

Stabilizing Consensus with Many Opinions

L. Becchetti¹, A. Clementi², E. Natale¹, F. Pasquale², and L. Trevisan³

¹*Sapienza* Università di Roma, becchett@dis.uniroma1.it, natale@di.uniroma1.it

²Università *Tor Vergata* di Roma, clementi@mat.uniroma2.it,
pasquale@mat.uniroma2.it

³U.C. Berkeley, luca@berkeley.edu

October 2, 2018

Abstract

We consider the following distributed consensus problem: Each node in a complete communication network of size n initially holds an *opinion*, which is chosen arbitrarily from a finite set Σ . The system must converge toward a consensus state in which all, or almost all nodes, hold the same opinion. Moreover, this opinion should be *valid*, i.e., it should be one among those initially present in the system. This condition should be met even in the presence of an adaptive, malicious adversary who can modify the opinions of a bounded number of nodes in every round.

We consider the *3-majority dynamics*: At every round, every node pulls the opinion from three random neighbors and sets his new opinion to the majority one (ties are broken arbitrarily). Let k be the number of valid opinions. We show that, if $k \leq n^\alpha$, where α is a suitable positive constant, the 3-majority dynamics converges in time polynomial in k and $\log n$ with high probability even in the presence of an adversary who can affect up to $o(\sqrt{n})$ nodes at each round.

Previously, the convergence of the 3-majority protocol was known for $|\Sigma| = 2$ only, with an argument that is robust to adversarial errors. On the other hand, no anonymous, uniform-gossip protocol that is robust to adversarial errors was known for $|\Sigma| > 2$.

Keywords: Distributed Consensus, Byzantine Agreement, Gossip Model, Majority Rules, Markov Chains.

1 Introduction

We study the following probabilistic, *synchronous* process on a complete network of n anonymous nodes: At the beginning, each node holds an “opinion” which is an element of an arbitrary finite set Σ . We call an opinion *valid* if it is held by at least one node at the beginning. Then, in each round, the following happens: 1) every node pulls the opinion from three random nodes and sets its new opinion to the majority one (ties are broken arbitrarily), and 2) an adaptive *dynamic adversary* can arbitrarily change the opinions of some nodes. We consider *F-dynamic adversaries* that, at every round, can change the opinions of up to F nodes, possibly introducing non-valid opinions.

Let the system start from any configuration having k valid opinions with $k \leq n^\alpha$ for some constant $\alpha < 1$ and consider any F -dynamic adversary with $F = \mathcal{O}(\sqrt{n}/(k^{5/2} \log n))$. We prove that the process converges to a configuration in which all but $O(\sqrt{n})$ nodes hold the same valid opinion within $O((k^2 \sqrt{\log n} + k \log n)(k + \log n))$ rounds, with high probability. So, this bounded adversary has no relevant chances to force the system to converge to non-valid opinions.

This shows that the *3-majority dynamics* provides an efficient solution to the *stabilizing-consensus* problem in the *uniform-gossip* model. Previously, this was known only for the binary case, i.e. $|\Sigma| = 2$, while for any $|\Sigma| \geq 3$, it has been an important open question for several years [3, 12]. Furthermore, still for any $|\Sigma| \geq 3$, $o(n)$ -time convergence of the 3-majority dynamics was open even in the absence of an adversary whenever the initial bias toward some plurality opinion is not large.

In the remainder of this section, we will describe in more detail the consensus problem and various network scenarios in which it is of interest, our result in this setting, and a comparison with previous related results.

1.1 Consensus (or Byzantine agreement)

The *consensus* problem in a distributed network is defined as follows: A collection of agents, each holding a piece of information (an element of a set Σ), interact with the goal of agreeing on one of the elements of Σ initially held by at least one agent, possibly in the presence of an adversary that is trying to disrupt the protocol. The consensus problem in the presence of an adversary (known as Byzantine agreement) is a fundamental primitive in the design of distributed algorithms [22, 24]. The goal is to design a distributed, local protocol that brings the system into a configuration that meets the following conditions: (1) *Agreement*: All non-corrupted nodes support the same opinion v ; (2) *Validity*: The opinion v must be a *valid* one, i.e., an opinion which was initially declared by at least one (non-corrupted) node; (3) *Termination*: Every non-corrupted node can correctly decide to stop running the protocol at some round.

Recently, there has been considerable interest in the design of consensus algorithms in models that severely restrict both communication and computation [3, 6, 12], both for efficiency consideration and because such models capture aspects of the way consensus is reached in social networks, biological systems, and other domains of interest in network science [2, 4, 8, 9, 14, 15, 16].

In particular, we assume an anonymous network in which nodes possess no unique IDs, nor do they have any static binding of their local link ports (i.e., nodes cannot keep track of *who sent what*). From the point of view of computation, the most restrictive setting is to assume that each node only has $\mathcal{O}(\log |\Sigma|)$ bits of memory available, i.e., it just suffices to store a constant number of opinions. We further assume that this bound extends to link bandwidth available in each round. Finally, communication capabilities are severely constrained and non-deterministic: Every node can communicate with at most a (small) constant number of random neighbors in each round. These constraints are well-captured by the *uniform-gossip* communication model

[10, 18, 19]: At every round, every node can exchange a (short) message (say, $\Theta(\log(|\Sigma|))$ bits) with each of at most h random neighbors, where h is a (small) absolute constant¹. A more recent, sequential variant of the uniform-gossip model is the *(random) population-protocols* model [3, 1, 2] in which, in each round, a single interaction between a pair of randomly selected nodes occurs.

The classic notion of consensus is too strong and unrealistic in the aforementioned distributed settings, that instead rely on *weaker* forms of consensus, deeply investigated in [3, 4, 5, 12]. In this paper, we consider a variant of the *stabilizing-consensus* problem [4] considered in [3]: There, a solution is required to converge to a stable *regime* in which the above three properties are guaranteed in a relaxed, still useful form². More precisely:

Definition 1.1. *A stabilizing almost-consensus protocol must ensure the following properties:*

- Almost agreement. *Starting from any initial configuration, in a finite number of rounds, the system must reach a regime of configurations where all but a negligible “bad” subset (i.e. having size $\mathcal{O}(n^\gamma)$ for constant $\gamma < 1$) of the nodes support the same opinion.*
- Almost validity. *The system is required to converge w.h.p. to an almost-agreement regime where all but a negligible bad set of nodes keep the same valid opinion.*
- Non termination. *In dynamic distributed systems, nodes represent simple and anonymous computing units which are not necessarily able to detect any global property.*
- Stability. *The convergence toward such a weaker form of agreement is only guaranteed to hold with high probability (in short, w.h.p.³) and only over a long period (i.e. for any arbitrarily-large polynomial number of rounds).*

The main result of this paper is on the convergence properties of the 3-majority dynamics in the uniform-gossip model in the presence of the adaptive F -dynamic adversary (defined above) and of the adaptive F -static adversary. In the latter, the adversary looks at the initial configuration, then changes the opinion of up to F nodes and, after that, no further adversary’s actions are allowed.

Theorem 1.2. *Let $k \leq n^\alpha$ for some constant $\alpha < 1$ and $F = \beta\sqrt{n}/(k^{\frac{5}{2}} \log n)$ for some constant $\beta > 0$. Starting from any initial configuration having k valid opinions, the 3-majority dynamics reaches a (valid) stabilizing almost-consensus in presence of any F -dynamic adversary within $\mathcal{O}((k^2\sqrt{\log n} + k \log n)(k + \log n))$ rounds, w.h.p.*

Moreover, the same bound on the convergence time holds in the presence of any F -static adversary with a larger bound on F , i.e., $F = n/k - \sqrt{kn \log n}$.

In [7], an $\Omega(k \log n)$ bound on the convergence-time of the 3-majority dynamics is derived (that holds even when the system starts from biased configurations): So, our bound is almost-tight whenever $k = \mathcal{O}(\text{polylog}(n))$.

Not assuming a large initial bias of the plurality opinion considerably complicates the analysis. Indeed, the major open challenge is the analysis from (almost) uniform configurations, where the system needs to break the initial symmetry in the absence of significant drifts towards any of the initial opinions. So far, this issue has never been analyzed even in the non-adversarial case. Moreover, the phase before symmetry breaking is the one in which the adversary has more chances to cause undesired behaviours: Long delays and/or convergence towards non-valid opinions. In Section 2, after providing some preliminaries, we shall discuss the above technical challenges.

¹In fact, $h = 1$ in the standard uniform-gossip model. It is easy to verify that all our results still hold in this more restricted model at the cost of a constant slow-down in convergence time and local memory size.

² These relaxed convergence properties are described in detail in Section 7 of [3].

³According to the standard definition, we say that a sequence of events \mathcal{E}_n , $n = 1, 2, \dots$ holds *with high probability* if $\mathbf{P}(\mathcal{E}_n) = 1 - \mathcal{O}(1/n^\lambda)$ for some positive constant $\lambda > 0$.

1.2 Previous results

Consensus problems in distributed systems have been the focus of a large body of work in several research areas, such as distributed computing [17], communication networks [25], social networks and voting systems [21, 27], distributed databases [10, 11], biological systems and Chemical Reaction Networks [9]. For brevity's sake, we here focus on results that are closest in spirit to our work.

In [3], the authors show that w.h.p. n agents that meet at random can reach valid stabilizing almost-consensus in $\mathcal{O}(n \log n)$ pairwise interactions against an $F = o(\sqrt{n})$ -bounded dynamic adversary. The adopted protocol is the well-studied third-state protocol [3, 23]. However, their analysis (and, thus, their result) only holds for the binary case and for the *population-protocol* model: At every round only one pair of nodes can interact. The authors left the existence of protocols for the multi-valued Byzantine case as a final open question [3]. In general, sequential processes are much easier to analyze than parallel ones (like those yielded by the uniform-gossip model): For instance, the resulting Markov chains are reversible [20] while those arising from parallel processes are non-reversible.

In the uniform-gossip model, in [12] the authors provide an analysis of the *3-median* rule, in which every node updates its value to the median of its random sample. They show that this dynamics converges to an almost-agreement configuration (which is even a good approximation of the global median) within $\mathcal{O}(\log k \cdot \log \log n + \log n)$ rounds, w.h.p. It turns out that, in the binary case, the median rule is equivalent to the 3-majority dynamics, thus their result implies that 3-majority is an $(F = \sqrt{n})$ -stabilizing consensus with $\mathcal{O}(\log n)$ convergence time. However, in the non-binary case, it requires Σ to be a totally-ordered set and the possibility to perform basic algebraic operations: This is a rather strong restriction in applications arising from social networks, voting-systems, and bio-inspired systems. More importantly, we emphasize that, even assuming an ordered opinion set (Σ, \leq) , the 3-median rule does not guarantee the crucial property of *validity* against both F -static (and, clearly, dynamic) adversaries even for very-small bounds on F (say $F = \text{polylog}(n)$).

We strongly believe that the validity property of consensus plays a crucial role in several realistic scenarios, such as monitoring sensor networks, bio-inspired dynamic systems, and voting systems [9, 21, 27].

More recently, the 3-majority rule in the multi-opinion case (i.e. for $|\Sigma| \geq 3$) has been studied for a stronger goal than consensus, namely, *stabilizing plurality consensus* [7]. In this task, the goal is to reach an almost-stable consensus towards the valid opinion initially supported by *the plurality* of the nodes. However, the initial configuration is assumed to have a large bias towards the plurality opinion. Then, let k be the number of valid opinions, and let s be the initial difference between the largest and the second-largest opinion: By strongly exploiting the assumption $s \geq \sqrt{kn \log n}$, the authors in [7] proved that, w.h.p., the system converges to the plurality opinion within time $\Theta(k \log n)$.

