Undecided State Dynamics with Many Opinions

Colin Cooper
King's College London
London, United Kingdom
colin.cooper@kcl.ac.uk

Frederik Mallmann-Trenn
King's College London
London, United Kingdom
Institute of Science and Technology Austria (ISTA)
Klosterneuburg, Austria
frederik.mallmann-trenn@kcl.ac.uk

Tomasz Radzik
King's College London
London, United Kingdom
tomasz.radzik@kcl.ac.uk

Nobutaka Shimizu
Institute of Science Tokyo
Tokyo, Japan
shimizu.n.ah@m.titech.ac.jp

Takeharu Shiraga
Chuo University
Tokyo, Japan
shiraga.076@g.chuo-u.ac.jp

Abstract

We study the Undecided-State Dynamics (USD), a fundamental consensus process in which each vertex holds one of k decided opinions or the undecided state. We consider both the gossip model and the population protocol model. Prior work established tight bounds on the consensus time of this process only for the regime $k = O(\sqrt{n}/(\log n)^2)$ (for the population protocol model) and $k = O((n/\log n)^{1/3})$ (for the gossip model), often under restrictive assumptions on the initial configuration.

In this paper, we obtain the first consensus-time guarantees for USD that hold for *arbitrary* $2 \le k \le n$ and for *arbitrary* initial configurations in both the gossip model and the population protocol model. In the gossip model, USD reaches consensus within $\widetilde{O}(\min\{k, \sqrt{n}\})$ synchronous rounds with probability $1 - p_\perp - n^{-c}$, where $p_\perp$ is the gossip-specific probability of collapsing to the all-undecided state in the first round. In the population protocol model, USD reaches consensus within $\widetilde{O}(\min\{kn, n^{3/2}\})$ asynchronous interactions with high probability. We also present lower bounds that match the upper bounds up to polylogarithmic factors for a specific initial configuration and show that our upper bounds are essentially optimal.

CCS Concepts

• **Theory of computation → Distributed algorithms**.

Keywords

consensus dynamics, undecided state dynamics, gossip model, population protocol model

ACM Reference Format:
Colin Cooper, Frederik Mallmann-Trenn, Tomasz Radzik, Nobutaka Shimizu, and Takeharu Shiraga. 2026. Undecided State Dynamics with Many Opinions. In *ACM Symposium on Principles of Distributed Computing (PODC '26), July 6–10, 2026, Egham, United Kingdom.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3796701.3815920

ACM ISBN 979-8-4007-2512-8/26/07
https://doi.org/10.1145/3796701.3815920

1 Introduction

The Undecided-State Dynamics (USD) [3, 6] is one of the simplest and most studied consensus dynamics on complete graphs. Each vertex holds an opinion in a finite set $\Sigma = [k] \cup \{\perp\}$, where $[k] := \{1, \dots, k\}$ are decided opinions and $\perp$ is a distinguished undecided opinion. When a vertex u interacts with a vertex v, the update follows a simple cancellation/adoption rule (with only u updating its opinion): if u and v hold different decided opinions in $[k]$, then u becomes $\perp$; if u is undecided and v is decided, then u adopts v's opinion; otherwise u keeps its opinion. This elementary rule (formalized in Section 3.1) already gives rise to rich and nontrivial behavior.

We analyze USD with k possible decided opinions on an n-vertex complete graph with self-loops[1] under two classical communication models. In the *gossip model* [28], time proceeds in synchronous rounds in which each vertex samples one random neighbor and simultaneously updates its opinion accordingly. In the *population protocol model* [2], interactions occur asynchronously: an ordered pair (u, v) is chosen uniformly at random with replacement, and u updates its opinion based on v's opinion.

A central quantity of interest is the *consensus time*: the number of rounds (in the gossip model) or pairwise interactions (in the population protocol model) until all vertices hold the same decided opinion in $[k]$. The all-$\perp$ configuration, although absorbing, does not constitute consensus and is regarded as failure. Consensus time therefore measures the time to reach agreement on a real opinion. In what follows, we review previous consensus time bounds on USD. For results on other relevant consensus dynamics, see Section 1.2.

USD was first introduced by Angluin et al. [3] for $k = 2$ in the population protocol model, where they showed that the consensus time is $O(n \log n)$ with high probability.[2] They (and Condon et al. [13]) further proved that USD with $k = 2$ reaches consensus with the initial majority with high probability if the more popular opinion has a sufficiently large initial advantage compared to the other opinions. Thus USD offered a critical performance improvement over the classic pull voting, both in terms of the consensus time (expected $\Theta(n^2)$ in pull voting [16, 26]) and the chance for the

[1] Thus, choosing a random neighbor is equivalent to choosing a vertex uniformly at random.
[2] A given event holds "with high probability," if it holds with probability $1 - O(n^{-c})$ for some constant $c > 0$.

majority opinion to win. For larger k, Amir et al. [1] analyzed USD for $2 \le k = O(\sqrt{n}/(\log n)^2)$ and proved that the consensus time is $O(kn \log n)$ with high probability for any initial configuration. Recently, El-Hayek et al. [20] obtained a matching lower bound of $\Omega(kn \log n)$ for all $k \le n^{1/2-\varepsilon}$, where ε is an arbitrary positive constant.

In the gossip model, USD was first analyzed by Becchetti et al. [6], who proved an $O(k \log n)$ upper bound for $2 \le k = O((n/\log n)^{1/3})$ under a strong assumption on the initial advantage of the most popular opinion. This assumption was later removed for $k = 2$ by Clementi et al. [12], who showed an $O(\log n)$ bound for arbitrary initial configurations. However, no general bound was known for any $k > 2$ without any assumptions on the initial configuration.

It is natural to conjecture that the consensus time in the gossip model should be $\widetilde{O}(k)$[3] for $2 \le k \ll \sqrt{n}$,[4] by comparison with the $\widetilde{O}(kn)$ upper bound of [1] for the population protocol model: a single synchronous round of the gossip model is roughly equivalent to n asynchronous interactions in the population protocol model. However, the analysis of [1] critically exploits the asynchronous nature of population protocols and does not transfer to the gossip model. Indeed, they explicitly identify the gossip-model bound as an open problem [1, Section 9].

Moreover, prior work provides a rather clear picture only in the regime $k \ll \sqrt{n}$: In the population protocol model, the best-known upper and lower bounds are $\widetilde{\Theta}(kn)$ in this range [1, 20], strongly suggesting that the consensus time scales linearly with k. In the gossip model, the only existing upper bounds for USD with $k > 2$ are also of order $\widetilde{O}(k)$, but they hold only under strong assumptions on the initial configuration and only for $k \ll n^{1/3}$ [6]. Beyond this threshold, however, the efficiency of USD is mysterious: the analytical techniques used before do not work for larger values of k and no general consensus-time bounds are available for any $k \ge \sqrt{n}$ in the population protocol model or any $k \ge n^{1/3}$ in the gossip model. This raises a natural and central question: does the linear-in-k behavior of the consensus time persist for all k, or does USD transition to a qualitatively different—possibly sublinear—regime once k exceeds $\sqrt{n}$? Addressing this question requires techniques that go well beyond those used in the $k < \sqrt{n}$ regime, and lies at the core of our contribution.

1.1 Our Contributions

We obtain the first general bounds on the consensus time of the Undecided-State Dynamics (USD) for an *arbitrary* number k of decided opinions and an *arbitrary* initial configuration in both the gossip model and the population protocol model. Our main results are summarized in the following informal theorem. See also Table 1.

Theorem 1.1 (Main Theorem). *Consider USD on n vertices starting from an arbitrary initial configuration[5] in $\{1, \dots, k, \perp\}^n$ other than the all-$\perp$ configuration, where the number of possible decided opinions k is an arbitrary integer between 2 and n (inclusive).*

- ***Gossip model.*** *In the gossip model, USD reaches consensus within $\widetilde{O}(\min\{k, \sqrt{n}\})$ synchronous rounds with probability $1-p_\perp-n^{-c}$ for some constant $c > 0$, where $p_\perp$ is the probability that all vertices hold $\perp$ after the first synchronous round.*
- ***Population protocol model.*** *In the population protocol model, USD reaches consensus within $\widetilde{O}(\min\{kn, n^{1.5}\})$ asynchronous interactions with high probability.*

[3]Throughout the paper, $\widetilde{O}(\cdot)$ hides polylogarithmic factors in n.
[4]"$k \ll f(n)$" abbreviates that $k = O(f(n)/\log^c n)$ for some constant $c > 0$.
[5]A configuration is an assignment of opinions to vertices.

Model	Work	Consensus Time	Range of k
PP	[3] [13]	$O(n \log n)^*$	$k = 2$
	[1]	$O(kn \log n)$	$k = O(\sqrt{n}/(\log n)^2)$
	Theorem 1.1	$O(n^{1.5}(\log n)^3)$	$2 \le k \le n$
Gossip	[6]	$O(k \log n)^*$	$k = O((n/\log n)^{1/3})$
	[12]	$O(\log n)$	$k = 2$
	Theorem 1.1	$O(k \log n)$	$k = O(\sqrt{n}/(\log n)^2)$
	Theorem 1.1	$O(\sqrt{n}(\log n)^3)$	$2 \le k \le n$

Table 1: Comparison of consensus time bounds for USD in the Population Protocol (PP) and Gossip models. An asterisk (∗) attached to a consensus-time bound indicates that the bound holds under an initial-gap assumption.

Our result for the gossip model says that, unless the process collapses to the all-$\perp$ configuration in the first round, it reaches consensus within $\widetilde{O}(\min\{k, \sqrt{n}\})$ rounds with high probability.

Regarding the hidden polylog(n) factor in the consensus time, we have the following bounds: In the gossip model, the consensus time is, with high probability, $O(\sqrt{n}(\log n)^3)$ for arbitrary $2 \le k \le n$, and $O(k \log n)$ if $k = O\left(\frac{\sqrt{n}}{(\log n)^2}\right)$ (the similar bound holds for the population protocol model). Our results hold for every $2 \le k \le n$ and for arbitrary initial configurations, including those with undecided vertices.

