A general result by Kempe et al. [33] shows how to compute a large class of functions in the uniform push model. However, their protocol requires the nodes to send slightly more complex messages than their sole current opinion, and its effectiveness heavily relies on a potential function argument that does not hold in the presence of noise.

To the best of our knowledge, we are the first considering the plurality consensus problem in the presence of noise.

2 Model and Formal Statement of our Results

In this section we formally define the communication model, the main definitions, the investigated problems and our contribution to them.

As discussed in Section 1.1, intuitively we look for protocols that are simple enough to be plausible communication strategies for primitive biological system. We believe that the computational investigation regarding biologically-feasible protocols is still too premature for a reasonable attempt to provide a *general* formal definition of what constitutes a *biologically feasible computation*. Hence, in the following we restrict our attention solely on the biological significance of the rumor-spreading and plurality consensus problems and the corresponding protocols that we consider.

Regarding the problems of multivalued rumor spreading and plurality consensus, while for practical reasons many experiments on collective behavior have been designed to investigate the binary-decision setting, the considered natural phenomena usually involve a decision among a large number of different options [18]: famous examples in the literature include cockroaches aggregating in a common site [8], and the house-hunting process of ant colonies when seeking a new site to relocate their nest [25] or of honeybee swarms when a portion of a strong colony branches from it in order to start a new one [40, 41]. Therefore, it is natural to ask what trade-offs and constraints are required by the extension of the results in [27] to the multivalued case.

Regarding the solution we consider, as illustrated in sections 2.3 and 3.1, we consider a natural generalization of the protocol given in [27], which is essentially an elementary combination of sampling and majority operations. These elementary operations have extensively been observed in the aforementioned experimental settings [18].

2.1 Communication Model and Definition of the Problems

The communication model we consider is essentially the *uniform push model* [19, 32, 38], where in each (synchronous) round each agent can send (*push)* a message to another agent chosen uniformly at random. This occurs without having the sender or the receiver learning about each other's identity. Note that it may happen that several agents push a message to the same node u at the same round. In the latter case we assume that the nodes receive them in a random order; we discuss this assumption in detail in Section 2.1.1.

We study the problems of rumor-spreading and plurality consensus. In both cases, we assume that nodes can support opinions represented by an integer in $[k] = \{1, \ldots, k\}$. Additionally, there may be *undecided* nodes that do not support any opinion, which represents nodes that are not actively aware that the system has started to solve the problem; thus, undecided nodes are not allowed to send any message before receiving any of them.

- In rumor spreading, initially, one node, called the source, has an opinion $m \in \{1, \ldots, k\}$, called the *correct opinion*. All the other nodes have no

opinion. The objective is to design a protocol insuring that, after a certain number of communication rounds, every node has the correct opinion m.

- In plurality consensus, initially, for every $i \in \{1, \ldots, k\}$, a set A_i of nodes have opinion i. The sets A_i, $i = 1, \ldots, k$, are pairwise disjoint, and their union does not need to cover all nodes, i.e., there may be some *undecided* nodes with no opinion initially. The objective is to design a protocol insuring that, after a certain number of communication rounds, every node has the plurality opinion, that is, the opinion m with relative majority in the initial setting (i.e., $|A_m| > |A_j|$ for any $j \neq m$).

Observe that the rumor-spreading problem is a special case of the plurality consensus problem with $|A_m| = 1$ and $|A_j| = 0$ for any $j \neq m$.

Following the guidelines of [27], we work under two constraints:

1. We restrict ourselves to protocols in which each node can only transmit opinions, i.e., every message is an integer in $\{1, \ldots, k\}$.

2. Transmissions are subject to noise, that is, for every round, and for every node u, if an opinion $i \in \{1, \ldots, k\}$ is transmitted to node u during that round, then node u will receive message $j \in \{1, \ldots, k\}$ with probability $p_{i,j} \geq 0$, where $\sum_{j=1}^{k} p_{i,j} = 1$.

The *noisy push model* is the uniform push model together with the previous two constraints. The probabilities $\{p_{i,j}\}_{i,j \in [k]}$ can be seen as a transition matrix, called the *noise matrix*, and denoted by $P = (p_{i,j})_{i,j \in [k]}$. The noise matrix in [27] is simply

$$P = \begin{pmatrix} \frac{1}{2} + \epsilon & \frac{1}{2} - \epsilon \\ \frac{1}{2} - \epsilon & \frac{1}{2} + \epsilon \end{pmatrix}. \tag{1}$$

2.1.1 The Reception of Simultaneous Messages

In the uniform push model, it may happen that several agents push a message to the same node u at the same round. In such cases, the model should specify whether the node receives all such messages, only one of them or neither of them. Which choice is better depends on the biological setting that is being modeled: if the communication between the agents of the system is an auditory or tactile signal, it could be more realistic to assume that simultaneous messages to the same node would "collide", and the node would not be able to grasp any of them. If, on the other hand, the messages represent visual or chemical signals (see e.g. [42, 30, 11, 10]), then it may be unrealistic to assume that nodes cannot receive more than one of such messages at the same round and besides, by a standard balls-into-bins argument (e.g. by applying Lemma 3), it follows that in the uniform push model at each round no node receives more than $\mathcal{O}(\log n)$ messages w.h.p. In this work we thus consider the model in which all messages are received, also because such assumption allows us to obtain simpler proofs than the other variants. We finally note that our protocol does not strictly need such assumption, since it only requires the nodes to collect a small random

sample of the received messages. However, since we look at the latter feature as a consequence of active choices of the nodes rather than some inherent property of the environment, we avoid to weaken the model to the point that it matches the requirements of the protocol.

2.1.2 On the Role of Synchronicity in the Result

An important aspect of many natural biological *computations* is their tolerance with respect to a high level of asynchrony. Following [27], in this work we tackle the noisy rumor-spreading and plurality consensus problems by assuming that agents are provided a *shared clock*, which they can employ to synchronize their behavior across different *phases* of a protocol. In [27, Section 3], it is shown how substantially relax the previous assumption by assuming, instead, that a source agent can broadcast a *starting signal* to the rest of the system to initiate the execution of the protocol. That is, a simple *rumor-spreading* procedure is employed to *awake* the agents which join the system asynchronously, and it is shown that with high probability the level of asynchrony (i.e. the largest difference among the agents' estimates of the time since the start of the execution of the protocol), is logarithmically bounded with high probability. It thus follows that their results can be generalized to the setting in which source agents initiate the execution of the protocol by waking up the rest of the system, with only a logarithmic overhead factor in the running time. We defer the reader to [27, Section 3] for formal details regarding this synchronization procedure. Our generalization of the results in [27] is independent from any aspect which concerns the aforementioned procedure. Hence, their relaxation holds for our results as well, with the same $\log n$ additional factor in the running time. It is an open problem to obtain a simple procedure to wake up the system while incurring a smaller overhead than the logarithmic one given by [27]. Finally, we remark that, more generally, the research on solving fundamental coordination problems such as plurality consensus in fully-asynchronous communication models such as population protocols is an active research area [23, 6]. We believe that obtaining analogous results to those provided here in a noisy version of population protocols is an interesting direction for future research.

2.2 Plurality Bias, and Majority Preservation

When time proceeds, our protocols will result in the proportion of nodes with a given opinion to evolve. Note that there might be nodes who do not support any opinion at time t. As mentioned in the previous section, we call such nodes *undecided*. We denote by $a^{(t)}$ the fraction of nodes supporting any opinion at time t and we call the nodes contributing to $a^{(t)}$ *opinionated*. Consequently, the fraction of undecided nodes at time t is $1 - a^{(t)}$. Let $c_i^{(t)}$ be the fraction of opinionated nodes in the system that support opinion $i \in [k]$ at the beginning of round t, so that $\sum_{i\in[k]} c_i^{(t)} = a^{(t)}$. Let $\hat{c}_i^{(t)}$ be the fraction of opinionated nodes which receive at least one message at time $t-1$ and support opinion $i \in [k]$ at the beginning of round t. We write $\mathbf{c}^{(t)} = (c_1^{(t)}, ..., c_k^{(t)})$ to denote the

opinion distribution of the opinions at time t. Similarly, let $\hat{\mathbf{c}}^{(t)} = (\hat{c}_1^{(t)}, ..., \hat{c}_k^{(t)})$. In particular, if every node would simply switch to the last opinion it received, then

$$\begin{aligned}\mathbb{E}[\hat{c}_i^{(t+1)} \mid \mathbf{c}^{(t)}] \\ = \sum_{j\in[k]} \Pr[\text{received } i \mid \text{original message is } j] \cdot \Pr[\text{original message is } j] \\ = \sum_{j\in[k]} c_j^{(t)} \cdot p_{j,i}.\end{aligned}$$

That is,

$$\mathbb{E}[\hat{\mathbf{c}}^{(t+1)} \mid \mathbf{c}^{(t)}] = \mathbf{c}^{(t)} \cdot P, \tag{2}$$

where P is the noise matrix. In particular, in the absence of noise, we have $P = I$ (the identity matrix), and if every node would simply copy the opinion that it just received, we had $\mathbb{E}[\hat{\mathbf{c}}^{(t+1)} \mid \mathbf{c}^{(t)}] = \mathbf{c}^{(t)}$. So, given the opinion distribution at round t, from the definition of the model it follows that the messages each node receives at round $t+1$ can equivalently be seen as being sent from a system without noise, but whose opinion distribution at round t is $\mathbf{c}^{(t)} \cdot P$.

Recall that m denotes the initially correct opinion, that is, the source's opinion in the rumor-spreading problem, and the initial plurality opinion in the plurality consensus problem. The following definition naturally extends the concept of *majority bias* in [27] to *plurality bias*.

Definition 1. *Let $\delta > 0$. An opinion distribution $\mathbf{c}$ is said to be* δ-biased toward opinion m *if $c_m - c_i \geq \delta$ for all $i \neq m$.*

In [27], each binary opinion that is transmitted between two nodes is flipped with probability at most $\frac{1}{2} - \epsilon$, with[3] $\epsilon = n^{-\frac{1}{4}+\eta}$ for an arbitrarily small $\eta > 0$. Thus, the noise is parametrized by ϵ. The smaller ϵ, the more noisy are the communications. We generalize the role of this parameter with the following definition.

Definition 2. *Let $\epsilon = \epsilon(n)$ and $\delta = \delta(n)$ be positive. A noise matrix P is said to be* (ϵ, δ)-majority-preserving ((ϵ, δ)-m.p.) with respect to opinion m *if, for every opinion distribution $\mathbf{c}$ that is δ-biased toward opinion m, we have $(\mathbf{c} \cdot P)_m - (\mathbf{c} \cdot P)_i > \epsilon\,\delta$ for all $i \neq m$.*

In the rumor-spreading problem, as well as in the plurality consensus problem, when we say that a noise matrix is (ϵ, δ)-m.p., we implicitly mean that it is (ϵ, δ)-m.p. with respect to the initially correct opinion. Because of the space constraints, we defer the discussion on the class of (ϵ, δ)-m.p. noise matrices in Section 4 (including its tightness w.r.t. theorems 1 and 2).

[3] For a discussion on what happens for other values of ϵ, see Appendix C.

2.3 Our Formal Results

We show that a natural generalization of the protocol in [27] solves the rumor spreading problem and the plurality consensus problem for an arbitrary number of opinions k. More precisely, using the protocol which we describe in Section 3.1, we can establish the following two results, whose proof can be found in Section 3.

Theorem 1. *Assume that the noise matrix P is (ϵ, δ)-m.p. with $\epsilon = \Omega(n^{-\frac{1}{4}+\eta})$ for an arbitrarily small constant $\eta > 0$ and $\delta = \Omega(\sqrt{\log n/n})$. The noisy rumor-spreading problem with k opinions can be solved in $O(\frac{\log n}{\epsilon^2})$ communication rounds, w.h.p., by a protocol using $O(\log\log n + \log\frac{1}{\epsilon})$ bits of memory at each node.*

Theorem 2. *Let S with $|S| = \Omega(\frac{1}{\epsilon^2}\log n)$ be an initial set of nodes with opinions in $[k]$, the rest of the nodes having no opinions. Assume that the noise matrix P is (ϵ, δ)-m.p. for some $\epsilon > 0$, and that S is $\Omega(\sqrt{\log n/|S|})$-majority-biased. The noisy plurality consensus problem with k opinions can be solved in $O(\frac{\log n}{\epsilon^2})$ communication rounds, w.h.p., by a protocol using $O(\log\log n + \log\frac{1}{\epsilon})$ bits of memory at each node.*

For $k = 2$, we get the theorems in [27] from the above two theorems. Indeed, the simple 2-dimensional noise matrix of Eq. (1) is ϵ-majority-biased. Note that, as in [27], the plurality consensus algorithm requires the nodes to known the size $|S|$ of the set S of opinionated nodes.

3 The Analysis

In this section we prove Theorem 1 and Theorem 2 by generalizing the analysis of Stage 1 given in [27] and by providing a new analysis of Stage 2. Note that the proof techniques required for the generalization to arbitrary k significantly depart from those in [27] for the case $k = 2$. In particular, our approach provides a general framework to rigorously deal with many kind of stochastic dependences among messages in the uniform push model.