Another version of binary stabilizing almost-consensus is the one studied by Yildiz et al in [27]: Here, corrupted nodes are *stubborn* agents of a social network who influence others but never change their opinions. They prove negative results under a generalized variant of the classic voter dynamics in the (Poisson-clock) population-protocol model.

2 The Process and its Analysis in a Nuthshell

Preliminaries. We assume a distributed system consisting of n nodes that communicate with each other over a complete graph via the synchronous *uniform-gossip* mechanism: In every round, each node can pull information from (at most) h random neighbors, where h is an

absolute constant (in this work, $h = 3$). At the onset, every node chooses an arbitrary item, called *opinion*, from an arbitrary finite set Σ . A simple dynamics for consensus is the *3-majority protocol* [7]:

In each round, every node samples three nodes uniformly at random (including itself and with repetitions) and revises its opinion according to the majority of the opinions it sees. If it sees three different opinions, it picks the first one.

Clearly, in the case of three different opinions, choosing the second or the third one would not make any difference, nor would choosing one of the observed opinions uniformly at random.

Since the communication graph is complete and nodes are anonymous, the overall system state at any round can be described by a *configuration* $\mathbf{c} := (c_1, \dots, c_{|\Sigma|})$, where the *support* c_i of opinion i is the number of nodes holding opinion i in that system's state. Given configuration \mathbf{c} , we say that an opinion i is *active* in \mathbf{c} if $c_i > 0$ and, for any set of active opinions $W \subseteq \Sigma$, we define $m(W) := \arg \min_{i \in W} c_i$. For any variable x of the process, we write $x^{(t)}$ if we are considering its value at round t and $X^{(t)}$ to denote the corresponding random variable. Furthermore, following [20], considered a configuration \mathbf{c} and a random variable X defined over the process, we write $\mathbf{P}_{\mathbf{c}}(X^{(t)} = x)$ for $\mathbf{P}(X^{(t)} = x \mid \mathbf{C}^{(0)} = \mathbf{c})$, i.e., to denote the probability distribution of the variable X when the system evolves for t consecutive rounds starting from configuration \mathbf{c} . Analogously, we write $\mathbf{E}_{\mathbf{c}}[X^{(t)}]$ for the associated conditional expectation.

The next lemma provides the expected number of nodes supporting a given opinion at round $t + 1$ (and a general upper bound to it), given the configuration at round t . The simple proof of the first equality is in [7]. It is also included in Appendix A to make the paper self-contained.

Lemma 2.1 (See [7]). *Let \mathbf{c} be the configuration at round t and let $W \subseteq \Sigma$ be the subset of active opinions in \mathbf{c} . Then, for any opinion $i \in W$,*

$$\mathbf{E} \left[C_i^{(t+1)} \mid \mathbf{C}^{(t)} = \mathbf{c} \right] = c_i \left(1 + \frac{c_i}{n} - \frac{\sum_{j \in W} c_j^2}{n^2} \right) \leq c_i \left(1 + \frac{c_i}{n} - \frac{1}{|W|} \right) \quad (1)$$

The above upper bound easily implies that opinions whose supports fall below the average $n/|W|$ decrease in expectation. This expected drift is a key-ingredient of our analysis and, as we will see in the next paragraph, it provides useful intuitions about the process. On the other hand, when \mathbf{c} is almost uniform, the above *drift* turns out to be negligible and symmetry breaking is due to the inherent variance of the random process.

Failed attempts. When the 3-majority dynamics starts from configurations that exhibit a large initial support bias between the largest and the second-largest opinions, the approach adopted in [7] successfully exploits the fact that the initial plurality is preserved throughout the evolution of the random process, with an expected positive drift that is also preserved, w.h.p. An intuition of this fact can be achieved from simple manipulations of (1). However, the aforementioned drift is only preserved if the largest opinion never changes (w.h.p.), *no matter which the second-largest opinion is*: a condition that is not met by uniform configurations. A promising attempt to cope with uniform configurations is to consider the r.v. $S^{(t)} = C_{\mathbb{M}(t)}^{(t)} - C_{2\mathbb{M}(t)}^{(t)}$ where $\mathbb{M}(t)$ and $2\mathbb{M}(t)$ are the r.v.s that take the index of (one of) the largest opinion and of (one of) the second-largest ones, respectively, in round t . For any *fixed* pair i, j , such that $c_i > c_j$, (1) implies that the difference $C_i^{(t+1)} - C_j^{(t+1)}$ in the next round is positive in expectation, so a suitable submartingale argument [20] seemed to work in order to show that the system (rather quickly) achieves a “sufficiently-large” bias toward the plurality as to allow

fast convergence. This approach would work if the *random* indices \mathbb{M} and $2\mathbb{M}$ maintained their initial values across the entire duration of the process. Unfortunately, starting from uniform configurations, in the next round, the expected difference between the *new* largest opinion and the *new* second largest one may have no positive drift at all. Roughly speaking, in the next round, the r.v. $C_{2\mathbb{M}(t+1)}^{(t+1)}$ can be much larger than the r.v. $C_{\mathbb{M}(t)}^{(t+1)}$.

A promising dynamics for the stabilizing almost-consensus problem is the one introduced in [12], in which nodes revise their opinions (assumed to be totally ordered) by taking the median between the currently held opinion and those held by two randomly sampled nodes. However, while we do not assume opinions to be integers (or totally ordered), their analysis strongly relies on the fact that the median opinion (or any good approximation of it) exhibits a strong increasing drift, even when starting from almost-uniform configuration, whereas no opinion is “special” to a majority rule when the starting configuration is uniform. The adoption of an inherently biased function as the median can have important consequences. To get an intuition, the reader may consider the following simple instance: $\Sigma = \{1, 2, 3\}$, with the system starting in configuration $c_1 = n/2, c_2 = 0, c_3 = n/2$. At the end of the first round, a static adversary changes the values of $F = \log n$ nodes, equally distributed in c_1 and c_3 , to value 2. The (non-valid) value 2 is the *global median* and some counting arguments show that, while values 1 and 3 have no positive expected drift, the median has an exponential expected drift that holds w.h.p. whenever $c_1, c_3 = \Theta(n)$. This might fool the system into the configuration in which $c_2 = n$, thus converging to a non-valid value.

Our New Approach: An Overview. Our analysis significantly departs from the above approaches. It is important to remark that, for $|\Sigma| \geq 3$, no analysis of the 3-majority dynamics with almost-uniform initial configurations is known, even in the simpler non-adversarial case. On the other hand, while simpler, the analysis of the non-adversarial case still has *per-se* interest and it requires to address some of the main technical challenges that also arise in the adversarial case. Section 3 will be thus devoted to the analysis of the non-adversarial case, while an outline is given in the paragraphs that follow.

When the configuration is (approximately) uniform, Lemma 2.1 says us that the process exhibits no significant drift toward any *fixed* opinion. Interestingly, things change if we consider the random variable $C_{\mathbb{m}}^{(t)}$, indicating the smallest opinion support at round t . Let $j \leq k$ be the number of active opinions in a given round t , we first prove that the expected value of $C_{\mathbb{m}}^{(t)}$ always exhibits a non-negligible negative drift:

$$\mathbf{E} \left[C_{\mathbb{m}}^{(t+1)} \mid \mathbf{C}^{(t)} = \hat{\mathbf{c}} \right] \leq c_{\mathbb{m}} - \varepsilon \frac{\sqrt{n}}{j^{3/2}}, \text{ for some constant } \varepsilon > 0 \quad (2)$$

This drift is essentially a consequence of Lemma 2.1 *and* of the standard deviation of r.v.s $C_i^{(t)}$ s (see the proof of Lemma 3.3). The analysis then proceeds along consecutive phases, each consisting of a suitable number of consecutive rounds. If the number of active opinions at the beginning of the generic phase is j , we prove that, with positive constant probability, $C_{\mathbb{m}}^{(t)}$ vanishes within the end of the phase, so that the next phase begins with (at most) $j - 1$ active opinions.

We clearly need a good bound on the length of a phase beginning with at most j opinions. To this aim, we derive a new upper bound - stated in Lemma 3.2 - on the *hitting time* of stochastic processes with expected drift that are defined by finite-state Markov chains [20]. Thanks to this result, we can use the negative drift in (2) to prove that, from any configuration with $j \leq k$ active opinions, $C_{\mathbb{m}}^{(t)}$ drops below the threshold $n/j - \sqrt{jn \log n}$ within $\mathcal{O}(\text{poly}(j, \log n))$ rounds, with constant positive probability: This “hitting” event represents the exit condition from the symmetry-breaking stage of the phase. Indeed, once it occurs, we can consider *any fixed* active

opinion i having support size c_i below the above threshold (thanks to the previous stage, we know that there is a good chance this opinion exists): We then show that C_i has a negative drift of order $\Omega(c_i/j)$. This allows us to prove that C_i drops from $n/j - \sqrt{jn \log n}$ to zero within $\mathcal{O}(\text{poly}(j, \log n))$ further rounds, with positive constant probability. This interval of rounds is the dropping stage of the phase.

Ideally, the process proceeds along k consecutive phases, indexed as $j = k, k-1, \dots, 2$, such that we are left with at most $j-1$ active opinions at the end of Phase j . In practice, we only have a constant probability that at least one opinion disappears during Phase j . However, using standard probabilistic arguments, we can prove that, w.h.p., for every j , the transition from j to $j-1$ active opinions takes a constant (amortized) number of phases, each requiring $\mathcal{O}(\text{poly}(j, \log n))$ rounds.

The presence of a dynamic, adaptive adversary makes the above analysis technically more complex. A major issue is that a different definition of *Phase* must be considered, since the adversary might permanently feed any opinion so that the latter never dies. So the number of active opinions might not decrease from one phase to the next one. Essentially, we need to manage the persistence of “small” (valid or not) opinions: The end of a phase is now characterized by one “big” valid color that becomes “small” and, moreover, we need to show that, in general, “small” colors never becomes “big”, no matter what the dynamic F -bounded adversary does. An informal description of the dynamic-adversary case is given in Subsection 4.2.

3 The 3-Majority Dynamics without Adversary

Let $\mathcal{C} \subseteq \Sigma$ be the subset of valid opinions, i.e. those supported by at least one node in the initial configuration, and denote by $k = |\mathcal{C}|$ its size. This section is devoted to the proof of the following result.

Theorem 3.1 (The Adversary-Free Case.). *Starting from any initial configuration with $k \leq n^{1/3-\varepsilon}$ active opinions, where $\varepsilon > 0$ is an arbitrarily-small constant, the 3-majority dynamics reaches consensus within $\mathcal{O}\left((k^2 \log^{1/2} n + k \log n)(k + \log n)\right)$ rounds, w.h.p.*

We first provide the lemmas required for the process analysis and then we give the formal proof of the above theorem.

The next lemma shows an upper bound on the time it takes a stochastic process with values in $N = \{0, 1, \dots, n\}$ to reach or exceed a target value m , under mild hypotheses on the process. We here give only an idea of the proof, the full proof is in Appendix A.

Lemma 3.2. *Let $\{X_t\}_t$ be a Markov chain with finite state space Ω , let $f : \Omega \rightarrow N$ be a function mapping states of the chain in non-negative integer numbers, and let $\{Y_t\}_t$ be the stochastic process over N defined by $Y_t = f(X_t)$. Let $m \in N$ be a “target value” and let*

$$\tau = \inf\{t \in \mathbb{N} : Y_t \geq m\}$$

be the random variable indicating the first time Y_t reaches or exceeds value m . Assume that, for every state $x \in \Omega$ with $f(x) \leq m-1$, it holds that

1. (Positive drift). $\mathbf{E}[Y_{t+1} | X_t = x] \geq f(x) + \lambda$ for some $\lambda > 0$
2. (Bounded jumps). $\mathbf{P}_x(Y_\tau \geq \alpha m) \leq \alpha m/n$, for some $\alpha > 1$.