Remark 1.2. *The probability $p_\perp$ of collapsing to the all-$\perp$ configuration in the first synchronous round is a phenomenon specific to the gossip model. Because all vertices update simultaneously, a single round can eliminate all decided opinions. In contrast, in the population protocol model only one vertex updates its opinion at each interaction, and an undecided vertex cannot be created unless two distinct decided opinions are present.*

The value of $p_\perp$ strongly depends on the initial distribution of opinions. For example, if $k = O(n/\log n)$, or if the initial configuration contains at least one but not all undecided vertices, then $p_\perp = n^{-\Omega(1)}$ (see Lemma 5.1). At the opposite extreme, if $k = n$ and each vertex initially holds a distinct decided opinion, then it becomes undecided in the first round with probability $1 - 1/n$, independently of the other vertices, and thus $p_\perp = (1-1/n)^n \approx 1/e$.

Matching lower bounds. We complement our upper bounds for the gossip model by proving a matching lower bound up to a polylogarithmic factor in n. Such a lower bound in the population protocol model is already given by the result of [20] in the regime of $k \le n^{1/2-\varepsilon}$ for any constant $\varepsilon > 0$.

Theorem 1.3 (matching lower bound for the gossip model). *For any sufficiently large n, any constant $0 < c < 1/2$ and any $2 \le k \le (1/2 - c)n$, there exists an initial configuration over exactly*

k decided opinions (i.e., each of the k opinions is supported by at least one vertex) such that USD in the gossip model requires $\widetilde{\Omega}(\min\{k, \sqrt{n}\})$ rounds to reach consensus with high probability.

1.2 Other Related Works

Previous works on USD consider the plurality consensus problem and they show that USD solves it correctly with high probability if the initially most popular opinion has a significant advantage over the others. Berenbrink et al. [11], Ghaffari and Parter [25] presented protocols based on USD in the gossip model and showed that their protocols solve the plurality consensus problem correctly with high probability within $\widetilde{O}(1)$ synchronous rounds with $|\Sigma| = O(k)$ states under the assumption that the most popular opinion has a margin of at least $\Omega(\sqrt{n \log n})$ vertices over any other opinion in the initial configuration. Recently, Bankhamer et al. [4] presented protocols having similar flavor based on USD in both the population protocol model and the gossip model. A remarkable feature of these protocols is that the consensus time is small for any $k \geq 2$ and any initial configuration, whereas the protocols of Berenbrink et al. [11], Ghaffari and Parter [25] require the initial margin assumption to bound the consensus time. There is a work on another variant of USD that deals with a preferred opinion [9] and under noisy channel [17] for $k = 2$.

A long line of works studied the consensus times of 3-Majority and related dynamics (e.g., 2-Choices) with $k > 2$ opinions on an n-vertex complete graph [7, 8, 10, 14, 21, 24, 30]. Recently, Cooper et al. [14] showed that 3-Majority in the asynchronous update model (i.e., at every round, a uniformly random vertex is allowed to update its opinion) reaches consensus within $\widetilde{O}(\min\{k, \sqrt{n}\})$ rounds. In the synchronous update model where all vertices simultaneously update their opinions at every round, Shimizu and Shiraga [30] showed that the consensus time of 2-Choices is $\widetilde{O}(k)$ and that of 3-Majority is $\widetilde{O}(\min\{k, \sqrt{n}\})$. Interestingly, this consensus time bound for 3-Majority is the same as the one for USD in the gossip model.

Several other relevant consensus dynamics have been considered in the literature. A model that it somewhat similar is the model proposed in [22], where initialy some nodes do not have an opinion, and adopt any opnion immedieatly. Afterwards, the process behaves exactly like the voter model. The deterministic majority model, where each agent sets its opinion to the majroity among all of its neighbors, has been studied in [27]. See [5] for more details on consensus dynamics.

2 Proof Overview

We outline the proof of Theorem 1.1 for the gossip model and compare it with arguments from previous work. For simplicity, we assume that $k \leq n/\log n$ throughout this section. This is only to ensure that with high probability USD does not fall into the all-$\perp$ configuration within one round; it can be relaxed to $k \leq n$ without much difficulty.

2.1 Process in a Nutshell

Following prior work on USD, we track the evolution of the number of vertices holding each decided opinion. For $t \geq 0$ and $i \in [k]$, define $\alpha_t(i)$ as the fraction of vertices holding opinion i at the beginning of round t. Then USD can be viewed as a Markov chain $(\alpha_t)_{t\geq 0}$ for $\alpha_t = (\alpha_t(1), \dots, \alpha_t(k))$ on the state space $\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}^k$. The key progress parameter is

$$\alpha_t^{\max} := \max_{i\in[k]} \alpha_t(i),$$

and the consensus is reached at the first time t when $\alpha_t^{\max} = 1$. Unless all vertices hold $\perp$, it holds that $\alpha_t^{\max} \geq 1/n$.

We also introduce the following quantities:

$$\beta_t := \sum_{i\in[k]} \alpha_t(i), \quad \psi_t := \beta_t(2\beta_t - 1) - \sum_{i\in[k]} \alpha_t(i)^2. \tag{1}$$

Note that β_t is the total fraction of vertices holding a decided opinion. Roughly speaking, ψ_t is a key parameter that controls the stability of USD. See Section 2.3 for more details.

Call an opinion i *weak at round t* if $\alpha_t(i) \leq 0.9 \cdot \alpha_t^{\max}$ and *strong at round t* if $\alpha_t(i) \geq 0.95 \cdot \alpha_t^{\max}$. Note that there is a margin between weak and strong opinions, which is crucial for our analysis. We decompose the analysis into four parts.

Lemma 2.1 (informal; see Lemmas 5.2 to 5.5, 5.9 and 5.10 for the formal statements)**.** *Consider USD in the gossip model such that $k \leq n/\log n$ and $\beta_0 \geq 1/n$.*

(I) *For some $T_\beta = O(\log n)$, with high probability, for all $t \geq T_\beta$, $\beta_t \geq 1/2 - o(\log n/\sqrt{n})$. Moreover, with high probability, for all $t \geq 1$, $\psi_t = o(\log n/\sqrt{n})$.*

(II) *Suppose that $\psi_0 = o(\log n/\sqrt{n})$ and $\beta_0 \geq 1/2 - o(\log n/\sqrt{n})$. For some $T_\alpha = O(\sqrt{n}(\log n)^3)$, with high probability, $\alpha_t^{\max} \geq \Omega\left(\frac{(\log n)^2}{\sqrt{n}}\right)$ holds for all $t \geq T_\alpha$.*

(III) *Suppose that $\psi_0 = o(\log n/\sqrt{n})$, $\beta_0 \geq 1/2 - o(\log n/\sqrt{n})$, and $\alpha_0^{\max} \geq \Omega\left(\frac{(\log n)^2}{\sqrt{n}}\right)$. Then, for some $T' = O(\log n/\alpha_0^{\max})$, there exists exactly one strong opinion at round T'.*

(IV) *Suppose that $\psi_0 = o(\log n/\sqrt{n})$, $\beta_0 \geq 1/2 - o(\log n/\sqrt{n})$, $\alpha_0^{\max} = \alpha_0(1) \geq \Omega\left(\frac{(\log n)^2}{\sqrt{n}}\right)$, and $\alpha_0(1) \geq (1 + c) \cdot \alpha_0(i)$ for some constant $c > 0$ and all $i \geq 2$. Then, USD reaches consensus within $O(\log n/\alpha_0^{\max})$ rounds with high probability.*

The four items above immediately imply that the consensus time is $O(\sqrt{n}(\log n)^3)$. Moreover, (I) ensures that with high probability, for some $t = O(\log n)$, $\beta_t = \Omega(1)$, implying that $\alpha_t^{\max} = \Omega(1/k)$. Thus, if $k \leq \sqrt{n}/(\log n)^2$, then (III) and (IV) imply the consensus time bound of $O(\log n/\alpha_0^{\max}) = O(k \log n)$. Therefore, the consensus time is $\widetilde{O}(\min\{k, \sqrt{n}\})$.

Our main technical contribution is to establish part (II). An important feature is that (II) holds regardless of how large k is, so our analysis broadens the range of k from $k \ll n^{1/3}$ (or $k \ll n^{1/2}$ for the population protocol model) in the previous works to $k \leq n/\log n$.

Establishing part (III) is also challenging. [1, Phases 1 to 3] proved (III), for $k = O(\sqrt{n}/(\log n)^2)$ for the population protocol model, but their argument crucially relies on the local update property of the population protocol model which cannot be applied to the gossip model.

Part (IV) was shown in an earlier work [6, Theorem 3.2] for $k = O((n/\log n)^{1/3})$, which (together with $\beta_0 = \Omega(1)$) implies $\alpha_0^{\max} \geq \frac{C \log^{1/3} n}{n^{1/3}}$, but we prove it under a weaker condition that

$\alpha_0^{\max} \geq \Omega\left(\frac{(\log n)^2}{\sqrt{n}}\right)$ by combining the argument of [6] with sharp concentration inequalities from [14, 30]. In this section, we outline the proofs of (I) to (III) and omit the proof of (IV) for brevity.

2.2 Difficulties in Extending Previous Techniques

Before describing the ideas for the proof of Lemma 2.1, we briefly discuss the techniques used in previous work and explain why they cannot be applied to USD directly.

(1) Breakdown of concentration-based analysis for large k. The key observation introduced by Becchetti et al. [6] is that β_t concentrates around $1/2$ and remains near this value for a substantial period of time. Using this stabilization property, they divided the USD process into several phases according to $\alpha_t^{\max}$ and proved how the gap between the most popular opinion and the second most popular opinion grows in each phase. This strategy is used in later work [1] to analyze USD in the population protocol model. Our proof is based on this idea.

Becchetti et al. [6] needed in their analysis a strong assumption that $k \ll n^{1/3}$ due to the limitation of the standard Chernoff bound. In this paper, we broaden the range of k to $k \ll \sqrt{n}$ by exploiting the wider martingale-based framework used in [14, 30]—particularly, we use Freedman's inequality together with additional technical ideas (see Section 2.6 for more details).

However, for $k \gg \sqrt{n}$, the tail bounds from Freedman's inequality do not suffice to yield high-probability bounds, and the analysis based on the concentration of $\alpha_t(i)$ and β_t as in [6] entirely breaks down.