3.1 Definition of the Protocol

We describe a rumor spreading protocol performing in two *stages*. Each stage is decomposed into a number of *phases*, each one decomposed into a number of *rounds*. During each phase of the two stages, the nodes apply the simple rules given below.

3.1.1 The rule during each phase of Stage 1.

Nodes that already support some opinion at the beginning of the phase push their opinion at each round of the phase. Nodes that do not support any opinion at the beginning of the phase but receive at least one opinion during the phase

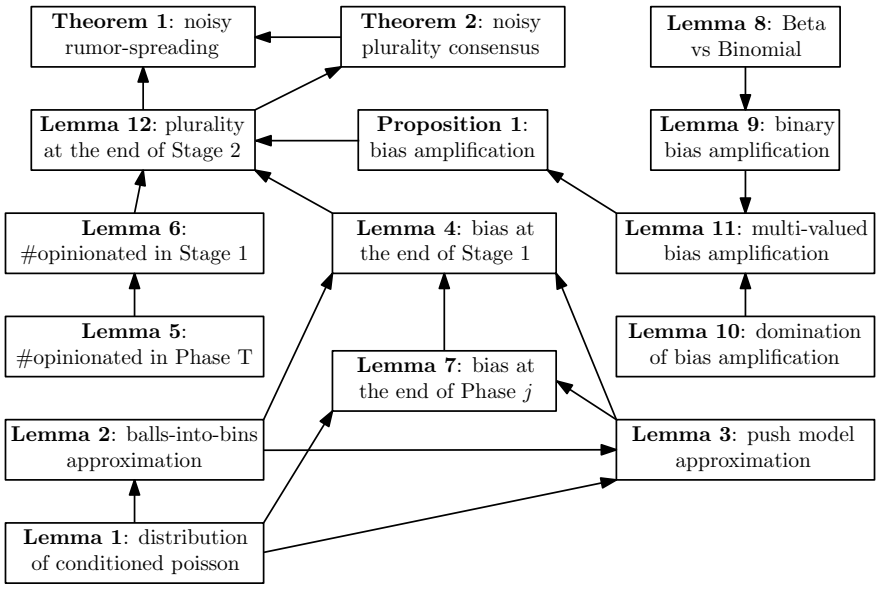

Figure 1: Diagrams of dependencies among the different parts of the analysis. Each box represents a statement proven in the analysis, and an arrow between two boxes u and v signifies that the statement of box u is employed in the proof of box v.

start supporting an opinion at the end of the phase, chosen u.a.r. (counting multiplicities) from the received opinions[4]. In other words, each node tries to acquire an opinion during each phase of Stage 1, and, as it eventually receives some opinions, it starts supporting one of them (chosen u.a.r.) from the beginning of the next phase. In particular, opinionated nodes never change their opinion during the entire stage.

More formally, let ϕ, β, and s be three constants satisfying $\phi > \beta > s$. The rounds of Stage 1 are grouped in $T+2$ phases with $T = \lfloor \log(n/(2s/\epsilon^2 \log n))/\log(\beta/\epsilon^2 + 1) \rfloor$. Phase 0 takes $s/\epsilon^2 \log n$ rounds, phase $T+1$ takes $\phi/\epsilon^2 \log n$ rounds, and each phase j with $1 \leq j \leq T$ takes β/ϵ^2 rounds. We denote with τ_j the end of the last round of phase j.

Let t_u be the first time in which u receives any opinion since the beginning of the protocol (with $t_u = 0$ for the source). Let j_u be the phase of t_u, and let $\mathrm{val}(u)$ be an opinion chosen u.a.r. by u among those that it receives during phase j_u[5]. During the first stage of the protocol each node applies the following rule.

[4]Note that, in the protocol considered in [27], the choice of each node's new opinion in both stages is based on the first messages received. In [27], in order to relax the synchronicity assumption that nodes share a common clock, they adopt the same sample-based variant of the rule that we adopt here.

[5]Note that, in order to sample u.a.r. one of them, u does not need to collect all the opinions it receives. A natural sampling strategy such as reservoir sampling can be used.

Rule of Stage 1. Each opinionated node u pushes opinion $\mathrm{val}(u)$ during each round of every phase $j = j_u + 1, ..., T + 1$.

3.1.2 The rule during each phase of Stage 2.

During each phase of Stage 2, every node pushes its opinion at each round of the phase. At the end of the phase, each node that received "enough" opinions takes a random sample[4] of them, and starts supporting the most frequent opinion in that sample (breaking ties u.a.r.).

More formally, the rounds of stage 2 are divided in $T' + 1$ phases with $T' = \lceil \log(\sqrt{n/\log n}) \rceil$. Each phase j, $0 \leq j \leq T' - 1$, has length 2ℓ with $\ell = \lceil c/\epsilon^2 \rceil$ for some large-enough constant $c > 0$, and phase T' has length $2\ell'$ with $\ell' = O(\epsilon^{-2} \log n)$. For any finite multiset A of elements in $\{1, \ldots, k\}$, and any $i \in \{1, \ldots, k\}$, let $\mathrm{occ}(i, A)$ be the number of occurrences of i in A, and let $\mathrm{mode}(A) = \{i \in \{1, \ldots, k\} \mid \mathrm{occ}(i, A) \geq \mathrm{occ}(j, A) \text{ for every } j \in \{1, \ldots, k\}\}$. We then define $\mathrm{maj}(A)$ as the most frequent value in A (breaking ties u.a.r.), i.e., $\mathrm{maj}(A)$ is the r.v. on $\{1, \ldots, k\}$ such that $\Pr(\mathrm{maj}(A) = i) = \mathbb{1}_{\{i \in \mathrm{mode}(A)\}}/|\mathrm{mode}(A)|$. Let $R_j(u)$ be the multiset of messages received by node u during phase j. During the second stage of the protocol each node applies the following rule.

Rule of Stage 2. During each phase j of length $2L$ of Stage 2 ($L = \ell$ or ℓ'), each node u pushes its current opinion at each round of the phase, and starts drawing a random uniform sample $\mathcal{S}(u)$ of size L from $R_j(u)$. Provided $|R_j(u)| \geq L$, at the end of the phase u changes its opinion to $\mathrm{maj}(\mathcal{S}(u))$.

Let us remark that the reason we require the use of sampling in the previous rule is that at a given round a node may receive much more messages than $2L$. Thus, if the nodes were to collect all the messages they receive, some of them would need much more memory than our protocol does. Finally, observe that overall both stages 1 and 2 take $O(\frac{1}{\epsilon^2} \log n)$ rounds.

3.2 Pushing Colored Balls-into-Bins

Before delving into the analysis of the protocol, we provide a framework to rigorously deal with the stochastic dependence that arises between messages in the uniform push model. Let process **O** be the process that results from the execution of the protocol of Section 3.1 in the uniform push model. In order to apply concentration of probability results that requires the involved random variables to be independent, we view the messages as balls, and the nodes as bins, and employ Poisson approximation techniques. More specifically, during each phase j of the protocol, let M_j be the set of messages that are sent to random nodes, and N_j be the set of messages sent *after* the noise has acted on them. (In other words, $N_j = \bigcup_u R_j(u)$). We prove that, at the end of phase j, we can equivalently assume that all the messages M_j have been sent to the nodes according to the following process.

Definition 3. *The balls-into-bins process* **B** *associated to phase j is the two-step process in which the nodes represent bins and all messages sent in the phase represent colored balls, with each color corresponding to some opinion. Initially, balls are colored according to M_j. At the first step, each ball of color $i \in \{1, \ldots, k\}$ is re-colored with color $j \in \{1, \ldots, k\}$ with probability $p_{i,j}$, independently of the other balls. At the second step all balls are thrown into the bins u.a.r. as in a balls-into-bins experiment.*

Claim 1. *Given the opinion distribution and the number of active nodes at the beginning of phase j, the probability distribution of the opinion distribution and the number of active nodes at the end of phase j in process* **O** *is the same as if the messages were sent according to process* **B**.

It is not hard to see that Claim 1 holds in the case of a single round. For more than one round, it is crucial to observe that the way each node u acts in the protocol depends only on the received messages $R_j(u)$, regardless of the order in which these messages are received. As an example, consider the opinion distribution in which one node has opinion 1, one other node has opinion 2, and all other nodes have opinion 3. Suppose that each node pushes its opinion for two consecutive rounds. Since, at each round, exactly one opinion 1 and exactly one opinion 2 are pushed, no node can receive two 1s during the first round and then two 2s during the second round, i.e. no node can possibly receive the sequence of messages "1,1,2,2" in this exact order. Instead, in process **B** such a sequence is possible.

Proof of Claim 1. In both process **B** and process **O**, at each round, the noise acts independently on each ball/message of a given color/opinion, according to the same probability distribution for that color/opinion. Then, in both processes, each ball/message is sent to some bin/node chosen u.a.r. and independently of the other balls/messages. Indeed, we can couple process **B** and process **O** by requiring that:

1. each ball/message is changed by the noise to the same color/value, and
2. each ball/message ends up in the same bin/node.

Thus, the joint probability distribution of the sets $\{R_j(u)\}_{u \in [n]}$ in process **O** is the same as the one given by process **B**.

Observe also that, from the definition of the protocol (see the rule of Stage 1 and Stage 2 in Section 3.1), it follows that each node's action depends only on the set $R_j(u)$ of received messages at the end of each phase j, and does not depend on any further information such as the actual order in which the messages are received during the phase.

Summing up the two previous observations, we get that if, at the end of each phase j, we generate the $R_j(u)$s according to process **B**, and we let the protocol execute according to them, then we indeed get the same stochastic process as process **O**. □

Now, one key ingredient in our proof is to approximate process **B** using the following process **P**.

Definition 4. *Given N_j, process* **P** *associated to phase j is the one-shot process in which each node receives a number of opinions i that is a random variable with distribution Poisson(h_i/n), where h_i is the number of messages in N_j carrying opinion i, and each Poisson random variable is independent of the others.*

Now we provide some results from the theory of Poisson approximation for balls-into-bins experiments that are used in Section 3.2. For a nice introduction to the topic, we refer to [36].

Lemma 1. *Let $\{X_j\}_{j\in[\tilde{n}]}$ be independent r.v. such that $X_j \sim$Poisson(λ_j). The vector $(X_1, ..., X_{\tilde{n}})$ conditional on $\sum_j X = \tilde{m}$ follows a multinomial distribution with $\tilde{m}$ trials and probabilities $(\frac{\lambda_1}{\sum_j \lambda_j}, ..., \frac{\lambda_{\tilde{n}}}{\sum_j \lambda_j})$.*

Lemma 2. *Consider a balls-into-bins experiment in which h colored balls are thrown in n bins, where h_i balls have color i with $i \in \{1, ..., k\}$ and $\sum_i h_i = h$. Let $\{X_{u,i}\}_{u\in\{1,...,n\},i\in\{1,...,k\}}$ be the number of i-colored balls that end up in bin u, let $f(x_{1,1}, ..., x_{n,1}, x_{n,2}, ..., x_{n,k}, z_1, ..., z_n)$ be a non-negative function with positive integer arguments $x_{1,1}, ..., x_{n,1}, x_{n,2}, ..., x_{n,k}, z_1, ..., z_n$, let $\{Y_{u,i}\}_{u\in\{1,...,n\},i\in\{1,...,k\}}$ be independent r.v. such that $Y_{u,i} \sim$Poisson(h_i/n) and let $Z_1, ..., Z_n$ be integer valued r.v. independent from the $X_{u,i}$s and $Y_{u,i}$s. Then*

$$\mathbb{E}\left[f\left(X_{1,1}, ..., X_{n,1}, X_{n,2}, ..., X_{n,k}, Z_1, ..., Z_n\right)\right] \leq e^k \sqrt{\prod_i h_i}\ \mathbb{E}\left[f\left(Y_{1,1}, ..., Y_{n,1}, Y_{n,2}, ..., Y_{n,k}, Z_1, ..., Z_n\right)\right].$$

Proof. To simplify notation, let $\bar{Z} = (Z_1, ..., Z_n)$, $\bar{X} = (X_{1,1}, ..., X_{n,1}, X_{n,2}, ..., X_{n,n})$, $\bar{Y} = (Y_{1,1}, ..., Y_{n,1}, Y_{n,2}, ..., Y_{n,n})$, $\bar{Y}_{\sum} = (\sum_{u=1}^n Y_{u,1}, ..., \sum_{u=1}^n Y_{u,k})$, $\lambda_i = h_i/n$, $\bar{\lambda} = (\lambda_1, ..., \lambda_k)$ and finally $\bar{x} = (x_1, ..., x_k)$ for any $x_1, ..., x_k$. Observe that, while $X_{u,i}$ and $X_{v,i}$ are clearly dependent, $X_{u,i}$ and $X_{v,j}$ with $i \neq j$ are stochastically independent (even if $u = v$). Indeed, the distribution of the r.v. $\{X_{u,i}\}_{u\in\{1,...,n\}}$ for each fixed i is multinomial with λ_i trials and uniform distribution on the bins. Thus, from Lemma 1 we have that $\{X_{u,i}\}_{u\in\{1,...,n\}}$are distributed as $\{Y_{u,i}\}_{u\in\{1,...,n\}}$ conditional on $\sum_{u=1}^n Y_{u,i} = \lambda_i$, that is

$$\mathbb{E}\left[f\left(\bar{Y}, \bar{Z}\right) \middle| \sum_{u=1}^n Y_{u,1} = \lambda_1, ..., \sum_{u=1}^n Y_{u,k} = \lambda_k\right] = \mathbb{E}\left[f\left(\bar{X}, \bar{Z}\right)\right].$$

Therefore, we have

$$\begin{aligned}\mathbb{E}\left[f\left(\bar{Y}, \bar{Z}\right)\right] &= \sum_{\bar{x}:x_1,...,x_k\geq 0} \mathbb{E}\left[f\left(\bar{Y}, \bar{Z}\right)\middle|\ \bar{Y}_{\sum} = \bar{x}\right] \Pr\left(\bar{Y}_{\sum} = \bar{x}\right) \\ &\geq \mathbb{E}\left[f\left(\bar{Y}, \bar{Z}\right)\middle|\ \bar{Y}_{\sum} = \bar{\lambda}\right] \Pr\left(\bar{Y}_{\sum} = \bar{\lambda}\right) = \mathbb{E}\left[f\left(\bar{X}, \bar{Z}\right)\right] \Pr\left(\bar{Y}_{\sum} = \bar{\lambda}\right) \\ &= \mathbb{E}\left[f\left(\bar{X}, \bar{Z}\right)\right] \prod_i \frac{h_i^{h_i}}{h_i!} e^{-h_i} \geq \mathbb{E}\left[f\left(\bar{X}, \bar{Z}\right)\right] \frac{e^{-k}}{\sqrt{\prod_i h_i}},\end{aligned}$$

where, in the last inequality, we use that, by Stirling's approximation, $a! \leq e\sqrt{a}(\frac{a}{e})^a$ for any $a > 0$. □

From Lemma 1 and Lemma 2, we get the following general result which says that if a generic event $\mathcal{E}$ holds w.h.p in process $\mathbf{P}$, it also holds w.h.p. in process $\mathbf{O}$.

