

The Minority Dynamics and the Power of Synchronicity

Luca Becchetti* Andrea Clementi† Francesco Pasquale† Luca Trevisan‡
 Robin Vacus§ Isabella Ziccardi‡

Abstract

We study the minority-opinion dynamics over a fully-connected network of n nodes with binary opinions. Upon activation, a node receives a sample of opinions from a limited number of neighbors chosen uniformly at random. Each activated node then adopts the opinion that is least common within the received sample.

Unlike all other known consensus dynamics, we prove that this elementary protocol behaves in dramatically different ways, depending on whether activations occur sequentially or in parallel. Specifically, we show that its expected consensus time is exponential in n under asynchronous models, such as asynchronous GOSSIP. On the other hand, despite its chaotic nature, we show that it converges within $O(\log^2 n)$ rounds with high probability under synchronous models, such as synchronous GOSSIP.

Finally, our results shed light on the bit-dissemination problem, that was previously introduced to model the spread of information in biological scenarios. Specifically, our analysis implies that the minority-opinion dynamics is the first *stateless* solution to this problem, in the parallel passive-communication setting, achieving convergence within a polylogarithmic number of rounds. This, together with a known lower bound for sequential stateless dynamics, implies a parallel-vs-sequential gap for this problem that is nearly quadratic in the number n of nodes. This is in contrast to all known results for problems in this area, which exhibit a linear gap between the parallel and the sequential setting.

1 Introduction

1.1 Dynamics and Consensus In distributed computing, the term *dynamics* is used to refer to distributed processes in which each node of a network updates its state on the basis of a simple update rule applied to messages received in the last round of communication; furthermore, the network is *anonymous*, meaning that nodes do not have distinguished names and incoming messages lack any identification of the sender.

The study of dynamics, of their global properties and of their applications is an active research topic that touches several scientific areas [2, 6, 7, 16, 18, 27]. Among other applications, dynamics are suitable to model the restrictions on computation and communication of IoT protocols, and to study the way information spreads and groups self-organize in biological models and in models of animal behavior.

In this paper we consider *opinion dynamics*. These are dynamics in which every node has an opinion, which is an element of a finite set, and possibly, additional state information. Over time, these opinions are affected by the information exchanged by nodes. Typically, one is interested in whether all nodes eventually reach a *consensus* configuration in which they all share the same opinion, and in how quickly this happens.

The *consensus* problem is a fundamental task in distributed computing [3, 15, 17, 29] and a problem of considerable interest in the study of dynamics. We provide the definition of a basic version of the problem.

DEFINITION 1. (CONSENSUS PROBLEM) *We say that a dynamics in which every node, in every round, holds an opinion solves the consensus problem if*

- *For every initial configuration of opinions, the dynamics reaches with probability 1 a configuration in which all nodes have the same opinion, and such opinion was held by at least one node in the initial configuration;*

*Sapienza University of Rome, Rome, Italy. becchetti@diag.uniroma1.it. Supported by the ERC Advanced Grant 788893 AMDROMA, PNRR MUR project PE0000013-FAIR”.

†Tor Vergata University of Rome, Rome, Italy. clementi/pasquale@mat.uniroma2.it. Partially supported by Spoke1 “FutureHPC & BigData” of the Italian Research Center on High-Performance Computing, Big Data and Quantum Computing (ICSC) funded by MUR Missione 4 Componente 2 Investimento 1.4: Potenziamento strutture di ricerca e creazione di “campioni nazionali” di R&S (M4C2-19) - Next Generation EU (NGEU).

‡Bocconi University, Milan, Italy. l.trevisan/isabella.ziccardi@unibocconi.it. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 834861).

§CNRS, IRIF, Paris, France. rvacus@irif.fr.

- Once all nodes have the same opinion, the configuration of opinions does not change in subsequent rounds.

The number of rounds until all nodes share the same opinion for the first time is called convergence time. For every initial configuration, convergence time is a random variable over the randomness of the dynamics.

Since dynamics are meant to capture agents with very limited computational and communication abilities, several models have been defined and analyzed that place restrictions on how communication takes place (e.g. LOCAL, various versions of GOSSIP, POPULATION PROTOCOL, etc. [4, 28, 31]). All dynamics we consider in this paper operate in a standard model in which communication in each round is restricted, so that each active node (that is, each node that updates its opinion in that round) is only able to receive messages from k randomly chosen neighbors (where k is a parameter of the model). This is called k -uniform GOSSIP PULL in the literature [12, 13, 31] and we will refer to it as k -PULL in the remainder.

Moreover, we will always assume that the underlying communication network is the complete graph on n nodes. We will be interested in two well-studied variants of this model [12], which differ for the schedule with which nodes are activated: in the *asynchronous sequential* k -PULL, in each round exactly one node becomes active, while in the *synchronous parallel* k -PULL, all nodes are active and their updates are simultaneous in each round.

In both settings described above, a solution to the *consensus* problem is given by the VOTER MODEL [1, 25, 28], which is the dynamics in which each active node receives the opinion of one random neighbor and then adopts that opinion as its own (thus the voter model operates as a 1-PULL dynamics).

It is known that, for every initial Boolean configuration of opinions, the VOTER MODEL's expected convergence time is $\Theta(n^2)$ in the asynchronous sequential setting [1, 25] and $\Theta(n)$ in the synchronous parallel setting [21].

In the k -MAJORITY dynamics, each active node receives the Boolean opinions of k random neighbors (k odd) and updates its opinion to the majority opinion among those k . This is a k -PULL dynamics, and the VOTER MODEL can be seen as implementing 1-MAJORITY. It is easy to see that k -MAJORITY also provides a solution to the *consensus* problem for every odd k , and it is known [16] that, for odd $k \geq 3$, in the synchronous parallel model, for every initial configuration, convergence time is at most $O(\log n)$ both in expectation and in high probability¹. It is folklore that for odd $k \geq 3$, the convergence time in the asynchronous sequential model is $O(n \log n)$. The k -MAJORITY dynamics has other desirable properties, such as being fault-tolerant and such that the consensus opinion is likely to be the majority, or close to the majority, of the original opinions. Moreover, bounds on the convergence time of 3-MAJORITY have been derived for the case of non-binary opinions in [6, 7, 10, 16, 19].

We note that in both above examples above, there is a $\Theta(n)$ gap between parallel and sequential convergence times. Moreover, a $\tilde{\Theta}(n)$ gap on the convergence time holds between sequential and parallel communication models for other forms of *consensus* as well [2, 7, 20].

This is natural, because we are applying the same update rule, but n times simultaneously in each round in the parallel case, and once per round in the sequential case. Indeed, in all the dynamics for which we have tight bounds both in the sequential and parallel models we see a $\tilde{\Theta}(n)$ gap between them. Because of synchronicity of the updates however, one round in the parallel model might potentially have different effects than n rounds in the sequential model, and we will return to this point in our main results.

1.2 The bit-dissemination problem More recently, inspired by the study of distributed biological systems, a variant of *consensus*, called the *bit-dissemination* problem has been considered in [11, 5, 23, 8].

DEFINITION 2. (BIT-DISSEMINATION PROBLEM) *In the bit-dissemination problem, every node, at every round, holds an opinion. One of the nodes, called the source, holds an opinion that it knows to be correct, and never changes it. The other nodes update their opinions according to the update rule and do not know the identity of the source. We say that a dynamics solves the bit-dissemination problem if:*

- For every initial configuration, with probability 1, the dynamics eventually reaches a configuration in which all nodes share the same opinion as the source;
- Once all nodes have the same opinion as the source, the configuration of opinions does not change in subsequent rounds.

¹We say that an event holds *with high probability* if it holds with probability $1 - n^{-\Omega(1)}$.

The number of rounds until all nodes have the same opinion as the source for the first time is called the convergence time.

This information-dissemination problem finds its main motivations in modeling communication processes that take place in biological systems [4, 32], where the correct information may include, e.g., knowledge about a preferred migration route [22, 26], the location of a food source [14], or the need to recruit agents for a particular task [30].

In these applications, it is well motivated to restrict to dynamics that use *passive-communication*, meaning that the only kind of message that a node can receive from a neighbor is the neighbor's current opinion. Note that the VOTER MODEL and the k -MAJORITY dynamics that we described above are examples of passive-communication opinion dynamics.

Another desirable property in these applications is for dynamics to be *self-stabilizing*. This means that not only does the dynamics converge to the desired stable configuration for every initial configuration of the opinions, but it also converges regardless of the initial contents of any additional state that the update rule relies on. The reason that this is called a self-stabilizing property is that if, at some point, an adversary corrupts the internal state of some nodes, convergence is still assured (because we could take the round in which this corruption happens as the first round).

We call an opinion dynamics *stateless* if the update rule by which an active node updates its opinion depends only on the messages received in that round, and if nodes do not keep any additional state information except for their opinions. Note that a stateless dynamics is always self-stabilizing because there is nothing for an adversary to corrupt. The VOTER MODEL and k -MAJORITY dynamics are examples of stateless dynamics.

In [23], Korman and Vacus discuss the difficulty of developing passive-communication and self-stabilizing dynamics for the *bit-dissemination* problem. To see why passive-communication is difficult to achieve for the *bit-dissemination* problem, consider that if, at some round, the opinion of the source is in the minority, then we want to use an update rule that increases the number of nodes with the minority opinion. On the contrary, if the opinion of the source is in the majority, then we want an update rule that decreases the number of nodes with the minority opinion. These two scenarios, however, are indistinguishable to nodes that are only allowed to see other nodes' opinions and that do not know the identity of the source.

Indeed, previous protocols for this problem exploit non-passive forms of communication and/or do not achieve self-stabilizing *bit-dissemination* [9, 11, 5]. Korman and Vacus [23] develop and analyze a protocol for *bit-dissemination* in the synchronous parallel k -PULL with passive communication in which every node, at every round, sees the opinion of $k = \Theta(\log n)$ random nodes (that is, communication follows the k -PULL model), and they show that their protocol converges with high probability in $O(\log^{5/2} n)$ rounds. Their protocol is not stateless, as it involves a trend-following update rule: roughly speaking, each node adopts the opinion whose number of occurrences in the messages of the current round has increased compared to the previous round, breaking ties arbitrarily. This requires each node to store the number of opinions of either type seen in the previous round, which take $\log k = \log \log n + O(1)$ bits of storage. Their protocol is self-stabilizing, and so it converges even if, in the first round, nodes start with adversarially chosen "false memories" of a previous round. Their work leaves open whether convergence in poly $\log n$ rounds for the *bit-dissemination* problem can be achieved, in the synchronous parallel k -PULL, by a passive-communication *stateless* dynamics, as is possible for the *consensus* problem.

Becchetti et al. [8] show that in the asynchronous sequential k -PULL, there exists an $\Omega(n^2)$ lower bound for the expected convergence time of *any* passive-communication stateless dynamics for the *bit-dissemination* problem. This result holds under the even stronger model that, in every round, the active node is given full access to the opinions of *all* other nodes.

Given that there is typically a $\tilde{\Theta}(n)$ gap between the convergence time of stateless passive-communication dynamics in the parallel versus sequential model, it would seem natural to conjecture a $\tilde{\Omega}(n)$ lower bound for the convergence time of *bit-dissemination* in the synchronous parallel passive-communication k -PULL for stateless dynamics, and that the memory of the past round used by Korman and Vacus would be essential to achieve poly $\log n$ convergence time. A bit surprisingly however, we show this is not the case, as we discuss below.

1.3 Our results We study the *minority* update rule, that defines a stateless dynamics in the passive-communication setting in the k -PULL communication model.

DEFINITION 3. (*k*-MINORITY DYNAMICS²) *In each round, each active node, upon seeing the opinions of k random neighbors, updates its opinion to the minority opinion among those held by the k neighbors. Ties are broken randomly.*

We will only consider binary opinions and, for simplicity, odd k , so that ties never occur. According to this rule, if a node sees unanimity among its k sampled neighbors, it adopts those neighbors' unanimous opinion. Otherwise, it adopts whatever opinion is held by the fewer neighbors.

In the *consensus* problem, for any initial configuration, this dynamics eventually, with probability 1, reaches a configuration in which all the nodes have the same opinion, and, in the *bit-dissemination* problem, it reaches with probability 1 the configuration in which all nodes agree with the source. This is because, in the Markov chain that describes the change of the configuration of opinions over rounds, the only absorbing states are those in which all nodes agree, and such configurations are reachable from any initial configuration.

Concerning the speed with which such convergence happens, consider the sequential setting first. In the asynchronous sequential model, if we start from a balanced configuration in which each opinion is held by 50% of the nodes, it is very difficult for the *k*-MINORITY dynamics to make progress, because as soon as an opinion gains more followers than the other, in subsequent rounds there will be a bias toward adopting the minority opinion and the configuration will drift back towards the initial 50%-50% configuration. Indeed, we can prove an exponential lower bound for the round complexity of *k*-MINORITY.