Remark 2.2. *In the gossip model, the standard concentration argument yields that the number of remaining decided opinions becomes with high probability at most* $\min(k, n\log n/k)$ *after the first round if* $\beta_0 = 1$ *(see [15, Section C]), i.e., if all vertices are initially decided. On the other hand, if we could extend the argument in [1] from the population protocol model to the gossip model, we would be able to derive an* $\widetilde{O}(k)$ *bound on the consensus time for* $k \leq O(\sqrt{n}/(\log n)^2)$. *Combining these two observations, we would be able to obtain a* $\widetilde{O}(\min\{k, \sqrt{n}\})$ *bound on the consensus time if either* $k \leq O(\sqrt{n}/(\log n)^2)$ *or* $k \geq \Omega(\sqrt{n}(\log n)^3)$ *holds. Thus even if it was possible to transfer somehow the analysis in [1] to the gossip model, the case of* $O(\sqrt{n}/(\log n)^2) \leq k \leq \Omega(\sqrt{n}(\log n)^3)$ *would still remain unclear. Dealing with k around* $\sqrt{n}$ *has been always challenging in the literature on USD (e.g., [4, Section 3.3]).*

(2) Limitations of the $k > \sqrt{n}$ techniques for 3-Majority. USD behaves similarly to the well-known consensus dynamics 3-Majority, which has been analyzed extensively for the *entire* range of $2 \leq k \leq n$ [7, 8, 10, 14, 24, 30]. In 3-Majority, a selected vertex chooses three neighbors uniformly at random and takes the majority opinion held by the three, breaking the tie in favor of the third selected neighbor. However, several structural and probabilistic features of USD make its analysis, especially in the regime $k \gg \sqrt{n}$, substantially more difficult.

We next explain why all known techniques that successfully handle the large-k regime in 3-Majority fail for USD. These approaches fall into two broad categories:

- **Coupling + majorization ([10, 14]).** The successful analyses for 3-Majority by [10, 14] rely on a coupling with Pull Voting together with a majorization argument to show that the number of surviving opinions rapidly drops to $O(\sqrt{n})$, after which the concentration-based analysis becomes effective.
- **Potential-based analysis ([30]).** Another recent approach by [30] is to design a potential function that exhibits an additive drift and keeps growing during the process. When the potential function reaches a certain level, the concentration-based analysis becomes effective.

Neither of the above two approaches transfers to USD. Unlike 3-Majority, USD cannot be related to Pull Voting through coupling due to the existence of undecided vertices. At the same time, the potential function used in [30] itself does not exhibit the additive drift property in USD, again due to undecided vertices. Specifically, the potential function used in [30] is defined by

$$\gamma_t = \sum_{i \in [k]} \alpha_t(i)^2. \tag{2}$$

Since $\alpha_t(i)$ given the configuration at round $t-1$ can be written as the sum of n independent random variables (in the gossip model), a standard moment calculation yields that, for 3-Majority,

$$\mathop{\mathbb{E}}_{t-1}[\gamma_t] = \gamma_{t-1} + \Omega(1/n),$$

where $\mathbb{E}_{t-1}[\cdot]$ is the expectation conditioned on the configuration at round $t-1$. However, in USD, again due to the emergence of undecided vertices, this inequality does not hold in general.

In summary, although several structural insights from previous work remain conceptually useful (e.g., the general ideas of potential-based arguments and concentration-based analysis), none of the existing analytical frameworks transfers to USD in a black-box way. However, as shown in Section 2.4, we can design another potential function $\widetilde{\gamma}_t$ that does exhibit an additive drift and keeps growing during the process.

2.3 Proof of Lemma 2.1 (I): Behavior of β_t and ψ_t (Section 5.2)

Now we outline the proof of the first item of Lemma 2.1, which states that (i) β_t becomes $1/2 - o(1)$ with high probability within $O(\log n)$ rounds and remains at least $1/2 - o(1)$ thereafter, and (ii) ψ_t is small with high probability for all $t \geq 1$, where ψ_t is defined in Eq. (1) and measures the change of β_t from round $t-1$ to t since

$$\mathop{\mathbb{E}}_{t-1}[\beta_t] = 2\beta_{t-1}(1-\beta_{t-1}) + \sum_{i \in [k]} \alpha_{t-1}(i)^2 \tag{3}$$

$$= \beta_{t-1} - \psi_{t-1}. \tag{4}$$

For example, if $\alpha_0(i) = 1/k$ for all $i \in [k]$ (i.e., the initial configuration is balanced), then $\psi_0 = 1 - 1/k$, which means that β_1 becomes $\beta_0 - \psi_0 = 1/k$ in expectation, i.e., β_1 drops significantly.

Interestingly, as shown in Lemma 4.6, after the first round the expectation of ψ_t is always small:

$$\mathop{\mathbb{E}}_{t-1}[\psi_t] \leq \frac{1}{n}.$$

Combined this with a concentration inequality, we obtain the "Moreover" part of Lemma 2.1(I). Thus, once β_t becomes $1/2 - o(1)$, it remains at least $1/2 - o(1)$ thereafter with high probability.

It remains to show that β_t becomes $1/2 - o(1)$ within $O(\log n)$ rounds with high probability. Suppose that $\beta_{t-1} = o(1)$. Then, by Eq. (3), we have $\mathbb{E}_{t-1}[\beta_t] \geq (2-o(1))\beta_{t-1}$. Therefore, β_t grows by a factor of $1+\Omega(1)$ in each round, which ensures that β_t becomes $\Omega(1)$ within $O(\log n)$ rounds with high probability. A similar argument yields that we can show that $\beta_t = 1/2-o(1)$ within $O(\log n)$ rounds with high probability.

2.4 Proof of Lemma 2.1 (II): Growth of $\alpha_t^{\max}$ (Sections 5.3 and 5.4)

We now outline the proof of the second item of Lemma 2.1. The key idea is to normalize the potential γ_t of Eq. (2) used for 3-Majority by β_t^2 to obtain the potential $\widetilde{\gamma}_t$ defined by

$$\widetilde{\gamma}_t := \begin{cases} \sum_{i\in[k]} \left(\frac{\alpha_t(i)}{\beta_t}\right)^2, & \beta_t > 0, \\ 0, & \text{otherwise.} \end{cases}$$

A reader familiar with the literature might notice that this quantity is somewhat similar to the *monochromatic distance* introduced in [6]. In our notation, the monochromatic distance is defined in [6] by

$$\mathrm{md}_t := \sum_{i\in[k]} \left(\frac{\alpha_t(i)}{\alpha_t^{\max}}\right)^2.$$

The difference between $\widetilde{\gamma}_t$ and md_t is that $\widetilde{\gamma}_t$ is normalized by β_t^2 while md_t is normalized by $(\alpha_t^{\max})^2$. This slight modification allows us to evaluate the expected growth of $\widetilde{\gamma}_t$, although $\widetilde{\gamma}_t$ is in a more complex form compared to γ_t. Using Lemma 2.1(I) that ensures $\beta_{t-1} = \Omega(1)$ and a Taylor approximation, we can show that the potential $\widetilde{\gamma}_t$ has an additive drift (Lemma 4.8):

$$\mathop{\mathbb{E}}_{t-1}[\widetilde{\gamma}_t] \approx \frac{\mathbb{E}_{t-1}[\gamma_t]}{\mathbb{E}_{t-1}[\beta_t]^2} \geq \widetilde{\gamma}_{t-1} + \Omega(1/n).$$

Using a variant of the optional stopping theorem, we can show that the potential $\widetilde{\gamma}_t$ reaches $\Omega((\log n)^2/\sqrt{n})$ within $\widetilde{O}(\sqrt{n})$ rounds with high probability.

Once $\widetilde{\gamma}_t$ reaches $\Omega((\log n)^2/\sqrt{n})$, we have

$$\alpha_t^{\max} \geq \sum_{i\in[k]} \alpha_t(i)^2 = \widetilde{\gamma}_t \cdot \underbrace{\beta_t^2}_{\Omega(1) \text{ from Lemma 2.1(I)}} \geq \Omega\Big((\log n)^2/\sqrt{n}\Big).$$

We then show that $\alpha_t^{\max}$ does not decrease too much thereafter. To this end, we introduce the normalized quantity

$$\widetilde{\alpha}_t^{\max} := \begin{cases} \alpha_t^{\max}/\beta_t, & \beta_t > 0, \\ 0, & \beta_t = 0, \end{cases}$$

which is well-behaved because $\beta_t = \Omega(1)$ throughout this part of the process. In Lemma 4.7, we will prove

$$\mathop{\mathbb{E}}_{t-1}[\widetilde{\alpha}_t^{\max}] \geq \left(1 - O\left(\frac{1}{n}\right)\right)\widetilde{\alpha}_{t-1}^{\max}, \tag{5}$$

which ensures that once $\widetilde{\alpha}_t^{\max}$ reaches $\Omega((\log n)^2/\sqrt{n})$, it remains above this threshold (up to a constant factor) for $\Theta(n)$ rounds with high probability.

2.5 Proof of Lemma 2.1 (III): Gap between Two Opinions (Sections 5.5 and 5.6)

Recall that an opinion i is weak at round t if $\alpha_t(i) \leq 0.9 \cdot \alpha_t^{\max}$ and strong at round t if $\alpha_t(i) \geq 0.95 \cdot \alpha_t^{\max}$. Fix two distinct non-weak opinions $i, j \in [k]$. The goal is to show that (i) either i or j becomes weak within $O(\log n/\alpha_0^{\max})$ rounds with high probability, and (ii) once an opinion becomes weak, it cannot become strong during the rest of the process.