Then, for every starting state $x \in \Omega$, it holds that

$$\mathbf{E}_x[\tau] \leq 2\alpha \frac{m}{\lambda}$$

Idea of the proof. From Hypothesis 1 it follows that $Z_t = Y_t - \lambda t$ is a *submartingale* that satisfies the hypotheses of the Doob's *Optional Stopping Theorem* [13] (see e.g. Corollary 17.8 in [20] or Theorem 10.10 in [26]), thus

$$0 \leq f(x) = \mathbf{E}_x [Z_0] \leq \mathbf{E}_x [Z_\tau] = \mathbf{E}_x [Y_\tau] - \lambda \mathbf{E}_x [\tau]$$

And from Hypothesis 2 it follows that $\mathbf{E}_x [Y_\tau] \leq 2\alpha m$. \square

We now exploit the above lemma in order to bound the time required by the *symmetry-breaking* stage.

Lemma 3.3 (Symmetry-breaking stage). *Let \mathbf{c} be any configuration with j active opinions. Within $t = \mathcal{O}\left(j^2 \log^{1/2} n\right)$ rounds it holds that*

$$\mathbf{P}_{\mathbf{c}} \left(\exists i \text{ such that } C_i^{(t)} \leq n/j - \sqrt{jn \log n} \right) \geq \frac{1}{2}$$

Sketch of Proof. Let J be the set of j active opinions in \mathbf{c} and let $\mathbf{C}^{(t)} = \left(C_i^{(t)} : i \in J \right)$ be the random variable indicating the opinion configuration at round t , where we assume $\mathbf{C}^{(0)} = \mathbf{c}$. Let $C_{\mathbf{m}}^{(t)} = \min \left\{ C_i^{(t)} : i \in J \right\}$ be the minimum among all $C_i^{(t)}$ s and consider the stochastic process $\{Y_t\}_t$ defined as $Y_t = \lfloor n/j \rfloor - C_{\mathbf{m}}^{(t)}$. Observe that Y_t takes values in $\{0, 1, \dots, \lfloor n/j \rfloor\}$ and it is a function of $\mathbf{C}^{(t)}$. We are interested in the first time Y_t becomes at least as large as $\sqrt{jn \log n}$, i.e.

$$\tau = \inf \left\{ t \in \mathbb{N} : Y_t \geq \sqrt{jn \log n} \right\}$$

We now show that $\{Y_t\}_t$ satisfies Hypotheses 1 and 2 of Lemma 3.2, with $\lambda = \varepsilon \sqrt{n}/j^{3/2}$, for a suitable constant $\varepsilon > 0$.

1. Let $\hat{\mathbf{c}} = (\hat{c}_i : i \in J)$ be any configuration with j active opinions such that $\hat{c}_{\mathbf{m}} > n/j - \sqrt{jn \log n}$. We want to prove that

$$\mathbf{E} \left[C_{\mathbf{m}}^{(t+1)} \mid \mathbf{C}^{(t)} = \hat{\mathbf{c}} \right] \leq c_{\mathbf{m}} - \varepsilon \frac{\sqrt{n}}{j^{3/2}} \quad (3)$$

Two cases may arise.

Case $\hat{c}_{\mathbf{m}} > n/j - 2\varepsilon \sqrt{n/j}$: Observe that, in this case, r.v.s $\left\{ C_i^{(t+1)} : i \in J \right\}$ conditional on $\left\{ \mathbf{C}^{(t)} = \hat{\mathbf{c}} \right\}$ have standard deviation $\Omega\left(\sqrt{n/j}\right)$. Moreover, they are binomial and negatively associated. Hence, by choosing ε small enough, from the Central Limit Theorem we have that

$$\mathbf{P} \left(i \in J \text{ exists such that } C_i^{(t+1)} \leq \frac{n}{j} - 6\varepsilon \cdot \sqrt{\frac{n}{j}} \right) \geq 1/2$$

We thus get

$$\mathbf{E} \left[C_{\mathbf{m}}^{(t+1)} \mid \mathbf{C}^{(t)} = \hat{\mathbf{c}} \right] \leq \frac{1}{2} \left(\frac{n}{j} - 6\varepsilon \cdot \sqrt{\frac{n}{j}} \right) + \frac{1}{2} \cdot \frac{n}{j} = \frac{n}{j} - 3\varepsilon \sqrt{\frac{n}{j}} \leq c_{\mathbf{m}} - \varepsilon \sqrt{\frac{n}{j}} \leq c_{\mathbf{m}} - \varepsilon \frac{\sqrt{n}}{j^{3/2}} \quad (4)$$

Case $\hat{c}_{\mathbf{m}} \leq n/j - 2\varepsilon \sqrt{n/j}$: Equation (3) easily follows from Lemma 2.1. Indeed, let $i \in J$ be an opinion such that $\hat{c}_i = \hat{c}_{\mathbf{m}}$, then

$$\begin{aligned} \mathbf{E} \left[C_{\mathbf{m}}^{(t+1)} \mid \mathbf{C}^{(t)} = \hat{\mathbf{c}} \right] &\leq \mathbf{E} \left[C_i^{(t+1)} \mid \mathbf{C}^{(t)} = \hat{\mathbf{c}} \right] \leq \hat{c}_i \left(1 + \frac{\hat{c}_i}{n} - \frac{1}{j} \right) \\ &\leq \hat{c}_i \left(1 - \frac{2\varepsilon}{\sqrt{n/j}} \right) \leq \hat{c}_i - \frac{\varepsilon \sqrt{n}}{j^{3/2}} = \hat{c}_{\mathbf{m}} - \varepsilon \frac{\sqrt{n}}{j^{3/2}} \end{aligned} \quad (5)$$

where we used the case's condition and the fact that $\hat{c}_i = \hat{c}_m \geq n/(2j)$.

2. Since random variables $\{C_i^{(t+1)} : i \in J\}$ conditional on the configuration at round t are binomial, it is possible to apply Chernoff bound (though with some care) to prove that

$$\mathbf{P}_{\mathbf{c}} \left(Y_{\tau} \geq \alpha \sqrt{jn \log n} \right) \leq \frac{1}{n}, \quad \text{for some constant } \alpha > 1 \quad (6)$$

Though this result seems intuitive, its formal proof is less obvious, since τ is a stopping time and thus itself a random variable. Lemma B.1 in Appendix B offers a formal proof of the above statement.

From (3) and (6), we have that $\{Y_t\}_t$ satisfies the hypotheses of Lemma 3.2 with $m = \sqrt{jn \log n}$ and $\lambda = \varepsilon \sqrt{n}/j^{3/2}$. Hence $\mathbf{E}_{\mathbf{c}}[\tau] < j^2 \sqrt{\log n}$ and, from Markov inequality, for $t = 2j^2 \sqrt{\log n}$, we finally get

$$\mathbf{P}_{\mathbf{c}} \left(\forall i \in J : C_i^{(t)} \geq n/j - \sqrt{jn \log n} \right) \leq \mathbf{P}_{\mathbf{c}} \left(\tau > 2j^2 \sqrt{\log n} \right) \leq \frac{1}{2}$$

□

We now provide the analysis of the *dropping* stage: More precisely, we show that, if the system starts with up to j active opinions and one of them (say i) is below the threshold $n/j - \sqrt{jn \log n}$, then i drops to the smaller threshold $j^2 \log n$ within $\mathcal{O}(j \log n)$ additional rounds. This bound can be proved w.h.p. since, in this regime, C_i is still sufficiently large to apply the Chernoff bound. This concentration result is not necessary to the purpose of proving Theorem 3.1, while it is a key ingredient in the analysis of the adversarial case (Theorem 4.2). The next lemma can be proved by standard concentration arguments - applied in an iterative way - on the r.v. $C_i^{(t)}$ (see Appendix B).

Lemma 3.4 (Dropping stage 1). *Let \mathbf{c} be any configuration with $j \leq n^{1/3-\varepsilon}$ active opinions, where $\varepsilon > 0$ is an arbitrarily-small positive constant, and such that an opinion i exists with $c_i \leq n/j - \sqrt{jn \log n}$. Within $t = \mathcal{O}(j \log n)$ rounds opinion i becomes $\mathcal{O}(j^2 \log n)$ w.h.p.*

In the next lemma we prove that once c_i becomes smaller than $n/(2j)$, then opinion i disappears within further $\mathcal{O}(j \log n)$ rounds with constant probability. We here give only an idea of the proof, the full proof is in Appendix B.

Lemma 3.5 (Dropping stage 2). *Let \mathbf{c} be any configuration with $j \leq n^{1/3-\varepsilon}$ active opinions, where $\varepsilon > 0$ is an arbitrarily-small positive constant, and such that an opinion i exists with $c_i \leq n/(2j)$. Within $t = \mathcal{O}(j \log n)$ rounds opinion i disappears with probability at least $1/2$.*

Idea of the proof. If $c_i \leq n/(2j)$ in configuration \mathbf{c} , then from Lemma 2.1 it follows that

$$\mathbf{E} \left[C_i^{(t+1)} \mid \mathbf{C}^{(t)} = \mathbf{c} \right] \leq c_i \left(1 - \frac{1}{2j} \right)$$

Moreover, since $C_i^{(t+1)}$ conditional on $\{\mathbf{C}^{(t)} = \mathbf{c}\}$ is binomial, if $j \leq n^{1/3-\varepsilon}$, from the Chernoff bound it follows that $\mathbf{P} \left(C_i^{(t+1)} > n/(2j) \mid \mathbf{C}^{(t)} = \mathbf{c} \right) \leq e^{-\Theta(n^\varepsilon)}$. Hence, it is easy to check that for any initial configuration \mathbf{c} with $c_i \leq n/(2j)$ the following recursive relation holds

$$\mathbf{E}_{\mathbf{c}} \left[C_i^{(t)} \right] \leq \left(1 - \frac{1}{2j} \right) \mathbf{E}_{\mathbf{c}} \left[C_i^{(t-1)} \right] + e^{-n^{\varepsilon/2}}$$

that for some $t = \mathcal{O}(j \log n)$ gives $\mathbf{E}_{\mathbf{c}} [C_i^{(t)}] \leq 1/2$. Since $C_i^{(t)}$ is a non-negative integer-valued r.v., the thesis then follows from the Markov inequality. \square

Proof of Theorem 3.1. From Lemmas 3.3, 3.4, and 3.5 it follows that from any configuration with $j \leq k$ active opinions, within $\mathcal{O}(k^2 \sqrt{\log n} + k \log n)$ rounds at least one of the opinions disappears with probability at least $1/4$. Thus, within $\mathcal{O}((k^2 \sqrt{\log n} + k \log n)(k + \log n))$ rounds, all opinions but one disappear w.h.p. \square

4 Convergence Time of 3-Majority with Adversary

In this section we consider the presence of a Byzantine adversary that can adaptively change the opinion of a bounded number of nodes in order to delay convergence time toward a valid consensus, or even worse, to let the system converge toward a non valid one. We consider two different adversarial strategies: A static one and a stronger, dynamic one.