THEOREM 1.1. *There exist initial configurations from which the expected convergence time of *k*-MINORITY in the asynchronous sequential *k*-PULL is $2^{\Omega(kn)}$.*

In the synchronous parallel communication model, the *k*-MINORITY dynamics exhibits chaotic behavior. If, for example, the dynamics starts from a balanced configuration, in the next round we may expect a slight imbalance due to random noise, let's say slightly more 0s than 1s. In the subsequent round, this will cause a drift in the opposite direction, and we will see more 1s than 0s, with an increased imbalance. The majority value will keep alternating, and the imbalance will keep increasing, until a round in which the imbalance becomes such that the minority opinion is held by $\ll n/k$ nodes. At that point the number of nodes with the minority opinion will grow by about a factor of k each time, and then the majority value will start alternating again, and so on.

The main technical contribution of this paper is that, despite this chaotic behavior, when $k = \Omega(\sqrt{n \log n})$, we are able to analyze the behavior of the minority dynamics, by dividing the set of possible configurations into a finite number of ranges, and by understanding how likely it is for the configuration to move across these ranges.

Our analysis shows that *k*-MINORITY converges in polylogarithmic time in the parallel setting, both when all nodes follow the protocol (*consensus* problem) and when one node is a source that never changes its opinion (*bit-dissemination* problem).

THEOREM 1.2. *If $185\sqrt{n \log n} \leq k \leq \frac{n}{2}$, then, in the synchronous parallel *k*-PULL, *k*-MINORITY solves the consensus and its convergence time is $O(\log n)$ in expectation and $O(\log^2 n)$ rounds w.h.p.*

This also provides an efficient (i.e. in a polylogarithmic number of rounds) solution to the *bit-dissemination* problem via a stateless passive-communication dynamics, answering a question left open by the work of Korman and Vacus [23].

THEOREM 1.3. *If $185\sqrt{n \log n} \leq k \leq \frac{n}{2}$, then, in the synchronous parallel *k*-PULL, *k*-MINORITY solves the bit-dissemination problem and its convergence time is $O(\log n)$ in expectation and $O(\log^2 n)$ rounds w.h.p.*

We mentioned above that for all the dynamics for which we have a tight analysis there is a $\tilde{\Theta}(n)$ gap between the sequential and parallel convergence time, at least in the context of stateless opinion dynamics with passive-communication.

Our results show a natural example of a dynamics for which this gap is exponential in n , going from $O(\log^2 n)$ to $2^{\Omega(kn)}$. Moreover, we have a natural problem, the *bit-dissemination* problem, for which there is a nearly-quadratic gap between parallel passive-communication stateless solutions, for which we show an $O(\log^2 n)$ upper bound, and sequential passive-communication stateless solutions, for which an $\Omega(n^2)$ lower bound was known.

Roadmap The rest of this paper is organized as follows. [Section 2](#) introduces notation and notions that will be used throughout the paper and then provides a self-contained overview of the main technical part, proving poly-logarithmic convergence time of k -MINORITY for the *consensus* and *bit-dissemination* problems. In more detail, in [Section 2.2](#), we provide the high-level proofs of [Theorem 1.2](#) and [Theorem 1.3](#), we state our main technical lemmas, we discuss the main ideas in our proofs, and we describe the main technical challenges that we have to overcome. Full proofs of all technical results above are given in [Section 3](#), presented in the same order as in [Section 2.2](#). [Section 4](#) presents the proof of [Theorem 1.1](#), based on the analysis of the birth-death chain of k -MINORITY. Finally, [Section 5](#) discusses some technical or more general questions that this work leaves open and that in our opinion deserve further investigation.

2 k -Minority in the Parallel Model

In this section, we analyze the k -MINORITY dynamics in the synchronous parallel communication model, over the complete graph K_n . We consider the case in which the sample size parameter k depends on n , i.e. $k = k(n)$ and we prove [Theorem 1.2](#) and [Theorem 1.3](#) stated in the previous section.

2.1 Notation and preliminaries We use the following Markov chain to describe the evolution of the k -MINORITY dynamics on K_n . Its state space is $\{0, \dots, \frac{n}{2}\} \times \{0, 1\}$, and, at every round $t \geq 0$, the state of the process at round t is given by the size $\{m_t\}_t$ and the opinion $\{\mathcal{O}_t\}_t$ of the minority. To get some intuition about the key quantities that govern the evolution of the process, we introduce the following notions.

DEFINITION 4. (WRONG NODES) *We say node v is wrong in round t if and only if: i) $m_t > 0$, and ii) v adopts the majority opinion at time $t + 1$. Observe that, if $m_t > 0$, the following events may cause v to be wrong in round t :*

$$\begin{aligned} \mathcal{A}_t(v) &= \{v \text{ samples more than } \frac{k}{2} \text{ nodes holding the minority opinion}\} \\ \mathcal{B}_t(v) &= \{v \text{ samples } k \text{ nodes holding the majority opinion}\}. \end{aligned}$$

We define $W_t = \sum_{v \in V} \mathbf{1}_{\mathcal{A}_t(v)}$ as the r.v. counting the number of nodes that sample more than $\frac{k}{2}$ nodes with the minority opinion in round t , and $U_t = \sum_{v \in V} \mathbf{1}_{\mathcal{B}_t(v)}$ as the r.v. counting the number of nodes that sample a unanimity of nodes holding the majority opinion in round t .

The following is a simple fact that relates m_{t+1} to the number of wrong nodes in the previous round:

FACT 2.1. *We deterministically have that*

$$m_{t+1} = \begin{cases} W_t + U_t, & \text{if } W_t + U_t < n/2, \\ n - W_t - U_t, & \text{otherwise.} \end{cases}$$

Moreover, $\mathcal{O}_{t+1} = 1 - \mathcal{O}_t$ in the first case and $\mathcal{O}_{t+1} = \mathcal{O}_t$ in the second.

Proof. In the first case, $n - W_t - U_t > \frac{n}{2}$ nodes will correctly assess and adopt the minority opinion, which will thus become majoritarian in the next round. The remaining (wrong) $W_t + U_t$ nodes will adopt the majority opinion, which will become minoritarian in round $t + 1$. I.e., $m_{t+1} = W_t + U_t$ and $\mathcal{O}_{t+1} = 1 - \mathcal{O}_t$ in this case. The second case is similar, this time with $m_{t+1} = n - W_t - U_t \leq \frac{n}{2}$ and $\mathcal{O}_{t+1} = \mathcal{O}_t$. \square

The next lemma clarifies that the events introduced in [Definition 4](#) play distinct roles in different regimes of the process. More precisely, they are *almost* mutually exclusive, in the sense that the probability that two nodes u and v exist, such that both $\mathcal{A}_t(u)$ and $\mathcal{B}_t(v)$ occur in any given round t , is essentially negligible.

LEMMA 2.1. *Assume $k \geq 5 \log n$. The following holds: i) if $m_t \leq \frac{n}{2} - n\sqrt{\frac{1.5 \log n}{k}}$, then $W_t = 0$ w.h.p.; ii) if $m_t \geq \frac{3n \log n}{k}$, then $U_t = 0$ w.h.p.*

[Lemma 2.1](#) is proved as part of [Lemma 3.1](#) (points (c) and (d)), while we next discuss some consequences. First, in every round t , w.h.p. we have either $W_t = 0$ or $U_t = 0$, which implies that the process m_t can essentially be

described by one of the two random variables U_t or W_t . Moreover, it implies that if $\frac{3n \log n}{k} \leq m_t \leq \frac{n}{2} - n\sqrt{\frac{1.5n \log n}{k}}$, then k -MINORITY achieves consensus in round $t + 1$, w.h.p. We call the above safe range for m_t the *green area*.

Thanks to the behaviour of variables U_t and W_t discussed above, we can identify a constant number of areas (see [Figure 1](#)), each defined by a specific range for m_t , showing that with constant probability, k -MINORITY reaches the green area (and thus achieves consensus in the next round w.h.p.) within a logarithmic number of rounds: This is the basic strategy towards proving [Theorem 1.2](#). From this, the same bound can be shown for the *bit-dissemination* problem ([Theorem 1.3](#)).

As simple as it may seem, implementing the above strategy is not straightforward. The reason is that the evolution of m_t is highly non-monotonic, exhibiting a behaviour that strongly depends on the area m_t belongs to. In particular, the process might jump over non-contiguous areas and might, in principle, jump back and forth between any two of them. Moreover, the presence of two fixed points for $\mathbf{E}[m_{t+1} | m_t]$ highlights the existence of two (unstable, as we shall see) equilibria, in which the process might get stuck. We also remark that in principle, one might attempt a proof strategy that differs from the one outlined by [Figure 1\(a\)](#). However, while this may be possible, it does not seem a trivial endeavour, since [Figure 1](#) actually seems to summarize “preferred” (or more likely) state transitions in the global Markov chain that describes the process.

2.2 Proof of main results In this section, we prove [Theorem 1.2](#) and [Theorem 1.3](#). For the sake of the analysis, we partition the set of all possible configurations into a small, constant number of *areas*, that in the sequel will be named *colors*, each of them identified by a specific interval of the range $\{0, 1, \dots, \frac{n}{2}\}$ of possible values for the random variable m_t . Informally, the warmer the color, the slower the convergence will be from the corresponding subset, as illustrated in [Figure 1](#). Formally, they are defined as follows:

$$\begin{aligned} \text{BLUE}_1 &= \left\{ m \in \mathbb{N} : 1 \leq m \leq \frac{n \log 2}{k} - \sqrt{\frac{n \log 2}{k}} \right\} \\ \text{RED} &= \left\{ m \in \mathbb{N} : \frac{n \log 2}{k} - \sqrt{\frac{n \log 2}{k}} \leq m \leq \frac{n \log 2}{k} + \sqrt{\frac{n \log 2}{k}} \right\} \\ \text{BLUE}_2 &= \left\{ m \in \mathbb{N} : \frac{n \log 2}{k} + \sqrt{\frac{n \log 2}{k}} \leq m \leq \frac{n}{k} \log \left(\frac{k}{4 \log n} \right) \right\} \\ \text{YELLOW} &= \left\{ m \in \mathbb{N} : \frac{n}{k} \log \left(\frac{k}{4 \log n} \right) \leq m \leq \frac{3n \log n}{k} \right\} \\ \text{GREEN} &= \left\{ m \in \mathbb{N} : \frac{3n \log n}{k} \leq m \leq \frac{n}{2} - n\sqrt{\frac{2 \log n}{k}} \right\} \\ \text{ORANGE} &= \left\{ m \in \mathbb{N} : \frac{n}{2} - n\sqrt{\frac{2 \log n}{k}} \leq m \leq \frac{n}{2} \right\} \end{aligned}$$

In order to understand the behaviour of the process in the different areas, it is useful to consider the function f represented on [Figure 1 \(b\)](#). Informally, f can be seen as an approximation of $m \mapsto \mathbf{E}[m_{t+1} | m_t = m]$. Formally, we define it as follows:

- when $m \in A = \text{BLUE}_1 \cup \text{RED} \cup \text{BLUE}_2$, we set

$$f(m) = \min\{\mathbf{E}[U_t | m_t = m], n - \mathbf{E}[U_t | m_t = m]\},$$

since in this range, $W_t = 0$ w.h.p.

- when $m \in B = \text{GREEN}$, we set $f(m) = 0$, since in this range, $m_{t+1} = 0$ w.h.p.
- when $m \in C = \text{ORANGE}$, we set $f(m) = \mathbf{E}[W_t | m_t = m]$, since in this range, $U_t = 0$ w.h.p.

Using these insights, we will show that, starting from any initial configuration, the process m_t follows a short path over the colored areas, eventually landing into the green area and then consensus within an additional step. This typical behaviour is summarized by [Figure 1 \(a\)](#). However, this path is somewhat chaotic, in the sense that m_t does not tend to evolve monotonically; instead, it can jump across non-adjacent areas over consecutive rounds.