To this end, we are interested in the gap between i and j, which is defined as

$$\delta_t := \alpha_t(i) - \alpha_t(j).$$

Fix a round $t - 1 \geq 1$. In Lemma 4.4, we will show that the expectation of δ_t given α_{t-1} is

$$\mathop{\mathbb{E}}_{t-1}[\delta_t] = \delta_{t-1}(\alpha_{t-1}(i) + \alpha_{t-1}(j) + 2(1 - \beta_{t-1})). \tag{6}$$

Our aim is to show that δ_t grows by a factor of $1 + \Omega(\alpha_{t-1}^{\max})$ in each round. However, this does not hold in general. For example, if $\beta_{t-1} = 1$ and $\alpha_{t-1}(i) = \Theta(1/k)$ for all $i \in [k]$, then a standard Chernoff bound argument yields that $\alpha_t(i) \approx \alpha_{t-1}(i)^2 = \Theta(1/k^2)$, yielding $\delta_t \ll \delta_{t-1}$. Thus we need to put some assumptions on the configuration at round $t - 1$, which are listed below:

- Suppose that both i and j are non-weak at round $t - 1$. Then, we have $\alpha_{t-1}(i) + \alpha_{t-1}(j) \geq 1.8 \cdot \alpha_{t-1}^{\max}$.
- Suppose that $\beta_{t-1} \geq 1/2 - o(1)$ and $\psi_{t-1} \leq o(\log n/\sqrt{n})$, which holds from Lemma 2.1(I). Since $\psi_{t-1} = \beta_{t-1}(2\beta_{t-1} - 1) - \gamma_{t-1}$ and $\gamma_{t-1} = \sum_i \alpha_{t-1}(i)^2 \leq \alpha_{t-1}^{\max} \cdot \sum_i \alpha_{t-1}(i) = \alpha_{t-1}^{\max} \cdot \beta_{t-1}$, we have

$$\begin{aligned} 2\beta_{t-1} - 1 &\leq \frac{\gamma_{t-1} + o(\log n/\sqrt{n})}{\beta_{t-1}} \\ &= \alpha_{t-1}^{\max} + o(\log n/\sqrt{n}). \end{aligned}$$

Applying these two items to Eq. (6), we have

$$\begin{aligned} \mathop{\mathbb{E}}_{t-1}[\delta_t] &= \delta_{t-1}\left(1 + \underbrace{\alpha_{t-1}(i) + \alpha_{t-1}(j)}_{\geq 1.8\cdot\alpha_{t-1}^{\max}} - \underbrace{(2\beta_{t-1} - 1)}_{\leq \alpha_{t-1}^{\max} + o(\log n/\sqrt{n})}\right) \\ &\geq \delta_{t-1}\Big(1 + 0.8 \cdot \alpha_{t-1}^{\max} - o(\log n/\sqrt{n})\Big) \\ &\geq \delta_{t-1}\big(1 + \Omega(\alpha_{t-1}^{\max})\big). \end{aligned} \tag{7}$$

In the last inequality, note that $\alpha_{t-1}^{\max} \geq \Omega((\log n)^2/\sqrt{n})$ by the assumption.

Finally, from Eq. (5) and $\beta_t = \Omega(1)$, we have that $\alpha_t^{\max} \geq \Omega(\alpha_0^{\max})$ for all $t = O(n)$ (in expectation). Thus, we obtain

$$\mathop{\mathbb{E}}_{t-1}[\delta_t] \geq \delta_{t-1}\big(1 + \Omega(\alpha_0^{\max})\big).$$

Even if we start with $\delta_0 = 0$ (i.e., $\alpha_0(i) = \alpha_0(j)$), by calculation (see Lemma 4.4), we have that its second moment satisfies

$$\mathop{\mathbf{Var}}_{t-1}[\delta_t] = \Omega\left(\frac{(1 - \beta_{t-1})^2(\alpha_{t-1}(i) + \alpha_{t-1}(j))}{n}\right) = \Omega\left(\frac{\alpha_0^{\max}}{n}\right),$$

and in particular

$$\mathop{\mathbb{E}}_{t-1}[\delta_t^2] \geq \delta_{t-1}^2 + \Omega\left(\frac{\alpha_0^{\max}}{n}\right).$$

Thus, by the optional stopping theorem combined with a widely-known argument from [18], we can show that $|\delta_t|$ becomes at least $\Omega(\sqrt{\log n/n})$ within $O(\log n/\alpha_0^{\max})$ rounds with high probability. Therefore, within $O(\log n/\alpha_0^{\max})$ rounds, either i or j becomes weak with high probability.

2.6 Concentration Bounds via Freedman's Inequality

To make the analysis of Sections 2.4 and 2.5 rigorous, we require concentration bounds for several stochastic quantities (e.g., δ_t). Our approach follows the Freedman-based framework developed in [14, 30], which applies Freedman's inequality (a Bernstein-type concentration inequality for martingales) to obtain tight concentration bounds for various quantities.

In the population protocol model, these inequalities apply directly because each update affects only one vertex. In the gossip model, however, a synchronous round aggregates n independent local changes. Based on the observation that the one-step difference (e.g., $\alpha_t(i)-\alpha_{t-1}(i)$) can be written as the sum of n independent random variables, [30] introduced the concept of *Bernstein condition*, which is a sufficient condition for the application of Freedman's inequality even without bounded single-step jumps; see Section 3.2.

However, the framework of [30] used for 3-Majority is insufficient for USD. For 3-Majority, the quantities of interest (e.g., $\alpha_t(i)$ or $\delta_t = \alpha_t(i) - \alpha_t(j)$) are linear in independent random variables, which allows us to write the one-step difference as the sum of independent random variables. In USD, however, the presence of undecided vertices destroys the clean drift structure of $\alpha_t(i)$, forcing us to analyze normalized quantities such as

$$\widetilde{\gamma}_t = \frac{\|\alpha_t\|_2^2}{\beta_t^2}, \qquad \widetilde{\alpha}_t^{\max} = \frac{\alpha_t^{\max}}{\beta_t}.$$

While β_t concentrates reasonably well, its variance typically dominates that of $\alpha_t(i)$, so concentration of the numerator and the denominator *separately* does not yield tight bounds for their ratio. Thus, the straightforward application of the Freedman-based framework is insufficient in its current form.

To overcome this difficulty, we apply a first-order Taylor expansion of $1/\beta_t$ around $\mathbb{E}_{t-1}[\beta_t]$:

$$\frac{1}{\beta_t} \geq \frac{1}{\mathbb{E}_{t-1}[\beta_t]} - \frac{\beta_t - \mathbb{E}_{t-1}[\beta_t]}{\mathbb{E}_{t-1}[\beta_t]^2}.$$

Multiplying by $\alpha_t(i)$ gives

$$\frac{\alpha_t(i)}{\beta_t} \geq \frac{\alpha_t(i)}{\mathbb{E}_{t-1}[\beta_t]} - \frac{\alpha_t(i)}{\mathbb{E}_{t-1}[\beta_t]^2}\left(\beta_t - \mathop{\mathbb{E}}_{t-1}[\beta_t]\right).$$

Both $\alpha_t(i)$ and β_t are sums of n independent random variables, so the right-hand side becomes a quadratic form in such sums. Using read-k concentration bounds [19, 23], we obtain sharp concentration for these quadratic forms, and hence for the normalized quantities themselves.

This resolves the main obstacle preventing the direct use of Freedman-style arguments for USD and provides the concentration guarantees needed throughout the analysis.

Organization. In the following sections, we introduce the formal definition and main analysis tools (Section 3), key quantities that will be used throughout the analysis (Section 4), and technical lemmas for the proof of the main theorem of the gossip model (Section 5). Due to the page limitation, we omit arguments of the population protocol model and proofs of all lemmas. See the full version [15] for the details.

3 Definitions and Main Analysis Tools

3.1 Model and Basic Notation

For $n \in \mathbb{N}$, let $[n] = \{1, 2, \dots, n\}$. For $p > 0$ and a vector $x \in \mathbb{R}^n$, let $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ denote the ℓ^p-norm of x. For a finite set S, by $x \sim S$ we mean that x is chosen uniformly at random from S. For $a, b \in \mathbb{R}$, we use the shorthand $a \wedge b = \min\{a, b\}$. We denote by $\mathbf{1}_{\mathcal{E}}$ the indicator random variable of an event $\mathcal{E}$, i.e., $\mathbf{1}_{\mathcal{E}} = 1$ if $\mathcal{E}$ occurs and 0 otherwise.

Throughout the paper, we fix a set of n vertices, denoted by V, with $|V| = n$. Each vertex holds an opinion from a finite alphabet $\Sigma = [k] \cup \{\perp\}$, where $k \in \mathbb{N}$ is a parameter and $\perp$ represents the undecided state. All configurations of the system are elements of Σ^V. The interaction rule (gossip or population protocol) will be specified later, but in all cases the opinion update is governed by the same USD update rule introduced below.

Definition 3.1 (USD Update Rule)**.** *The* USD update rule *is a function* update: $\Sigma \times \Sigma \to \Sigma$ *defined by*

$$\mathsf{update}(\sigma_1, \sigma_2) = \begin{cases} \perp & \text{if } |\{\sigma_1, \sigma_2, \perp\}| = 3, \\ \sigma_2 & \text{if } \sigma_1 = \perp, \\ \sigma_1 & \text{otherwise.} \end{cases}$$

Note that, the condition $|\{\sigma_1, \sigma_2, \perp\}| = 3$ means that both σ_1 and σ_2 are decided and they are distinct.

Using the update rule above, we define USD in the gossip model. In this model, every vertex updates simultaneously in each round by interacting with an independently chosen random neighbor.

Definition 3.2 (Gossip USD)**.** *The* gossip USD *is the discrete-time Markov chain* $(\mathrm{opn}_t)_{t\geq 0}$ *on the state space* Σ^V. *Given* $\mathrm{opn}_t \in \Sigma^V$, *the next configuration* $\mathrm{opn}_{t+1} \in \Sigma^V$ *is obtained by the following procedure: For each vertex* $u \in V$, *independently select a random vertex* $v \sim V$ *and set* $\mathrm{opn}_{t+1}(u) = \mathsf{update}(\mathrm{opn}_t(u), \mathrm{opn}_t(v))$.

The main performance measure of USD is the *consensus time*, defined by $\tau_{\mathrm{cons}} = \inf\left\{t \geq 0 : \exists a \in [k], \forall u \in V, \mathrm{opn}_t(u) = a\right\}$. Note that $\tau_{\mathrm{cons}} = \infty$ if for some $t \geq 0$, $\mathrm{opn}_t(u) = \perp$ for all $u \in V$.

We sometimes use $\mathcal{F}_t$ to denote the natural filtration generated by the process $(\mathrm{opn}_s)_{s\leq t}$. Using this notation, we often use abbreviations $\Pr_{t-1}[\cdot]$, $\mathbb{E}_{t-1}[\cdot]$, and $\mathbf{Var}_{t-1}[\cdot]$ for $\Pr[\cdot \mid \mathcal{F}_{t-1}]$, $\mathbb{E}[\cdot \mid \mathcal{F}_{t-1}]$, and $\mathbf{Var}[\cdot \mid \mathcal{F}_{t-1}]$, respectively. For a random variable X_t defined for step t in USD, we often use the term "X_t conditioned on round $t-1$" to denote this random variable conditioned on $\mathcal{F}_{t-1}$.