Lemma 3. *Given the opinion distribution and the number of active nodes at the beginning of a fixed phase j, let $\mathcal{E}$ be an event that, at the end of that phase, holds with probability at least $1 - n^{-b}$ in process $\mathbf{P}$, for some $b > (k \log h)/(2 \log n)$ with $h = \sum_i h_i$.*[6] *Then, at the end of phase j, $\mathcal{E}$ holds w.h.p. also in process $\mathbf{O}$.*

Proof. Thanks to Claim 1, it suffices to prove that, at the end of phase j, $\mathcal{E}$ holds w.h.p. in process $\mathbf{B}$.

Let $\bar{\mathcal{E}}$ be the complementary event of $\mathcal{E}$. Let $h = |M_j|$ be the number of balls that are thrown in process $\mathbf{B}$ associated to phase j, where h_i balls have color i with $i \in \{1, ..., k\}$ and $\sum_i h_i = h$. Let $\{X_{u,i}\}_{u \in \{1,...,n\}, i \in \{1,...,k\}}$ be the number of i-colored balls that end up in bin u, let $\{Y_{u,i}\}_{u \in \{1,...,n\}, i \in \{1,...,k\}}$ be the independent r.v. of process $\mathbf{P}$ such that $Y_{u,i} \sim$Poisson(h_i/n) and let $Z_1, ..., Z_n$ be integer valued r.v. independent from the $X_{u,i}$s and $Y_{u,i}$s.

Fix any realization of N_j, i.e. any re-coloring of the balls in the first step of process $\mathbf{B}$. By choosing f in Lemma 2 as the binary r.v. indicating whether event $\bar{\mathcal{E}}$ has occurred, where $\bar{\mathcal{E}}$ is a function of the r.v. $X_{1,1}, ..., X_{n,1}$, $X_{n,2}, ..., X_{n,k}$, $Z_1, ..., Z_n$, we get

$$\Pr\left(\bar{\mathcal{E}}\left(X_{1,1}, ..., X_{n,k}, Z_1, ..., Z_n\right) \middle| N_j\right) \leq e^k \sqrt{\prod_i h_i} \Pr\left(\bar{\mathcal{E}}\left(Y_{1,1}, ..., Y_{n,k}, Z_1, ..., Z_n\right) \middle| N_j\right). \quad (3)$$

Thus, from Eq. (3), the Inequality of arithmetic and geometric means and the hypotheses on the probability of $\mathcal{E}$, we get

$$\begin{aligned}\Pr\left(\bar{\mathcal{E}}\left(X_{1,1}, ..., X_{n,k}, Z_1, ..., Z_n\right) \middle| N_j\right) &\leq e^k \sqrt{\prod_i h_i} \Pr\left(\bar{\mathcal{E}}\left(Y_{1,1}, ..., Y_{n,k}, Z_1, ..., Z_n\right) \middle| N_j\right) \\ &\leq e^k \left(\frac{h}{k}\right)^{\frac{k}{2}} \Pr\left(\bar{\mathcal{E}}\left(Y_{1,1}, ..., Y_{n,k}, Z_1, ..., Z_n\right) \middle| N_j\right)\end{aligned}$$

Finally, let $\mathcal{N}$ be the set of all possible realizations of N_j. By the law of total

[6] Note that, if N_j is not yet fixed, the parameters h_i of process $\mathbf{P}$ associated to phase j are random variables. However, if the opinion distribution and the number of active nodes at the beginning of phase j are given, then $h = \sum_i h_i = |N_j| = |M_j|$ is fixed.

probability over $\mathcal{N}$, we get that

$$\sum_{s\in\mathcal{N}} \Pr\left(\bar{\mathcal{E}}\left(X_{1,1},...,X_{n,k},Z_1,...,Z_n\right)\,\middle|\,N_j=s\right)\Pr\left(N_j=s\right)$$
$$\leq e^k\left(\frac{h}{k}\right)^{\frac{k}{2}}\sum_{s\in\mathcal{N}} \Pr\left(\bar{\mathcal{E}}\left(Y_{1,1},...,Y_{n,k},Z_1,...,Z_n\right)\,\middle|\,N_j=s\right)\Pr\left(N_j=s\right)$$
$$\leq e^k\left(\frac{h}{k}\right)^{\frac{k}{2}}\Pr\left(\bar{\mathcal{E}}\left(Y_{1,1},...,Y_{n,k},Z_1,...,Z_n\right)\right)$$
$$\leq \frac{e^k}{k^{\frac{k}{2}}}\,h^{\frac{k}{2}}n^{-b}\leq \frac{e^k}{e^{\frac{k}{2}\log k}}\,e^{\frac{k}{2}\log h}e^{-b\log n}\leq e^{k-\frac{k}{2}\log k+\frac{k}{2}\log h-b\log n}\leq n^{-\Theta(1)},$$

where in the last line we used the hypotheses on the probability of $\bar{\mathcal{E}}$. □

We now analyze the two stages of our protocol, starting with Stage 1. Note that, in the following two sections, the statements about the evolution of the process refer to process **O**.

3.3 Stage 1

The rule of Stage 1 is aimed at guaranteeing that, w.h.p., the system reaches a target opinion distribution from which the rumor-spreading problem becomes an instance of the plurality consensus problem. More precisely, we have the following.

Lemma 4. *Stage 1 takes $O(\frac{1}{\epsilon^2}\log n)$ rounds, after which w.h.p. all nodes are active and $\mathbf{c}^{(\tau_{T+1})}$ is δ-biased toward the correct opinion, with $\delta=\Omega(\sqrt{\log n/n})$.*

Proof. The fact that an undecided node becomes opinionated during a phase only depends on whether it gets a message during that phase, regardless of the value of such messages. Hence, the proof that, w.h.p., $a^{(\tau_{T+1})}=1$ is reduced to the analysis of the rule of Stage 1 as an information spreading process. First, by carefully exploiting the Chernoff bound and Lemma 3, we can establish Claim 2 and Claim 3 below:

Claim 2. *W.h.p., at the end of phase 0, we have $s/\epsilon^2\log n/3n\leq a^{(\tau_0)}\leq s/\epsilon^2\log n/n$.*

Claim 3. *W.h.p., at the end of phase j, $1\leq j\leq T$, we have*

$$(\beta/\epsilon^2+1)^j a^{(\tau_0)}/8\leq a^{(\tau_j)}\leq(\beta/\epsilon^2+1)^j a^{(\tau_0)}.$$

Proof of Claim 2 and Claim 3. The probability that, in the process **O**, an undecided node becomes opinionated at the end of phase j is $1-(1-\frac{1}{n})^h$ where h is the number of messages sent during that phase. In process **P**, this probability is $1-e^{-\frac{h}{n}}$. By using that $e^{\frac{x}{1+x}}\leq 1+x\leq e^x$ for $|x|<1$ we see that $1-e^{-\frac{h}{n}}\leq 1-(1-1/n)^h\leq 1-e^{-\frac{h}{n-1}}$. Thus, we can prove Claim 2 and Claim 3

for process **P** by repeating essentially the same calculations as in the proofs of Claim 2.2 and 2.4 in [27]. Since the Poisson distributions in process **P** are independent, we can apply the Chernoff bound as claimed in [27]. Finally, we can prove that the statements hold also for process **O** thanks to Lemma 2. □

From the previous two claims, and by the definition of T we get the following.

Lemma 5. *W.h.p., at the end of phase T, we have $a^{(\tau_{T+1})} = \Omega((\beta/\epsilon^2 + 1)^T a^{(\tau_0)}) = \Omega(\epsilon^2)$.*

Finally, from Lemma 5, an application of the Chernoff bound gives us the following.

Lemma 6. *W.h.p., at the end of Stage 1, all nodes are opinionated.*

As for the fact that, w.h.p., $\mathbf{c}^{(\tau+1)}$ is a δ-biased opinion distribution with $\delta = \Omega(\sqrt{\log n/n})$, we can prove the following.

Lemma 7. *W.h.p., at the end of each phase j of Stage 1, we have an $(\epsilon/2)^j$-biased opinion distribution.*

Proof. We prove the lemma by induction on the phase number. The case $j = 1$ is a direct application of Lemma 16 to $c_m^{(\tau_1)} - c_i^{(\tau_1)}$ $(i \neq m)$, where the number of opinionated nodes is given by Claim 2, and, where the independence of the r.v. follows from the fact that each node that becomes opinionated in the first phase has necessarily received the messages from the source-node. Now, suppose that the lemma holds up to phase $j-1 \leq T$. Let $S_j = \{u|\ j_u = j\}$ be the set of nodes that become opinionated during phase j. Recall the definition of M_j and N_j from Section 3.2, and observe that $|M_j| = |N_j| = (\tau_j - \tau_{j-1})\, n \cdot a^{(\tau_{j-1})}$, and that the number of times opinion i occurs in M_j is $|M_j|\, c_i^{(\tau_{j-1})}$. Let us identify each message in M_j with a distinct number in $1, ..., |M_j|$, and let $\{X_w(i)\}_{w\in\{1,...,|M_j|\}}$ be the binary r.v. such that $X_w(i) = 1$ if and only if w is i after the action of the noise. The frequency of opinion i in N_j is $\frac{1}{|N_j|}\sum_{w=1}^{|N_j|} X_w(i)$.

Thanks to Lemma 3, it suffices to prove the lemma for process **P**. By definition, in process **P**, for each i, the number of messages with opinion i that each node receives conditional on N_j follows a Poisson$(\frac{1}{n}\sum_{w=1}^{|N_j|} X_w\,(i))$ distribution. Each node u that becomes opinionated during phase j gets at least one message during the phase. Thus, from Lemma 1, the probability that u gets opinion i conditional on N_j is

$$\frac{\sum_{w=1}^{|N_j|} X_w\,(i)}{\sum_{i=1}^{k}\sum_{w=1}^{|N_j|} X_w\,(i)} = \frac{1}{|N_j|}\sum_{w=1}^{|N_j|} X_w\,(i)\,.$$

Since opinionated nodes never change opinion during Stage 1, the bias of $\mathbf{c}^{(\tau_j)}$ is at least the minimum between the bias of $\mathbf{c}^{(\tau_{j-1})}$ and the bias among the newly

opinionated nodes in S_j. Hence, we can apply the Chernoff bound to the nodes in S_j to prove that the bias at the end of phase j is, w.h.p.[7],

$$\Pr\left(c_m^{(\tau_j)} - c_i^{(\tau_j)} \middle| N_j\right) \geq \left(\frac{1}{|N_j|}\sum_{w=1}^{|N_j|} X_w(m) - \frac{1}{|N_j|}\sum_{w=1}^{|N_j|} X_w(i)\right)\left(1-\tilde{\delta}_j\right), \quad (4)$$

where $\tilde{\delta}_j = O(\sqrt{\log n/|S_j|})$.