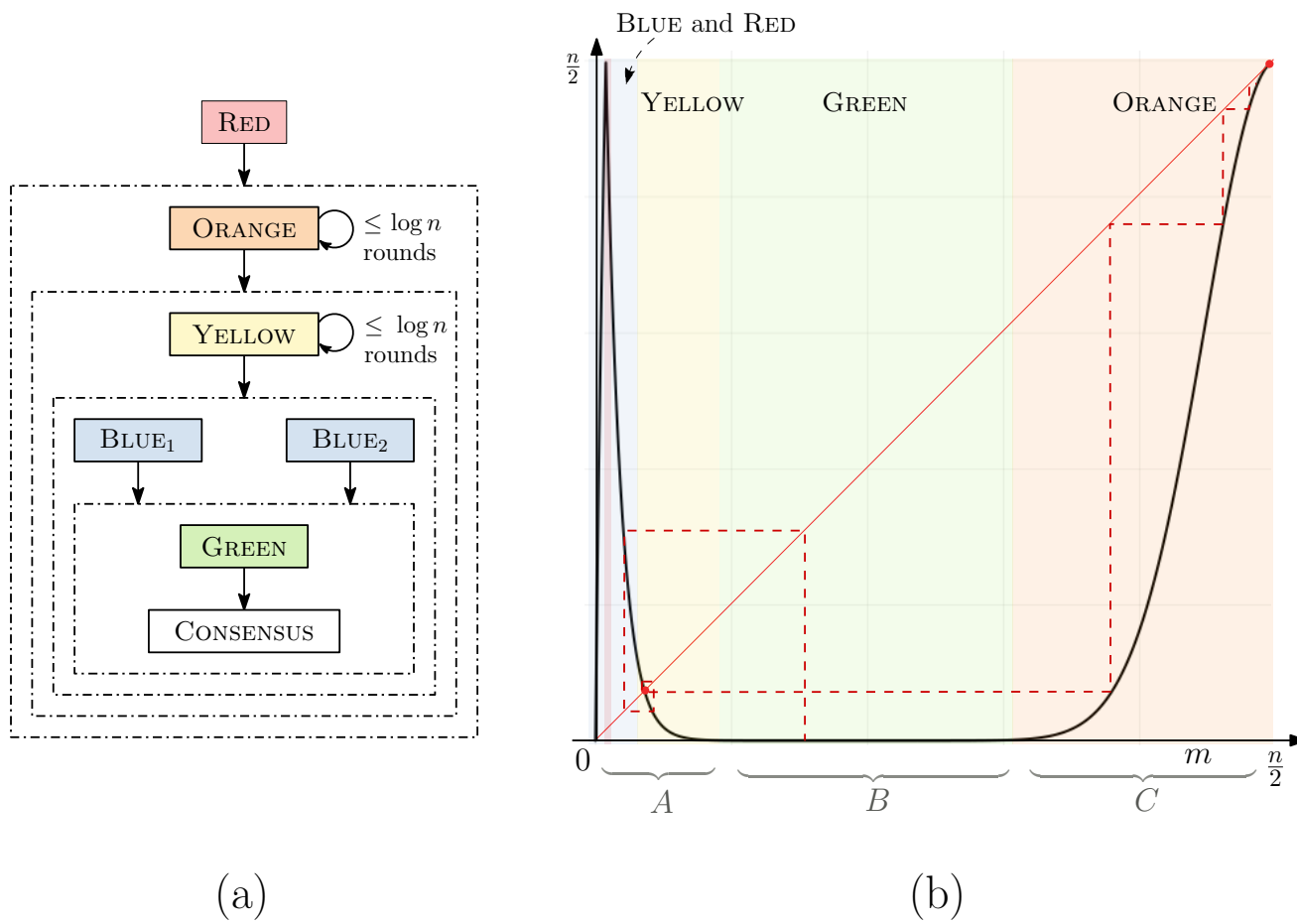


Figure 1: (a) Roadmap for the proof of **Theorem 1.2**. Unless explicitly indicated, transitions occur with constant probability in a constant number of rounds. Next to self-loops, we give the number of rounds the process spends in corresponding areas. (b) Function f , satisfying $f(m) \approx \mathbf{E}[m_{t+1} \mid m_t = m]$, and the colored areas. The red circles are the fixed points of $f(m)$, i.e. $m_1 = f(m_1)$ and $m_2 = f(m_2) = \frac{n}{2}$. The red dashed track represent some iteration of $f(m)$, starting with a configuration close to $\frac{n}{2}$ – illustrating the non-monotonic, chaotic behavior of the process.

Intermediate Results In order to formalize the aforementioned arguments, we state a few intermediate lemmas, which we will prove in subsequent sections.

LEMMA 2.2. (RED AREA) For any $t \geq 1$ and $m \in [\frac{n}{2}]$, $\Pr[m_{t+1} \notin \text{RED} \mid m_t = m] \geq c$, for a suitable positive constant c .

The proof of Lemma 2.2 uses a simple variance argument relying on the modest width of the red area, and is deferred to Section 3.1. The next lemma, however, poses some technical challenges.

LEMMA 2.3. (ORANGE AREA) For any $t_0 \geq 1$ and $m_{t_0} \in \text{ORANGE}$, let

$$T = \inf\{s > t_0 : m_s \notin \text{ORANGE}\}.$$

Then, for a suitable positive constant c ,

$$\Pr[\{T - t_0 \leq \log n\} \cap \{m_T \notin \text{RED} \cup \text{ORANGE}\} \mid m_{t_0} = m] \geq c.$$

The proof of Lemma 2.3 is given in Section 3.3. The first problem is that the ORANGE area contains the fixed point $m_t = \frac{n}{2}$. To prove that the process leaves this area within $O(\log n)$ rounds, we need to show that it tends to drift away from $\frac{n}{2}$. In Lemma 3.4 and Lemma 3.5, we indeed prove that $m_{t+1} \leq \frac{n}{2} - 1.5 \frac{\alpha}{\sqrt{n}}$ with probability at least $1 - e^{-0.1\alpha^2} - \frac{1}{n^2}$, whenever $m_t = \frac{n}{2} - \frac{\alpha}{\sqrt{n}}$. This is an exponentially increasing drift occurring with increasing probability. This fact allows us to show that the process leaves the ORANGE area with constant probability, within $O(\log n)$ rounds.

The second problem we face is that, upon leaving the ORANGE area, the process might in principle fall back to the RED one. In order to prove that this event does not occur with constant probability, we need to proceed with care, in order to somehow capture two different conditions under which m_t may have left the ORANGE area, each of which needs a distinct characterization.

LEMMA 2.4. (YELLOW AREA) For any $t_0 \geq 1$ and $m_{t_0} \in \text{YELLOW}$, let

$$T = \inf\{s > t_0 : m_s \notin \text{YELLOW}\}.$$

Then, for a suitable positive constant c ,

$$\Pr[\{T \leq \log n\} \cap \{m_{t_0+T} \in \text{BLUE}_1 \cup \text{BLUE}_2 \cup \text{GREEN}\} \mid m_{t_0} = m] \geq c.$$

The proof of Lemma 2.4, given in Section 3.3, proceeds along lines similar to those of Lemma 2.3, albeit with a few differences that do not allow a simple reuse of the arguments given for the ORANGE case. Again, we have a fixed point \bar{m} in the YELLOW area, but one that we do not compute exactly. This time, the roles of W_t and U_t are reversed, since we are in a regime in which $W_t = 0$ w.h.p. Moreover, while in the previous case the drift was towards decreasing values for m_t , the values of m_t are considerably smaller (i.e., $O(\sqrt{n})$) and the drift potentially harder to characterize in the YELLOW area. In particular, while drifting away from \bar{m} , $m_t - \bar{m}$ can change sign over consecutive rounds. For this reason, it proved necessary to quantify drift using $|m_t - \bar{m}|$ in the proof of Lemma 3.6. Finally, some extra care is needed to address the fact that this time, we are interested in the probability that, upon leaving the YELLOW area, the process avoids both the RED and ORANGE ones.

The two remaining cases address the BLUE and the GREEN area. Their analysis is easy and relies on simple concentration arguments.

LEMMA 2.5. (BLUE AREA) For any $t \geq 1$ and $m \in \text{BLUE}_1 \cup \text{BLUE}_2$, $\Pr[m_{t+1} \in \text{GREEN} \mid m_t = m] \geq c$, for a suitable positive constant c .

LEMMA 2.6. (GREEN AREA) For any $t \geq 1$, $m \in \text{GREEN}$ and $\ell \in \{0, 1\}$,

$$\Pr[\{m_{t+1} = 0\} \cap \{\mathcal{O}_{t+1} = 1 - \ell\} \mid m_t = m, \mathcal{O}_t = \ell] \geq c,$$

for a suitable positive constant c .³

³Note that if $m_t = 0$, \mathcal{O}_t is the opinion opposite to the majority.

Proof of Theorem 1.2. Recall that by definition of the protocol, for every agent i , if all samples collected by i in round t are equal to its opinion, then agent i maintains the same opinion in round $t + 1$. Therefore, if $m_t = 0$, we have $m_s = 0$ for every $s > t$, i.e., consensus has been reached permanently. Moreover, Lemmas 2.2 to 2.6 prove the existence of two constants $c_1, c_2 > 0$, such that for every $t \geq 0$ and $m \in [\frac{n}{2}]$,

$$(2.1) \quad \Pr [m_{t+c_1 \log n} = 0 \mid m_t = m] > c_2.$$

Let $T = \inf\{s \in \mathbb{N}, m_s = 0\}$ and $x \in \mathbb{N}$. Since (2.1) holds for every t and since $\{m_t\}_t$ is a Markov chain, we have $\Pr [T < x \cdot c_1 \log n] > 1 - (1 - c_2)^x$. We can thus take a constant c_3 such that, for $x \geq c_3 \log n$, we have $\Pr [T < c_1 c_3 \log^2 n] > 1 - \frac{1}{n^2}$, which concludes the proof of Theorem 1.2. \square

The following Lemma, whose proof is given in Section 3.7, yields Theorem 1.3.

LEMMA 2.7. (BIT-DISSEMINATION PROBLEM) *If ℓ is the opinion of the source agent, then*

$$\Pr [m_{t+2} = 0 \mid m_t = 1, \mathcal{O}_t = \ell] \geq c,$$

for a suitable positive constant c .

Proof of Theorem 1.3. Consider the k -MINORITY dynamics in the presence of a source agent. In this case, Theorem 1.2 implies the existence two constants $c_1, c_2 > 0$ such that, for every t and $m \in [\frac{n}{2}]$, $\Pr [m_{t+c_1 \log n} \leq 1 \mid m_t = m] > c_2$. Moreover, Lemma 2.7 implies a constant $c_3 > 0$ for which

$$\Pr [m_{t+2+c_1 \log n} = 0 \mid m_t = m] > c_3.$$

Setting $T = \inf\{s \in \mathbb{N}, m_s = 0\}$ and arguing as in the proof of Theorem 1.2, we have $m_T = 0$ w.h.p. for some $T = O(\log^2 n)$. \square

3 k -Minority in the Parallel Case: Technical Details

We gave an overview of the proofs of Theorem 1.2 and Theorem 1.3 in Section 2. This section contains the proofs of all technical lemmas that were used there, some of which highlight important properties of the k -MINORITY dynamics, some of which are not trivial to capture, due to the chaotic nature of the process.

3.1 Preliminary lemmas We begin by proving a set of simple results describing the behavior of the random variables W_t and U_t , which will be used throughout the analysis. Our proofs often use two standard probabilistic tools, namely, *Chernoff Bound* and the *Reverse Chernoff Bound*, in the versions presented in Theorem A.1 and Lemma A.1 of the Appendix, respectively. The following is a general statement that also includes the results singled out in Lemma 2.1 of Section 2.1. Its proof is deferred to Appendix B.

LEMMA 3.1. *Assume $k \geq 5 \log n$. Then, for any $m \in \{0, \dots, \frac{n}{2}\}$, we have the following results*

- (a) $\frac{n}{4} e^{-4k(\frac{1}{2} - \frac{m}{n})^2} \leq \mathbf{E}[W_t \mid m_t = m] \leq n e^{-2k(\frac{1}{2} - \frac{m}{n})^2}$;
- (b) $\mathbf{E}[U_t \mid m_t = m] = n \left(1 - \frac{m}{n}\right)^k$;
- (c) For any fixed $m_t = m$ with $m \geq \frac{3n \log n}{k}$, $U_t = 0$ w.h.p.
- (d) For any fixed $m_t = m$ with $m \leq \frac{n}{2} - n \sqrt{\frac{1.5 \log n}{k}}$, $W_t = 0$ w.h.p.
- (e) For any fixed $m_t = m$ with $m \leq \frac{n}{3k}$, $U_t \geq 0.6n$ and $\mathcal{O}_{t+1} = \mathcal{O}_t$ w.h.p.
- (f) For any fixed $m_t = m$ with $\frac{2n}{k} \leq m \leq \frac{n}{2} - n \sqrt{\frac{1.5 \log n}{k}}$, $U_t \leq 0.2n$ and $\mathcal{O}_{t+1} = 1 - \mathcal{O}_t$ w.h.p.

The following lemma is a simple concentration result applied to the variables U_t , $n - U_t$ and W_t , and gives necessary conditions so that the value of these random variables is close to their expectation.

LEMMA 3.2. Consider an interval $I = [a, b] \cap \mathbb{N}$ such that $b \geq 4a$ and $b \geq 32 \log n$ and fix a value $m_t = m$.⁴ Let X be any of the random variables in $\{U_t, n - U_t, W_t\}$. For any $m \in \{0, \dots, \frac{n}{2}\}$, if $\mathbf{E}[X | m_t = m] \in I$, then $X \in I$ with constant probability. Moreover, considered any constant $\gamma \geq \frac{3}{2}$, if $a \geq 36 \log n$ then $X \in \gamma I$ w.h.p., where $\gamma I = [\frac{a}{\gamma}, \gamma b] \cap \mathbb{N}$.