3.2 Bernstein Condition

The Bernstein condition [30] provides uniform control on the moment generating function of a random variable and will be used throughout to obtain concentration bounds for the quantities appearing in USD, such as $\alpha_t(i)$ and their normalized forms.

Definition 3.3 (Bernstein condition; [30])**.** *Let* $D, s \geq 0$. *A random variable* X *satisfies the* (D, s)*-Bernstein condition if for all* $\lambda \in \mathbb{R}$

with $|\lambda|D < 3$, $\mathbb{E}[e^{\lambda X}] \le \exp\left(\frac{\lambda^2 s/2}{1-|\lambda|D/3}\right)$. *It satisfies the* one-sided (D,s)-Bernstein condition *if the same bound holds for all* $\lambda \ge 0$ *with* $\lambda D < 3$.

We will use the basic closure properties of the Bernstein condition ([30, Lemma 3.4]) repeatedly later. In addition, we will need Bernstein bounds for sums of random variables that may not be independent, namely, the sum of a read-ℓ family (see [15, Lemma 3.8] for details).

3.3 Freedman's Inequality and Drift Analysis

Drift analysis allows us to bound the hitting time of a stochastic process once we control both its expected evolution and the concentration of its increments. We briefly recall the basic mechanism.

Let $(X_t)_{t\ge 0}$ be a process with $\mathbb{E}_{t-1}[X_t] \ge X_{t-1} + R$ whenever $X_{t-1} < a$, for some $R > 0$ and threshold $a > X_0$, and define the stopping time $\tau := \inf\{t \ge 0 : X_t \ge a\}$. Consider the stopped process $Y_t := X_{t\wedge\tau} - R(t \wedge \tau)$, which is a submartingale. If the increments $Y_t - Y_{t-1}$ satisfy a Bernstein condition, Freedman's inequality yields, with high probability, $Y_T \ge Y_0 - \varepsilon = X_0 - \varepsilon$ for all sufficiently large T. On the event $\{T < \tau\}$, we have $X_T < a$ and hence $X_T - RT = Y_T \ge X_0 - \varepsilon$, which implies

$$T \le \frac{a - X_0 + \varepsilon}{R}. \tag{8}$$

Thus τ is at most this value with high probability. This is the standard way in which positive drift and concentration combine to yield upper bounds on hitting times; see, e.g., [29]. The following lemma, introduced in [30], summarizes the above discussion, which is based on the Freedman's inequality.

Lemma 3.4 (Lemma 3.5 of [30])**.** *Let* $(X_t)_{t\in\mathbb{N}_0}$ *be a sequence of random variables and let* $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ *be a filtration such that* X_t *is* $\mathcal{F}_t$*-measurable for all* $t \ge 0$. *Let* τ *be a stopping time with respect to* $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$. *Let* $D, s \ge 0$ *and* $R \in \mathbb{R}$ *be parameters. Suppose the following condition holds for any* $t \ge 1$*: conditioned on* $\mathcal{F}_{t-1}$,

- $(C1)$ $\mathbf{1}_{\tau>t-1}(\mathbb{E}_{t-1}[X_t] - X_{t-1} - R) \le 0$,
- $(C2)$ $\mathbf{1}_{\tau>t-1}(X_t - X_{t-1} - R)$ *satisfies one-sided* (D,s)*-Bernstein condition.*

For a parameter $h > 0$, *define stopping times*

$\tau_X^+ := \inf\{t \ge 0 \colon X_t \ge X_0 + h\}$ *and* $\tau_X^- := \inf\{t \ge 0 \colon X_t \le X_0 - h\}$.

Then, we have the following:

- *(i) Suppose* $R \ge 0$. *Then, for any* $h, T > 0$ *such that* $z := h - R\cdot T > 0$, *we have* $\Pr\left[\tau_X^+ \le \min\{T,\tau\}\right] \le \exp\left(-\frac{z^2/2}{sT+(zD)/3}\right)$.
- *(ii) Suppose* $R < 0$. *Then, for any* $h, T > 0$ *such that* $z := (-R)\cdot T - h > 0$, *we have* $\Pr\left[\min\{\tau_X^-,\tau\} > T\right] \le \exp\left(-\frac{z^2/2}{sT+(zD)/3}\right)$.

4 Key Quantities

In this section, we introduce the quantities that will be used throughout our analysis of USD. For each quantity, we establish its conditional expectation, conditional variance, and a one-sided Bernstein condition. These properties form the backbone of the drift arguments used later.

Since the definitions are intentionally consolidated here, the presentation is dense; readers are encouraged to refer back to this section as needed when following the proofs in Section 5.

Definition 4.1 (Key Quantities)**.** *Let* $(\mathrm{opn}_t)_{t\ge 0}$ *be either the gossip USD or the population protocol USD. For each round* $t \ge 0$, *define:*

- *For each* $i \in [k]$, *let* $\alpha_t(i) := \frac{1}{n}\left|\{u \in V : \mathrm{opn}_t(u) = i\}\right|$, *and let* $\beta_t := \sum_{i\in[k]} \alpha_t(i)$ *denote the total fraction of vertices holding a decided opinion.*
- $\gamma_t := \|\alpha_t\|_2^2 = \sum_{i\in[k]} \alpha_t(i)^2$.
- $\psi_t := \beta_t(2\beta_t - 1) - \gamma_t$.
- $\widetilde{\gamma}_t := \gamma_t/\beta_t^2$. *If* $\beta_t = 0$, *we define* $\widetilde{\gamma}_t = 0$.
- $\alpha_t^{\max} := \|\alpha_t\|_\infty = \max_{i\in[k]} \alpha_t(i)$.
- $\widetilde{\alpha}_t^{\max} := \alpha_t^{\max}/\beta_t$. *If* $\beta_t = 0$, *we define* $\widetilde{\alpha}_t^{\max} = 0$.
- *For* $i, j \in [k]$ *and* $\varepsilon \in [0,1]$, *define* $\delta_t^{(\varepsilon)}(i,j) := \alpha_t(i) - (1+\varepsilon)\alpha_t(j)$ *and* $\delta_t(i,j) := \delta_t^{(0)}(i,j)$. *If* i *and* j *are clear from the context, we write* δ_t *or* $\delta_t^{(\varepsilon)}$ *instead of* $\delta_t(i,j)$ *or* $\delta_t^{(\varepsilon)}(i,j)$, *respectively.*
- *Let* $I_t := \min\{i \in [k] : \alpha_t(i) = \alpha_t^{\max}\}$ *be the most popular opinion at round* t. *For* $i \in [k]$, *let* $\eta_t(i) := \delta_t^{(c_\eta)}(I_t, i) = \alpha_t^{\max} - (1+c_\eta)\alpha_t(i)$, *where the constant* $c_\eta > 0$ *is fixed in Definition 5.8.*

4.1 Basic Properties in the Gossip Model

We now present the basic properties (expectations, variances, and Bernstein conditions) of the key quantities introduced in Definition 4.1 in the gossip model. These results are applied repeatedly in the phase analysis. All proofs of the lemmas in this section are deferred to [15, Section B], as they are routine calculations or applications of lemmas in Section 3.2.

Lemma 4.2 (Basic properties for $\alpha_t(i)$ in the gossip model)**.** *For the gossip USD, the quantity* $\alpha_t(i)$ *satisfies the following:*

1. *(Expectation)* $\mathbb{E}_{t-1}[\alpha_t(i)] = \alpha_{t-1}(i)(\alpha_{t-1}(i) + 2(1-\beta_{t-1}))$.
2. *(Variance)* $\mathbf{Var}_{t-1}[\alpha_t(i)] \le \frac{\alpha_{t-1}(i)}{n}$ *and* $\mathbf{Var}_{t-1}[\alpha_t(i)] \ge \frac{(1-\beta_{t-1})^2\alpha_{t-1}(i)}{n}$ *hold.*
3. *(Bernstein condition)* $\alpha_t(i) - \mathbb{E}_{t-1}[\alpha_t(i)]$ *conditioned on round* $t-1$ *satisfies* $\left(\frac{1}{n}, \frac{\alpha_{t-1}(i)}{n}\right)$*-Bernstein condition.*

Lemma 4.3 (Basic properties for β_t in the gossip model)**.** *For the gossip USD, the quantity* β_t *satisfies the following:*

1. *(Expectation)* $\mathbb{E}_{t-1}[\beta_t] = 2\beta_{t-1}(1-\beta_{t-1}) + \gamma_{t-1} = \beta_{t-1} - \psi_{t-1}$.
2. *(Variance)* $\mathbf{Var}_{t-1}[\beta_t] = \frac{(\beta_{t-1}-\gamma_{t-1})(1-\beta_{t-1}+\gamma_{t-1}) + \gamma_{t-1}^2 - \|\alpha_{t-1}\|_3^3}{n}$. *In particular,* $\mathbf{Var}_{t-1}[\beta_t] \le \frac{\beta_{t-1}}{n}$.
3. *(Bernstein condition)* $\beta_t - \mathbb{E}_{t-1}[\beta_t]$ *conditioned on round* $t-1$ *satisfies* $\left(\frac{1}{n}, \frac{\beta_{t-1}}{n}\right)$*-Bernstein condition.*

Lemma 4.4 (Basic properties for $\delta_t^{(\varepsilon)}$ in the gossip model)**.** *For the gossip USD, the quantity* $\delta_t^{(\varepsilon)} := \delta_t^{(\varepsilon)}(i,j)$ *satisfies the following:*

1. *(Expectation)*

$$\mathop{\mathbb{E}}_{t-1}\left[\delta_t^{(\varepsilon)}\right] = \delta_{t-1}^{(\varepsilon)}(\alpha_{t-1}(i) + \alpha_{t-1}(j) + 2(1-\beta_{t-1})) + \varepsilon\alpha_{t-1}(i)\alpha_{t-1}(j).$$

2. *(Variance)* $\mathbf{Var}_{t-1}\left[\delta_t\right] \ge \frac{(1-\beta_{t-1})^2}{n}(\alpha_{t-1}(i) + \alpha_{t-1}(j))$.