4.1 The F -static adversary

At the end of the first round, once every node has fixed his own initial opinion, the adversary looks at the configuration and arbitrarily replaces the opinion of at most $F = n/k - \sqrt{kn \log n}$ nodes with an arbitrary opinion in Σ . Then the protocol starts the process and no further adversary's actions are allowed. Since any opinion the adversary may introduce has size less than $n/k - \sqrt{kn \log n}$, as a simple consequence of the dropping stage (see Lemmas 3.4 and 3.5), the static adversarial case easily reduces to the non-adversarial one. We thus get the following

Corollary 4.1. *Let $k \leq n^\alpha$ for some constant $\alpha < 1$ and $F = n/k - \sqrt{kn \log n}$. Starting from any initial configuration having k opinions, the 3-majority protocol reaches a stabilizing almost-consensus in presence of any F -static adversary within $\mathcal{O}(k^2 \sqrt{\log n} + k \log n)(k + \log n)$ rounds, w.h.p.*

4.2 The F -dynamic adversary

The actions of this adversary over the studied process can be described as follows. At the end of *every round* t , after nodes have updated their opinions (i.e. once the configuration $\mathbf{C}^{(t)} = \mathbf{c}^{(t)}$ is realized), the F -dynamic adversary looks at the current opinion configuration and replaces the opinion of up to F nodes with any opinion in Σ .

In what follows we consider an F -dynamic adversary with $F \leq \beta \sqrt{n} / (k^{\frac{5}{2}} \log n)$ for a suitable positive constant β . As we will show in the proof of Lemma C.7, this bound on F turns out to be almost tight for guaranteeing that the process converges to an almost-consensus regime in polynomial time, w.h.p.

The presence of the adversary requires us to distinguish between valid and non valid opinions. So, we recall that the set of valid opinions $\mathcal{C} \subseteq \Sigma$ is the subset of active opinions in the initial configuration and we observe that, in the reminder of this section, k denotes the number of valid opinions, i.e., $k := |\mathcal{C}|$.

We are now ready to state our main result in the presence of the dynamic adversary (its full proof is given in Appendix C).

Theorem 4.2 (The Dynamic-Adversary Case.). *Let $k \leq n^\alpha$ for some constant $\alpha < 1$ and $F = \beta \sqrt{n} / (k^{\frac{5}{2}} \log n)$ for some constant $\beta > 0$. Starting from any initial configuration having k opinions, the 3-majority reaches a (valid) stabilizing almost-consensus in presence of any F -dynamic adversary within $\mathcal{O}((k^2 \sqrt{\log n} + k \log n)(k + \log n))$ rounds, w.h.p.*

Idea of the Proof. We here provide a description of the main technical differences w.r.t. the analysis for the non-adversarial case.

As discussed in the overview of the process analysis, the adversary can introduce “small” non-valid opinions and it can keep small valid opinions active that would otherwise disappear (as shown in Section 3). These facts lead us to the problem of managing “small” opinions: The rigorous definition of small opinion is determined by the minimal negative drift for $C_m^{(t)}$ we derived in the proof of Lemma 3.3 (see (5)).

Let $S := \{i \mid c_i \leq \gamma\sqrt{n}/k^{\frac{3}{2}}\}$ be the set of *small opinions* for some constant $\gamma > \beta$, and let its complement $B := \bar{S} = \{i \mid c_i > \gamma\sqrt{n}/k^{\frac{3}{2}}\}$ be the set of *big opinions*.

It turns out that we cannot use the definition of the (end of) phase adopted in the non-adversarial case: At least one (valid) opinion dies. Wlog, let us assume that, at the beginning, all the k valid opinions are big. Then the new phase j is an interval of consecutive rounds, in each of which exactly j big valid opinions are present. The new goal is to show that at the end of phase j , one of the j big colors will get small and, moreover, this color (and no other small color) will never get big. In the symmetry-breaking stage of each phase, we thus need to show that the negative drift of $C_m^{(t)}$ (notice that the latter now denotes the minimum among the j big colors) cannot be opposed by the actions of the F -dynamic adversary, provided that $F \leq \beta\sqrt{n}/(k^{\frac{5}{2}} \log n)$. This fact (stated in Lemma C.7) is obtained via two different technical steps: i) A new bound on the expected negative drift for $C_m^{(t)}$ that considers both the presence of small good opinions and the adversary’s opposing action (this result is formalized in Lemma C.3); ii) A novel use of Lemma 3.2 on the hitting time of random processes in order to bound the expected time of the symmetry-breaking stage. We in fact need to define a new stopping condition that also includes some “bad” event: Some small (valid or not) color become big. We then show that bad stopping events never happen along the entire process, w.h.p. (this is essentially guaranteed by Lemma C.4).

The dropping stage of phase j is now defined as the interval of rounds in which $C_m^{(t)}$ drops from the symmetry-breaking threshold $n/j - \sqrt{jn \log n}$ to the size of small colors i.e. $\gamma\sqrt{n}/k^{\frac{3}{2}}$. Similarly to the non-adversary case, we can here fix the big opinion i that is dropped below the symmetry-breaking threshold and look at its negative drift derived from Lemma C.3. The drift is strong enough to tolerate the actions of the F -bounded adversary and implies an $O(j \log n)$ bound on the time required by this second stage of phase j . This stage’s analysis is given in Lemma C.8.

Finally, after k phases, we are left with one (valid) opinion that accounts for $n - O(\sqrt{n})$ nodes, while the remaining nodes can have any (possibly non valid) opinion and reflect the presence of the adversary. In fact, this is what happens with high probability. □

5 Future Works

We strongly believe that our upper bound on the convergence time of the 3-majority dynamics is not tight w.r.t. k . The factor $\Omega(k^3)$ seems to be not necessary: We believe that at least a factor k can be saved. To this aim, we would need to show that “more” opinions get small during a phase. This number should also depend on the current number of big colors. Another idea would be that of (also) considering the growth of the maximal opinion. Unfortunately, differently from the minimal opinion (see (2) in Section 2), we don’t have any good bound on the expected drift for the maximal opinion that holds from *any* configuration. So, we don’t see how to efficiently adapt our approach without this crucial ingredient.

References

- [1] Dana Angluin, James Aspnes, and Eisenstat David. Stably computable predicates are semilinear. In *In Proc. of the 25th Ann. ACM SIGACT-SIGOPS Symp. on Principles of Distributed Computing (PODC'06)*, pages 292–299. ACM, 2006.
- [2] Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, and Peralta René. Computation in networks of passively mobile finite-state sensors. *Distributed Computing*, 18(4):235–253, 2006.
- [3] Dana Angluin, James Aspnes, and David Eisenstat. A Simple Population Protocol for Fast Robust Approximate Majority. *Distributed Computing*, 21(2):87–102, 2008. (Preliminary version in DISC'07).
- [4] Dana Angluin, Michael J. Fischer, and Hong Jiang. Stabilizing consensus in mobile networks. In *Proc. of Distributed Computing in Sensor Systems (DCOSS'06)*, volume 4026 of *LNCS*, pages 37–50, 2006.
- [5] James Aspnes. Faster Randomized Consensus with an Oblivious Adversary. In *Proc. of the 31st Ann. ACM SIGACT-SIGOPS Symp. on Principles of Distributed Computing (PODC'12)*, pages 1–8. ACM, 2012.
- [6] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Riccardo Silvestri. Plurality Consensus in the Gossip Model. In *Proc. of the 26th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'15)*, pages 371–390. SIAM, 2015.
- [7] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, Riccardo Silvestri, and Luca Trevisan. Simple dynamics for plurality consensus. In *Proc. of the 26th ACM Symp. on Parallelism in Algorithms and Architectures (SPAA'14)*, pages 247–256. ACM, 2014.
- [8] Ohad Ben-Shahar, Shlomi Dolev, Andrey Dolgin, and Michael Segal. Direction election in flocking swarms. In *Proc. of the 6th Int. Workshop on Foundations of Mobile Computing (DIALM-POMC'10)*, pages 73–80. ACM, 2010.
- [9] Luca Cardelli and Attila Csikász-Nagy. The Cell Cycle Switch Computes Approximate Majority. *Scientific Reports*, Vol. 2, 2012.
- [10] Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. Epidemic algorithms for replicated database maintenance. In *Proc. of the 6th Ann. ACM Symposium on Principles of Distributed Computing (PODC'12)*, pages 1–12. ACM, 1987.
- [11] Martin Dietzfelbinger, Andreas Goerdt, Michael Mitzenmacher, Andrea Montanari, Rasmus Pagh, and Michael Rink. Tight thresholds for cuckoo hashing via XORSAT. In *Proc. of the 37th Int. Coll. on Automata, Languages, and Programming (ICALP'10)*, volume 6198 of *LNCS*, pages 213–225. Springer, 2010.
- [12] Benjamin Doerr, Leslie A. Goldberg, Lorenz Minder, Thomas Sauerwald, and Christian Scheideler. Stabilizing consensus with the power of two choices. In *Proc. of the 23rd Ann. ACM Symp. on Parallelism in Algorithms and Architectures (SPAA'11)*, pages 149–158. ACM, 2011.
- [13] Joseph L. Doob. *Stochastic Processes*. John Wiley & Sons Inc., 1953.

- [14] David Doty. Timing in chemical reaction networks. In *Proc. of 25th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA '14)*, pages 772–784. SIAM, 2014.
- [15] Ofer Feinerman, Bernhard Haeupler, and Amos Korman. Breathe Before Speaking: Efficient Information Dissemination Despite Noisy, Limited and Anonymous Communication. In *Proc. of the ACM Symposium on Principles of Distributed Computing (PODC '14)*. ACM, 2014.
- [16] 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 B: Biological Sciences*, 357(1427):1567–1583, 2002.
- [17] Seth Gilbert and Dariusz Kowalski. Distributed agreement with optimal communication complexity. In *Proc. of 21st Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA '10)*, pages 965–977. SIAM, 2010.
- [18] Richard Karp, Christian Schindelhauer, Scott Shenker, and Berthold Vocking. Randomized rumor spreading. In *Proc. of the 41th Ann. IEEE Symp. on Foundations of Computer Science (FOCS'00)*, pages 565–574. IEEE, 2000.
- [19] David Kempe, Alin Dobra, and Johannes Gehrke. Gossip-Based Computation of Aggregate Information. In *Proc. of 43rd Ann. IEEE Symp. on Foundations of Computer Science (FOCS'03)*, pages 482–491. IEEE, 2003.
- [20] David Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.
- [21] Elchanan Mossel, Joe Neeman, and Omer Tamuz. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multi-Agent Systems*, 28(3):408–429, 2014.
- [22] Marshall Pease, Robert Shostak, and Leslie Lamport. Reaching agreement in the presence of faults. *Journal of the ACM*, 27(2):228–234, 1980.
- [23] Etienne Perron, Dinkar Vasudevan, and Milan Vojnovic. Using Three States for Binary Consensus on Complete Graphs. In *Proc. of the 28th IEEE Conf. on Computer Communications (INFOCOM'09)*, pages 2527–1535. IEEE, 2009.
- [24] Michael O. Rabin. Randomized byzantine generals. In *Proc. of the 24th Ann. Symp. on Foundations of Computer Science (SFCS'83)*, pages 403–409. IEEE, 1983.
- [25] Yongxiang Ruan and Yasamin Mostofi. Binary Consensus with Soft Information Processing in Cooperative Networks. In *Proc. of the 47th IEEE Conf. on Decision and Control (CDC'08)*, pages 3613–3619. IEEE, 2008.
- [26] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.
- [27] Mehmet E. Yildiz, Asuman E. Ozdaglar, Daron Acemoglu, Amin Saberi, and Anna Scaglione. Binary Opinion Dynamics with Stubborn Agents. *ACM Trans. Econ. Comput.*, 4(1), 2013.