Proof. We first remark that X can always be written as the sum of n independent Bernoulli random variables. Let M_I be the median point of I , i.e. $M_I = a + \frac{b-a}{2}$. Note that, since $b \geq 4a$ and $b \geq 32 \log n$, we have $M_I \geq \frac{b}{2} \geq 16 \log n$. We consider two cases and, without loss of generality, we prove the lemma for $\gamma = \frac{3}{2}$.

First, assume that $\mathbf{E}[X | m_t = m] \leq M_I$. In this case, the use of Chernoff's Bound implies that $\Pr[X \leq 1.5M_I | m_t = m] \geq 1 - e^{-M_I/8} \geq 1 - \frac{1}{n^2}$, since $M_I \geq 16 \log n$ and $b \geq 4a$, $1.5M_I \leq b$. If $a \geq 36 \log n$, we further have $\Pr[X \geq \frac{2}{3}\mathbf{E}[X | m_t = m] | m_t = m] \geq 1 - e^{-a/18} \geq 1 - \frac{1}{n^2}$. If no condition on a is given, the reverse Chernoff bound guarantees the existence of a constant $c > 0$, such that $\Pr[X \geq a] \geq \Pr[X \geq \mathbf{E}[X | m_t = m] | m_t = m] \geq c$, where the first inequality follows from the assumption $\mathbf{E}[X | m_t = m] \in I$. Therefore, in the absence of constraints on a , we have $\Pr[X \in I | m_t = m] \geq \Pr[a \leq X \leq 1.5M_I | m_t = m] > c - \frac{1}{n^2}$. If, on the other hand, $a \geq 36 \log n$, then $\Pr[X \in \frac{3}{2}I | m_t = m] \geq \Pr[\frac{2}{3}\mathbf{E}[X | m_t = m] \leq X \leq 1.5M_I | m_t = m] \geq 1 - \frac{2}{n^2}$.

Next, assume that $\mathbf{E}[X | m_t = m] \geq M_I$. Similarly to the previous case, Chernoff bound gives $\Pr[X \geq 0.5M_I] \geq 1 - e^{-M_I/8} \geq 1 - \frac{1}{n^2}$ and, since $b \geq 4a$, we have $0.5M_I \geq a$. Moreover, since $b \geq 32 \log n$, we also have $\Pr[X \leq \frac{3}{2}b] \geq 1 - e^{-M_I/8} \geq 1 - \frac{1}{n^2}$. In addition, the reverse Chernoff bound gives the existence of a constant $c' > 0$, such that $\Pr[X \leq b] \leq \Pr[X \leq \mathbf{E}[X | m_t = m]] > c'$, where the first inequality follows from the fact that $\mathbf{E}[X | m_t = m] \in I$. In conclusion, we have $\Pr[X \in I] \geq \Pr[0.5M_I \leq X \leq \mathbf{E}[X | m_t = m]] \geq c' - \frac{1}{n^2}$ and $\Pr[X \in \frac{3}{2}I | m_t = m] \geq \Pr[0.5M_I \leq X \leq \frac{3}{2}b | m_t = m] \geq 1 - \frac{2}{n^2}$. \square

3.2 The red area The width of the Red area is small enough that given $m_t, m_{t+1} \notin \text{RED}$ with at least constant probability. The following result formalizes this idea and relies on a simple anti-concentration result.

LEMMA 3.3. Let X be any random variable in $\{U_t, n - U_t, W_t, n - W_t\}$. For any value $m \in \{0, \dots, \frac{n}{2}\}$, if $\mathbf{E}[X | m_t = m] \in \text{RED}$ then $X \in \text{BLUE}_1 \cup \text{BLUE}_2$ with constant probability.

Proof. We first remark that X can always be written as a sum of n i.i.d. Bernoulli random variables and, since $\mathbf{E}[X | m_t = m] \in \text{RED}$, we have $\mathbf{E}[X | m_t = m] \leq \frac{n}{2}$, thus meeting the assumptions of the reverse Chernoff bound applied to X under the conditioning $m_t = m$. Moreover, since $k \leq 0.5n$, we have $\frac{n \log 2}{k} \geq 1$. By definition, the RED area has width $2\sqrt{\frac{n \log 2}{k}}$ approximately, so that $\mathbf{E}[X | m_t = m]$ lies within distance at most $\sqrt{\mathbf{E}[X | m_t = m]}$ from the border of RED. For this reason, the reverse Chernoff bound implies that X can deviate at least $\sqrt{\mathbf{E}[X | m_t = m]}$ from its expectation, with constant probability. This concludes the proof. \square

The proof of Lemma 2.2 is a simple application of previous results.

Proof of Lemma 2.2. Assume that, for $m_t = m$, $\mathbf{E}[X | m_t = m] \in \text{RED}$ for a random variable $X \in \{U_t, n - U_t, W_t, n - W_t\}$. Then, Lemma 3.3 implies $X \in \text{BLUE}_1 \cup \text{BLUE}_2$ with constant probability. This, together with Lemma 3.1(c) and Lemma 3.1(d), implies $m_{t+1} \in \text{BLUE}_1 \cup \text{BLUE}_2$ with constant probability.

Next, suppose that $\mathbf{E}[X | m_t = m] \notin \text{RED}$ for every choice of $X \in \{U_t, n - U_t, W_t, n - W_t\}$. Consider $\bar{X} = \operatorname{argmin}_{X \in \{U_t, n - U_t, W_t, n - W_t\}} \mathbf{E}[X | m_t = m]$: we obviously have $\mathbf{E}[\bar{X} | m_t = m] \leq \frac{n}{2}$ and $\mathbf{E}[\bar{X} | m_t = m] \notin \text{RED}$. We can thus apply Lemma 3.2 to the intervals $I_1 = \text{BLUE}_1$ and $I_2 = \text{BLUE}_2 \cup \text{YELLOW} \cup \text{GREEN} \cup \text{ORANGE}$. This, together with Lemma 3.1(c) and Lemma 3.1(d), proves that $\Pr[m_{t+1} \notin \text{RED} | m_t = m] \geq \Pr[\bar{X} \in I_1 \cup I_2 | m_t = m] > c$, for a constant $c > 0$. \square

3.3 The orange area The goal of this section is to prove Lemma 2.3. In this area, we have the presence of the fixed point $\frac{n}{2}$ for the expectation of the process, so that $\mathbf{E}[m_{t+1} | m_t = \frac{n}{2}] = \frac{n}{2}$. In the remainder of this section, we prove that the aforementioned fixed point is unstable. The main ingredients of our proof are the following: i) first, even if $m_t = \frac{n}{2}$, m_{t+1} can deviate by $\Omega(\sqrt{n})$ from its conditional expectation ($\frac{n}{2}$) thanks to the variance

⁴To the purpose of our analysis, we are interested in the case $I \subseteq [0, n]$, but the result is more general.

of the process; ii) assuming instead that $m_t = \frac{n}{2} - \Omega(\sqrt{n})$, we are able to show a drift in the process, proving that with constant probability, it leaves the ORANGE area within an at most logarithmic number of steps. The following lemma quantifies the drift mentioned in point ii) above.

LEMMA 3.4. *Let $m = \frac{n}{2} - \alpha\sqrt{n}$, where $1 \leq \alpha \leq \frac{\sqrt{n}}{4}$. Then, $\mathbf{E}[W_t \mid m_t = m] \leq \frac{n}{2} - 1.8\alpha\sqrt{n}$.*

Proof. Fix an agent i and consider the random variable X indicating the number of agents sampled by i and holding the minority opinion. Our goal is to prove that $\mathbf{Pr}[X > \frac{k}{2} \mid m_t = \frac{n}{2} - \alpha\sqrt{n}] \leq \frac{1}{2} - \frac{\alpha}{1.1\sqrt{n}}$. For simplicity, we omit the conditioning in the rest of the proof. We have from Claim 2,

$$(3.2) \quad \mathbf{Pr}[X = \lceil \frac{k}{2} \rceil] \leq \mathbf{Pr}[X = \lfloor \frac{k}{2} \rfloor] \leq \binom{k}{\lfloor \frac{k}{2} \rfloor} \left(\frac{1}{2}\right)^k \leq \frac{4^{k/2}}{\sqrt{\pi(k/2)} \cdot 2^k} \leq \frac{1}{\sqrt{k}}.$$

Denote by Y the number of agents sampled by i that hold the majority opinion, so that $Y = k - X$. Since $X \sim \text{Bin}(k, 1 - \frac{m}{n})$ and $Y \sim \text{Bin}(k, \frac{m}{n})$, it is easy to see that

$$\mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil + 1] \leq \left(\frac{\frac{m}{n}}{1 - \frac{m}{n}}\right)^2 \mathbf{Pr}[Y \geq \lceil \frac{k}{2} \rceil + 1].$$

Letting $\gamma = \frac{2\alpha}{\sqrt{n}}$ and $\beta = \left(\frac{1-\gamma}{1+\gamma}\right)^2$, we can rewrite the inequality above as follows

$$\mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil + 1] \leq \beta \cdot \mathbf{Pr}[Y \geq \lceil \frac{k}{2} \rceil + 1].$$

Since $X = k - Y$, $\mathbf{Pr}[Y \geq x] = \mathbf{Pr}[X \leq k - x]$ for every $x \in [k]$, so that

$$\mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil + 1] \leq \beta \cdot (1 - \mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil] - \mathbf{Pr}[X = \lfloor \frac{k}{2} \rfloor]).$$

Moreover, (3.2) and the above inequality imply

$$\begin{aligned} \mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil] &\leq \beta \cdot (1 - \mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil]) - \beta \cdot \mathbf{Pr}[X = \lfloor \frac{k}{2} \rfloor] + \mathbf{Pr}[X = \lceil \frac{k}{2} \rceil] \\ &\leq \left(\frac{1-\gamma}{1+\gamma}\right)^2 \cdot (1 - \mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil]) + \mathbf{Pr}[X = \lfloor \frac{k}{2} \rfloor] (1 - \beta) \\ &\leq \beta \cdot (1 - \mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil]) + \frac{1}{\sqrt{k}} (1 - \beta) \\ &\leq \beta \cdot (1 - \mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil]) + \frac{6\gamma}{\sqrt{k}}, \end{aligned}$$

where the last inequality holds whenever $\gamma < 0.3$ (which, in our case, is true since $\gamma = \frac{2\alpha}{\sqrt{n}}$ and $\alpha < \frac{\sqrt{n}}{4}$). Finally, simple calculus yields

$$\mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil] \leq \frac{\left(\frac{1-\gamma}{1+\gamma}\right)^2 + \frac{6\gamma}{\sqrt{k}}}{1 + \left(\frac{1-\gamma}{1+\gamma}\right)^2} \leq \frac{1}{2} - \frac{\gamma}{1.1} + \frac{6\gamma}{\sqrt{k}} \leq \frac{1}{2} - 0.92\gamma \leq \frac{1}{2} - 1.8 \cdot \frac{\alpha}{\sqrt{n}}.$$

where the first inequality holds whenever $\gamma < 0.3$. Hence, noting that $\mathbf{E}[W_t \mid m_t = m] = n \cdot \mathbf{Pr}[X \geq \lceil \frac{k}{2} \rceil]$ the claim follows. \square

LEMMA 3.5. *For any $t_0 \geq 1$ and $m \in \text{ORANGE}$, let*

$$T = \inf\{s > t_0 : m_s \notin \text{ORANGE}\}.$$

Then, $\mathbf{Pr}[T - t_0 \leq \log n] \mid m_{t_0} = m \geq c$, for a suitable positive constant c .

Proof. For every t , define α_t so that $m_t = \frac{n}{2} - \alpha_t \sqrt{n}$, i.e., $\alpha_t = (\frac{n}{2} - m_t) \frac{1}{\sqrt{n}}$. Note that, if $m_t \in \text{ORANGE}$, then $\alpha_t \in \left[0, \sqrt{\frac{2n \log n}{k}}\right]$. We can prove the following:

(i) If $\alpha_{t_0} \leq 3$, then $\alpha_{t_0+1} \geq 3$ with constant probability.

(ii) For every $t > t_0$ and $3 \leq \alpha \leq \frac{\sqrt{n}}{4}$, we have $\Pr[\alpha_{t+1} > 1.5\alpha \mid \alpha_t = \alpha] \geq 1 - e^{-0.1\alpha^2} - \frac{1}{n^2}$.

We prove (i) and (ii) below, after showing that they imply [Lemma 3.5](#).