Moreover, note that

$$\mathbb{E}\left[\frac{1}{|N_j|}\sum_{w=1}^{|N_j|} X_w(i) \middle| \mathbf{c}^{(\tau_{j-1})}, a^{(\tau_{j-1})}\right] = \left(\mathbf{c}^{(\tau_{j-1})} \cdot P\right)_i.$$

Furthermore, (conditional on $\mathbf{c}^{(\tau_{j-1})}$ and $a^{(\tau_{j-1})}$) the r.v. $\{X_w(i)\}_{w\in\{1,\ldots,|N_j|\}}$ are independent. Thus, for each $i \neq m$, from Claim 3, and by applying the Chernoff bound on $\sum_{w=1}^{|N_j|} X_w(m)$, and on $\sum_{w=1}^{|N_j|} X_w(i)$, we get that w.h.p.

$$\frac{1}{|N_j|}\sum_{w=1}^{|N_j|} X_w(m) - \frac{1}{|N_j|}\sum_{w=1}^{|N_j|} X_w(i) \geq (1-\delta_j)\,2^{-j+1}\epsilon^j, \quad (5)$$

where $\delta_j = O(\sqrt{\log n/|N_j|})$.

From Claim 2 and Claim 3, it follows that $\tilde{\delta}_j, \delta_j \leq \frac{1}{4}$ w.h.p. Thus by putting together Eq. (4) and (5) via the chain rule, we get that, w.h.p.,

$$c_m^{(\tau_j)} - c_i^{(\tau_j)} \geq \left(1-\tilde{\delta}_j\right)(1-\delta_j)\,2^{-j+1}\epsilon^j \geq \left(\frac{\epsilon}{2}\right)^j.$$

□

Lemma 7 implies that, w.h.p., we get a bias $\epsilon^{T+2} = \Omega(\sqrt{\log n/n})$ at the end of Stage 1, which completes the proof of Lemma 4. □

3.4 Stage 2

As proved in the previous section, w.h.p., all nodes are opinionated at the end of Stage 1, and the final opinion distribution is $\Omega(\sqrt{\log n/n})$-biased. Now, we have that the rumor-spreading problem is reduced to an instance of the plurality consensus problem. The purpose of Stage 2 is to progressively amplify the initial bias until all nodes support the plurality opinion, i.e. the opinion originally held by the source node.

During the first T' phases, it is not hard to see that, by taking c large enough, a fraction arbitrarily close to 1 of the nodes receives at least ℓ messages, w.h.p.

[7] We remark that Eq. (4) concerns the value of $\Pr(c_m^{(\tau_j)} - c_i^{(\tau_j)}|N_j)$, which is a random variable.

Each node u in such fraction changes its opinion at the end of the phase. With a slight abuse of notation, let $\mathrm{maj}_\ell(u) = \mathrm{maj}\left(\mathcal{S}(u)\right)$ be u's new opinion based on the $\ell = |\mathcal{S}(u)|$ randomly sampled received messages. We show that, w.h.p., these new opinions increase the bias of the opinion distribution toward the plurality opinion by a constant factor > 1.

For the sake of simplicity, we assume that ℓ is odd (see Appendix B for details on how to remove this assumption).

Proposition 1. *Suppose that, at the beginning of phase j of Stage 2 with $0 \leq j \leq T' - 1$, the opinion distribution is δ-biased toward m. In process $\mathbf{P}$, if a node u changes its opinion at the end of the phase, then, for any $i \neq m$, we have*

$$\Pr\left(\mathrm{maj}_\ell(u) = m\right) - \Pr\left(\mathrm{maj}_\ell(u) = i\right) \geq \sqrt{\frac{2\ell}{\pi}} \frac{g(\delta, \ell)}{e^{(k-2)\ln 4}}, \tag{6}$$

where

$$g\left(\delta, \ell\right) = \begin{cases} \delta(1-\delta^2)^{\frac{\ell-1}{2}} & \text{if } \delta < \frac{1}{\sqrt{\ell}}, \\ \sqrt{1/\ell}\left(1-\sqrt{1/\ell}\right)^{\frac{\ell-1}{2}} & \text{if } \delta \geq \frac{1}{\sqrt{\ell}}. \end{cases}$$

First, we prove Eq. (6) for $k = 2$. We then obtain the general case by induction. The proof for $k = 2$ is based on a known relation between the cumulative distribution function of the binomial distribution, and the cumulative distribution function of the beta distribution. This relation is given by the following lemma.

Lemma 8. *Given $p \in (0, 1)$ and $0 \leq j \leq \ell$ it holds*

$$\sum_{j < i \leq \ell} \binom{\ell}{i} p^i \left(1-p\right)^{\ell-i} = \binom{\ell}{j+1} (j+1) \int_0^p z^j \left(1-z\right)^{\ell-j-1} dz.$$

Proof. By integrating by parts, for $j < \ell - 1$ we have

$$\begin{aligned} &\binom{\ell}{j+1} (j+1) \int_0^p z^j \left(1-z\right)^{\ell-j-1} dz \\ &\qquad = \binom{\ell}{j+1} p^{j+1} \left(1-p\right)^{\ell-j-1} \\ &\qquad - \binom{\ell}{j+1} (\ell-j-1) \int_0^p z^{j+1} \left(1-z\right)^{\ell-j-2} dz \\ &= \binom{\ell}{j+1} p^{j+1} \left(1-p\right)^{\ell-j-1} - \binom{\ell}{j+2} (j+2) \int_0^p z^{j+1} \left(1-z\right)^{\ell-j-2} dz, \end{aligned} \tag{7}$$

where, in the last equality, we used the identity

$$\binom{\ell}{j} (\ell - j) = \binom{\ell}{j+1} (j+1).$$

Note that when $j = \ell - 1$, Eq. (6) becomes

$$p^\ell = \ell \int_0^p z^{\ell-1} dz.$$

Hence, we can unroll the recurrence given by Eq. (7) to obtain

$$\begin{aligned}\binom{\ell}{j+1}(j+1)\int_0^p z^j(1-z)^{\ell-j-1}\,dz &= \sum_{j<i\leq\ell-1}\binom{\ell}{i}p^i(1-p)^{\ell-i}+\ell\int_0^p z^{\ell-1}dz\\ &= \sum_{j<i\leq\ell}\binom{\ell}{i}p^i(1-p)^{\ell-i}\end{aligned}$$

concluding the proof. □

Lemma 8 allows us to express the survival function of a binomial sample as an integral. Thanks to it, we can prove Proposition 1 when $k = 2$.

Lemma 9. *Let $\mathbf{c} = (c_1, c_2)$ be a δ-biased opinion distribution during Stage 2. In process $\mathbf{P}$, for any node u, we have* $\Pr(\mathrm{maj}_\ell(u) = m) - \Pr(\mathrm{maj}_\ell(u) = 3-m) \geq \sqrt{2\ell/\pi}\cdot g(\delta,\ell)$.

Proof. Without loss of generality, let $m = 1$. Let $X_1^{(\ell)}$ be a r.v. with distribution $Bin(\ell, p_1)$, and let $X_2^{(\ell)} = \ell - X_1^{(\ell)}$. By using Lemma 8, we get

$$\begin{aligned}&\Pr(\mathrm{maj}_\ell(u)=1)-\Pr(\mathrm{maj}_\ell(u)=2)\\ &=\Pr\left(X_1^{(\ell)}>X_2^{(\ell)}\right)-\Pr\left(X_2^{(\ell)}>X_1^{(\ell)}\right)\\ &=\sum_{\lceil\frac{\ell}{2}\rceil\leq i\leq\ell}\binom{\ell}{i}p_1^i p_2^{\ell-i}-\sum_{\lceil\frac{\ell}{2}\rceil\leq i\leq\ell}\binom{\ell}{i}p_1^{\ell-i}p_2^i\\ &=\sum_{\lceil\frac{\ell}{2}\rceil\leq i\leq\ell}\binom{\ell}{i}p_1^i(1-p_1)^{\ell-i}-\sum_{\lceil\frac{\ell}{2}\rceil\leq i\leq\ell}\binom{\ell}{i}p_1^{\ell-i}(1-p_1)^i\\ &=\binom{\ell}{\lceil\frac{\ell}{2}\rceil}\left\lceil\frac{\ell}{2}\right\rceil\left(\int_0^{p_1}z^{\lfloor\frac{\ell}{2}\rfloor}(1-z)^{\lfloor\frac{\ell}{2}\rfloor}\,dz-\int_0^{p_2}z^{\lfloor\frac{\ell}{2}\rfloor}(1-z)^{\lfloor\frac{\ell}{2}\rfloor}\,dz\right).\end{aligned}$$

By setting $t = z - \frac{1}{2}$, and rewriting $p_1 = \frac{p_1-p_2}{2}+\frac{1}{2}$ and $p_2 = \frac{p_2-p_1}{2}+\frac{1}{2}$ we obtain

$$\begin{aligned}&\Pr(\mathrm{maj}_\ell(u)=1)-\Pr(\mathrm{maj}_\ell(u)=2)\\ &=\binom{\ell}{\lceil\frac{\ell}{2}\rceil}\left\lceil\frac{\ell}{2}\right\rceil\left(\int_0^{p_1}z^{\lfloor\frac{\ell}{2}\rfloor}(1-z)^{\lfloor\frac{\ell}{2}\rfloor}\,dz-\int_0^{p_2}z^{\lfloor\frac{\ell}{2}\rfloor}(1-z)^{\lfloor\frac{\ell}{2}\rfloor}\,dz\right)\\ &=\binom{\ell}{\lceil\frac{\ell}{2}\rceil}\left\lceil\frac{\ell}{2}\right\rceil\left(\int_{-\frac{1}{2}}^{\frac{p_1-p_2}{2}}\left(\frac{1}{4}-t^2\right)^{\lfloor\frac{\ell}{2}\rfloor}dt-\int_{-\frac{1}{2}}^{-\frac{p_1-p_2}{2}}\left(\frac{1}{4}-t^2\right)^{\lfloor\frac{\ell}{2}\rfloor}dt\right)\\ &=\binom{\ell}{\lceil\frac{\ell}{2}\rceil}\left\lceil\frac{\ell}{2}\right\rceil\int_{-\frac{p_1-p_2}{2}}^{\frac{p_1-p_2}{2}}\left(\frac{1}{4}-t^2\right)^{\lfloor\frac{\ell}{2}\rfloor}dt.\end{aligned}$$

For any $t \in (-\frac{y}{2}, \frac{y}{2}) \subseteq (-\frac{p_1-p_2}{2}, \frac{p_1-p_2}{2})$, it holds

$$\left(\frac{1}{4} - t^2\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor} \geq \left(\frac{1-y^2}{4}\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor}.$$

Thus, for any $y \in (-p_1 + p_2, p_1 - p_2)$ we have

$$\int_{-\frac{p_1-p_2}{2}}^{\frac{p_1-p_2}{2}} \left(\frac{1}{4} - t^2\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor} dt \geq y \left(\frac{1-y^2}{4}\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor}. \tag{8}$$

The r.h.s. of Eq. (8) is maximized w.r.t. $y \in (-p_1 + p_2, p_1 - p_2)$ when

$$y = \min\left\{p_1 - p_2, \frac{1}{\sqrt{2\left\lfloor \frac{\ell}{2} \right\rfloor + 1}}\right\} = \min\left\{p_1 - p_2, \frac{1}{\sqrt{\ell}}\right\}.$$

Hence, for $p_1 - p_2 < \frac{1}{\sqrt{\ell}}$, we get

$$\begin{aligned}
\int_{-\frac{p_1-p_2}{2}}^{\frac{p_1-p_2}{2}} \left(\frac{1}{4} - t^2\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor} dt &\geq (p_1 - p_2)\left(\frac{1-(p_1-p_2)^2}{4}\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor} \\
&= 2^{-\ell+1}(p_1 - p_2)\left(1-(p_1-p_2)^2\right)^{\frac{\ell-1}{2}} \\
&= 2^{-\ell+1} g(p_1 - p_2, \ell).
\end{aligned}$$

For $p_1 - p_2 \geq \frac{1}{\sqrt{\ell}}$ we get

$$\int_{-\frac{p_1-p_2}{2}}^{\frac{p_1-p_2}{2}} \left(\frac{1}{4} - t^2\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor} dt \geq \frac{2^{-\ell+1}}{\sqrt{\ell}}\left(1 - \frac{1}{\ell}\right)^{\frac{\ell-1}{2}} = 2^{-\ell+1} g(p_1 - p_2, \ell).$$

By using the fact that g is a non-decreasing function w.r.t. its first argument, we obtain

$$\begin{aligned}
\Pr(\mathrm{maj}_\ell(u) = 1) - \Pr(\mathrm{maj}_\ell(u) = 2) &= \binom{\ell}{\lceil \frac{\ell}{2} \rceil}\left\lceil \frac{\ell}{2} \right\rceil \int_{-\frac{p_1-p_2}{2}}^{\frac{p_1-p_2}{2}} \left(\frac{1}{4} - t^2\right)^{\left\lfloor \frac{\ell}{2} \right\rfloor} dt \\
&\geq \binom{\ell}{\lceil \frac{\ell}{2} \rceil}\left\lceil \frac{\ell}{2} \right\rceil 2^{-\ell+1} g(p_1 - p_2, \ell) \\
&\geq \binom{\ell}{\lceil \frac{\ell}{2} \rceil}\left\lceil \frac{\ell}{2} \right\rceil 2^{-\ell+1} g(\delta, \ell).
\end{aligned}$$

Finally, by using the bounds $\binom{2r}{r} \geq \frac{2^{2r}}{\sqrt{\pi r}} e^{\frac{1}{9r}}$ (see Lemma 13), and $e^x \geq 1 - x$ together with the identity[8]

$$\binom{\ell}{\lceil \frac{\ell}{2} \rceil}\left\lceil \frac{\ell}{2} \right\rceil = \binom{\ell}{\frac{\ell+1}{2}}\frac{\ell+1}{2} = \binom{\ell-1}{\frac{\ell-1}{2}}\ell,$$

[8]Recall that we are assuming that ℓ is odd.

we get

$$\begin{aligned}
&\Pr\left(\mathrm{maj}_\ell(u)=1\right)-\Pr\left(\mathrm{maj}_\ell(u)=2\right)\\
&\qquad\geq \binom{\ell}{\lceil\frac{\ell}{2}\rceil}\left\lceil\frac{\ell}{2}\right\rceil 2^{-\ell+1}g\left(\delta,\ell\right)\\
&\qquad\geq \frac{2^{\ell-1}}{\sqrt{\pi\frac{\ell-1}{2}}}e^{\frac{2}{9(\ell-1)}}\ell\cdot 2^{-\ell+1}g\left(\delta,\ell\right)\\
&\geq\sqrt{\frac{2\ell}{\pi}}\left(1-\frac{2}{9\left(\ell-1\right)}\right)\left(1-\frac{1}{\ell}\right)^{-\frac{1}{2}}\cdot g\left(\delta,\ell\right)\\
&\qquad\geq\sqrt{\frac{2\ell}{\pi}}\cdot g\left(\delta,\ell\right),
\end{aligned}$$

concluding the proof. □

Next we show how to lower bound the above difference with a much simpler expression.