A Preliminary Results

Proof of Lemma 2.1 According to the 3-majority protocol, a node u gets opinion i if it chooses 3 times opinion i , or if it chooses two times i and one time a different opinion, or if it chooses the first time opinion i and then, the second and third time, two different distinct opinions. Hence, if we denote by $X_{i,u}^{(t)}$ the indicator random variable of the event “Node u gets opinion i at time t ”, we have that

$$\begin{aligned} \mathbf{P}\left(X_{i,u}^{(t+1)} = 1 \mid \mathbf{C}^{(t)} = \mathbf{c}\right) &= \left(\frac{c_i}{n}\right)^3 + 3\left(\frac{c_i}{n}\right)^2 \left(\frac{n-c_i}{n}\right) + \left(\frac{c_i}{n}\right) \left[1 - \left(\frac{\sum_{\ell \in S}^k c_\ell^2}{n^2} + 2\left(\frac{c_i}{n}\right)\left(\frac{n-c_i}{n}\right)\right)\right] \\ &= \left(\frac{c_i}{n^3}\right) \left(n^2 + c_i n - \sum_{\ell \in S}^k c_\ell^2\right) \end{aligned}$$

Then the inequality in (1) is obtained by observing that the sum $\sum_{\ell \in S} c_\ell^2$ is minimized for $c_\ell = n/|S|$. □

Proof of Lemma 3.2 Consider the stochastic process $Z_t = Y_t - \lambda t$ and observe that for any state $x \in \Omega$ with $f(x) \leq m - 1$ it holds that

$$\begin{aligned} \mathbf{E}[Z_{t+1} \mid X_t = x] &= \mathbf{E}[Y_{t+1} \mid X_t = x] - \lambda(t+1) \\ &\geq f(x) + \lambda - \lambda(t+1) \\ &\geq f(x) - \lambda t \end{aligned}$$

where in the inequality we used Hypotheses 1. Thus Z_t is a *submartingale* up to the stopping time τ , i.e. $\mathbf{E}[Z_{t+1} \mid X_t] \geq Z_t$ for any $t < \tau$. Moreover, since $|Y_t| \leq n$ the *jumps* of Z_t can be bounded by a value independent of t

$$|Z_{t+1} - Z_t| = |Y_{t+1} - \lambda(t+1) - Y_t + \lambda t| \leq n + \lambda$$

and it is easy to see that Hypotheses 1 implies $\mathbf{E}_x[\tau] < \infty$, thus we can apply *Doob's Optional Stopping Theorem* [13] (see also, e.g., Corollary 17.8 in [20] and Theorem 10.10 in [26]). It then follows that $\mathbf{E}_x[Z_\tau] \geq \mathbf{E}_x[Z_0] = f(x)$ and, since $\mathbf{E}_x[Z_\tau] = \mathbf{E}_x[Y_\tau] - \lambda \mathbf{E}_x[\tau]$, we have that

$$\mathbf{E}_x[\tau] \leq \frac{\mathbf{E}_x[Y_\tau] - f(x)}{\lambda} \leq \frac{\mathbf{E}_x[Y_\tau]}{\lambda}$$

Finally, we get

$$\begin{aligned} \mathbf{E}_0[Y_\tau] &= \sum_{j=1}^n j \mathbf{P}_0(Y_\tau = j) \\ &= \sum_{j=1}^{\lfloor \alpha m \rfloor} j \mathbf{P}_0(Y_\tau = j) + \sum_{j=\lfloor \alpha m \rfloor + 1}^n j \mathbf{P}_0(Y_\tau = j) \\ &\leq (\alpha m) + n \mathbf{P}_0(Y_\tau > \alpha m) \leq 2(\alpha m) \end{aligned}$$

where in the last inequality we used Hypothesis 2. □

B Proofs for the Non-Adversarial Case

Proof of Lemma 3.4 We first prove that the decreasing rate of C_i depends on its value at the end of the previous round. More formally, if we are in a configuration satisfying the hypotheses of the lemma:

$$\mathbf{P} \left(C_i^{(t)} > c_i^{(t-1)} \left(1 - \frac{1}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) \right) \right) = \mathbf{P} \left(C_i^{(t)} > c_i^{(t-1)} \left(1 - \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) \right) (1 + \delta) \right),$$

where $\delta = \frac{1}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) / 1 - \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right)$

Using Lemma 2.1 and applying Chernoff bound we have:

$$\begin{aligned} \mathbf{P} \left(C_i^{(t)} > c_i^{(t-1)} \left(1 - \frac{1}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) \right) \right) &\leq \exp \left\{ -\frac{\delta^2}{3} \left(1 - \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) \right) c_i^{(t-1)} \right\} \\ &= \exp \left\{ -\frac{\delta}{3} \left(\frac{1}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) \right) c_i^{(t-1)} \right\} \\ &< \exp \left\{ -\frac{1}{3} \left(\frac{1}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) \right)^2 c_i^{(t-1)} \right\} \\ &= n^{-\Theta(1)}. \end{aligned} \tag{7}$$

The second equality in (7) follows from the definition of δ , while the third inequality follows by (upper) bounding the denominator of δ by 1, which is always possible since $c_i/n - 1/j < 0$ from the hypotheses. Finally, to prove the last equality, we used the fact that $c_i \geq j^2 \log n$ and that the function $x(1-x)^2$ is decreasing iff $x \in (1/3, 1)$, with $x = jc_i/n$.

Finally, we can iteratively apply (7) as long as we have at most j active opinions and $C_i^{(t)}$ keeps not smaller than $j^2 \log n$. By standard concentration arguments we get that the time to reach this threshold is $\mathcal{O}(j \log n)$, w.h.p. \square

Proof of Lemma 3.5 Let J be the set of active opinions. By conditioning on all the configurations $\hat{\mathbf{c}} = (\hat{c}_\ell : \ell \in J)$ that the system can take at round $t-1$, we can bound the expectation of $C_i^{(t)}$ as follows

$$\begin{aligned} \mathbf{E}_{\mathbf{c}} [C_i^{(t)}] &= \sum_{\hat{\mathbf{c}}} \mathbf{E} [C_i^{(t)} | \mathbf{C}^{(t-1)} = \hat{\mathbf{c}}] \mathbf{P}_{\mathbf{c}} (\mathbf{C}^{(t-1)} = \hat{\mathbf{c}}) \\ &\leq \left(1 - \frac{1}{2j} \right) \sum_{\hat{\mathbf{c}} : \hat{c}_i \leq n/(2j)} \hat{c}_i \cdot \mathbf{P}_{\mathbf{c}} (\mathbf{C}^{(t-1)} = \hat{\mathbf{c}}) + n \cdot \sum_{\hat{\mathbf{c}} : \hat{c}_i > n/(2j)} \mathbf{P}_{\mathbf{c}} (\mathbf{C}^{(t-1)} = \hat{\mathbf{c}}) \\ &\leq \left(1 - \frac{1}{2j} \right) \mathbf{E}_{\mathbf{c}} [C_i^{(t-1)}] + n \cdot \mathbf{P}_{\mathbf{c}} \left(C_i^{(t-1)} > \frac{n}{2j} \right) \end{aligned}$$

where we used that, for any configuration $\hat{\mathbf{c}}$ with $\hat{c}_i \leq n/(2j)$, Lemma 2.1 gives the bound $\mathbf{E} [C_i^{(t)} | \mathbf{C}^{(t-1)} = \hat{\mathbf{c}}] \leq \hat{c}_i \left(1 - \frac{1}{2j} \right)$. Moreover, if $j \leq n^{1/3-\varepsilon}$, from Chernoff bound it follows that

$$\mathbf{P} \left(C_i^{(t)} > \frac{n}{2j} | \mathbf{C}^{(t-1)} = \hat{\mathbf{c}} \right) \leq e^{-\Theta(n^\varepsilon)}$$

for any such configuration $\hat{\mathbf{c}}$. Hence, for any t we have that $\mathbf{P}_{\mathbf{c}} \left(C_i^{(t)} > \frac{n}{2j} \right) \leq te^{-\Theta(n^\varepsilon)}$. Indeed,

$$\begin{aligned} \mathbf{P}_{\mathbf{c}} \left(C_i^{(t)} > \frac{n}{2j} \right) &\leq \mathbf{P}_{\mathbf{c}} \left(\exists \bar{t} = 1, \dots, t : C_i^{(\bar{t})} > \frac{n}{2j} \wedge C_i^{(\bar{t}-1)} \leq \frac{n}{2j} \right) \\ &\leq \sum_{\bar{t}=1}^t \mathbf{P}_{\mathbf{c}} \left(C_i^{(\bar{t})} > \frac{n}{2j} \wedge C_i^{(\bar{t}-1)} \leq \frac{n}{2j} \right) \\ &= \sum_{\bar{t}=1}^t \sum_{\hat{\mathbf{c}}: \hat{c}_i \leq n/(2j)} \mathbf{P} \left(C_i^{(\bar{t})} > \frac{n}{2j} \mid \mathbf{C}^{(\bar{t}-1)} = \hat{\mathbf{c}} \right) \mathbf{P}_{\mathbf{c}} \left(\mathbf{C}^{(\bar{t}-1)} = \hat{\mathbf{c}} \right) \\ &\leq te^{-\Theta(n^\varepsilon)} \end{aligned}$$

Thus for any $t = \text{poly}(n)$ the following recursive relation holds

$$\mathbf{E}_{\mathbf{c}} \left[C_i^{(t)} \right] \leq \left(1 - \frac{1}{2j} \right) \mathbf{E}_{\mathbf{c}} \left[C_i^{(t-1)} \right] + e^{-n^{\varepsilon/2}}$$

And it gives

$$\mathbf{E}_{\mathbf{c}} \left[C_i^{(t)} \right] \leq \left(1 - \frac{1}{2j} \right)^t \frac{n}{2j} + e^{-n^{\varepsilon/3}}$$

Hence, for $t = 2j(\log n + 1)$ we have that $\mathbf{E}_{\mathbf{c}} \left[C_i^{(t)} \right] \leq 1/2$ and since $C_i^{(t)}$ takes non-negative integer values, the thesis follows from Markov inequality. \square

Lemma B.1. *Let \mathbf{c} be any configuration with j active opinions. Consider the stochastic process $\{Y_t\}_t$ defined as $Y_t = \left\lfloor \frac{n}{j} \right\rfloor - C_{\mathbf{m}}^{(t)}$ and define the stopping time $\tau = \inf \{t \in \mathbb{N} : Y_t \geq \sqrt{jn \log n}\}$. Then:*

$$\mathbf{P}_{\mathbf{c}} \left(Y_\tau > \alpha \sqrt{jn \log n} \right) \leq \frac{1}{n}.$$

Proof. First of all, $\mathbf{E}_{\mathbf{c}}[\tau] < \infty$, since $C_{\mathbf{m}}^{(t)}$ has a negative drift (see the proof of Lemma 3.3). Next, from the definition of Y_t :

$$\begin{aligned} \mathbf{P}_{\mathbf{c}} \left(Y_\tau > \alpha \sqrt{jn \log n} \right) &= \mathbf{P}_{\mathbf{c}} \left(C_{\mathbf{m}}^{(\tau)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \right) \\ &= \mathbf{P}_{\mathbf{c}} \left(\exists \ell : C_\ell^{(\tau)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \right) \\ &\leq \sum_{\ell=1}^j \mathbf{P}_{\mathbf{c}} \left(C_\ell^{(\tau)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \right). \end{aligned}$$