Assume (i) and (ii) hold. We define the events $\mathcal{E}_t = \{t \geq T\} \cup \{\alpha_{t+1} > 1.5\alpha_t\}$, $\mathcal{F} = \{\alpha_{t_0+1} \geq 3\}$ and $\mathcal{E} = \bigcap_{t=t_0+1}^{t_0+1+\log n} \mathcal{E}_t \cap \mathcal{F}$. We note that, conditioning on $m_{t_0} = m$, we have $\{T \leq \log n\} \subseteq \mathcal{E}$. Indeed, this follows since, if $\alpha_{t_0+1} \geq 3$ and $\alpha_{t+1} > 1.5\alpha_t$ for every $t = t_0 + 1, \dots, t_1$, then a straightforward calculation shows that $\alpha_{t_1} \geq \sqrt{\frac{2n \log n}{k}}$ for $t_1 < \log n$. Recalling that (i) implies $\Pr[\mathcal{F}] > c_1$ for a suitable constant c_1 we have

$$(3.3) \quad \Pr[T - t_0 \leq \log n \mid m_{t_0} = m] \geq \Pr[\mathcal{E} \mid m_{t_0} = m] \geq c_1 \prod_{t=t_0+1}^{t_0+1+\log n} \Pr[\mathcal{E}_t \mid \bigcap_{j=t_0+1}^{t-1} \mathcal{E}_j \cap \mathcal{F}, m_{t_0} = m].$$

Now note that, from our definition of the \mathcal{E}_t 's, we have $\Pr[\mathcal{E}_t \mid \bigcap_{j=t_0+1}^{t-1} \mathcal{E}_j \cap \mathcal{F}, m_{t_0} = m] = 1$ if $T = t'$ for some $t' \in \{t_0 + 1, \dots, t - 1\}$, while $\Pr[\mathcal{E}_t \mid \bigcap_{j=t_0+1}^{t-1} \mathcal{E}_j \cap \mathcal{F}, m_{t_0} = m] \geq \left(1 - e^{-0.1\alpha^2} - \frac{1}{n^2}\right)$ if $t' < T$ for $t' \in \{t_0 + 1, \dots, t - 1\}$ and $\alpha_t = \alpha$. To finish the proof, define $\beta_t = 3 \cdot (1.5)^{t-t_0-1}$. Then, for every t

$$\begin{aligned} \Pr[\mathcal{E}_t \mid \bigcap_{j=t_0+1}^{t-1} \mathcal{E}_j \cap \mathcal{F}, m_{t_0} = m] &\geq \Pr[\alpha_{t+1} > \beta_{t+1} \mid \bigcap_{j=t_0+1}^{t-1} \alpha_{j+1} > 1.5\alpha_j, t < T] \\ &\geq 1 - e^{-0.1\beta_t^2} - \frac{1}{n^2}, \end{aligned}$$

where the first inequality follows since $\bigcap_{j=t_0+1}^{t-1} \alpha_{j+1} > 1.5\alpha_j$ implies $\alpha_t > \beta_t$, while the second follows from this fact and from (ii). Finally, (3.3) and the inequality above give

$$\begin{aligned} \Pr[T - t_0 \leq \log n \mid m_{t_0} = m] &\geq \Pr[\mathcal{E} \mid m_{t_0} = m] \geq c_1 \prod_{t=t_0+1}^{t_0+1+\log n} \left(1 - e^{-0.1\beta_t^2} - \frac{1}{n^2}\right) \\ &\geq c_1 \left(1 - \sum_{t=t_0+1}^{t_0+1+\log n} e^{-0.1\beta_t^2} - \frac{1}{n}\right) \geq c_1 \left(1 - e^{-0.1 \cdot 9} - \sum_{t=1}^{+\infty} 3^{-t} - \frac{1}{n}\right) \geq c, \end{aligned}$$

for a suitable positive constant c , with the third inequality following from the definition of β_t .

Proof of (i) We consider two cases. Assume first that $\mathbf{E}[W_t \mid \alpha_{t_0} = \alpha] \leq \frac{n}{2} - 3\sqrt{n}$. In this case, [Lemma 3.2](#) implies that with constant probability, we $W_{t_0} \leq \frac{n}{2} - 3\sqrt{n}$, whence $m_{t_0+1} \leq \frac{n}{2} - 3\sqrt{n}$ and $\alpha_{t_0+1} \geq 3$. Assume next that $\frac{n}{2} - 3\sqrt{n} \leq \mathbf{E}[W_t \mid \alpha_{t_0} = \alpha] \leq \frac{n}{2}$. In this case, we apply the reverse Chernoff Bound to show that with constant probability, W_{t_0+1} deviates at least $3\sqrt{\mathbf{E}[W_t \mid \alpha_t = \alpha]}$ from its expectation, thus giving $\alpha_{t_0+1} \geq 3$.

Proof of (ii) By definition, we have that $m_t = \frac{n}{2} - \alpha_t \sqrt{n}$. Now, [Lemma 3.4](#) implies

$$\mathbf{E}[W_t \mid \alpha_t = \alpha] \leq \frac{n}{2} - 1.8\alpha\sqrt{n}.$$

Therefore, Chernoff's bound ([Theorem A.1](#)), [Lemma 3.1\(c\)](#) and a union bound give

$$\Pr\left[U_t = 0 \text{ and } W_t \leq \frac{n}{2} - 1.5\alpha_t\sqrt{n} \mid \alpha_t = \alpha\right] \geq 1 - e^{-0.1\alpha_t^2} - \frac{1}{n^2}.$$

Hence, with the above probability we have $m_{t+1} = W_t$, so that

$$\Pr[\alpha_{t+1} > 1.5\alpha \mid \alpha_t = \alpha] = \Pr\left[m_{t+1} \leq \frac{n}{2} - 1.5\alpha\sqrt{n} \mid \alpha_t = \alpha\right] \geq 1 - e^{-0.1\alpha^2} - \frac{1}{n^2}.$$

□

We now proceed with the proof of [Lemma 2.3](#).

Proof of Lemma 2.3. Consider the stopping time T' defined as follows:

$$T' = \inf\{s > t_0 : \mathbf{E}[W_s | m_s] \notin \text{ORANGE or } m_s \notin \text{ORANGE}\}.$$

Roughly speaking, T' is the first time the expectation of the process, or the process, leaves the ORANGE area. Of course, $T' \leq T$. Considering T' instead of T is useful to prove that, at time $T = T' + 1$, $m_T \notin \text{ORANGE} \cup \text{RED}$ with constant probability. Consider the events

$$A_t = \{m_t \in \text{ORANGE}\} \cap \{\mathbf{E}[W_t | m_t] \in \text{ORANGE}\} \quad \text{and} \quad B_t = \{\frac{n}{3} \leq W_t \leq \frac{3n}{4}\} \cap \{U_t = 0\},$$

and let

$$\mathcal{E}_t = (A_t)^C \cup B_t \cup \{t \geq T'\}, \quad \mathcal{E}_1 = \bigcap_{t=t_0}^{t_0+\log n} \mathcal{E}_t \quad \text{and} \quad \mathcal{E}_2 = \{T' - t_0 \leq \log n\}.$$

We have:

$$\begin{aligned} \Pr[\mathcal{E}_1^C | m_{t_0} = m] &\leq \sum_{t=t_0}^{t_0+\log n} \Pr[\mathcal{E}_t^C | m_{t_0} = m] \\ &\leq \sum_{t=t_0}^{t_0+\log n} \Pr[(B_t)^C | A_t, t < T', m_{t_0} = m] \\ &\leq \frac{2 \log n}{n^2}, \end{aligned}$$

where the last inequality follows from [Lemma 3.2](#) applied to the interval $I = \text{ORANGE}$, by observing that $\frac{3}{2}I \subseteq [\frac{n}{3}, \frac{3n}{4}]$ and from [Lemma 3.1\(c\)](#). Since $T' \leq T$, [Lemma 3.5](#) shows the existence of a constant $c_2 > 0$ such that $\Pr[\mathcal{E}_2 | m_{t_0} = m] > c_2$, whence $\Pr[\mathcal{E}_1 \cap \mathcal{E}_2 | m_{t_0} = m] > c_2 - \frac{1}{n} > 0$. We assume that the event $\mathcal{E}_1 \cap \mathcal{E}_2$ holds in the remainders of the proof, remarking that the event $\mathcal{E}_1 \cap \mathcal{E}_2$ is independent of what happens during round $T' + 1$.

We now distinguish two cases, according to the realization of T' . First, we assume that $m_{T'-1} = \gamma \in \text{ORANGE}$ and that $\mathbf{E}[W_{T'} | m_{T'-1} = \gamma] \notin \text{ORANGE}$. If $\mathbf{E}[W_{T'} | m_{T'-1} = \gamma] \notin \text{RED} \cup \text{ORANGE}$, [Lemma 3.2](#) applied to intervals $I_1 = \text{BLUE}_1$ and $I_2 = \text{BLUE}_2 \cup \text{YELLOW} \cup \text{GREEN}$ implies that $W_{T'} \notin \text{RED} \cup \text{ORANGE}$ and $W_{T'} \leq n/2$ with constant probability. If otherwise $\mathbf{E}[W_{T'} | m_{T'-1} = \gamma] \in \text{RED}$, [Lemma 3.3](#) implies $W_{T'} \in \text{BLUE}_1 \cup \text{BLUE}_2$ with constant probability. Since $m_{T'-1} = \gamma \in \text{ORANGE}$, the two previous statements, together with [Lemma 3.1\(c\)](#) implying $U_{T'} = 0$ w.h.p., proving that $m_{T'+1} \notin \text{RED} \cup \text{ORANGE}$ with constant probability, and hence with constant probability $m_T \notin \text{RED} \cup \text{ORANGE}$ and $T = T' + 1$.

Next, assume that, for $\gamma \in \text{ORANGE}$, $\mathbf{E}[W_{T'} | m_{T'-1} = \gamma] \in \text{ORANGE}$ and $m_{T'} \notin \text{ORANGE}$. In this case, the event $\mathcal{E}_1 \cap \mathcal{E}_2$ implies $\frac{n}{3} \leq W_{T'} \leq \frac{3n}{4}$ and $U_{T'} = 0$. Since $m_{T'+1} = \min\{W_{T'}, n - W_{T'}\}$, this implies that $m_{T'+1} \notin \text{RED}$, concluding that $T = T' + 1$ and so $m_T = m_{T'+1} \notin \text{RED} \cup \text{ORANGE}$ with constant probability. \square

3.4 The yellow area The goal of this section is to prove [Lemma 2.4](#). The main difficulty in the analysis of this area is the presence of a fixed point $\bar{m} \in \text{YELLOW}$ for the expectation of the process. Notice that, according to [Fact 2.1](#) and [Lemma 2.1](#), in the whole YELLOW area it holds that $m_{t+1} = U_t$, w.h.p., for the range of k that we consider. Hence, we here define \bar{m} as the point such that $\mathbf{E}[U_t | m_t = \bar{m}] = \bar{m}$. In the following lemma, we show that such a fixed point is unstable: we first estimate the initial deviation from \bar{m} using the variance of the process, then we show that the distance between m_t and the fixed point grows by a factor $\Omega(\log n)$ in each round, with a sufficiently large probability. We use this argument to derive an upper bound on the time it takes the process to leave the YELLOW area.

LEMMA 3.6. *For any $t_0 \geq 1$, let T be the random variable indicating the first time larger than t_0 such that the process is out of the YELLOW area,*

$$T = \inf\{s > t_0 : m_s \notin \text{YELLOW}\}.$$

For every $m \in \text{YELLOW}$, it holds that $\Pr[T \leq \log n | m_{t_0} = m] \geq c$, for a suitable positive constant c .

Proof. Let $f(x) = \mathbf{E}[U_t | m_t = x] = n \left(1 - \frac{x}{n}\right)^k + n \left(\frac{x}{n}\right)^k$. We have that f is strictly decreasing in YELLOW and, moreover, that $f(x) > x$ if $x = \frac{n}{k} \log\left(\frac{k}{4 \log n}\right)$ and that $f(x) < x$ if $x = \frac{3n \log n}{k}$. These three facts imply that f admits a unique fixed point $\bar{m} \in \text{YELLOW}$ satisfying $f(\bar{m}) = \bar{m}$.

Let $\Delta_t = |m_t - \bar{m}|$. We prove the following in the remainder:

- (i) Conditioning on $\Delta_{t_0} \leq \sqrt{\frac{n}{k}}$, we have that $\Delta_{t_0+1} \geq \sqrt{\frac{n}{k}}$ with constant probability.
- (ii) For every $t > t_0$, we have $\Pr\left[\Delta_{t+1} > \frac{\log n}{16} \cdot \Delta_t \mid \Delta_t \geq \sqrt{\frac{n}{k}}, t < T\right] \geq 1 - n^{-\frac{1}{33}}$.