(3) *(Bernstein condition) The difference* $\delta_t^{(\varepsilon)} - \mathbb{E}_{t-1}\left[\delta_t^{(\varepsilon)}\right]$ *conditioned on round* $t-1$ *satisfies* $\left(\frac{2(1+\varepsilon)}{n}, s\right)$*-Bernstein condition for* $s = \frac{2}{n}\left(\alpha_{t-1}(i) + (1+\varepsilon)^2\alpha_{t-1}(j)\right)$

Lemma 4.5 (Basic properties for γ_t in the gossip model). *For the gossip USD, the quantity* γ_t *satisfies the following:*

(1) *(Expectation Upper Bound)* $\mathbb{E}_{t-1}[\gamma_t] \le 10\gamma_{t-1}$.
(2) *(Expectation Lower Bound)* $\beta_{t-1}^2 \mathbb{E}_{t-1}[\gamma_t] \ge \mathbb{E}_{t-1}[\beta_t]^2\gamma_{t-1} + \frac{\beta_{t-1}^3(1-\beta_{t-1})^2}{n}$.
(3) *(Bernstein condition)* $\gamma_t - \mathbb{E}_{t-1}[\gamma_t]$ *conditioned on round* $t-1$ *satisfies* $\left(\frac{2}{n}, \frac{20\gamma_{t-1}}{n}\right)$*-Bernstein condition.*

Lemma 4.6 (Basic properties for ψ_t in the gossip model). *For the gossip USD, the quantity* ψ_t *satisfies the following:*

(1) *(Expectation)* $\mathbb{E}_{t-1}[\psi_t] \le \frac{\beta_{t-1}}{n}$.
(2) *(Bernstein condition)* $\psi_t - \mathbb{E}_{t-1}[\psi_t]$ *conditioned on round* $t-1$ *satisfies* $\left(\frac{24}{n}, \frac{400\beta_{t-1}}{n}\right)$*-Bernstein condition.*

Lemma 4.7 (Basic properties for $\widetilde{\alpha}_t^{\max}$ in the gossip model). *For the gossip USD, the quantity* $\widetilde{\alpha}_t^{\max}$ *satisfies the following: Let* $C > 0$ *be a sufficiently large constant. Suppose that* $\beta_{t-1} \ge 1/2 - o(1)$ *and* $\psi_{t-1} \le o(1)$*. Then,*

(1) *(Expectation)* $\mathbb{E}_{t-1}[\widetilde{\alpha}_t^{\max}] \ge \widetilde{\alpha}_{t-1}^{\max}\left(1 + \frac{\alpha_{t-1}^{\max} - \gamma_{t-1}/\beta_{t-1}}{2(1-\beta_{t-1}) + \gamma_{t-1}/\beta_{t-1}}\right) - \frac{9\alpha_{t-1}^{\max}}{n}$.
(2) *(Bernstein condition)* $\widetilde{\alpha}_{t-1}^{\max}\left(1 + \frac{\alpha_{t-1}^{\max} - \gamma_{t-1}/\beta_{t-1}}{2(1-\beta_{t-1}) + \gamma_{t-1}/\beta_{t-1}}\right) - \widetilde{\alpha}_t^{\max} - \frac{9\alpha_{t-1}^{\max}}{n}$ *conditioned on round* $t-1$ *satisfies one-sided* $\left(\frac{C}{n}, \frac{C\alpha_{t-1}^{\max}}{n}\right)$*-Bernstein condition.*

Lemma 4.8 (Basic properties for $\widetilde{\gamma}_t$ in the gossip model). *For the gossip USD, the quantity* $\widetilde{\gamma}_t$ *satisfies the following: Suppose* $\beta_{t-1} \ge 1/2 - o(1)$*,* $\psi_{t-1} \le o(1)$*, and* $\gamma_{t-1} \le o(1)$*. Then,* $\mathbb{E}_{t-1}[\widetilde{\gamma}_t] \ge \widetilde{\gamma}_{t-1} + \frac{1}{12n}$.

4.2 Stopping Times

In the remainder of the paper we frequently refer to the first time at which one of the key quantities introduced above crosses a given threshold. Since these stopping times are used uniformly in both the gossip model and the population–protocol model, we introduce them here in a unified way.

For a stochastic process $(Z_t)_{t\ge 0}$ and a threshold $\theta \in \mathbb{R}$, we consider the upward and downward hitting times of the form $\tau_Z^+ := \inf\{t \ge 0 : Z_t \ge \theta\}$ and $\tau_Z^- := \inf\{t \ge 0 : Z_t \le \theta\}$. We will consider such stopping times for β_t, $\alpha_t^{\max}$, δ_t, $\widetilde{\gamma}_t$, etc. These stopping times serve as canonical milestones in our drift and concentration arguments later on.

Whenever a stopping time refers to a quantity that depends on a pair of opinions (e.g., δ_t), the stopping time is understood to be taken with respect to that fixed pair unless otherwise stated.

Definition 4.9 (Stopping Times). *Consider quantities defined in Definition 4.1.*

- *(Stopping times for* β_t*) For a parameter* $x > 0$*, define the stopping times* $\tau_\beta^+(x) = \inf\left\{t \ge 0 : \beta_t \ge \frac{1}{2} - x\right\}$ *and* $\tau_\beta^-(x) = \inf\left\{t \ge 0 : \beta_t < \frac{1}{2} - x\right\}$*. By default, we set the parameter* x *to be any positive function* $x_\beta = x_\beta(n)$ *satisfying* $x_\beta(n) = \omega\left(\sqrt{\log n/n}\right)$ *and* $x_\beta(n) = o(\log n/\sqrt{n})$ *(for example,* $x_\beta = (\log n)^{3/5}/\sqrt{n}$*). We sometimes abbreviate* $\tau_\beta^+(x_\beta)$ *and* $\tau_\beta^-(x_\beta)$ *to* τ_β^+ *and* τ_β^-*, respectively.*
- *(Stopping times for* ψ_t*) For a parameter* $x > 0$*, define the stopping times* $\tau_\psi^+(x) = \inf\{t \ge 0 : \psi_t > x\}$ *and* $\tau_\psi^-(x) = \inf\{t \ge 0 : \psi_t \le x\}$*. By default, we set the parameter* x *to be a function* $x = x_\psi(n)$ *satisfying* $x_\psi = x_\beta/4$*. We sometimes abbreviate* $\tau_\psi^+(x_\psi)$ *and* $\tau_\psi^-(x_\psi)$ *to* τ_ψ^+ *and* τ_ψ^-*, respectively.*
- *(Stopping time for* γ_t*) For a parameter* $x_\gamma \in (0,1)$*, define the stopping time* $\tau_\gamma^+ := \inf\left\{t \ge 0 : \gamma_t \ge x_\gamma\right\}$*. Throughout this paper, we fix* $x_\gamma = (\log n)^2/\sqrt{n}$*.*
- *(Stopping times for* $\alpha_t^{\max}$ *and* $\widetilde{\alpha}_t^{\max}$*) Define the stopping times* $\tau_{\max}^{\uparrow} := \inf\{t \ge 0 : \alpha_t^{\max} \ge (1 + c_{\max}^{\uparrow})\alpha_0^{\max}\}$ *and* $\tau_{\max}^{\downarrow} := \inf\{t \ge 0 : \alpha_t^{\max} \le (1 - c_{\max}^{\downarrow})\alpha_0^{\max}\}$ *for constants* $c_{\max}^{\uparrow}, c_{\max}^{\downarrow} \in (0,1)$*. For constants* $c_{\widetilde{\max}}^{\uparrow}, c_{\widetilde{\max}}^{\downarrow} \in (0,1)$*, define the stopping times* $\tau_{\widetilde{\max}}^{\uparrow} := \inf\left\{t \ge 0 : \widetilde{\alpha}_t^{\max} \ge (1 + c_{\widetilde{\max}}^{\uparrow})\widetilde{\alpha}_0^{\max}\right\}$ *and* $\tau_{\widetilde{\max}}^{\downarrow} := \inf\left\{t \ge 0 : \widetilde{\alpha}_t^{\max} \le (1 - c_{\widetilde{\max}}^{\downarrow})\widetilde{\alpha}_0^{\max}\right\}$*. By default, we set* $c_{\max}^{\uparrow} = c_{\max}^{\downarrow} = c_{\widetilde{\max}}^{\uparrow} = c_{\widetilde{\max}}^{\downarrow} = 0.1$*.*
- *(Stopping times for* δ_t*) For constants* $c_\delta^{\uparrow}, c_\delta^{\downarrow} \in (0,1)$*, let* $\tau_\delta^{\uparrow} := \inf\{t \ge 0 : \delta_t \ge (1 + c_\delta^{\uparrow})\delta_0\}$ *and* $\tau_\delta^{\downarrow} := \inf\{t \ge 0 : \delta_t \le (1 - c_\delta^{\downarrow})\delta_0\}$ *be the stopping times for* δ_t *growing up and down multiplicatively. The constants* $c_\delta^{\uparrow}$ *and* $c_\delta^{\downarrow}$ *are defined as appropriate within the proofs.*
  For a threshold parameter $x > 0$*, define the stopping time* $\tau_\delta^+(x) := \inf\{t \ge 0 : |\delta_t| \ge x\}$*. By default, we set the parameter* x *to be* $x = c_\delta^+/\sqrt{n}$*, where* c_δ^+ *is a positive constant defined as appropriate within the proofs. We sometimes abbreviate* $\tau_\delta^+(x_\delta)$ *to* τ_δ^+*, respectively.*
- *(Stopping times for* η_t*) For a parameter* $x > 0$ *and an opinion* $j \in [k]$*, define the stopping times* $\tau_\eta^+(x) := \inf\{t \ge 0 : \eta_t(j) \ge x\}$ *and* $\tau_\eta^-(x) := \inf\{t \ge 0 : \eta_t(j) < x\}$*. By default, we set the parameter* x *to be any positive function* $x = x_\eta(n)$ *satisfying* $x_\eta(n) = \omega\left(x_\beta(n)\right)$ *and* $x_\eta(n) = o(\log n/\sqrt{n})$ *(for example,* $x_\eta = (\log n)^{4/5}/\sqrt{n}$*). We sometimes abbreviate* $\tau_\eta^+(x_\eta)$ *and* $\tau_\eta^-(x_\eta)$ *to* τ_η^+ *and* τ_η^-*, respectively.*

5 Analysis for the Gossip Model

In this section we list our key lemmas for the proof of Theorem 1.1 for the gossip model. The structure of the argument follows the outline given in Section 2: each subsection implements one of the steps described there, in the same order and with the corresponding technical tools. Because every component of the overview requires its own drift and concentration analysis, the proofs of them, which can be found in the full version, are necessarily long; however, its organization mirrors the conceptual roadmap of Section 2, and the reader may consult that discussion to track how each ingredient fits into the overall proof.