Next, from the definition of the stopping time τ :

$$\begin{aligned} \mathbf{P}_{\mathbf{c}} \left(C_\ell^{(\tau)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \right) &= \sum_{t=1}^{\infty} \mathbf{P}_{\mathbf{c}} \left(C_\ell^{(t)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \wedge \tau = t \right) = \\ &= \sum_{t=1}^{\infty} \mathbf{P}_{\mathbf{c}} \left(C_\ell^{(t)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \wedge C_{\mathbf{m}}^{(t)} \leq \left\lfloor \frac{n}{j} \right\rfloor - \sqrt{jn \log n} \mid \bigwedge_{s=1}^{t-1} C_{\mathbf{m}}^{(s)} > \left\lfloor \frac{n}{j} \right\rfloor - \sqrt{jn \log n} \right) \\ &\quad \cdot \mathbf{P}_{\mathbf{c}} \left(\bigwedge_{s=1}^{t-1} C_{\mathbf{m}}^{(s)} > \left\lfloor \frac{n}{j} \right\rfloor - \sqrt{jn \log n} \right) = \\ &= \sum_{t=1}^{\infty} \mathbf{P}_{\mathbf{c}} \left(C_\ell^{(t)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \mid \bigwedge_{s=1}^{t-1} C_{\mathbf{m}}^{(s)} > \left\lfloor \frac{n}{j} \right\rfloor - \sqrt{jn \log n} \right) \\ &\quad \cdot \mathbf{P}_{\mathbf{c}} \left(\bigwedge_{s=1}^{t-1} C_{\mathbf{m}}^{(s)} > \left\lfloor \frac{n}{j} \right\rfloor - \sqrt{jn \log n} \right) \end{aligned} \tag{8}$$

where the last equality follows since $(C_\ell^{(t)} < \lfloor \frac{n}{j} \rfloor - \alpha\sqrt{jn \log n})$ implies $(C_m^{(t)} < \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n})$.

We next consider $\mathbf{P}_c \left(\bigwedge_{s=1}^{t-1} C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right)$. We can write:

$$\begin{aligned} \mathbf{P}_c \left(\bigwedge_{s=1}^{t-1} C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) &= \prod_{s=1}^{t-1} \mathbf{P}_c \left(C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \middle| \bigwedge_{r=1}^{s-1} C_m^{(r)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) \\ &= \prod_{s=1}^{t-1} \mathbf{P}_c \left(C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \middle| C_m^{(s-1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) \end{aligned}$$

where the last equality follows since the 3-majority process is Markovian. We next give an upper bound on $\mathbf{P}_c \left(C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \middle| C_m^{(s-1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right)$:

$$\begin{aligned} \mathbf{P}_c \left(C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \middle| C_m^{(s-1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) &= \\ &= \sum_{\hat{c}: \hat{c}_m > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n}} \mathbf{P}_{\hat{c}} \left(C_m^{(1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) \cdot \mathbf{P}_c \left(\mathbf{C}^{(s-1)} = \hat{c} \middle| C_m^{(s-1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) \leq \\ &\leq \sum_{\hat{c}: \hat{c}_m > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n}} \mathbf{P}_{\hat{c}} \left(C_l^{(1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) \cdot \mathbf{P}_c \left(\mathbf{C}^{(s-1)} = \hat{c} \middle| C_m^{(s-1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right), \end{aligned}$$

where $l = \arg \hat{c}_m$ (ties broken arbitrarily). We can give an upper bound on $\mathbf{P}_{\hat{c}} \left(C_l^{(1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right)$ using a “reverse” Chernoff bound⁴. In particular, it is possible to show that

$$\mathbf{P}_{\hat{c}} \left(C_l^{(1)} > (1 - \delta) \mathbf{E}_{\hat{c}} [C_l^{(1)}] \right) \leq 1 - e^{-\beta \delta^2 \mathbf{E}_{\hat{c}} [C_l^{(1)}]}$$

for a suitable constant β . We use $\delta = \sqrt{jn \log n} / \mathbf{E}_{\hat{c}} [C_l^{(1)}]$ and note that $n/2j \leq \mathbf{E}_{\hat{c}} [C_l^{(1)}] \leq n/j$, so that

$$\mathbf{P}_{\hat{c}} \left(C_l^{(1)} > (1 - \delta) \mathbf{E}_{\hat{c}} [C_l^{(1)}] \right) \leq 1 - e^{-4\beta j^2 \log n}. \quad (9)$$

Saturating with respect to \hat{c} yields

$$\mathbf{P}_c \left(C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \middle| C_m^{(s-1)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) \leq 1 - e^{-4\beta j^2 \log n}.$$

On the other hand, using standard concentration techniques and recalling that $\mathbf{E}_{\hat{c}} [C_l^{(1)}] \leq n/2j$, we can prove that:

$$\mathbf{P}_c \left(C_\ell^{(t)} < \lfloor \frac{n}{j} \rfloor - \alpha\sqrt{jn \log n} \middle| \bigwedge_{s=1}^{t-1} C_m^{(s)} > \lfloor \frac{n}{j} \rfloor - \sqrt{jn \log n} \right) \leq e^{-\frac{\alpha^2}{6} j^2 \log n} \quad (10)$$

Next, substituting (9) and (10) into (8), the result follows after simple calculations and by suitably choosing α . \square

⁴A folklore example with complete proofs can be found at <http://csttheory.stackexchange.com/questions/14471/reverse-chernoff>

C Convergence Time of 3-Majority with Adversary

In this section we give the formal definitions of the notions introduced in Section 4.2 and we provide the proof of the lemmas stated in the proof outline of Theorem 4.2.

The actions of the dynamic adversary over the studied process can be formalized as follows.

Definition C.1. *At the end of every round t , after nodes have updated their opinions (i.e. once the configuration $\mathbf{C}^{(t)} = \mathbf{c}^{(t)}$ is realized), the F -dynamic adversary looks at the current opinion configuration and replaces the opinion of up to F nodes with any opinion in Σ . We define $\tilde{\mathbf{C}}^{(t)}$ as the configuration that results from the adversary's action on $\mathbf{c}^{(t)}$ and $D_i^{(t)} = D_i^{(t)}(\mathbf{c}^{(0)}, \tilde{\mathbf{c}}^{(0)}, \dots, \mathbf{c}^{(t-1)}, \tilde{\mathbf{c}}^{(t-1)}, \mathbf{c}^{(t)})$ as the r.v. corresponding to the number of nodes that the adversary adds or removes from c_i (note that $\sum_{i \in \Sigma} |D_i| \leq 2F$) at the end of the t -th round, based on all the past history of the process, i.e.*

$$\tilde{\mathbf{C}}^{(t)} = \left(C_1^{(t)} + D_1^{(t)}, \dots, C_{|\Sigma|}^{(t)} + D_{|\Sigma|}^{(t)} \right)$$

We consider an F -dynamic adversary with $F \leq \beta\sqrt{n}/(k^{\frac{5}{2}} \log n)$ for a suitable positive constant β . As we will show in the proof of Lemma C.7, this bound on F turns out to be almost tight for guaranteeing that the process converges to an almost-consensus regime in polynomial time, w.h.p.

The presence of the adversary requires us to distinguish between valid and non valid opinions. So, we recall that the set of valid opinions $\mathcal{C} \subseteq \Sigma$ is the subset of active opinions in the initial configuration and we observe that, in the remainder of this section, k denotes the number of valid opinions, i.e., $k := |\mathcal{C}|$. As discussed in the overview of the process analysis (Section 2), the adversary can introduce “small” non-valid opinions and it can keep active small valid opinions that would otherwise disappear (as shown in Section 3). These facts lead us to the problem of managing “small” opinions, whose formal definition can be given as follows.

Definition C.2. *Let $S := \{i \mid c_i \leq \gamma\sqrt{n}/k^{\frac{3}{2}}\}$ be the set of the small opinions, where γ is some constant such that $\gamma > \beta$, and let its complement $B := \bar{S} = \{i \mid c_i > \gamma\sqrt{n}/k^{\frac{3}{2}}\}$ be the set of the big opinions.*

We now re-state Theorem 4.2, whose idea of proof is outlined in Section 4.2.

Theorem 4.2. *Let $k \leq n^\alpha$ for some constant $\alpha < 1$ and $F = \beta\sqrt{n}/(k^{\frac{5}{2}} \log n)$ for some constant $\beta > 0$. Starting from any initial configuration having k opinions, the 3-majority reaches a (valid) stabilizing almost-consensus in presence of any F -dynamic adversary within $\mathcal{O}((k^2\sqrt{\log n} + k \log n)(k + \log n))$ rounds, w.h.p.*

Sketch of proof. From Lemmas C.7 and C.8 it follows that from any configuration with $j \leq k$ active opinions, within $\mathcal{O}(k^2\sqrt{\log n} + k \log n)$ rounds at least one of the opinions becomes small with probability at least $1/2$. Thus within $\mathcal{O}((k^2\sqrt{\log n} + k \log n)(k + \log n))$ rounds all opinions but one become small w.h.p. \square

In order to prove Lemma C.7 and Lemma C.8, we need the following three lemmas.

Lemma C.3. *Let $\tilde{\mathbf{c}}$ be any configuration such that $|B| \leq j$ and $\sum_{i \in \bar{C}} \tilde{c}_i^{(t)} \leq \gamma\sqrt{n}/k^{\frac{3}{2}}$. For some constant $\alpha > 0$, for any opinion i such that $\tilde{c}_i \geq \gamma\sqrt{n}/k^{\frac{3}{2}}$, it holds*

$$\mathbf{E} \left[C_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] \leq \tilde{c}_i \left(1 - \frac{1}{j} + \frac{\tilde{c}_i + \alpha\sqrt{n/k}}{n} \right) \quad (11)$$

$$\mathbf{E} \left[\tilde{C}_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] \leq \tilde{c}_i \left(1 - \min \left\{ \frac{1}{j} - \frac{\tilde{c}_i + \alpha\sqrt{n/k}}{n}, \frac{1}{2} \left(\frac{1}{j} - \frac{\tilde{c}_i}{n} \right) \right\} \right) \quad (12)$$

Proof. Similarly to the proof of Lemma 2.1 we have

$$\begin{aligned}
\mathbf{E} \left[C_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] &\leq \tilde{c}_i \left(1 + \frac{\tilde{c}_i}{n} - \frac{\sum_j \tilde{c}_j^2}{n^2} \right) \leq \tilde{c}_i \left(1 + \frac{\tilde{c}_i}{n} - \frac{\sum_{j \in B} \tilde{c}_j^2}{n^2} \right) \\
&\leq \tilde{c}_i \left(1 + \frac{\tilde{c}_i}{n} - \frac{\sum_{j \in B} \left(\frac{n - (k-j+1)\gamma\sqrt{n/k}^{\frac{3}{2}}}{j} \right)^2}{n^2} \right) \\
&\leq \tilde{c}_i \left(1 + \frac{\tilde{c}_i}{n} - \frac{\sum_{j \in B} (n - \alpha/4\sqrt{n/k})^2}{j^2 n^2} \right) \\
&\leq \tilde{c}_i \left(1 + \frac{\tilde{c}_i}{n} - \frac{1}{j} + \frac{\alpha/2\sqrt{n/k}}{jn} \right) \leq \tilde{c}_i \left(1 - \frac{n/j - \tilde{c}_i - \alpha/2\sqrt{n/k}}{n} \right)
\end{aligned}$$

Taking into account any possible action of the adversary, we thus get that

$$\begin{aligned}
\mathbf{E} \left[\tilde{C}_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] &= \mathbf{E} \left[C_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] + \mathbf{E} \left[D_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] \\
&\leq \tilde{c}_i \left(1 - \frac{n/j - \tilde{c}_i - \alpha/2\sqrt{n/k}}{n} \right) + F \\
&\leq \tilde{c}_i \left(1 - \frac{n/j - \tilde{c}_i}{n} + \frac{2 \max \left\{ \alpha/2\sqrt{n/k}, Fn/\tilde{c}_i \right\}}{n} \right). \tag{13}
\end{aligned}$$

By distinguishing the cases $\tilde{c}_i \geq n/(3j)$ or $\tilde{c}_i < n/(3j)$, from (13) we get (12). \square

In Lemma C.4 we prove the following key-properties of the process in the presence of the dynamic adversary: w.h.p., it is *never* the case that

1. if in a given round a valid opinion is small then it gets big at a later time, i.e. $S^{(t-1)} \subseteq S^{(t)}$;
2. the size of the overall set of non valid opinions grows beyond $\gamma\sqrt{n}/k^{\frac{3}{2}}$, i.e. $\sum_{i \in \bar{C}} c_i \leq \gamma\sqrt{n}/k^{\frac{3}{2}}$.