Lemma 10. *In process* $\mathbf{P}$*, during Stage 2, for any node* u*,* $\Pr(\mathrm{maj}_\ell(u)=m)-\Pr(\mathrm{maj}_\ell(u)=3-m)$ *is at least* $\Pr(X_1^{(\ell)}>X_2^{(\ell)},...,X_k^{(\ell)})-\Pr(X_i^{(\ell)}>X_1^{(\ell)},...,X_{i-1}^{(\ell)},X_{i+1}^{(\ell)},...,X_k^{(\ell)})$*, where* $\bar{X}^{(\ell)}=(X_1^{(\ell)},...,X_k^{(\ell)})$ *follows a multinomial distribution with* ℓ *trials and probability distribution* $\mathbf{c}\cdot P$*.*

Proof. Without loss of generality, let $m=1$. Let $\mathbf{x}=(x_1,...,x_k)$ denote a generic vector with positive integer entries such that $\sum_{j=1}^k x_j=\ell$, let $W(\mathbf{x})$ be the set of the greatest entries of $\mathbf{x}$, and, for $j\in\{1,i\}$, let

- $A_j^{(!)}=\{\mathbf{x}\,|\,W(\mathbf{x})=\{j\}\}$,
- $A_j^{(=)}=\{\mathbf{x}\,|\,1,i\in W(\mathbf{x})\}$,
- $A_1^{(\neq)}=\{\mathbf{x}\,|\,1\in W(\mathbf{x})\wedge i\notin W(\mathbf{x})\wedge|W(\mathbf{x})|>1\}$ and
- $A_i^{(\neq)}=\{\mathbf{x}\,|\,i\in W(\mathbf{x})\wedge 1\notin W(\mathbf{x})\wedge|W(\mathbf{x})|>1\}$.

It holds

$$\begin{aligned}\Pr\left(\mathrm{maj}_\ell(u)=j\right) &= \sum_{\mathbf{x}\in A_j^{(!)}} \Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)\Pr\left(\mathrm{maj}_\ell(u)=j\middle|\,\bar{X}^{(\ell)}=\mathbf{x}\right)\\ &+ \sum_{\mathbf{x}\in A_j^{(=)}} \Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)\Pr\left(\mathrm{maj}_\ell(u)=j\middle|\,\bar{X}^{(\ell)}=\mathbf{x}\right)\\ &+ \sum_{\mathbf{x}\in A_j^{(\neq)}} \Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)\Pr\left(\mathrm{maj}_\ell(u)=j\middle|\,\bar{X}^{(\ell)}=\mathbf{x}\right)\\ &= \sum_{\mathbf{x}\in A_j^{(!)}} \Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)+\sum_{\mathbf{x}\in A_j^{(=)}}\frac{\Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)}{|W(\mathbf{x})|}\\ &+ \sum_{\mathbf{x}\in A_j^{(\neq)}}\frac{\Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)}{|W(\mathbf{x})|} \end{aligned} \tag{9}$$

Let

$$\sigma(\mathbf{x})=(x_i,...,x_{i-1},x_1,x_{i+1},...,x_k)$$

be the vector function that swaps the entries x_1 and x_i in $\mathbf{x}$. σ is clearly a bijection between the sets $A_1^{(!)}$,$A_1^{(=)}$,$A_1^{(\neq)}$ and $A_i^{(!)}$, $A_i^{(=)}$, $A_i^{(\neq)}$, respectively, namely

$$\sigma: A_1^{(!)} \twoheadrightarrow A_i^{(!)},\ \sigma: A_1^{(=)} \twoheadrightarrow A_i^{(=)},\ \sigma: A_1^{(\neq)} \twoheadrightarrow A_i^{(\neq)}$$

where $\twoheadrightarrow$ denotes a bijection.

Moreover, for all $\mathbf{x}\in A_j^{(=)}$, it holds

$$\Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)=\Pr\left(\bar{X}^{(\ell)}=\sigma(\mathbf{x})\right).$$

Therefore

$$\sum_{\mathbf{x}\in A_1^{(=)}}\Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right)=\sum_{\sigma(\mathbf{x})\in A_1^{(=)}}\Pr\left(\bar{X}^{(\ell)}=\sigma(\mathbf{x})\right)=\sum_{\mathbf{x}\in A_i^{(=)}}\Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right). \tag{10}$$

Furthermore, for all $\mathbf{x}\in A_1^{(\neq)}$, we have

$$\begin{aligned}\Pr\left(\bar{X}^{(\ell)}=\mathbf{x}\right) &= \binom{\ell}{x_1\ ...\ x_k}p_1^{x_1}\dots p_i^{x_i}\dots p_k^{x_k}\\ &> \binom{\ell}{x_1\ ...\ x_k}p_1^{x_i}\dots p_i^{x_1}\dots p_k^{x_k}\\ &= \Pr\left(\bar{X}_1^{(\ell)}=\sigma(\mathbf{x})\right), \end{aligned} \tag{11}$$

where $\sigma(\mathbf{x}) \in A_i^{(\neq)}$. From Eq. (11) we thus have that

$$\sum_{\mathbf{x}\in A_1^{(\neq)}} \Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right) > \sum_{\sigma(\mathbf{x})\in A_1^{(\neq)}} \Pr\left(\bar{X}^{(\ell)} = \sigma\left(\mathbf{x}\right)\right) = \sum_{\mathbf{x}\in A_i^{(\neq)}} \Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right). \quad (12)$$

From Eq. (9), (10) and (12) we finally get

$$\begin{aligned}
&\Pr\left(\mathrm{maj}_\ell(u) = 1\right) - \Pr\left(\mathrm{maj}_\ell(u) = i\right) \\
&\quad = \sum_{\mathbf{x}\in A_1^{(!)}} \Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right) + \sum_{\mathbf{x}\in A_1^{(=)}} \frac{\Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right)}{|W\left(\mathbf{x}\right)|} \\
&\quad + \sum_{\mathbf{x}\in A_1^{(\neq)}} \frac{\Pr\left(\bar{X}^{(l)} = \mathbf{x}\right)}{|W\left(\mathbf{x}\right)|} - \sum_{\mathbf{x}\in A_i^{(!)}} \Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right) \\
&\quad - \sum_{\mathbf{x}\in A_i^{(=)}} \frac{\Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right)}{|W\left(\mathbf{x}\right)|} - \sum_{\mathbf{x}\in A_i^{(\neq)}} \frac{\Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right)}{|W\left(\mathbf{x}\right)|} \\
&\quad \geq \sum_{\mathbf{x}\in A_1^{(!)}} \Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right) - \sum_{\mathbf{x}\in A_i^{(!)}} \Pr\left(\bar{X}^{(\ell)} = \mathbf{x}\right) \\
&\quad = \Pr\left(W(\bar{X}^{(\ell)}) = \{X_1^{(\ell)}\}\right) - \Pr\left(W(\bar{X}^{(\ell)}) = \{X_i^{(\ell)}\}\right),
\end{aligned}$$

concluding the proof of Lemma 10. □

Intuitively, Lemma 10 says that the set of events in which a tie occurs among the most frequent opinions in the node's sample of observed messages does not favor the probability that the node picks the wrong opinion. Thus, by avoiding considering those events, we get a lower bound on $\Pr(\mathrm{maj}_\ell(u) = 1) - \Pr(\mathrm{maj}_\ell(u) = i)$.

Thanks to Lemma 10, the proof of Eq. (6) reduces to proving the following.

Lemma 11. *For any fixed k, and with $\bar{X}$ defined as in Lemma 10, we have*

$$\Pr(X_1^{(\ell)} > X_2^{(\ell)}, ..., X_k^{(\ell)}) - \Pr(X_i^{(\ell)} > X_1^{(\ell)}, ..., X_{i-1}^{(\ell)}, X_{i+1}^{(\ell)}, ..., X_k^{(\ell)}) \geq \sqrt{2\ell/\pi}\, \frac{g\left(\delta, \ell\right)}{4^{k-2}}. \quad (13)$$

Proof. We prove Eq. (13) by induction. Lemma 9 provides us with the base case for $k = 2$. Let us assume that, for $k \leq \kappa$, Eq. (13) holds. For $k = \kappa + 1$, by

using the law of total probability, we have

$$\begin{aligned}
&\Pr\left(X_1^{(\ell)} > X_2^{(\ell)}, ..., X_{\kappa+1}^{(\ell)}\right) - \Pr\left(X_i^{(\ell)} > X_1^{(\ell)}, ..., X_{i-1}^{(\ell)}, X_{i+1}^{(\ell)}, ..., X_{\kappa+1}^{(\ell)}\right) \\
&\geq \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \Pr\left(X_1^{(\ell)} > X_2^{(\ell)}, ..., X_{\kappa+1}^{(\ell)} \middle| X_{\kappa+1}^{(\ell)} = h\right) \Pr\left(X_{\kappa+1}^{(\ell)} = h\right) \\
&- \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \Pr\left(X_i^{(\ell)} > X_1^{(\ell)}, ..., X_{i-1}^{(\ell)}, X_{i+1}^{(\ell)}, ..., X_{\kappa+1}^{(\ell)} \middle| X_{\kappa+1}^{(\ell)} = h\right) \Pr\left(X_{\kappa+1}^{(\ell)} = h\right).
\end{aligned} \tag{14}$$

Now, $\arg\max_j\{X_j^{(\ell)}\} = X_i^{(\ell)}$ and $X_{\kappa+1}^{(\ell)} \leq \left\lfloor \frac{\ell}{\kappa+1} \right\rfloor$ together imply $X_i^{(\ell)} > X_{\kappa+1}^{(\ell)}$. Thus, in the r.h.s. of Eq. (14), we have

$$\Pr\left(X_1^{(\ell)} > X_2^{(\ell)}, ..., X_{\kappa+1}^{(\ell)} \middle| X_{\kappa+1}^{(\ell)} = h\right) = \Pr\left(X_1^{(l)} > X_2^{(\ell)}, ..., X_{\kappa}^{(\ell)} \middle| X_{\kappa+1}^{(\ell)} = h\right)$$

and

$$\begin{aligned}
&\Pr\left(X_i^{(\ell)} > X_1^{(\ell)}, ..., X_{i-1}^{(\ell)}, X_{i+1}^{(\ell)}, ..., X_{\kappa+1}^{(\ell)} \middle| X_{\kappa+1}^{(\ell)} = h\right) \\
&= \Pr\left(X_i^{(\ell)} > X_1^{(\ell)}, ..., X_{i-1}^{(\ell)}, X_{i+1}^{(\ell)}, ..., X_{\kappa}^{(\ell)} \middle| X_{\kappa+1}^{(\ell)} = h\right).
\end{aligned}$$