We prove (i) and (ii) later, while we now show that they imply that $\{T \leq \log n\}$ with constant probability. Indeed, if we define $\mathcal{E}_t = \{t \geq T\} \cup \left\{\Delta_{t+1} > \frac{\log n}{16} \Delta_t\right\}$, $\mathcal{F} = \left\{\Delta_{t_0+1} \geq \sqrt{\frac{n}{k}}\right\}$ and $\mathcal{E} = \bigcap_{t=t_0+1}^{t_0+1+\log n} \mathcal{E}_t \cap \mathcal{F}$, we have that, conditioning on $m_{t_0} = m$, $\{T \leq \log n\} \subseteq \mathcal{E}$. Indeed, if there exists some t_1 such that $t_0 \leq t_1 \leq t_0 + \log n$ for which,

$$\Delta_{t_1} > \max\left\{\frac{3n \log n}{k} - \bar{m}, \bar{m} - \frac{n}{k} \log\left(\frac{k}{4 \log n}\right)\right\},$$

we have $m_{t_1} \notin \text{YELLOW}$. Moreover, we have

$$(3.4) \quad \Pr[\mathcal{E} \mid m_{t_0} = m] = \prod_{t=t_0+1}^{t_0+1+\log n} \Pr[\mathcal{E}_t \mid \bigcap_{j=t_0+1}^{t-1} \mathcal{E}_j \cap \mathcal{F}, m_{t_0} = m] \Pr[\mathcal{F} \mid m_{t_0} = m].$$

From (i), we have $\Pr[\mathcal{F} \mid m_{t_0} = m] > c_1$ for a suitable positive constant c_1 , while (ii) implies that, for any t such that $t_0 + 1 \leq t \leq \log n + t_0 + 1$,

$$\Pr[\mathcal{E}_t \mid \bigcap_{j=t_0+1}^{t-1} \mathcal{E}_j \cap \mathcal{F}, m_{t_0} = m] \geq \Pr\left[\Delta_{t+1} > \frac{\log n}{16} \Delta_t \mid \Delta_t \geq \sqrt{\frac{n}{k}}, t < T\right] \geq 1 - n^{-\frac{1}{33}}.$$

Hence, from (3.4) and the inequality below, $\Pr[\mathcal{E} \mid m_{t_0} = m] \geq \left(1 - n^{-\frac{1}{33}}\right)^{\log n} \cdot c_1 \geq c$, for a suitable positive constant c .

Proof of (i) First, consider the case in which $\mathbf{E}[U_{t_0} \mid m_{t_0} = m] \geq \bar{m}$ and recall that $m \in \text{YELLOW}$. From Lemma 3.1(b) we have $\mathbf{E}[U_{t_0} \mid m_t = m] \leq \frac{4n \log n}{k}$ and, from the reverse and standard Chernoff bounds we have that, with constant probability, $\bar{m} + \sqrt{\bar{m}} \leq U_{t_0} \leq \frac{n}{2}$. Since $\bar{m} \geq \frac{n}{k}$, this implies $m_{t_0+1} \geq \bar{m} + \sqrt{\frac{n}{k}}$. Similarly, we can prove that, when $\mathbf{E}[U_{t_0} \mid m_{t_0} = m] \leq \bar{m}$, we have that $m_{t_0+1} \leq \bar{m} - \sqrt{\frac{n}{k}}$ with constant probability. Hence, with constant probability we have $\Delta_{t_0+1} \geq \sqrt{\frac{n}{k}}$, which concludes the proof of (i).

Proof of (ii) Since $\Delta_t \geq \sqrt{\frac{n}{k}}$, we can write $\Delta_t = \alpha \sqrt{\frac{n}{k}}$, for some $\alpha \geq 1$. Note that since $\bar{m} \in \text{YELLOW}$ and $k \geq \log n \sqrt{n}$, it satisfies

$$(3.5) \quad \bar{m} \geq \frac{n}{k} \log\left(\frac{k}{3 \log n}\right) \geq \frac{n \log n}{3k}.$$

First, consider the case when $m_t = \ell \geq \bar{m}$, i.e., $m_t = \bar{m} + \Delta_t$. We have

$$(3.6) \quad \begin{aligned} \mathbf{E}[U_t \mid m_t = \ell] &\leq \bar{m} \cdot e^{-\frac{k}{n} \Delta_t} = \bar{m} e^{-\alpha \sqrt{\frac{k}{n}}} \leq \bar{m} \cdot \max\left\{0.3, 1 - \frac{\alpha}{2} \sqrt{\frac{k}{n}}\right\} \\ &\leq \max\left\{0.3\bar{m}, \bar{m} - \frac{\log n}{8} \alpha \sqrt{\frac{n}{k}}\right\} \end{aligned}$$

$$(3.7) \quad \leq \bar{m} - \frac{\log n}{8} \Delta_t,$$

where the first inequality follows from the fact that $e^{-\frac{1}{x}} \leq \max\{0.3, 1 - \frac{1}{2x}\}$ for each $x \geq 0$ and the inequalities (3.6) and (3.7) follow from (3.5). Hence, $\mathbf{E}[\bar{m} - U_t \mid m_t = m] \geq \frac{\log n}{8} \Delta_t$ and so, from using the multiplicative Chernoff bound, we have

$$(3.8) \quad \Pr\left[\bar{m} - U_t \leq \frac{\log n}{16} \Delta_t \mid m_t = \ell\right] \leq e^{-\frac{\log n}{64} \Delta_t} \leq n^{-\frac{1}{32}},$$

where the last inequality follows since $k \leq 0.5n$. We proved that $U_t \leq \bar{m} - \frac{\log n}{16} \Delta_t$ with probability at least $1 - n^{-\frac{1}{32}}$ and, hence, using [Lemma 3.1\(d\)](#) we have $m_{t+1} = U_t$ with probability at least $1 - n^{-\frac{1}{33}}$, whence $\Delta_{t+1} \geq \frac{\log n}{16} \Delta_t$ follows.

Now, consider the case in which $m_t = \ell \leq \bar{m}$, and hence $m_t = \bar{m} - \Delta_t$. In this case, for each $\ell \in \text{YELLOW}$ we have

$$(3.9) \quad \mathbf{E}[U_t \mid m_t = \ell] \geq ne^{-k\frac{\ell}{n} - k(\frac{\ell}{n})^2} \geq \bar{m}e^{\frac{k}{n}\Delta_t} e^{-k(\frac{\ell}{n})^2}$$

$$(3.10) \quad \geq \bar{m}e^{\frac{1}{2}\alpha\sqrt{\frac{k}{n}}}$$

$$(3.11) \quad \geq \bar{m} + \frac{\bar{m}\alpha}{2}\sqrt{\frac{k}{n}}$$

$$(3.12) \quad \geq \bar{m} + \frac{\log n}{6}\alpha\sqrt{\frac{n}{k}}$$

where (3.9) follows from [Claim 1](#), (3.10) follows from the fact that $\ell \in \text{YELLOW}$ and so $k(\frac{\ell}{n})^2 \leq \frac{1}{2}\alpha\sqrt{\frac{k}{n}}$, (3.11) follows from the fact that $e^{\frac{1}{x}} \geq 1 + \frac{1}{x}$ for any $x > 0$. Finally, (3.12) follows from (3.5). Next, from the multiplicative Chernoff bound we have

$$(3.13) \quad \Pr\left[U_t - \bar{m} \leq \frac{\log n}{8}\Delta_t \mid m_t = \ell\right] \leq e^{-\frac{\log n}{16}\Delta_t} \leq n^{-\frac{1}{8}},$$

where the last inequality follows since $k \leq 0.5n$. Moreover, from the definition of [YELLOW](#), [Lemma 3.1\(f\)](#) implies that $U_t \leq 0.2n$ w.h.p. From a union bound with (3.13), we have that $\bar{m} + \frac{\log n}{8}\Delta_t \leq U_t \leq 0.2n$ with probability at least $1 - n^{-\frac{1}{8}}$. Hence, from [Lemma 3.1\(d\)](#) we have $m_{t+1} = U_t$ w.h.p., so that $m_{t+1} \geq \bar{m} + \frac{\log n}{8}\Delta_t$ with probability at least $1 - n^{-\frac{1}{10}}$. We thus conclude that, also in this case, $\Delta_{t+1} > \frac{\log n}{16}\Delta_t$ with probability at least $1 - n^{-\frac{1}{32}}$. \square

We now proceed with the proof of [Lemma 2.4](#).

Proof of Lemma 2.4. In order to prove the lemma, we introduce the stopping time $T' \leq T$, which can be seen (in some situations) as the time just before T :

$$T' = \inf\{s > t_0 : \mathbf{E}[U_s \mid m_s] \notin \text{YELLOW} \text{ or } m_s \notin \text{YELLOW}\}.$$

In other words, T' is the first time the expectation of the process, or the process itself, leaves [YELLOW](#). Considering T' instead of T is helpful to prove that, with constant probability, the minority does not land in the [RED](#) area in round $T = T' + 1$. Consider the following events

$$A_t = \{m_t \in \text{YELLOW}\} \cup \{\mathbf{E}[U_t \mid m_t] \in \text{YELLOW}\},$$

$$B_t = \{U_t \in \text{BLUE}_2 \cup \text{YELLOW} \cup \text{GREEN}\} \cap \{W_t = 0\},$$

and

$$\mathcal{E}_t = (A_t)^C \cup B_t \cup \{t \geq T'\}, \quad \mathcal{E}_1 = \bigcap_{t=t_0}^{t_0+\log n} \mathcal{E}_t \quad \text{and} \quad \mathcal{E}_2 = \{T' \leq \log n\}.$$

We have

$$\begin{aligned} \Pr[\mathcal{E}_1^C \mid m_{t_0} = m] &\leq \sum_{t=t_0}^{t_0+\log n} \Pr[\mathcal{E}_t^C \mid m_{t_0} = m] \\ &\leq \sum_{t=t_0}^{t_0+\log n} \Pr[(B_t)^C \mid A_t, t < T', m_{t_0} = m] \\ &\leq \frac{2\log n}{n^2}, \end{aligned}$$

where the last inequality follows from [Lemma 3.2](#) taking the interval $I = \text{YELLOW}$ and noticing that $2I \subseteq \text{BLUE}_2 \cup \text{YELLOW} \cup \text{GREEN}$, from [Lemma 3.1\(d\)](#) and a union bound.

From [Lemma 3.6](#) and noticing that $T' \leq T$, we have that there exists a constant $c_2 > 0$ such that $\Pr[\mathcal{E}_2 \mid m_{t_0} = m] \geq c_2$. Hence, we have $\Pr[\mathcal{E}_1 \cap \mathcal{E}_2 \mid m_{t_0} = m] \geq c_2 - \frac{1}{n} > 0$, whenever n is sufficiently large. In the rest of the proof, we assume that the event $\mathcal{E}_1 \cap \mathcal{E}_2$ holds. We remark that the event $\mathcal{E}_1 \cap \mathcal{E}_2$ is independent of what happens in round $T' + 1$.

We now consider two cases, according to the realization of T' . First, we assume that $m_{T'-1} = \gamma \in \text{ORANGE}$ and that $\mathbf{E}[U_{T'} \mid m_{T'-1} = \gamma] \notin \text{YELLOW}$. In this case, we note that, since $\gamma \in \text{YELLOW}$, from [Lemma 3.1\(c\)](#) and the strong Markov property, we have

$$\frac{1}{n^2} \leq \mathbf{E}[U_{T'} \mid m_{T'-1} = \gamma] \leq \frac{3n \log n}{k}.$$

Since we assumed $\mathbf{E}[U_{T'} \mid m_{T'-1} = \gamma] \notin \text{YELLOW}$, it follows that $\mathbf{E}[U_{T'} \mid m_{T'-1}] \in \text{BLUE}_1 \cup \text{RED} \cup \text{BLUE}_2 \cup \text{GREEN}$. We first consider the case in which $\mathbf{E}[U_{T'} \mid m_{T'-1} = \gamma] \in \text{RED}$. From [Lemma 3.3](#), we have $U_{T'} \in \text{BLUE}_1 \cup \text{BLUE}_2$ with constant probability, while [Lemma 3.1\(d\)](#) implies that $m_{T'+1} \in \text{BLUE}_1 \cup \text{BLUE}_2$, thus proving that, in this case, $T = T + 1$ and that $m_{T'+1} = m_T \in \text{BLUE}_1 \cup \text{BLUE}_2$ with constant probability. Now assume that $\mathbf{E}[U_{T'} \mid m_{T'-1}] \in \text{BLUE}_1 \cup \text{BLUE}_2 \cup \text{GREEN}$. We can apply [Lemma 3.2](#) to the intervals $I_1 = \text{BLUE}_1$, $I_2 = \text{BLUE}_2$ and $I_3 = \text{GREEN}$, obtaining that, with constant probability, $U_{T'} \in \text{BLUE}_1 \cup \text{BLUE}_2 \cup \text{GREEN}$. Considering also [Lemma 3.1\(d\)](#), we have that $T = T' + 1$ and that $U_{T'} = m_{T'+1} = m_T \in \text{BLUE}_1 \cup \text{BLUE}_2 \cup \text{GREEN}$.

Suppose next that T' is such that $m_{T'-1} = \gamma \in \text{YELLOW}$, $\mathbf{E}[U_{T'} \mid m_{T'-1} = \gamma] \in \text{YELLOW}$ but $m_{T'} \notin \text{YELLOW}$. In this case, the event $\mathcal{E}_1 \cap \mathcal{E}_2$ implies that $T = T' + 1$ and $U_{T'} = m_{T'+1} = m_T \in \text{BLUE}_1 \cup \text{GREEN}$. \square

3.5 The blue area

In this section, we prove [Lemma 2.5](#).