Lemma C.4. *If $\tilde{c}^{(t)}$ is such that $\sum_{i \in \bar{C}} \tilde{c}_i^{(t)} \leq \gamma\sqrt{n}/k^{\frac{3}{2}}$, then $\sum_{i \in \bar{C}} \tilde{C}_i^{(t+1)} \leq \gamma\sqrt{n}/k^{\frac{3}{2}}$ and $S^{(t)} \subseteq S^{(t+1)}$, w.h.p.*

Proof. From Lemma C.3, for each $i \in S^{(t)}$ we have that

$$\mathbf{E} \left[C_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] \leq \tilde{c}_i \left(1 + \frac{\tilde{c}_i}{n} - \frac{1}{k} \right).$$

From a direct application of the Chernoff bound to $C_i^{(t+1)}$, and taking into account any possible action of the adversary, we thus get that w.h.p.

$$\tilde{C}_i^{(t+1)} = C_i^{(t+1)} + D_i^{(t+1)} \leq \gamma \frac{\sqrt{n}}{k^{\frac{3}{2}}} \left(1 - \frac{1}{4k} \right) + F \leq \gamma \frac{\sqrt{n}}{k^{\frac{3}{2}}},$$

that is, $i \in S^{(t)}$ w.h.p. Analogously, we have

$$\mathbf{E} \left[\sum_{i \in \bar{C}^{(t)}} C_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] \leq \sum_{i \in \bar{C}} \tilde{c}_i^{(t)} \left(1 + \frac{\tilde{c}_i}{n} - \frac{1}{k} \right) \leq \gamma \frac{\sqrt{n}}{k^{\frac{3}{2}}} \left(1 - \frac{1}{2k} \right),$$

and then, by applying the Chernoff bound, we get that w.h.p.

$$\sum_{i \in \bar{C}^{(t)}} \tilde{C}_i^{(t+1)} = \sum_{i \in \bar{C}^{(t)}} C_i^{(t+1)} + \sum_i D_i^{(t+1)} \leq \gamma \frac{\sqrt{n}}{k^{\frac{3}{2}}} \left(1 - \frac{1}{4k}\right) + F \leq \gamma \frac{\sqrt{n}}{k^{\frac{3}{2}}},$$

concluding the proof. \square

Lemma C.5. *Let $\tilde{\mathbf{c}}$ be any configuration such that $|B| = j$ and $\sum_{i \in \bar{C}} \tilde{c}_i \leq \gamma \sqrt{n}/k^{\frac{3}{2}}$. Consider the stochastic process $\{\tilde{Y}_t\}_t$ defined as $\tilde{Y}_t = \lfloor \frac{n}{j} \rfloor - \tilde{C}_m^{(t)}$ and define the stopping time*

$$\tau = \inf\{t \in \mathbb{N} : \tilde{Y}_t \geq \sqrt{jn \log n} \vee \left(\sum_{i \in \bar{C}} \tilde{C}_i \geq \gamma \sqrt{n}/k^{\frac{3}{2}}\right) \vee (S^{(t-1)} \not\subseteq S^{(t)})\}.$$

Then, it holds that

$$\mathbf{P}_{\mathbf{c}} \left(\tilde{Y}_\tau > \alpha \sqrt{jn \log n} \right) \leq \frac{1}{n}.$$

Sketch of proof. The proof of this Lemma follows from minor modifications of the proof of Lemma B.1. In particular, the argument is based on the following observations:

1. The event defining the stopping time τ is in this case

$$\mathcal{E}^{(t)} = \left(\tilde{Y}_t \geq \sqrt{jn \log n} \right) \vee \left(\sum_{i \in \bar{C}} \tilde{C}_i \geq \gamma \sqrt{n}/k^{\frac{3}{2}} \right) \vee (S^{(t-1)} \not\subseteq S^{(t)}).$$

The negated of this event is

$$\neg \mathcal{E}^{(t)} = \left(\tilde{Y}_t \leq \sqrt{jn \log n} \right) \wedge \left(\sum_{i \in \bar{C}} \tilde{C}_i \leq \gamma \sqrt{n}/k^{\frac{3}{2}} \right) \wedge (S^{(t-1)} \subseteq S^{(t)}),$$

which implies the event $\left(\tilde{Y}_t \leq \sqrt{jn \log n} \right)$.

2. Proceeding like in the proof of Lemma B.1, we can write an expression that is similar to (8), with the generic conditioning event

$$C_m^{(s)} > \left\lfloor \frac{n}{j} \right\rfloor - \sqrt{jn \log n},$$

replaced by $\neg \mathcal{E}^{(s)}$. The conditioned event

$$C_\ell^{(t)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n},$$

is instead replaced by the event

$$\left(C_\ell^{(t)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n} \right) \wedge \mathcal{E}^{(t)}.$$

Now, note that the event

$$C_\ell^{(t)} < \left\lfloor \frac{n}{j} \right\rfloor - \alpha \sqrt{jn \log n},$$

again implies $\mathcal{E}^{(t)}$. Hence, we can still write (8), from which the proof requires some non-hard adaptations w.r.t. the case without adversary. \square

Since the adversary, at time t , may decide what to do based on the full history of the process up to time t , the stochastic process $\{\tilde{\mathbf{C}}^{(t)}\}_t$ may not be a Markov process anymore. Thus, we need a more general version of Lemma 3.2.

Lemma C.6. *Let $\{X_t\}_t$ be a discrete time stochastic process with a finite state space Ω , let $f_t : \Omega^t \rightarrow \mathbb{N}$ be a function mapping histories of the process in non-negative integer numbers, and let $\{Y_t\}_t$ be the stochastic process over \mathbb{N} defined by $Y_t = f_t(X_0, \dots, X_t)$. Let $m \in \mathbb{N}$ be a “target value”, let $A \subseteq \Omega$ be an arbitrary subset of states, and let*

$$\tau = \inf\{t \in \mathbb{N} : Y_t \geq m \text{ or } X_t \notin A\}$$

be the random variable indicating the first time X_t exits from set A or Y_t reaches or exceeds value m . Assume that, for every sequence of states $x_0, \dots, x_t \in A$ with $f_t(x_0, \dots, x_t) \leq m - 1$, it holds that

1. (Positive drift). $\mathbf{E}[Y_{t+1} | X_0 = x_0, \dots, X_t = x_t] \geq f_t(x_0, \dots, x_t) + \lambda$ for some $\lambda > 0$
2. (Bounded jumps). $\mathbf{P}_x(Y_\tau \geq \alpha m) \leq \alpha m/n$, for some $\alpha > 1$.

Then, for every starting state $x \in A$, it holds that

$$\mathbf{E}_x[\tau] \leq 2\alpha \frac{m}{\lambda}$$

Proof. The proof is a straight adaptation of the proof of Lemma 3.2, in which we take into account the full history of the process.

Consider the stochastic process $Z_t = Y_t - \lambda t$. For any sequence of states $x_0, \dots, x_t \in A$ with $f_t(x_0, \dots, x_t) \leq m - 1$ it holds that

$$\begin{aligned} \mathbf{E}[Z_{t+1} | X_0 = x_0, \dots, X_t = x_t] &= \mathbf{E}[Y_{t+1} | X_0 = x_0, \dots, X_t = x_t] - \lambda(t+1) \\ &\geq f_t(x_0, \dots, x_t) + \lambda - \lambda(t+1) \\ &\geq f_t(x_0, \dots, x_t) - \lambda t \end{aligned}$$

where in the inequality we used Hypotheses 1. Thus, Z_t is a *submartingale* up to the stopping time τ . Moreover, since $|Y_t| \leq n$ then $|Z_{t+1} - Z_t| \leq n + \lambda$ and, together with Hypotheses 1 this implies $\mathbf{E}_x[\tau] < \infty$. Thus, we can apply *Doob's Optional Stopping Theorem* [13]. It follows that $\mathbf{E}_x[Z_\tau] \geq \mathbf{E}_x[Z_0] = f_0(x)$ and, since $\mathbf{E}_x[Z_\tau] = \mathbf{E}_x[Y_\tau] - \lambda \mathbf{E}_x[\tau]$, we have that

$$\mathbf{E}_x[\tau] \leq \frac{\mathbf{E}_x[Y_\tau] - f_0(x)}{\lambda} \leq \frac{\mathbf{E}_x[Y_\tau]}{\lambda}$$

Finally, we get

$$\mathbf{E}_0[Y_\tau] = \sum_{j=1}^{\lfloor \alpha m \rfloor} j \mathbf{P}_0(Y_\tau = j) + \sum_{j=\lfloor \alpha m \rfloor + 1}^n j \mathbf{P}_0(Y_\tau = j) \leq (\alpha m) + n \mathbf{P}_0(Y_\tau > \alpha m) \leq 2(\alpha m)$$

where in the last inequality we used Hypothesis 2. □

Now, we can exploit Lemma C.6 to bound the time required by the symmetry-breaking stage: We show that, from any configuration with j big opinions that satisfies condition \mathcal{H} , within $\mathcal{O}(j^2 \log^{1/2} n)$ rounds there exists a opinion supported by at most $n/j - \sqrt{jn \log n}$ nodes with probability at least $1/2$.

Lemma C.7 (Symmetry-breaking stage). *Let $\tilde{\mathbf{c}}$ be any configuration such that $|B| = j$ and $\sum_{i \in \bar{C}} \tilde{c}_i \leq \gamma\sqrt{n}/k^{\frac{3}{2}}$. Within $t = \mathcal{O}\left(j^2 \log^{1/2} n\right)$ rounds, with probability at least $1/2$ it holds that*

$$|B| = j, \quad \sum_{i \in \bar{C}} \tilde{C}_i \leq \gamma\sqrt{n}/k^{\frac{3}{2}} \quad \text{and} \quad \exists i \in B^{(t)} \quad \text{such that} \quad \tilde{C}_i^{(t)} \leq n/j - \sqrt{jn \log n}$$

Proof. We proceed by adapting the proof of Lemma 3.3. Let $\tilde{\mathbf{C}}^{(0)} = \tilde{\mathbf{c}}$ be the initial configuration. Let us consider the stochastic process $\{\tilde{Y}_t\}_{t \geq 0}$ defined as

$$\tilde{Y}_t = \left\lfloor \frac{n}{j} \right\rfloor - \tilde{C}_m^{(t)}$$

where $\tilde{C}_m^{(t)} = \min\{\tilde{C}_i^{(t)} : i \in B^{(t)}\}$. We are interested in the time step

$$\tau = \inf\{t \in \mathbb{N} : \tilde{Y}_t \geq (\sqrt{jn \log n}) \vee \left(\sum_{i \in \bar{C}} \tilde{C}_i \geq \gamma\sqrt{n}/k^{\frac{3}{2}}\right) \vee (S^{(t-1)} \not\subseteq S^{(t)})\}$$

Now we show that $\{\tilde{Y}_t\}_t$ satisfies the Hypotheses 1 and 2 of Lemma 3.2 with $A = \left(\sum_{i \in \bar{C}} \tilde{C}_i \leq \gamma\sqrt{n}/k^{\frac{3}{2}}\right) \vee (S^{(t-1)} \subseteq S^{(t)})$ and $\lambda = \varepsilon\sqrt{n}/j^{3/2}$, for a suitable constant $\varepsilon > \alpha$.