Moreover, $X^{(\ell)}$ follows a multinomial distribution with parameters $\mathbf{p}$ and ℓ. Thus $X_k^{(\ell)} = h$ implies that the remaining entries $X_1^{(\ell)}, ..., X_{k-1}^{(\ell)}$ follow a multinomial distribution with $l-h$ trials, and distribution $(\frac{p_1}{1-p_k}, ..., \frac{p_{k-1}}{1-p_k})$. Let $Y^{(\ell-h)} = (Y_1^{(\ell-h)}, ..., Y_{k-1}^{(\ell-h)})$ be the distribution of $X_1^{(\ell)}, ..., X_{k-1}^{(\ell)}$ conditional on $X_k^{(\ell)} = h$. From Eq. (14) we get

$$\begin{aligned}
&\Pr\left(X_1^{(\ell)} > X_2^{(\ell)}, ..., X_{\kappa+1}^{(\ell)}\right) - \Pr\left(X_i^{(\ell)} > X_1^{(\ell)}, ..., X_{i-1}^{(\ell)}, X_{i+1}^{(\ell)}, ..., X_{\kappa+1}^{(\ell)}\right) \\
&\geq \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \Pr\left(X_1^{(\ell)} > X_2^{(\ell)}, ..., X_{\kappa}^{(l)} \middle| X_{\kappa+1}^{(\ell)} = h\right) \Pr\left(X_{\kappa+1}^{(\ell)} = h\right) \\
&- \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \Pr\left(X_i^{(\ell)} > X_1^{(\ell)}, ..., X_{i-1}^{(\ell)}, X_{i+1}^{(\ell)}, ..., X_{\kappa}^{(\ell)} \middle| X_{\kappa+1}^{(\ell)} = h\right) \Pr\left(X_{\kappa+1}^{(\ell)} = h\right) \\
&\geq \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \left(\Pr\left(Y_1^{(\ell-h)} > Y_2^{(\ell-h)}, ..., Y_{\kappa}^{(\ell-h)}\right) - \right. \\
&\left. - \Pr\left(Y_i^{(l-h)} > Y_1^{(l-h)}, ..., Y_{i-1}^{(l-h)}, Y_{i+1}^{(l-h)}, ..., Y_{\kappa}^{(l-h)}\right)\right) \Pr\left(X_{\kappa+1}^{(l)} = h\right).
\end{aligned} \tag{15}$$

Now, using the inductive hypothesis on the r.h.s. of Eq. (15) we get

$$\begin{aligned}
&\sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \left(\Pr\left(Y_1^{(\ell-h)} > Y_2^{(\ell-h)}, ..., Y_\kappa^{(\ell-h)} \right) \right. \\
&\left. - \Pr\left(Y_i^{(\ell-h)} > Y_1^{(\ell-h)}, ..., Y_{i-1}^{(\ell-h)}, Y_{i+1}^{(\ell-h)}, ..., Y_\kappa^{(\ell-h)} \right) \right) \Pr\left(X_{\kappa+1}^{(\ell)} = h \right) \\
&\geq \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \left(\sqrt{\frac{2\ell - 2h}{\pi}} \frac{g\left(\delta, \ell - h\right)}{4^{\kappa-2}} \right) \Pr\left(X_{\kappa+1}^{(\ell)} = h \right) \\
&\geq \sqrt{\frac{2\ell}{\pi}} \frac{g\left(\delta, \ell\right)}{4^{\kappa-2}} \cdot \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \sqrt{1 - \frac{h}{\ell}} \Pr\left(X_{\kappa+1}^{(\ell)} = h \right),
\end{aligned}$$

where, in the last inequality, we used the fact that g is a non-increasing function w.r.t. the second argument (see Lemma 15).

It remains to show that

$$\sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \sqrt{1 - \frac{h}{l}} \Pr\left(X_{\kappa+1}^{(\ell)} = h \right) \geq \frac{1}{4}.$$

Let $W_{\kappa+1}^{(\ell)}$ be a r.v. with probability distribution $Bin(\ell, \frac{1}{\kappa+1})$. Since $X_{\kappa+1}^{(\ell)} \sim Bin(\ell, p_{\kappa+1})$ with $p_{\kappa+1} \leq \frac{1}{\kappa+1}$, a standard coupling argument (see for example [22, Exercise 1.1.]), enables to show that

$$\Pr\left(X_{\kappa+1}^{(\ell)} \leq h \right) \geq \Pr\left(W_{\kappa+1}^{(\ell)} \leq h \right).$$

Hence, we can apply the central limit theorem (Lemma 14) on $W_{\kappa+1}^{(\ell)}$, and get that, for any $\tilde{\epsilon} \leq \frac{2-\sqrt{3}}{4}$, there exists some fixed constant ℓ_0 such that, for $\ell \geq \ell_0$, we have

$$\Pr\left(X_{\kappa+1}^{(\ell)} \leq \frac{\ell}{\kappa+1} \right) \geq \Pr\left(W_{\kappa+1}^{(\ell)} \leq \frac{\ell}{\kappa+1} \right) \geq \left(\frac{1}{2} - \tilde{\epsilon} \right). \tag{16}$$

By using Eq. (16), for $\ell \geq \ell_0$ we finally get that

$$\begin{aligned}
\sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \sqrt{1 - \frac{h}{\ell}} \Pr\left(X_{\kappa+1}^{(\ell)} = h \right) &\geq \sqrt{1 - \frac{\left\lfloor \frac{\ell}{\kappa+1} \right\rfloor}{\ell}} \cdot \sum_{h=0}^{\lfloor \frac{\ell}{\kappa+1} \rfloor} \Pr\left(X_{\kappa+1}^{(\ell)} = h \right) \\
&\geq \sqrt{1 - \frac{2}{\kappa+1}} \cdot \Pr\left(X_{\kappa+1}^{(\ell)} \leq \frac{\ell}{\kappa+1} \right) \\
&\geq \sqrt{\frac{\kappa-1}{\kappa+1}} \cdot \left(\frac{1}{2} - \tilde{\epsilon} \right) \geq \sqrt{\frac{1}{3}} \cdot \left(\frac{1}{2} - \tilde{\epsilon} \right) \geq \frac{1}{4},
\end{aligned}$$

concluding the proof that

$$\Pr\left(\mathrm{maj}_\ell(u)=1\right)-\Pr\left(\mathrm{maj}_\ell(u)=i\right)\geq\sqrt{\frac{2\ell}{\pi}}\frac{g\left(\delta,\ell\right)}{e^{(k-2)\ln 4}}.$$

□

By using Proposition 1, we can then prove Lemma 12.

Lemma 12. *W.h.p., at the end of Stage 2, all nodes support the initial plurality opinion.*

Proof. Let $\delta=\Omega(\sqrt{\log n/n})$ be the bias of the opinion distribution at the beginning of a generic phase $j<T'$ of Stage 2. Thanks to Proposition 1, by choosing the constant c of the phase length large enough, in process **P** we get that $\Pr\left(\mathrm{maj}_\ell(u)=m\right)-\Pr\left(\mathrm{maj}_\ell(u)=i\right)\geq\alpha\delta$ for some constant $\alpha>1$ (provided that $\delta\leq 1/2$). Hence, by applying Lemma 16 in Appendix A with $\theta=\frac{\alpha}{4}\delta$, we get $\Pr(c_m^{(\tau_j)}-c_i^{(\tau_j)}\leq\alpha\delta/2)\leq\exp(-(\alpha\delta)^2n/16)\leq n^{-\tilde{\alpha}}$ for some constant $\tilde{\alpha}$ that is large enough to apply Lemma 2. Therefore, until $\delta\geq 1/2$, in process **P** we have that $c_m^{(\tau_j)}-c_i^{(\tau_j)}\geq\alpha\delta/2$ holds w.h.p. From the previous equation it follows that, after T' phases, the protocol has reached an opinion distribution with a bias greater than $1/2$. Thus, by a direct application of Lemma 16 and Lemma 2 to $c_m^{(\tau_{T'})}-c_i^{(\tau_{T'})}$, we get that, w.h.p., $c_m^{(\tau_{T'})}-c_i^{(\tau_{T'})}=1$, concluding the proof. □

Finally, the time efficiency claimed in Theorem 1 and Theorem 2 directly follows from Lemma 12, while the required memory follows from the fact that in each phase each node needs only to count how many times it has received each opinion, i.e. to count up to at most $O(\frac{1}{\epsilon^2}\log n)$ w.h.p.

4 On the Notion of (ϵ,δ)-Majority-Preserving Matrix

In this section we discuss the notion of (ϵ,δ)-m.p. noise matrix introduced by Definition 2. Let us consider Eq. (2). The matrix P represents the "perturbation" introduced by the noise, and so $(\mathbf{c}\cdot P)_m-(\mathbf{c}\cdot P)_i$ measures how much information the system is losing about the correct opinion m, in a single communication round. An (ϵ,δ)-m.p. noise matrix is a noise matrix that preserves at least an ϵ fraction of bias, provided the initial bias is at least δ. The (ϵ,δ)-m.p. property essentially characterizes the amount of noise beyond which some coordination problems cannot be solved without further hypotheses on the nodes' knowledge of the matrix P. To see why this is the case, consider an (ϵ,δ)-m.p. noise matrix for which there is a δ-biased opinion distribution $\tilde{\mathbf{c}}$ such that $(\tilde{\mathbf{c}}\cdot P)_m-(\tilde{\mathbf{c}}\cdot P)_i<0$ for some opinion i. Given opinion distribution $\tilde{\mathbf{c}}$, *from each node's perspective, opinion m does not appear to be the most frequent*

opinion. Indeed, the messages that are received are more likely to be i than m. Thus, plurality consensus cannot be solved from opinion distribution $\tilde{\mathbf{c}}$.

Observe that verifying whether a given matrix P is (ϵ, δ)-m.p. with respect to opinion m consists in checking whether for each $i \neq m$ the value of the following linear program is at least $\epsilon\delta$:

$$\begin{aligned} &\text{maximize } (P \cdot \mathbf{c})_m - (P \cdot \mathbf{c})_i \\ &\text{subject to } \sum_j c_j = 1, \\ &\qquad \text{and } \forall j,\, c_j \geq 0, c_m - c_j - \delta \geq 0. \end{aligned}$$

We now provide some negative and positive examples of (ϵ, δ)-m.p. noise matrices. First, we note that a natural matrix property such as being diagonally dominant does not imply that the matrix is (ϵ, δ)-m.p. For example, by multiplying the following diagonally dominant matrix by the δ-biased opinion distribution $\mathbf{c} = (1/2 + \delta, 1/2 - \delta, 0)^\intercal$, we see that it does not even preserve the majority opinion at all when $\epsilon, \delta < 1/6$:

$$\begin{pmatrix} \frac{1}{2}+\epsilon & 0 & \frac{1}{2}-\epsilon \\ \frac{1}{2}-\epsilon & \frac{1}{2}+\epsilon & 0 \\ 0 & \frac{1}{2}-\epsilon & \frac{1}{2}+\epsilon \end{pmatrix}.$$

On the other hand, the following natural generalization of the noise matrix in [27] (see Eq. (1)), is (ϵ, δ)-m.p. for every $\delta > 0$ with respect to any opinion:

$$(P)_{i,j} = p_{i,j} = \begin{cases} \frac{1}{k} + \epsilon & \text{if } i = j, \\ \frac{1}{k} - \frac{\epsilon}{k-1} & \text{otherwise.} \end{cases}$$

More generally, let P be a noise matrix such that

$$(P)_{i,j} = \begin{cases} p & \text{if } i = j, \\ q_l \leq q_{i,j} \leq q_u & \text{otherwise,} \end{cases} \tag{17}$$

for some positive numbers p, q_u and q_l. Since

$$\begin{aligned} (P\mathbf{c})_m - (P\mathbf{c})_i &= pc_m + \sum_{j \neq m} q_{j,m} c_j - pc_i - \sum_{j \neq i} q_{j,i} c_j \\ &\geq p(c_m - c_i) + \sum_{j \neq m} q_l c_j - \sum_{j \neq i} q_u c_j \\ &\geq p(c_m - c_i) + q_l(1 - c_m) - q_u(1 - c_i) \\ &\geq p(c_m - c_i) + q_l - q_l c_m - q_u + q_u c_i \\ &\geq p(c_m - c_i) - q_u(c_m - c_i) - (q_u - q_l) \\ &\geq (p - q_u)(c_m - c_i) - (q_u - q_l) \\ &\geq (p - q_u)\delta - (q_u - q_l). \end{aligned} \tag{18}$$

By defining $\epsilon = (p - q_u)/2$, we get that the last line in Eq. (18) is greater than $\epsilon\delta$ iff $(p - q_u)\delta/2 \geq (q_u - q_l)$, which gives a sufficient condition for any matrix of the form given in Eq. (17) for being (ϵ, δ)-m.p.

5 Conclusion

In this paper, we solved the general version of rumor spreading and plurality consensus in biological systems. That is, we have solved these problems for an arbitrarily large number k of opinions. We are not aware of realistic biological contexts in which the number of opinions might be a function of the number n of individuals. Nevertheless, it could be interesting, at least from a conceptual point to view, to address rumor spreading and plurality consensus in a scenario in which the number of opinions varies with n. This appears to be a technically challenging problem. Indeed, extending the results in the extended abstract of [27] from 2 opinions to any constant number k of opinions already required to use complex tools. Yet, several of these tools do not apply if k depends on n. This is typically the case of Proposition 1. We let as an open problem the design of stochastic tools enabling to handle the scenario where $k = k(n)$.

Acknowledgments. We thank the anonymous reviewers of an earlier version of this work for their constructive criticisms and comments, which were of great help in improving the results and their presentation.

References

[1] Abdullah, M. A., and Draief, M. Global majority consensus by local majority polling on graphs of a given degree sequence. *Discrete Applied Mathematics 180* (2015), 1–10.

[2] Afek, Y., Alon, N., Bar-Joseph, Z., Cornejo, A., Haeupler, B., and Kuhn, F. Beeping a maximal independent set. *Distributed computing 26*, 4 (2013), 195–208.

[3] Afek, Y., Alon, N., Barad, O., Barkai, N., Bar-Joseph, Z., and Hornstein, E. A biological solution to a fundamental distributed computing problem. *Science 331*, 6014 (2011), 183–185.