Proof of Lemma 2.5. We first assume that $m \in \text{BLUE}_1$. From [Lemma 3.1\(b\)](#) and since $m \leq \frac{n \log 2}{k} - \sqrt{\frac{n \log 2}{k}}$, we obtain the following bound:

$$(3.14) \quad \mathbf{E}[U_t \mid m_t = m] \geq e^{-k \frac{m}{n} - k \left(\frac{m}{n}\right)^2} \geq \frac{n}{2} e^{\sqrt{\frac{k \log 2}{n}}} e^{-k \left(\frac{\log 2}{k} - \sqrt{\frac{\log 2}{kn}}\right)^2}$$

$$(3.15) \quad \begin{aligned} &\geq \frac{n}{2} e^{\frac{1}{2} \sqrt{\frac{k \log 2}{n}}} \\ &\geq \frac{n}{2} + \frac{n}{4} \sqrt{\frac{k \log 2}{n}}, \end{aligned}$$

where (3.14) follows from [Claim 1](#), and (3.15) follows from the fact that $e^{\frac{1}{x}} \geq 1 + \frac{1}{x}$ for any $x > 0$. On the other side, from [Lemma 3.1\(b\)](#), we have that, since $m_t \geq 1$, $\mathbf{E}[U_t \mid m_t = m] \leq n e^{-\frac{k}{n}} \leq n - \frac{k}{3} \leq n - \frac{n \log n}{k}$, since $k \geq 185\sqrt{n \log n}$. Hence, we proved that $\mathbf{E}[n - U_t \mid m_t = m] \in \text{GREEN}$ and, from [Lemma 3.2](#) applied to $I = \text{GREEN}$ and [Lemma 3.1\(d\)](#), we have that $m_{t+1} = n - U_t \in \text{GREEN}$ with constant probability.

We now assume that $m \in \text{BLUE}_2$, and hence $\frac{n \log 2}{k} + \sqrt{\frac{n \log 2}{k}} \leq m \leq \frac{n}{k} \log \left(\frac{k}{4 \log n}\right)$. In this case, from [Lemma 3.1\(b\)](#), we have that

$$\begin{aligned} \mathbf{E}[U_t \mid m_t = m] &\leq n e^{-\frac{k}{n} m} \leq \frac{n}{2} e^{-\sqrt{\frac{k \log 2}{n}}} + \frac{1}{n^2} \\ &\leq \frac{n}{2} - \frac{n}{5} \sqrt{\frac{k \log 2}{n}}, \end{aligned}$$

where the last inequality follows since $e^{-1/x} \leq 1 - \frac{1}{2x}$ for any $x \geq 1$. On the other side,

$$\begin{aligned} \mathbf{E}[U_t \mid m_t = m] &\geq n e^{-k \frac{m}{n} - k \left(\frac{m}{n}\right)^2} \geq \frac{4n \log n}{k} e^{-\log^2 \left(\frac{k}{4 \log n}\right)^{\frac{1}{k}}} \\ &\geq \frac{3n \log n}{k}, \end{aligned}$$

Hence, $\mathbf{E}[U_t \mid m_t = m] \in \text{GREEN}$ and, from [Lemma 3.2](#) applied to the interval $I = \text{GREEN}$ and [Lemma 3.1\(d\)](#), we have that $m_{t+1} = U_t \in \text{GREEN}$ with constant probability. \square

3.6 The green area

Proof of Lemma 2.6. By definition of GREEN, we can apply Lemma 3.1(c) and Lemma 3.1(d), obtaining that $m_{t+1} = 0$ w.h.p. We notice that, in this case, $\mathcal{O}_{t+1} = 1 - \mathcal{O}_t$, since every node adopts the ex-minority opinion. \square

3.7 The Bit-Dissemination problem

Proof of Lemma 2.7. By assumption, $m_t = m = 1$ and the opinion of the source agent is $\ell = \mathcal{O}_t$. Since, from the fact that $k \leq 0.5n$, we have that $m = 1 \leq \frac{n}{3k}$, we have from Lemma 3.1(e) $\mathcal{O}_{t+1} = \mathcal{O}_t = \ell$ w.h.p. and, from Lemma 2.5, that $m_{t+1} \in \text{GREEN}$ with constant probability. Finally, from Lemma 2.6, we have that $m_{t+2} = 0$ and that $\mathcal{O}_t = 1 - \ell$, and therefore the opinion of the majority is the same opinion of the source. \square

4 k -Minority in the Sequential Communication Model

In this section, we analyze the k -MINORITY process on the uniform-random sequential model, proving Theorem 1.1.

4.1 Proof of Theorem 1.1 In this section, we will denote by X_t the number of agents with opinion 1 in round t . In addition, we will use the following notations for birth-death chains on $\{0, \dots, n\}$:

$$\begin{aligned} p_i &= \Pr[X_{t+1} = i + 1 \mid X_t = i], \\ q_i &= \Pr[X_{t+1} = i - 1 \mid X_t = i], \\ r_i &= \Pr[X_{t+1} = i \mid X_t = i], \text{ and} \\ \tau_{i,j} &= \mathbf{E}[\inf\{t \in \mathbb{N}, X_t = j\} \mid X_0 = i]. \end{aligned}$$

We start by recalling a classical lower bound on the expected time needed for a birth-death chain to travel from state 0 to n .

LEMMA 4.1. *Consider any birth-death chain on $\{0, \dots, n\}$. For $1 \leq i \leq j \leq n$, let $a_i = q_i/p_{i-1}$ and $a(i : j) = \prod_{k=i}^j a_k$. Then, $\tau_{0,n} \geq \sum_{1 \leq i < j \leq n} a(i : j)$.*

Proof (for the sake of completeness). Let $w_0 = 1$ and for $i \in \{1, \dots, n\}$, let $w_i = 1/a(1 : i)$. The following result is well-known (see, e.g., Eq. (2.13) in [24]). For every $\ell \in \{1, \dots, n\}$,

$$\tau_{\ell-1,\ell} = \frac{1}{q_\ell w_\ell} \sum_{i=0}^{\ell-1} w_i.$$

Thus,

$$\tau_{\ell-1,\ell} = \frac{1}{q_\ell} \sum_{i=0}^{\ell-1} \frac{a(1 : \ell)}{a(1 : i)} = \frac{1}{q_\ell} \sum_{i=1}^{\ell} a(i : \ell) \geq \sum_{i=1}^{\ell} a(i : \ell).$$

Eventually, we can write

$$\tau_{0,n} = \sum_{\ell=1}^n \tau_{\ell-1,\ell} \geq \sum_{1 \leq i < j \leq n} a(i : j),$$

which concludes the proof of Lemma 4.1. \square

Next, we identify some regions of the configuration space where k -MINORITY has a very strong drift toward the balanced configuration $X_t = n/2$.

LEMMA 4.2. *Let α, β such that $0 < \alpha < \beta < 1/2$. There exists a constant $c = c(\alpha, \beta) > 0$, such that for every n large enough, for every $i \in [\alpha n, \beta n]$,*

$$\Pr[X_{t+1} = i + 1 \mid X_t = i] \geq \frac{1 - \beta}{2} \quad \text{and} \quad \Pr[X_{t+1} = n - i - 1 \mid X_t = n - i] \geq \frac{1 - \beta}{2};$$

and

$$\Pr[X_{t+1} = i - 1 \mid X_t = i] \leq \exp(-ck) \quad \text{and} \quad \Pr[X_{t+1} = n - i + 1 \mid X_t = n - i] \leq \exp(-ck).$$

Proof. Let $i \in [\alpha n, \beta n]$. Let A_t the event: “In round t , the activated agent sees a unanimity of 0”. Since $i \geq \alpha n$,

$$\Pr[A_t \mid X_t = i] \leq (1 - \alpha)^k.$$

Let B_t the event: “In round t , the activated agent sees a minority of 0”. Let Y be a random variable following a binomial distribution with parameters (k, β) . Since $i \leq \beta n$, by the additive Chernoff bound ([Theorem A.1](#)),

$$\begin{aligned} \Pr[B_t \mid X_t = i] &\leq \Pr[Y \geq k/2] \\ &= \Pr[Y \geq \beta k + k(1/2 - \beta)] \\ &\leq \exp\left(\frac{-2k^2(1/2 - \beta)^2}{k}\right) \\ &= \exp\left(-2k(1/2 - \beta)^2\right). \end{aligned}$$

Let C_t the event: “In round t , the activated agent adopts opinion 0”. By construction of the Minority protocol, $C_t \subseteq A_t \cup B_t$. Taking the union bound, we get

$$\Pr[C_t \mid X_t = i] \leq (1 - \alpha)^k + \exp\left(-2k(1/2 - \beta)^2\right).$$

By setting

$$c = \min\left\{\log\left(\frac{1}{1 - \alpha}\right), 2\left(\frac{1}{2} - \beta\right)^2\right\} > 0,$$

we obtain

$$\Pr[C_t \mid X_t = i] \leq 2 \exp(-ck).$$

Therefore, we have

$$\Pr[X_{t+1} = i - 1 \mid X_t = i] = \frac{i}{n} \cdot \Pr[C_t \mid X_t = i] \leq \beta \cdot 2 \exp(-ck) \leq \exp(-ck).$$

Moreover, for k large enough,

$$\Pr[X_{t+1} = i + 1 \mid X_t = i] = \frac{n - i}{n} \cdot (1 - \Pr[C_t \mid X_t = i]) \geq (1 - \beta)(1 - 2 \exp(-ck)) \geq \frac{1 - \beta}{2}.$$

By symmetry of k -MINORITY, we obtain the same bounds for the case $X_t = n - i$, which concludes the proof of [Lemma 4.2](#). \square

Eventually, we can combine the above results into a lower bound on the convergence time of k -MINORITY in the sequential setting.

Proof of Theorem 1.1. Let $m = n/6$. Let $Z = (Z_s)$ be a birth-death chain on $\{0, \dots, m\}$, defined as follows: for every $\delta \in \{-1, 0, 1\}$, for every $i \in \{1, \dots, m - 1\}$,

$$\Pr[Z_{s+1} = i + \delta \mid Z_s = i] = \Pr\left[X_{t+1} = \frac{n}{2} + m + i + \delta \mid X_t = \frac{n}{2} + m + i\right].$$

Note that, by symmetry, this is also equal to $\Pr[X_{t+1} = \frac{n}{2} - m - i - \delta \mid X_t = \frac{n}{2} - m - i]$. Intuitively, making one step *forward* as Z_s , corresponds to making one step *away from* $n/2$ as X_t . In particular, with this construction, we have that

$$\mathbf{E}[\inf\{t \in \mathbb{N}, |X_t - n/2| \geq 2m\} \mid X_0 = n/2] \geq \mathbf{E}[\inf\{s \in \mathbb{N}, Z_s = m\} \mid Z_0 = 0].$$

By [Lemma 4.2](#) applied with $\alpha = 1/6$, $\beta = 2/6$, there exists a constant $c > 0$ such that for every $i \in \{1, \dots, m\}$, Z satisfies $p_i \leq \exp(-ck)$ and $q_i \geq 1/3$. As a consequence, for every $i \in \{1, \dots, m\}$, $q_i/p_{i-1} \geq \exp(ck)/3$. Finally, by [Lemma 4.1](#),

$$\mathbf{E}[\inf\{s \in \mathbb{N}, Z_s = m\} \mid Z_0 = 0] \geq \sum_{1 \leq i < j \leq n} a(i : j) \geq \left(\frac{\exp(ck)}{3}\right)^{n-1} = \exp(\Omega(nk)),$$

which concludes the proof of [Theorem 1.1](#). \square

5 Open Questions

We see two main questions that deserve future work. The first question is more technical and concerns the role of parameter k , i.e. the size of the sample used by nodes to apply the k -MINORITY rule. Despite our proofs of [Theorem 1.2](#) and [Theorem 1.3](#) require $k = \tilde{\Omega}(\sqrt{n})$ to achieve $O(\text{polylog } n)$ convergence time, preliminary experimental results suggest that $\text{polylog } n$ convergence time is possible even when $k = \Omega(\text{polylog } n)$, and it would be interesting to establish such bounds rigorously or prove that actual bounds might actually be worse. So far, we are able to prove that if $k = O(1)$ then any passive-communication stateless dynamics in the k -PULL model requires $n^{\Omega(1)}$ rounds to solve the bit-dissemination problem.

The second, more general question concerns the gap between the synchronous parallel and the asynchronous sequential models in *consensus* dynamics, or possibly in dynamics for other fundamental distributed tasks such as *broadcast*. A $\tilde{O}(n)$ gap between the two models is so common that, in order to analyze a parallel dynamics with (presumed) convergence time $\text{polylog } n$, it is common to first analyze its sequential implementation, to establish an $O(n \text{ polylog } n)$ bound on converge time, and then to adapt the analysis to the parallel setting. Our results show a natural, simple dynamics for which the usual linear gap does not hold, showing that it is not possible to establish a black-box result, whereby results for the asynchronous sequential setting immediately carry over to the parallel synchronous case. We believe it would be very interesting to derive necessary and/or sufficient conditions under which a linear gap is guaranteed to exist.