1. Let $\tilde{\mathbf{c}}$ be any configuration such that $\tilde{c}_m > n/j - \sqrt{jn \log n}$. Now we prove that

$$\mathbf{E} \left[\tilde{C}_m^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] \leq \tilde{c}_m - \varepsilon \frac{\sqrt{n}}{j^{3/2}} \quad (14)$$

Case $\tilde{c}_m > n/j - 2\varepsilon\sqrt{n/j}$: Observe that, in this case, random variables $\{C_i^{t+1} : i \in B\}$ have standard deviation is $\Omega(\sqrt{n/j})$. Moreover they are binomial and negatively associated. Hence, by choosing ε small enough, from the Central Limit Theorem we have that

$$\mathbf{P} \left(i \in B \text{ exists such that } C_i^{(t+1)} \leq \frac{n}{j} - 6\varepsilon \cdot \sqrt{\frac{n}{j}} \right) \geq 1/2$$

We thus get

$$\begin{aligned} \mathbf{E} \left[\tilde{C}_m^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] &\leq \frac{1}{2} \left(\frac{n}{j} - 6\varepsilon \cdot \sqrt{\frac{n}{j}} \right) + \frac{1}{2} \cdot \frac{n}{j} + \frac{\beta\sqrt{n}}{k^{\frac{5}{2}} \log n} \\ &= \frac{n}{j} - 2\varepsilon\sqrt{\frac{n}{j}} + \frac{\beta\sqrt{n}}{k^{\frac{5}{2}} \log n} \leq \tilde{c}_m - \varepsilon\sqrt{\frac{n}{j}} \leq \tilde{c}_m - \varepsilon \frac{\sqrt{n}}{j^{3/2}} \end{aligned} \quad (15)$$

Case $\tilde{c}_m \leq n/j - 2\varepsilon\sqrt{n/j}$: Equation (14) easily follows from Lemma C.3. Indeed, let $i \in B$ be a opinion such that $\tilde{c}_i = \tilde{c}_m$, then

$$\begin{aligned} \mathbf{E} \left[\tilde{C}_m^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] &\leq \mathbf{E} \left[\tilde{C}_i^{(t+1)} \mid \tilde{\mathbf{C}}^{(t)} = \tilde{\mathbf{c}} \right] \leq \tilde{c}_i \left(1 + \frac{\tilde{c}_i + \alpha\sqrt{n/k}}{n} - \frac{1}{j} \right) \\ &\leq \tilde{c}_i \left(1 - \frac{2\varepsilon}{\sqrt{nj}} + \frac{\alpha}{\sqrt{kn}} \right) \leq \tilde{c}_i - \frac{\varepsilon\sqrt{n}}{j^{3/2}} = \tilde{c}_m - \varepsilon \frac{\sqrt{n}}{j^{3/2}} \end{aligned}$$

where we used the case's condition and $\tilde{c}_i = \tilde{c}_m \geq n/(2j)$.

2. Since random variables $\{\tilde{C}_i^{(t)} : i \in B^{(t)}\}$ are binomial conditional on the configuration at round $t-1$, from the Chernoff bound it follows that

$$\mathbf{P}_{\tilde{\mathbf{c}}} \left(\tilde{Y}_\tau \geq \alpha\sqrt{jn \log n} \right) \leq \frac{1}{n}, \quad \text{for some constant } \alpha > 1 \quad (16)$$

See Lemma C.5 for the formal statement of the last fact.

From (14) and (16) we have that $\{\tilde{Y}_t\}_t$ satisfies the hypotheses of Lemma C.6 with $m = \sqrt{jn \log n}$, $\lambda = \varepsilon \sqrt{n}/j^{3/2}$ and $A = (\sum_{i \in \bar{C}} \tilde{C}_i \leq \gamma \sqrt{n}/k^{3/2}) \vee (S^{(t-1)} \subseteq S^{(t)})$. Moreover, by iteratively applying Lemma C.4, we have that, for any $t = \mathcal{O}(n^2)$, it holds w.h.p. that $(\sum_{i \in \bar{C}} \tilde{C}_i^{(t)} \leq \gamma \sqrt{n}/k^{3/2}) \vee (S^{(t-1)} \subseteq S^{(t)})$. Thus, from Markov's inequality, for $t = 2j^2 \sqrt{\log n}$, we have that

$$\begin{aligned} \mathbf{P}_{\bar{c}}(\forall i \in B, (C_i^{(t)} \leq n/j - \sqrt{jn \log n}) \wedge (\sum_{i \in \bar{C}} \tilde{C}_i^{(t)} \leq \gamma \sqrt{n}/k^{3/2}) \wedge (S^{(0)} \subseteq S^{(t)})) \\ \geq \mathbf{P}_{\bar{c}}(\hat{\tau} \leq 2j^2 \sqrt{\log n}) \geq \frac{1}{3} \end{aligned}$$

where $\hat{\tau} = \inf\{t \in \mathbb{N} : \tilde{Y}_t \geq \sqrt{jn \log n}\}$. □

Once the smallest opinion goes below the average size of the big opinions by a certain small amount, we can prove that the process push it in the set of small opinions w.h.p.

Lemma C.8 (Dropping stage). *Assume that, at round t' , $\tilde{c}^{(t')}$ is such that $\sum_{i \in \bar{C}} c_i \leq \gamma \sqrt{n}/k^{3/2}$, $|B^{(t')}| = j$, and an $i \in B^{(t')}$ exists such that $\gamma \sqrt{n}/k^{3/2} \leq c_i^{(t')} \leq n/j - \sqrt{kn \log n}$. Then, w.h.p., a round $t'' = t' + \mathcal{O}(k \log n)$ exists such that $\sum_{i \in \bar{C}} \tilde{C}_i^{(t'')} \leq \gamma \sqrt{n}/k^{3/2}$, $i \in S^{(t'')}$ and $|B^{(t'')}| \leq j - 1$.*

Proof. By iteratively applying Lemma C.4, we have that, w.h.p., for each $t \in \{t', \dots, t'' - 1\}$ it holds $\sum_{i \in \bar{C}} \tilde{C}_i^{(t)} \leq \gamma \sqrt{n}/k^{3/2}$ and $i \in S^{(t)}$.

To prove that $|B^{(t'')}| \leq j - 1$, we first prove that, for each round $t \in \{t' + 1, \dots, t''\}$, w.h.p. $\tilde{C}_i^{(t)}$ decreases by a certain extent that depends on $\tilde{c}_i^{(t-1)}$, regardless of what the adversary does.

Let $\psi = (1/j - (\tilde{c}_i^{(t-1)} + \alpha \sqrt{n/k})/n)$. If we are in a configuration satisfying the hypotheses of the lemma, we have

$$\mathbf{P}\left(C_i^{(t)} > \tilde{c}_i^{(t-1)} \left(1 - \frac{\psi}{2}\right)\right) = \mathbf{P}\left(C_i^{(t)} > \tilde{c}_i^{(t-1)} (1 - \psi(1 + \delta))\right),$$

where $\delta = \frac{1}{2}\psi/(1 - \psi)$. Thus, using Lemma C.3 and applying the Chernoff bound we have

$$\mathbf{P}\left(C_i^{(t)} > \tilde{c}_i^{(t-1)} \left(1 - \frac{\psi}{2}\right)\right) \leq \exp\left\{-\frac{\delta^2}{3}\psi \tilde{c}_i^{(t-1)}\right\} < \exp\left\{-\frac{1}{3}\left(\frac{1}{2}\psi\right)^2 \tilde{c}_i^{(t-1)}\right\} = n^{-\Theta(1)}, \quad (17)$$

where the second inequality follows from the definition of δ and the fact that its denominator is smaller than 1, and the equality in (17) follows by minimizing $\psi^2 \tilde{c}_i^{(t-1)}$ for $\gamma \sqrt{n}/k^{3/2} \leq c_i^{(t')} \leq n/j - \sqrt{kn \log n}$.

It follows that, w.h.p.

$$\tilde{C}_i^{(t)} = C_i^{(t)} + D_i^{(t)} \leq \tilde{c}_i^{(t-1)} \left(1 - \frac{\psi}{2}\right) + F \leq \tilde{c}_i^{(t-1)} \quad (18)$$

Thus, w.h.p., we can iteratively apply (18) until $\tilde{c}_i^{(t-1)} \leq \gamma \sqrt{n}/k^{3/2}$. We next prove that this happens within $\mathcal{O}(k \log n)$ rounds, w.h.p. Interestingly, showing that, within $\mathcal{O}(k \log n)$ rounds, C_i decreases to a constant fraction of its value at the beginning of the dropping stage

does not seem obvious. For this reason, we consider the evolution of the displacement $\frac{n}{j} - C_i$, which seems analytically more tractable. To this purpose, note that (17) implies that, w.h.p.

$$\begin{aligned}
\frac{n}{j} - C_i^{(t)} &\geq \frac{n}{j} - c_i^{(t-1)} + \frac{c_i^{(t-1)}}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)} + \alpha\sqrt{n/k}}{n} \right) \\
&= \frac{n}{j} - c_i^{(t-1)} + \frac{c_i^{(t-1)}}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \right) \left(1 - \frac{\alpha\sqrt{n/k}}{1/j - c_i^{(t-1)}/n} \right) \\
&= \frac{n}{j} - c_i^{(t-1)} + \frac{c_i^{(t-1)}}{2} \left(\frac{1}{j} - \frac{c_i^{(t-1)}}{n} \left(1 + \frac{\alpha}{\log n} \right) \right) \\
&= \left(\frac{n}{j} - c_i^{(t-1)} \right) \left(1 + \alpha_1 \frac{c_i^{(t-1)}}{2n} \right)
\end{aligned} \tag{19}$$

for some constant $\alpha_1 > 0$, where in the first equality of (19) we have used that $n/j - c_i^{(t-1)} \geq \sqrt{kn \log n}$.

We can now conclude the proof of Lemma C.8. We first prove that $C_i \leq n/(2j)$ within $\mathcal{O}(k \log n)$ steps, w.h.p. To this purpose, note that $\frac{n}{j} - c_i \geq \sqrt{kn \log n}$ at the beginning of the droppign stage from the hypotheses. Furthermore, for some positive constants α_2 and α_3 , as long as $C_i \geq \alpha_3 n/j$ it holds $1 + \alpha_1 c_i/n \geq 1 + \alpha_2/j$. Hence, after $\mathcal{O}(k \log n)$ steps, w.h.p. we have $\frac{n}{j} - c_i \geq (1 - \alpha_3) \frac{n}{j}$, which in turn implies $c_i \leq \alpha_3 n/j$. Once $c_i \leq \alpha_3 n/j$, using again (19) we have that C_i decreases by a factor $1 - \Omega(1/j)$ in every round w.h.p. By standard concentration arguments we obtain that eventually $c_i \leq \gamma \sqrt{n}/k^{\frac{3}{2}}$ within $\mathcal{O}(k \log n)$ more steps, w.h.p. \square