[4] Angluin, D., Aspnes, J., and Eisenstat, D. A simple population protocol for fast robust approximate majority. *Distributed Computing 21*, 2 (2008), 87–102.

[5] Angluin, D., Aspnes, J., Eisenstat, D., and Ruppert, E. The computational power of population protocols. *Distributed Computing 20*, 4 (2007), 279–304.

[6] D. Alistarh, J. Aspnes, and R. Gelashvili Space-Optimal Majority in Population Protocols. In *Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2018, pp. 2221–2239.

[7] Aspnes, J., and Ruppert, E. An introduction to population protocols. In *Middleware for Network Eccentric and Mobile Applications*. Springer, 2009, pp. 97–120.

[8] Ame, J.-M., Rivault, C. and Deneubourg, J.-L. Cockroach aggregation based on strain odour recognition. *Animal Behaviour*, vol. 68, no. 4, pp. 793–801, Oct. 2004.

[9] Becchetti, L., Clementi, A., Natale, E., Pasquale, F., and Silvestri, R. Plurality consensus in the gossip model. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms* (2015), SIAM, pp. 371–390.

[10] Becchetti, L., Clementi, A., Natale, E., Pasquale, F., Silvestri, R., and Trevisan, L. Simple dynamics for plurality consensus. *Distributed computing*, vol. 30 no. 4 (2016), 1–14.

[11] Ben-Shahar, O., Dolev, S., Dolgin, A., and Segal, M. Direction election in flocking swarms. In *Proceedings of the 6th International Workshop on Foundations of Mobile Computing* (2010), ACM, pp. 73–80.

[12] Berenbrink, P., Friedetzky, T., Kling, P., Mallmann-Trenn, F., and Wastell, C. Plurality Consensus in Arbitrary Graphs: Lessons Learned from Load Balancing, in *Proceedings of the 24th Annual European Symposium on Algorithms* (2016), vol. 57, p. 10:1–10:18.

[13] Boczkowski, L., Korman, A., and Natale, E. Minimizing Message Size in Stochastic Communication Patterns: Fast Self-Stabilizing Protocols with 3 bits. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms* (2017), SIAM, pp. 2540–2559.

[14] Boczkowski, L., Korman, A., and Natale, E. Limits for Rumor Spreading in Stochastic Populations. In *Proceedings of the 9th Innovations in Theoretical Computer Science Conference* (2018), vol. 94, p. 49:1–49:21.

[15] Cardelli, L., and Csikász-Nagy, A. The cell cycle switch computes approximate majority. *Scientific Reports 2* (2011), 656–656.

[16] Chazelle, B. Natural algorithms. In *Proceedings of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms* (2009), SIAM, pp. 422–431.

[17] Cooper, C., Elsässer, R., and Radzik, T. The power of two choices in distributed voting. In *Automata, Languages, and Programming*, vol. 8573 of *Lecture Notes in Computer Science*. Springer, 2014, pp. 435–446.

[18] Conradt, L. and Roper, T. J. Group decision-making in animals. *Nature*, vol. 421, no. 6919, pp. 155–158, Jan. 2003.

[19] Demers, A., Greene, D., Hauser, C., Irish, W., Larson, J., Shenker, S., Sturgis, H., Swinehart, D., and Terry, D. Epidemic algorithms for replicated database maintenance. In *Proceedings of the 6th Annual ACM Symposium on Principles of Distributed Computing* (1987), ACM, pp. 1–12.

[20] Doerr, B., Goldberg, L. A., Minder, L., Sauerwald, T., and Scheideler, C. Stabilizing consensus with the power of two choices. In *Proceedings of the 23th annual ACM symposium on Parallelism in algorithms and architectures* (2011), ACM, pp. 149–158.

[21] Draief, M., and Vojnovic, M. Convergence speed of binary interval consensus. *SIAM J. on Control and Optimization 50*, 3 (2012), 1087–1109.

[22] Dubhashi, D. P., and Panconesi, A. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

[23] Elsässer, R., Friedetzky, T., Kaaser, D., Mallmann-Trenn, F., and Trinker, H. Brief Announcement: Rapid Asynchronous Plurality Consensus In *Proceedings of the 37th ACM Symposium on Principles of Distributed Computing* (2017), ACM, pp. 363–365.

[24] El Gamal, A., and Kim, Y.-H. *Network information theory*. Cambridge University Press, 2011.

[25] Franks, N. R., Dornhaus, A., Fitzsimmons, J. P. and Stevens, M. Speed versus accuracy in collective decision making. *Proc. Biol. Sci.*, vol. 270, no. 1532, pp. 2457–2463, Dec. 2003.

[26] Feinerman, O., Haeupler, B., and Korman, A. Breathe before speaking: Efficient information dissemination despite noisy, limited and anonymous communication. In *Proceedings of the 34th ACM Symposium on Principles of Distributed Computing* (2014), ACM, pp. 114–123. Extended abstract of [27].

[27] Feinerman, O., Haeupler, B., and Korman, A. Breathe before speaking: Efficient information dissemination despite noisy, limited and anonymous communication. *Distributed Computing* vol. 30 no. 5 (2015), 1–17.

[28] Giakkoupis, G., Berenbrink, P., Friedetzky, T., and Kling, P. Efficient Plurality Consensus, or: The benefits of cleaning up from time to time. in *Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming* (2016) vol. 55, p. 136:1–136:14.

[29] Ghaffari, M., and Parter, M. A Polylogarithmic Gossip Algorithm for Plurality Consensus. In *Proceedings of the 36th ACM Symposium on Principles of Distributed Computing* 2016, ACM, pp. 117–126.

[30] Franks, N. R., Pratt, S. C., Mallon, E. B., Britton, N. F., and Sumpter, D. J. Information flow, opinion polling and collective intelligence in house–hunting social insects. *Philosophical Transactions of the Royal Society of London B: Biological Sciences 357*, 1427 (2002), 1567–1583.

[31] Jung, K., Kim, B. Y., and Vojnović, M. Distributed ranking in networks with limited memory and communication. In *Proceedings of the 2012 IEEE International Symposium on Information Theory* (2012), IEEE, pp. 980–984.

[32] Karp, R., Schindelhauer, C., Shenker, S., and Vocking, B. Randomized rumor spreading. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (2000), IEEE, pp. 565–574.

[33] Kempe, D., Dobra, A., and Gehrke, J. Gossip-based computation of aggregate information. In *Proceedings of the 44st Annual Symposium on Foundations of Computer Science* (2003), IEEE, pp. 482–491.

[34] Korman, A., Greenwald, E., and Feinerman, O. Confidence sharing: an economic strategy for efficient information flows in animal groups. *PLoS computational biology 10*, 10 (2014), e1003862–e1003862.

[35] Land, M. and Belew, R. No Perfect Two-State Cellular Automata for Density Classification Exists. *Physical Review Letters*, vol. 74, no. 25, pp. 5148–5150, Jun. 1995.

[36] Mitzenmacher, M., and Upfal, E. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis.* Cambridge University Press, 2005.

[37] Perron, E., Vasudevan, D., and Vojnovic, M. Using three states for binary consensus on complete graphs. In *Proc. of 28th IEEE INFOCOM* (2009).

[38] Pittel, B. On spreading a rumor. *SIAM J. Appl. Math. 47*, 1 (Mar. 1987), 213–223.

[39] Robbins, H. A remark on stirling's formula. *American Mathematical Monthly* (1955), 26–29.

[40] T. D. Seeley and S. C. Buhrman. Group decision making in swarms of honey bees. *Behav. Ecol. Sociobiol.*, vol. 45, no. 1, pp. 19–31, Jan. 1999.

[41] T. D. Seeley and P. K. Visscher. Quorum sensing during nest-site selection by honeybee swarms. *Behav. Ecol. Sociobiol.*, vol. 56, no. 6, pp. 594–601, Oct. 2004.

[42] Sumpter, D. J., Krause, J., James, R., Couzin, I. D., and Ward, A. J. Consensus decision making by fish. *Current Biology 18*, 22 (2008), 1773–1777.

APPENDIX

A Technical Tools

Lemma 13. *For any integer $r \geq 1$ it holds*

$$\frac{2^{2r}}{\sqrt{\pi r}} e^{\frac{1}{9r}} \leq \binom{2r}{r} \leq \frac{2^{2r}}{\sqrt{\pi r}} e^{\frac{1}{8r}}.$$

Proof. By using Stirling's approximation [39]

$$\sqrt{2\pi r}\left(\frac{r}{e}\right)^r e^{\frac{1}{12r+1}} \leq r! \leq \sqrt{2\pi r}\left(\frac{r}{e}\right)^r e^{\frac{1}{12r}},$$

we have

$$\binom{2r}{r} = \frac{(2r)!}{(r!)^2} \geq \frac{\sqrt{2\pi 2r}\left(\frac{2r}{e}\right)^{2r} e^{\frac{1}{12r+1}}}{\left(\sqrt{2\pi r}\left(\frac{r}{e}\right)^r e^{\frac{1}{12r}}\right)^2} = \frac{\sqrt{4\pi r}\left(\frac{2r}{e}\right)^{2r} e^{\frac{2}{12r+1}}}{2\pi r\left(\frac{r}{e}\right)^{2r} e^{\frac{1}{24r}}}$$
$$= \frac{2^{2r}}{\sqrt{\pi r}} e^{\frac{2}{12r+1}-\frac{1}{24r}} \geq \frac{2^{2r}}{\sqrt{\pi r}} e^{\frac{1}{9r}}.$$

The proof of the upper bound is analogous (swap $e^{\frac{1}{12r+1}}$ and $e^{\frac{1}{12r}}$ in the first inequality). □

Lemma 14. *Let $X_1, ..., X_n$ be a random sample from a Bernoulli(p) distribution with $p \in (0,1)$ constant, and let $Z \sim N(0,1)$. It holds*

$$\lim_{n\to\infty} \sup_{z\in\mathbb{R}} \left| \Pr\left(\frac{\sum_{i=1}^n X_i - pn}{\sqrt{n}} \leq z\right) - \Pr\left(Z \leq \frac{z}{\sqrt{p(1-p)}}\right)\right| = 0.$$

Lemma 15. *The function*

$$g(x,y) = \begin{cases} x\left(1-x^2\right)^{\frac{y-1}{2}} & \text{if } x < \frac{1}{\sqrt{y}}, \\ \frac{1}{\sqrt{y}}\left(1-\frac{1}{y}\right)^{\frac{y-1}{2}} & \text{if } x \geq \frac{1}{\sqrt{y}}, \end{cases}$$

with $x \in [0,1]$ and $y \in [1,+\infty)$ is non-decreasing w.r.t. x and non-increasing w.r.t. y.

Proof. To show that $g(x,y)$ is non-decreasing w.r.t. x, observe that

$$\frac{\partial}{\partial x} g(x,y) = \left(\left(1-x^2\right)^{\frac{y-1}{2}} - 2x^2\left(\frac{y-1}{2}\right)\left(1-x^2\right)^{\frac{y-1}{2}}\right)$$

for $x < y^{-\frac{1}{2}} < 1$, and

$$\left(1-x^2\right)^{\frac{y-1}{2}} - 2x^2\left(\frac{y-1}{2}\right)\left(1-x^2\right)^{\frac{y-1}{2}} \geq 0$$

for $x < y^{-\frac{1}{2}}$.

To show that $g(x,y)$ is non-increasing w.r.t. y, observe that this is true for $x < y^{-\frac{1}{2}}$. For $x \geq y^{-\frac{1}{2}}$, since

$$\frac{\partial}{\partial y}\left(\log y^{-\frac{1}{2}} + \frac{y-1}{2}\log\left(1-\frac{1}{y}\right)\right) = \frac{\partial}{\partial y}\left(\frac{y-1}{2}\log\left(y-1\right) - \frac{y}{2}\log y\right) \leq 0,$$

we have

$$\frac{\partial}{\partial y} g\left(x,y\right) = \frac{\partial}{\partial y}\exp\left\{\log y^{-\frac{1}{2}} + \frac{y-1}{2}\log\left(1-\frac{1}{y}\right)\right\} \leq 0,$$

concluding the proof. □

Lemma 16. *Let $\{X_t\}_{t\in[n]}$ be n i.i.d. random variables such that*

$$X_t = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } r, \\ -1 & \text{with probability } q. \end{cases}$$

with $p + r + q = 1$. It holds

$$\Pr\left(\sum_i X_t \leq (1-\theta)\,\mathbb{E}\left[\sum_i X_t\right] - \theta n\right) \leq \exp\left(-\frac{\theta^2}{4}\left(\mathbb{E}\left[\sum_i X_t\right] + n\right)\right).$$

Proof. Let us define the r.v.

$$Y_t = \frac{X_t + 1}{2}. \tag{19}$$

We can apply the Chernoff-Hoeffding bound to Y_t (see Theorem 1.1 in [22]), obtaining

$$\Pr\left(\sum_i Y_i \leq (1-\theta)\,\mathbb{E}\left[\sum_i Y_i\right]\right) \leq \exp\left(-\frac{\theta^2}{2}\mathbb{E}\left[\sum_i Y_i\right]\right)$$

for any $\theta \in (0,1)$. Substituting Eq. (19) we have

$$\begin{aligned}\Pr\left(\sum_i X_t + n \leq (1-\theta)\left(\mathbb{E}\left[\sum_i X_t\right] + n\right)\right) \\ = \Pr\left(\sum_i X_t \leq (1-\theta)\,\mathbb{E}\left[\sum_i X_t\right] - \theta n\right) \\ \leq \exp\left(-\frac{\theta^2}{4}\left(\mathbb{E}\left[\sum_i X_t\right] + n\right)\right),\end{aligned}$$

concluding the proof. □

B Removing the Parity Assumption on ℓ

The next lemma shows that, for $k = 2$, the increment of bias at the end of each phase of Stage 2 in the process $\mathbf{P}$ is non-decreasing in the value of ℓ, regardless of its parity. In particular, since Proposition 1 is proven by induction, and since the value of ℓ affects only the base case, the next lemma implies also the same kind of monotonicity for general k.