Acknowledgments

The authors wish to thank Amos Korman and Emanuele Natale for very helpful discussions on the topic.

References

- [1] David Aldous. Interacting particle systems as stochastic social dynamics. *Bernoulli*, 19(4):1122–1149, 2013.
- [2] D. Angluin, J. Aspnes, and D. Eisenstat. A simple population protocol for fast robust approximate majority. *Distributed Computing*, 21(2):87–102, 2008. (Preliminary version in DISC’07).
- [3] J. Aspnes. Faster randomized consensus with an oblivious adversary. In *Proceedings of the 31st Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC’12)*, pages 1–8. ACM, 2012.
- [4] James Aspnes and Eric Ruppert. An introduction to population protocols. In Benoît Garbinato, Hugo Miranda, and Luís Rodrigues, editors, *Middleware for Network Eccentric and Mobile Applications*, pages 97–120. Springer, Berlin, Heidelberg, 2009.
- [5] Paul Bastide, George Giakkoupis, and Hayk Saribekyan. Self-stabilizing clock synchronization with 1-bit messages. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA ’21)*, pages 2154–2173. Society for Industrial and Applied Mathematics, 2021.
- [6] L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan. Stabilizing consensus with many opinions. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA ’16)*. SIAM, 2016.
- [7] Luca Becchetti, Andrea Clementi, and Emanuele Natale. Consensus dynamics: An overview. *ACM SIGACT News*, 51(1):58–104, 2020.
- [8] Luca Becchetti, Andrea E. F. Clementi, Amos Korman, Francesco Pasquale, Luca Trevisan, and Robin Vacus. On the role of memory in robust opinion dynamics. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI ’23)*, pages 29–37, 2023.
- [9] Michael Ben-Or, Danny Dolev, and Ezra N. Hoch. Fast self-stabilizing byzantine tolerant digital clock synchronization. In *Proceedings of the twenty-seventh ACM symposium on Principles of distributed computing (PODC ’08)*, pages 385–394, 2008. Association for Computing Machinery.
- [10] Petra Berenbrink, Andrea E. F. Clementi, Robert Elsässer, Peter Kling, Frederik Mallmann-Trenn, and Emanuele Natale. Ignore or comply?: On breaking symmetry in consensus. In *Proceedings of the 36th ACM Symposium on Principles of Distributed Computing (PODC’17)*, pages 335–344, 2017.
- [11] Lucas Boczkowski, Amos Korman, and Emanuele Natale. Minimizing message size in stochastic communication patterns: fast self-stabilizing protocols with 3 bits. In *Proceedings of the 2017 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA ’17)*, Proceedings, pages 2540–2559. Society for Industrial and Applied Mathematics, January 2017.
- [12] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE transactions on information theory*, 52(6):2508–2530, 2006.
- [13] Andrea Clementi, Luciano Gualà, Emanuele Natale, Francesco Pasquale, Giacomo Scornavacca, Luca Trevisan, et al. Consensus vs broadcast, with and without noise. In *11th Innovations in Theoretical Computer Science Conference (ITCS ’20)*, 2020.

- [14] Iain D. Couzin, Christos C. Ioannou, Güven Demirel, Thilo Gross, Colin J. Torney, Andrew Hartnett, Larissa Conradt, Simon A. Levin, and Naomi E. Leonard. Uninformed individuals promote democratic consensus in animal groups. *Science*, 334(6062):1578–1580, December 2011. Publisher: American Association for the Advancement of Science.
- [15] M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.
- [16] B. Doerr, L. A. Goldberg, L. Minder, T. Sauerwald, and C. Scheideler. Stabilizing consensus with the power of two choices. In *Proceedings of the 23rd Annual ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '11)*, pages 149–158. ACM, 2011.
- [17] Shlomi Dolev, Ronen I. Kat, and Elad Michael Schiller. When consensus meets self-stabilization. *J. Comput. Syst. Sci.*, 76(8):884–900, 2010.
- [18] O. Feinerman, B. Haeupler, and A. Korman. Breathe before speaking: Efficient information dissemination despite noisy, limited and anonymous communication. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC '14)*, pages 114–123. ACM, 2014.
- [19] Mohsen Ghaffari and Johannes Lengler. Nearly-tight analysis for 2-choice and 3-majority consensus dynamics. In *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing (PODC'18)*, pages 305–313. ACM, 2018.
- [20] Mohsen Ghaffari and Merav Parter. A polylogarithmic gossip algorithm for plurality consensus. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, pages 117–126, 2016.
- [21] Yehuda Hassin and David Peleg. Distributed Probabilistic Polling and Applications to Proportionate Agreement. *Information and Computation*, 171(2):248–268, December 2001.
- [22] R. A. Johnstone, S. R. X. Dall, Nigel R. Franks, Stephen C. Pratt, Eamonn B. Mallon, Nicholas F. Britton, and David J. T. Sumpter. Information flow, opinion polling and collective intelligence in house-hunting social insects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1427):1567–1583, November 2002. Publisher: Royal Society.
- [23] Amos Korman and Robin Vacus. Early adapting to trends: Self-stabilizing information spread using passive communication. In *ACM Symposium on Principles of Distributed Computing*, pages 235–245. ACM, 2022.
- [24] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*, volume 107. American Mathematical Soc., 2017.
- [25] T. Liggett. *Interacting particle systems*, volume 276. Springer, 2012.
- [26] M. Lindauer. Communication among the honeybees and stingless bees of india. *Bee World*, 38(1):3–14, January 1957.
- [27] E. Mossel, J. Neeman, and O. Tamuz. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multi-Agent Systems*, 28(3):408–429, 2014.
- [28] David Peleg. *Distributed computing: a locality-sensitive approach*. SIAM, 2000.
- [29] M. O. Rabin. Randomized byzantine generals. In *Proceedings of the 24th Annual Symposium on Foundations of Computer Science (SFCS '83)*, pages 403–409. IEEE, 1983.
- [30] Nitzan Razin, Jean-Pierre Eckmann, and Ofer Feinerman. Desert ants achieve reliable recruitment across noisy interactions. *Journal of The Royal Society Interface*, 10(82):20130079, May 2013. Publisher: Royal Society.
- [31] Devavrat Shah et al. Gossip algorithms. *Foundations and Trends® in Networking*, 3(1):1–125, 2009.
- [32] Tamás Vicsek and Anna Zafeiris. Collective motion. *Physics Reports*, 517(3):71–140, August 2012.

A Tools

THEOREM A.1. (CHERNOFF'S INEQUALITY) Let $X = \sum_{i=1}^n X_i$, where X_i with $i \in [n]$ are independently distributed in $[0, 1]$. Let $\mu = \mathbf{E}[X]$ and $\mu_- \leq \mu \leq \mu_+$. Then:

- for every $t > 0$

$$\Pr[X > \mu_+ + t] \leq e^{-2t^2/n} \quad \text{and} \quad \Pr[X < \mu_- - t] \leq e^{-2t^2/n};$$

- for $\varepsilon > 0$

$$\Pr[X > (1 + \varepsilon)\mu_+] \leq e^{-\frac{\varepsilon^2}{2}\mu_+} \quad \text{and} \quad \Pr[X < (1 - \varepsilon)\mu_-] \leq e^{-\frac{\varepsilon^2}{2}\mu_-}.$$

LEMMA A.1. (REVERSE CHERNOFF BOUND) Let X_1, \dots, X_n be i.i.d. Bernoulli random variables. Let $X = \sum_{i=1}^n X_i$ and let X_i such that $\mathbf{E}[X_i] = p$, and $\mu = \mathbf{E}[X] = np$. Then, if $p \leq 1/4$, for any $t \geq 0$

$$\Pr[X \geq \mu + t] \geq \frac{1}{4} e^{-\frac{2t^2}{\mu}}.$$

If $p \leq 1/2$, then for any $0 \leq t \leq n(1 - 2p)$,

$$\Pr[X > t + \mu] \geq \frac{1}{4} e^{-\frac{2t^2}{\mu}}.$$

Furthermore, for any $\delta \in (0, \frac{1}{2}]$ such that $\delta^2\mu \geq 3$,

$$\Pr[X \geq (1 + \delta)\mu] \geq e^{-9\delta^2\mu} \quad \text{and} \quad \Pr[X \leq (1 - \delta)\mu] \geq e^{-9\delta^2\mu}.$$

CLAIM 1. If $x/n \leq 0.6$, then for n large enough, the following inequalities hold:

$$\exp\left(-2k \frac{x}{n}\right) \leq \exp\left(-k \frac{x}{n} - k \left(\frac{x}{n}\right)^2\right) \leq \left(1 - \frac{x}{n}\right)^k \leq \exp\left(-k \frac{x}{n}\right).$$

CLAIM 2. (CENTRAL BINOMIAL COEFFICIENT) For every large enough n , it holds

$$\frac{1}{2} \cdot \frac{4^n}{\sqrt{\pi n}} \leq \binom{2n}{n} \leq \frac{4^n}{\sqrt{\pi n}}.$$

B Missing Proofs

Proof of Lemma 3.1. We prove each item separately.

Proof of (a). Let $X \in \{0, \dots, k\}$ the number of samples corresponding to the minority opinion that node i receives in round t . Since $m_t = m$, each sample corresponds to the minority opinion with probability m/n . Therefore, we have $\mathbf{E}[X \mid m_t = m] = k \cdot \frac{m}{n} = \frac{k}{2} - k \left(\frac{1}{2} - \frac{m}{n}\right)$. Hence, by Chernoff Bound and the Reverse Chernoff Bound, we have

$$\frac{1}{4} e^{-4k \left(\frac{1}{2} - \frac{m}{n}\right)^2} \leq \Pr\left[X \geq \left\lceil \frac{k}{2} \right\rceil \mid m_t = m\right] \leq e^{-2k \left(\frac{1}{2} - \frac{m}{n}\right)^2}.$$

Now, given $m_t = m$, since W_t is the sum of n i.i.d. Bernoulli random variables taking value 1 when $\{X \geq \lceil k/2 \rceil\}$, we have that (a) holds.

Proof of (b). The generic agent i samples only the majority opinion with probability $(1 - m/n)^k$. Summing over all agents and taking the expectation concludes the proof of (b).

Proof of (c). If $m_t = m \geq \frac{3n \log n}{k}$, (b) implies that $\mathbf{E}[U_t \mid m_t = m] \leq \frac{2}{n^2}$. By Markov's inequality, $\Pr[U_t = 0 \mid m_t = m] \geq 1 - \frac{2}{n^2}$, which concludes the proof of (c).

Proof of (d). Since $m_t = m \leq \frac{n}{2} - n \sqrt{\frac{1.5 \log n}{k}}$, we have $\left(\frac{1}{2} - \frac{m}{n}\right)^2 \geq \frac{1.5 \log n}{k}$, whence (a) implies $\mathbf{E}[W_t \mid m_t = m] \leq \frac{1}{n^2}$. By Markov's inequality, $\Pr[W_t = 0 \mid m_t = m] \geq 1 - \frac{1}{n^2}$, which concludes the proof of (d).

Proof of (e). If $m_t = m \leq \frac{n}{3k}$, (b) yields $\mathbf{E}[U_t \mid m_t = m] \geq n \left(1 - \frac{m}{n}\right)^k \geq ne^{-\frac{1+k}{n}m} \geq 0.69n$. Then, using Chernoff's bound we have $\mathbf{Pr}[U_t \leq 0.6n \mid m_t = m] \leq e^{-2 \cdot (0.09)^2 n} \leq \frac{1}{n^2}$. Moreover, since $m \leq \frac{n}{3k}$, (d) gives $W_t = 0$ w.h.p., with no agent seeing a unanimity that corresponds to \mathcal{O}_t . Thus, by definition of the protocol, any agent seeing unanimity adopts opinion $1 - \mathcal{O}_t$. Since $U_t \geq 0.6n$, this implies that a minority of agents adopt opinion \mathcal{O}_t , hence $\mathcal{O}_{t+1} = \mathcal{O}_t$, which concludes the proof of (e).

Proof of (f). If $m_t = m$ and $\frac{2n}{k} \leq m \leq \frac{n}{2} - n\sqrt{\frac{1.5 \log n}{k}}$, (b) gives $\mathbf{E}[U_t \mid m_t = m] \leq ne^{-\frac{k}{n}m} \leq 0.15n$. Hence, using Chernoff's bound we have $\mathbf{Pr}[U_t \geq 0.2n \mid m_t = m] \leq e^{-2(0.05)^2 n} \leq \frac{1}{n^2}$. Moreover, from (d) we have $W_t = 0$ w.h.p. The latter event implies that agents adopt the majority opinion if and only if they sample unanimity. Hence, conditioning on the fact that $U_t \leq 0.2n$, we have that $\mathcal{O}_{t+1} = 1 - \mathcal{O}_t$, which concludes the proof of (f). \square