5.1 Probability of Failure at the First Round

In this subsection, we shall consider the configuration after the first synchronous update. Specifically, we obtain the following lemma, which gives the probability that the dynamics fails at the first round.

Lemma 5.1. *Consider USD in the gossip model. Let $p_\perp = \Pr[\beta_1 = 0]$ be the probability that all vertices hold $\perp$ after the first round. Then, we have the following:*

(1) If $\beta_0 = 1$, then $p_\perp = \prod_{i\in[k]} (1-\alpha_0(i))^{n\alpha_0(i)} \le \exp(-n\gamma_0) \le \exp(-n/k)$.
(2) If $0 < \beta_0 < 1$, then $\Pr[\beta_1 = 0] \le \frac{1}{n}$.

5.2 Behavior of the Fraction of Decided Vertices

From Lemma 5.1 (and [15, Lemma C.1]), we know that the fraction of undecided vertices is 0 with probability $p_\perp$ or otherwise lies between $1/n$ and $\gamma_0 \log n$ with high probability. In this subsection we show that when the latter occurs, then it holds with high probability that $\beta_t \ge 1/2 - o(1)$ for all $\Omega(\log n) \le t \le n^{O(1)}$; Lemma 5.2. In particular, this allows us to treat β_t as $\Omega(1)$ for all such t.

In addition, we prove Lemma 5.3 regarding the potential $\psi_t = \beta_t(2\beta_t - 1) - \gamma_t$, which is related to a lower bound of β_t. Note that by definition, $\mathbb{E}_{t-1}[\beta_t] = \beta_{t-1} - \psi_{t-1}$.

Lemma 5.2 (Growth of β_t)**.** *We have the following:*

(1) Let $C > 0$ be an arbitrary constant. Suppose that $\beta_0 > 0$. Then, for some $T = O(\log n)$, $\beta_T \ge \frac{C\log n}{n}$ with probability at least $1 - O\left(\frac{\log n}{n}\right)$.
(2) For some $T = O(\log n)$,
$$\Pr\left[\tau_\beta^+ > T\right] \le T\left(\exp(-\Omega(n\beta_0)) + \exp\left(-\Omega(nx_\beta^2)\right)\right).$$
(3) Suppose that $\psi_0 \le x_\psi$ and $\beta_0 \ge 1/2 - x_\beta$ hold. Then, for any $T \ge 1$,
$$\Pr\left[\tau_\beta^- \le T \text{ and } \tau_\psi^+ > T\right] \le T\exp\left(-\Omega(nx_\beta^2)\right).$$

Lemma 5.3 (Decay of ψ_t)**.** *We have the following:*

(1) $\Pr\left[\tau_\psi^- > 1\right] \le \exp\left(-\Omega(nx_\psi^2)\right)$ for any initial configuration.
(2) Suppose $\psi_0 \le x_\psi$. Then, $\Pr\left[\tau_\psi^+ \le T\right] \le T\exp\left(-\Omega(nx_\psi^2)\right)$ for any $T \ge 1$.

5.3 Behavior of Squared ℓ^2 Norm of Normalized Population

In this subsection, we track the change of the quantity $\widetilde{\gamma}_t = \|\alpha_t\|_2^2/\beta_t^2$ defined in Definition 4.1. We show that $\widetilde{\gamma}_t$ becomes $\Omega((\log n)^2/\sqrt{n})$ within $\widetilde{O}(\sqrt{n})$ rounds with high probability. Recall the definition of $x_\gamma = (\log n)^2/\sqrt{n}$ in Definition 4.9.

Lemma 5.4 (Growth of γ_t)**.** *Suppose $\beta_0 \ge 1/2 - x_\beta$ and $\psi_0 \le x_\psi$. Then, for some $T = O(nx_\gamma \log n)$, $\Pr\left[\tau_\gamma^+ > T \text{ or } \min\{\tau_\beta^-, \tau_\psi^+\} \le T\right] \le n^{-\Omega(1)}$.*

The intuition behind Lemma 5.4 is as follows. First, we derive an upper bound on the expected value of $\tau = \min\{\tau_\gamma^+, \tau_\beta^-, \tau_\psi^+\}$ in terms of $\mathbb{E}[\widetilde{\gamma}_\tau]$. The key technique is to combine the additive drift of $\widetilde{\gamma}$ with the optional stopping theorem. Second, to obtain an upper bound on $\mathbb{E}[\tau]$, we provide an upper bound of $\mathbb{E}[\widetilde{\gamma}_\tau]$.

5.4 Behavior of Maximum Population

In this subsection, we track the change of the maximum fractional population $\alpha_t^{\max} = \max_{i\in[k]} \alpha_t(i)$ and its normalized version $\widetilde{\alpha}_t^{\max} = \alpha_t^{\max}/\beta_t$. Note that, from Lemma 5.4, we know that $\alpha_t^{\max} = \Omega(\widetilde{\alpha}_t^{\max})$ becomes $\Omega((\log n)^2/\sqrt{n})$ within $\widetilde{O}(\sqrt{n})$ rounds with high probability. The aim of this subsection is to show that, with high probability, (i) $\widetilde{\alpha}_t^{\max} \ge (1 - c_{\widetilde{\max}}^{\downarrow})\widetilde{\alpha}_0^{\max}$ holds for all $t \le O(n\widetilde{\alpha}_0^{\max}/\log n)$ (see Lemma 5.5, Item 1), and (ii) $\alpha_t^{\max} \le (1 + c_{\max}^{\uparrow})\alpha_0^{\max}$ holds for all $t \le O(1/\alpha_0^{\max})$ (see Lemma 5.5, Item 2).

Lemma 5.5 (Key properties of $\widetilde{\alpha}_t^{\max}$ and $\alpha_t^{\max}$)**.** *Suppose that $\psi_0 \le x_\psi$ and $\beta_0 \ge 1/2 - x_\beta$. We have the following:*

(1) For any $T \le C_{5.5(1)}n$, $\Pr\left[\tau_{\widetilde{\max}}^{\downarrow} \le T \text{ and } \min\{\tau_\beta^-, \tau_\psi^+\} > T\right] \le T\exp\left(-\Omega(\frac{\widetilde{\alpha}_0^{\max} n}{T})\right)$. Here, $C_{5.5(1)} := c_{\widetilde{\max}}^{\downarrow}/36$ is a positive constant.
(2) Suppose $\alpha_0^{\max} = \omega(\log n/\sqrt{n})$. Let $C_{5.5(2)} = \frac{c_{\max}^{\uparrow}}{6(1+c_{\max}^{\uparrow})^2}$ be a positive constant. Then, for any $T \le \frac{C_{5.5(2)}}{\alpha_0^{\max}}$, $\Pr\left[\tau_{\max}^{\uparrow} \le \min\{T, \tau_\beta^-\}\right] \le k\exp\left(-\Omega\left(\frac{\alpha_0^{\max} n}{T}\right)\right)$.

5.5 Behavior of Gap between Two Opinions

For two opinions $i, j \in [k]$, recall $\delta_t = \delta_t(i,j) = \alpha_t(i) - \alpha_t(j)$ is the gap between the two opinions at time t (Definition 4.1). We define the *weak* opinion and its stopping time as follows:

Definition 5.6 (Weak Opinion)**.** *For a constant $c_{\mathrm{weak}} \in (0, 1/2)$ and an opinion $i \in [k]$, define $\tau_i^{\mathrm{weak}} := \inf\{t \ge 0 : \alpha_t(i) \le (1 - c_{\mathrm{weak}})\alpha_t^{\max}\}$. By default, we set $c_{\mathrm{weak}} = 0.1$. We call that an opinion $i \in [k]$ is* weak *at time t if $\alpha_t(i) \le (1 - c_{\mathrm{weak}})\alpha_t^{\max}$.*

The main results of this subsection is the following lemma. Intuitively, they show that if both i and j are non-weak, then: (i) the gap $\delta_t(i,j)$ between them grows to at least $\Omega(\sqrt{\log n/n})$ within $O(\log n/\alpha_0^{\max})$ rounds (Lemma 5.7 (Item 1)), and (ii) if the initial gap $\delta_0(i,j)$ is at least $\Omega(\sqrt{\log n/n})$, then j becomes weak within $O(\log n/\alpha_0^{\max})$ rounds (Lemma 5.7 (Item 2)).

Lemma 5.7 (Either of two non-weak opinions becomes weak)**.** *Let $i, j \in [k]$ be an arbitrary pair of two non-weak opinions. Suppose that $\beta_0 \ge 1/2 - x_\beta$, $\psi_0 \le x_\psi$, and $\alpha_0^{\max} \ge \omega(\log n/\sqrt{n})$. We have the following:*

(1) Let C be an arbitrary positive constant. Write $x = \sqrt{C\log n/n}$ for convenience. Then, for some $T = O(\log n/\alpha_0^{\max})$, we have
$$\Pr\left[\min\left\{\tau_\delta^+(x), \tau_i^{\mathrm{weak}}, \tau_j^{\mathrm{weak}}\right\} > T \text{ or } \min\{\tau_{\widetilde{\max}}^{\downarrow}, \tau_\beta^-, \tau_\psi^+\} \le T\right] \le n^{-10}.$$
(2) Suppose $\delta_0(i,j) \ge \sqrt{C\log n/n}$ for a sufficiently large constant $C > 0$. Then, for some $T = O\left(\log n/\alpha_0^{\max}\right)$,
$$\Pr\left[\tau_j^{\mathrm{weak}} > T \text{ or } \min\{\tau_{\widetilde{\max}}^{\downarrow}, \tau_\beta^-, \tau_\psi^+\} \le T\right] \le n^{-10}.$$

5.6 Emergence of Unique Strong Opinion

In this section, we show that once an opinion becomes weak, it cannot become *strong*—a key concept defined in Definition 5.8—for a sufficiently long time. Recall $\eta_t(j) = \delta_t^{(c_\eta)}(I_t, j) = \alpha_t^{\max} - (1 +$

$c_\eta)\alpha_t(j)$ and stopping times for $\eta_t(j)$ as defined in Definitions 4.1 and 4.9. We define the strong opinion and the stopping time for the emergence and persistence of the unique strong opinion, τ^+_{us} and τ^-_{us}:

Definition 5.8 (Strong Opinion). *We call that an opinion $j \in [k]$ is strong at time t if $\alpha_t(j) \geq (1-c_{\mathrm{strong}})\alpha^{\max}_t$. By default, we set $c_{\mathrm{strong}} = 0.05 < c_{\mathrm{weak}} = 0.1$. Set $c_\eta := \frac{c_{\mathrm{strong}}}{1-c_{\mathrm{strong}}}$. For a positive parameter x, define $\tau^-_{us}(x) := \inf\{t \geq 0 : \min_{j\neq I_t} \eta_t(j) < x\}$ and $\tau^+_{us}(x) := \inf\{t \geq 0 : \min_{j\neq I_t} \eta_t(j) \geq x\}$.*

The fundamental relationships between the notions of weak, strong, and the value of $\eta_t(j)$ are as follows:

(1) For any weak opinion j, we have $\eta_t(j) \geq \frac{c_{\mathrm{weak}}-c_{\mathrm{strong}}}{1-c_{\mathrm{strong}}}\alpha^{\max}_t$.
(2) The opinion j is not strong if $\eta_t(j) > 0$.