Lemma 17. *Let $k = 2$, $a = 1$, let ℓ be odd, and let $(\mathbf{c} \cdot P)_1 \geq (\mathbf{c} \cdot P)_2$. The rule of Stage 2 of the protocol is such that*

$$\begin{aligned} \Pr\left(\mathrm{maj}_\ell(u) = 1\right) &= \Pr\left(\mathrm{maj}_{\ell+1}(u) = 1\right) \leq \Pr\left(\mathrm{maj}_{\ell+2}(u) = 1\right), \\ \Pr\left(\mathrm{maj}_\ell(u) = 2\right) &= \Pr\left(\mathrm{maj}_{\ell+1}(u) = 2\right) \geq \Pr\left(\mathrm{maj}_{\ell+2}(u) = 2\right). \end{aligned} \tag{20}$$

Proof. To simplify notation, let $p_1 = (\mathbf{c} \cdot P)_1$ and $p_2 = (\mathbf{c} \cdot P)_2$. By definition, we have

$$\begin{aligned} \Pr\left(\mathrm{maj}_\ell(u) = 1\right) &= \Pr\left(X_1^{(\ell)} \geq \left\lceil \frac{\ell}{2} \right\rceil\right), \\ \Pr\left(\mathrm{maj}_{\ell+1}(u) = 1\right) &= \Pr\left(X_1^{(\ell+1)} > \frac{\ell+1}{2}\right) + \frac{1}{2}\Pr\left(X_1^{(\ell+1)} = \frac{\ell+1}{2}\right), \\ \Pr\left(\mathrm{maj}_{\ell+2}(u) = 1\right) &= \Pr\left(X_1^{(\ell+2)} \geq \left\lceil \frac{\ell+2}{2} \right\rceil\right), \end{aligned}$$

where $X_1^{(\ell)}$, $X_1^{(\ell+1)}$ and $X_1^{(\ell+2)}$ are binomial r.v. with probability p_1 and number of trials ℓ, $\ell+1$, and $\ell+2$, respectively. We can view $X_1^{(\ell)}$, $X_1^{(\ell+1)}$, and $X_1^{(\ell+2)}$ as the sum of ℓ, $\ell+1$ and $\ell+2$ $Bernoulli(p_1)$ r.v., respectively. In particular, let Y and Y' be independent r.v. with distribution $Bernoulli(p_1)$. We can couple $X_1^{(\ell)}$, $X_1^{(\ell+1)}$ and $X_1^{(\ell+2)}$ as follows:

$$X_1^{(\ell+1)} = X_1^{(\ell)} + Y$$

and

$$X_1^{(\ell+2)} = X_1^{(\ell+1)} + Y'$$

Since ℓ is odd, observe that if $X_1^{(\ell)} > \lceil \frac{\ell}{2} \rceil$, then $\mathrm{maj}_\ell(u) = 1$ regardless of the value of Y, and similarly if $X_1^{(\ell)} < \lceil \frac{\ell}{2} \rceil$ then $\mathrm{maj}_\ell(u) = 2$. Thus we have

$$\begin{aligned} &\Pr\left(\mathrm{maj}_{\ell+1}(u) = 1\right) = \sum_{i=1}^{\ell} \Pr\left(\mathrm{maj}_{\ell+1}(u) = 1 \middle| X_1^{(l)} = i\right) \Pr\left(X_1^{(\ell)} = i\right) \\ &= \sum_{i > \lceil \frac{\ell}{2} \rceil}^{\ell} \Pr\left(X_1^{(\ell)} = i\right) + \Pr\left(\mathrm{maj}_{\ell+1}(u) = 1 \middle| X_1^{(\ell)} = \left\lceil \frac{\ell}{2} \right\rceil\right) \Pr\left(X_1^{(\ell)} = \left\lceil \frac{\ell}{2} \right\rceil\right) \\ &+ \Pr\left(\mathrm{maj}_{\ell+1}(u) = 1 \middle| X_1^{(\ell)} = \left\lfloor \frac{\ell}{2} \right\rfloor\right) \Pr\left(X_1^{(\ell)} = \left\lfloor \frac{\ell}{2} \right\rfloor\right). \end{aligned} \tag{21}$$

As for the last two terms in the previous equation, we have that

$$\Pr\left(\mathrm{maj}_{\ell+1}(u)=1\middle|\, X_1^{(\ell)}=\left\lceil\frac{\ell}{2}\right\rceil\right)=\Pr\left(Y=1\right)+\Pr\left(Y=0\right)\frac{1}{2}, \tag{22}$$

and

$$\Pr\left(\mathrm{maj}_{\ell+1}(u)=1\middle|\, X_1^{(\ell)}=\left\lfloor\frac{l}{2}\right\rfloor\right)=\Pr\left(Y=1\right)\frac{1}{2}. \tag{23}$$

Moreover, by a direct calculation one can verify that

$$\Pr\left(X_1^{(\ell)}=\left\lceil\frac{\ell}{2}\right\rceil\right)=\frac{\Pr\left(Y=0\right)}{\Pr\left(Y=1\right)}\Pr\left(X_1^{(\ell)}=\left\lfloor\frac{\ell}{2}\right\rfloor\right). \tag{24}$$

From Eq. (22), (23) and (24) it follows that

$$\begin{aligned}
&\Pr\left(\mathrm{maj}_{\ell+1}(u)=1\middle|\, X_1^{(\ell)}=\left\lceil\frac{\ell}{2}\right\rceil\right)\Pr\left(X_1^{(\ell)}=\left\lceil\frac{\ell}{2}\right\rceil\right)\\
&+\Pr\left(\mathrm{maj}_{\ell+1}(u)=1\middle|\, X_1^{(\ell)}=\left\lfloor\frac{\ell}{2}\right\rfloor\right)\Pr\left(X_1^{(\ell)}=\left\lfloor\frac{\ell}{2}\right\rfloor\right)\\
&=\left(\Pr\left(Y=1\right)+\Pr\left(Y=0\right)\cdot\frac{1}{2}\right)\Pr\left(X_1^{(\ell)}=\left\lceil\frac{\ell}{2}\right\rceil\right)+\left(\Pr\left(Y=1\right)\cdot\frac{1}{2}\right)\Pr\left(X_1^{(\ell)}=\left\lfloor\frac{\ell}{2}\right\rfloor\right)\\
&=\Pr\left(X_1^{(\ell)}=\left\lceil\frac{\ell}{2}\right\rceil\right)\left(\Pr\left(Y=1\right)+\frac{\Pr\left(Y=0\right)}{2}+\frac{\Pr\left(Y=1\right)}{2}\frac{\Pr\left(Y=0\right)}{\Pr\left(Y=1\right)}\right)\\
&=\Pr\left(X_1^{(\ell)}=\left\lceil\frac{\ell}{2}\right\rceil\right).
\end{aligned} \tag{25}$$

By plugging Eq. (25) in Eq. (21) we get

$$\Pr\left(\mathrm{maj}_{\ell}(u)=1\right)=\Pr\left(\mathrm{maj}_{\ell+1}(u)=1\right).$$

The proof that

$$\Pr\left(\mathrm{maj}_{\ell}(u)=2\right)=\Pr\left(\mathrm{maj}_{\ell+1}(u)=1\right)$$

is analogous, proving the first part of Eq. (20).

As for the second part, observe that if $X_1^{(\ell+1)}>\frac{\ell+1}{2}$, then $\mathrm{maj}_{\ell+2}(u)=1$ regardless of the value of Y', and similarly if $X_1^{(\ell+1)}<\frac{\ell+1}{2}$ then $\mathrm{maj}_{\ell+2}(u)=2$. Observe also that

$$\Pr\left(\mathrm{maj}_{\ell+2}(u)=1\middle|\, X_1^{(\ell+1)}=\frac{l+1}{2}\right)=\Pr\left(Y=1\right)=p_1.$$

Because of the previous observations and the hypothesis that $p_1\geq\frac{1}{2}$, we have

that

$$\begin{aligned}
&\Pr\left(\mathrm{maj}_{\ell+2}(u)=1\right)\\
&=\sum_{i=0}^{\ell}\Pr\left(\mathrm{maj}_{\ell+2}(u)=1\,\middle|\,X_1^{(\ell+1)}=i\right)\Pr\left(X_1^{(\ell+1)}=i\right)\\
&=\sum_{i>\frac{\ell+1}{2}}^{l}\Pr\left(X_1^{(\ell+1)}=i\right)+\Pr\left(\mathrm{maj}_{\ell+2}(u)=1\,\middle|\,X_1^{(\ell+1)}=\frac{\ell+1}{2}\right)\Pr\left(X_1^{(\ell+1)}=\frac{\ell+1}{2}\right)\\
&=\sum_{i>\frac{\ell+1}{2}}^{\ell}\Pr\left(X_1^{(\ell+1)}=i\right)+p_1\cdot\Pr\left(X_1^{(\ell+1)}=\frac{\ell+1}{2}\right)\\
&\geq\sum_{i>\frac{\ell+1}{2}}^{\ell}\Pr\left(X_1^{(\ell+1)}=i\right)+\frac{1}{2}\Pr\left(X_1^{(\ell+1)}=\frac{\ell+1}{2}\right)=\Pr\left(\mathrm{maj}_{\ell+1}(u)=1\right).
\end{aligned}\tag{26}$$

The proof of

$$\Pr\left(\mathrm{maj}_{\ell+2}(u)=2\right)\leq\Pr\left(\mathrm{maj}_{\ell+2}(u)=2\right)$$

is the same up to the inequality in (26), whose direction is reversed because $p_2\leq\frac{1}{2}$. □

C Rumor spreading with $\epsilon=\Theta(n^{-\frac{1}{4}-\eta})$

In [27] it is shown that at the end of Stage 1 the bias toward the correct opinion is at least $\epsilon^{T+2}/2$ and, at the beginning of Stage 2, they assume a bias toward the correct opinion of $\Omega(\sqrt{\log n/n})$. In this section, we show that, when $\epsilon=\Theta(n^{-\frac{1}{4}-\eta})$ for some $\eta\in(0,1/4)$, the protocol considered by [27] and us cannot solve the rumor-spreading and the plurality consensus problem in time $\Theta(\log n/\epsilon^2)$.

First, observe that when $\epsilon=\Theta(\sqrt{\log n/n})$ the length of the first phase of Stage 1 is $\Theta\left(\log n/\epsilon^2\right)=\Omega(n\log n)$, which implies that, w.h.p., each node gets at least one message from the source during the first phase. Thus, thanks to our analysis of Stage 2 we have that when $\epsilon=\Theta(\sqrt{\log n/n})$ the protocol effectively solves the rumor-spreading problem, w.h.p., in time $\Theta(\log n/\epsilon^2)$.

In general, for $\epsilon<n^{-1/2-\eta}$ for some constant $\eta>0$, if we adopt the second stage right from the beginning (which means that the source node sends ϵ^{-2} messages), we get that, w.h.p., all nodes receive at least $\log n/(\epsilon^2 n)$ messages. Thus, by a direct application of Lemma 16, after the first phase we get an $\sqrt{\log n/n}$-biased opinion distribution, w.h.p., and Stage 2 correctly solves the problem according to Theorem 2.

However, when $\epsilon=\Theta(n^{-\frac{1}{4}-\eta})$ for some $\eta>0$, from Claim 2 and Lemma 7 we have that, after phase 0 in opinion distribution $\mathbf{c}$, at most $\mathcal{O}\left(\log n/\epsilon^2\right)=\mathcal{O}(n^{\frac{1}{2}+2\eta}\log n)$ nodes are opinionated, and $\mathbf{c}$ is $\frac{\epsilon}{2}$-biased. Each node that gets

opinionated in phase 1 receives a message pushed from some node of $\mathbf{c}$, and, because of the noise, the value of this message is distributed according to $\mathbf{c}^{(\tau_0)} \cdot P$. It follows that $\mathbf{c}$ is an $\epsilon^2/2$-biased opinion distribution with $\epsilon^2 = n^{-\frac{1}{2}-2\eta}$ which is much smaller than the $\Omega(\sqrt{\log n/n})$ bound required for the second stage.

We believe that no minor modification of the protocol proposed here can correctly solve the noisy rumor-spreading problem when $\epsilon = \Theta(n^{-\frac{1}{4}-\eta})$ in time $\mathcal{O}\left(\log n/\epsilon^2\right)$.