In summary, if an opinion j is weak, then $\eta_t(j) = \Omega(\alpha^{\max}_t) > 0$ (see Item 1), and if $\eta_t(j) > 0$, then j is not a strong opinion (Item 2). Therefore, if $\min_{j\neq I_t} \eta_t(j) \geq x$ (with $x > 0$), it means that all opinions other than I_t are not strong; that is, I_t is the unique strong opinion at time t. Note that Item 1 follows from $\eta_t(j) = \alpha^{\max}_t - \frac{1}{1-c_{\mathrm{strong}}}\alpha_t(j) \geq \alpha^{\max}_t - \frac{1-c_{\mathrm{weak}}}{1-c_{\mathrm{strong}}}\alpha^{\max}_t$, while Item 2 follows from the inequality $\frac{\alpha_t(j)}{\alpha^{\max}_t} < \frac{1}{1+c_\eta} = 1 - c_{\mathrm{strong}}$.

We show in this subsection that, with high probability, (i) there exists exactly one strong opinion within $O(\log n/\alpha^{\max}_0)$ rounds (Item 1 of Lemma 5.9), and (ii) any non-strong opinions keep non-strong for a sufficiently long time (Item 2 of Lemma 5.9).

Lemma 5.9 (Unique Strong Opinion Lemma). *Suppose that $\psi_0 \leq x_\psi$, $\beta_0 \geq 1/2-x_\beta$, and $\alpha^{\max}_0 = \omega(\log n/\sqrt{n})$. Then, we have the following:*

(1) $\Pr\left[\tau^+_{us}(x_\eta) > T \text{ or } \min\{\tau^{\downarrow}_{\widetilde{\max}}, \tau^-_\beta, \tau^+_\psi\} \leq T\right] \leq O(n^{-10})$ holds for some $T = O(\log n/\alpha^{\max}_0)$.
(2) Suppose $\eta_0(j) \geq x_\eta$ for all $j \neq I_0$. Let $C_{5.5(1)} = c^{\downarrow}_{\widetilde{\max}}/36$ be a positive constant that is defined in Lemma 5.5 (Item 1). Then, $\Pr\left[\min\{\tau^-_{us}(x_\eta/2), \tau^{\downarrow}_{\widetilde{\max}}, \tau^-_\beta, \tau^+_\psi\} \leq T\right] \leq n^{-10}$ holds for any $T \leq C_{5.5(1)}n$.

5.7 Towards Consensus

We now turn to establishing an upper bound on the consensus time. Specifically, we show that once there remains exactly one strong opinion, the process will reach consensus within $O(\log n/\alpha^{\max}_0)$ rounds with high probability.

Lemma 5.10 (Unique strong opinion leads to consensus). *Suppose $\psi_0 \leq x_\psi$, $\beta_0 \geq 1/2 - x_\beta$, $\alpha^{\max}_0 = \omega(\log n/\sqrt{n})$, and $\eta_0(j) \geq x_\eta$ for all $j \neq I_0$. Then, $\Pr[\tau_{\mathrm{cons}} > T] \leq 1/n$ holds for some $T = O\big(\log n/\alpha^{\max}_0\big)$.*

5.8 Putting All Together

Lemma 5.11. *Let $C > 0$ be any constant. For some*

$$T = \begin{cases} O(\log n) & \left(\text{if } k \leq \frac{C\sqrt{n}}{(\log n)^2}\right), \\ O\big(\sqrt{n}(\log n)^3\big) & (\text{otherwise}), \end{cases} \tag{9}$$

we have $\Pr\left[\psi_T \leq x_\psi \text{ and } \beta_T \geq \frac{1}{2} - x_\beta \text{ and } \alpha^{\max}_T \geq \frac{(\log n)^{1.5}}{\sqrt{n}}\right] \geq 1 - p_\perp - O\Big(\frac{\log n}{n}\Big)$.

Proof. Combining Lemmas 5.2 to 5.4, we have the following:

(1) From Lemma 5.3 (Item 1) and the definition of $p_\perp$, $\psi_1 \leq x_\psi$ and $\beta_1 \geq 1/n$ hold with probability at least $1 - p_\perp - n^{-\Omega(1)}$.
(2) From Lemma 5.2 (Items 1 and 2), for some $T = O(\log n)$, $\beta_{1+T} \geq 1/2 - x_\beta$ holds with probability at least $1 - O(\log n/n)$.
(3) From Lemma 5.3 (Item 2), $\max_{t\in[n^2]} \psi_t \leq x_\psi$ holds with probability at least $1 - n^{-10}$.
(4) From Lemma 5.4, for some $T' = O\big(\sqrt{n}(\log n)^3\big)$, $\gamma_{1+T+T'} \geq (\log n)^2/\sqrt{n}$, $\psi_{1+T+T'} \leq x_\psi$, and $\beta_{1+T+T'} \geq 1/2 - x_\beta$ with probability at least $1 - n^{-\Omega(1)}$.

Since $\alpha^{\max}_t \geq \gamma_t/\beta_t \geq \gamma_t$ holds for any t, combining Items 1 to 4, we obtain the claim for general k.

For the case where $k = O\big(\sqrt{n}/(\log n)^2\big)$, for any t, $\alpha^{\max}_t \geq \beta_t/k = \Omega(\beta_t(\log n)^2/\sqrt{n})$ holds. Hence, from Items 1 to 3, we obtain the claim. □

Lemma 5.12. *Suppose that $\psi_0 \leq x_\psi$, $\beta_0 \geq 1/2 - x_\beta$, and $\alpha^{\max}_0 = \omega(\log n/\sqrt{n})$. Then, $\tau_{\mathrm{cons}} = O\big(\log n/\alpha^{\max}_0\big)$ with high probability.*

Proof. We have the following:

- From Lemma 5.9 (Item 1), for some $T_1 = O\big(\log n/\alpha^{\max}_0\big)$, we have $\min_{j\neq I_{T_1}} \eta_{T_1}(j) \geq x_\eta$, $\beta_{T_1} \geq 1/2 - x_\beta$, $\psi_{T_1} \leq x_\psi$, and $\alpha^{\max}_{T_1} = \Omega(\alpha^{\max}_0) = \omega(\log n/\sqrt{n})$ with probability at least $1 - n^{-10}$.
- From Lemma 5.10, for some $T_2 = O\big(\log n/\alpha^{\max}_0\big)$, we have $\tau_{\mathrm{cons}} \leq T_1 + T_2$ with probability at least $1 - 1/n$.

Thus, we obtain the claim. □

Proof of Theorem 1.1. From Lemma 5.11, we have that $\psi_T \leq x_\psi$ and $\beta_T \geq 1/2 - x_\beta$ and $\alpha^{\max}_T \geq (\log n)^{1.5}/\sqrt{n}$ hold with high probability for some T as defined in Eq. (9) (Lemma 5.11). Then, from Lemma 5.12, we reach a consensus within additional $O\Big(\frac{\log n}{\alpha^{\max}_T}\Big) = O\Big(\min\Big\{k\log n, \sqrt{n/\log n}\Big\}\Big)$ rounds with high probability. Here, we use $\alpha^{\max}_T \geq \frac{\beta_T}{k} = \Omega(k)$ if k is small and $\alpha^{\max}_T \geq (\log n)^{1.5}/\sqrt{n}$ if k is large. Therefore, the consensus time is bounded by

$$\tau_{\mathrm{cons}} \leq T + O\Big(\min\Big\{k\log n, \sqrt{n/\log n}\Big\}\Big) = \widetilde{O}(\min\{k, \sqrt{n}\})$$

and obtain the claim. □

Acknowledgments

Nobutaka Shimizu is supported by JSPS KAKENHI Grant Number 23K16837. Takeharu Shiraga is supported by JSPS KAKENHI Grant Number 23K16840, and JST CRONOS Grant Number JPMJCS24K2. Colin Cooper is supported by a Mercator Fellowship from DFG Project 491453517 at the University of Hamburg. We thank the anonymous reviewers for their helpful comments and suggestions.

References

[1] Talley Amir, James Aspnes, Petra Berenbrink, Felix Biermeier, Christopher Hahn, Dominik Kaaser, and John Lazarsfeld. 2023. Fast Convergence of k-Opinion Undecided State Dynamics in the Population Protocol Model. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)* (Orlando, FL, USA) *(PODC '23)*. Association for Computing Machinery, New York, NY, USA, 13–23. doi:10.1145/3583668.3594589

[2] Dana Angluin, James Aspnes, Zoë Diamadi, Michael J Fischer, and René Peralta. 2006. Computation in Networks of Passively Mobile Finite-State Sensors. *Distributed Computing* 18, 4 (2006), 235–253. doi:10.1007/s00446-005-0138-3

[3] Dana Angluin, James Aspnes, and David Eisenstat. 2008. A simple population protocol for fast robust approximate majority. *Distributed Computing* 21 (2008), 87–102. Issue 2. doi:10.1007/s00446-008-0059-z

[4] Gregor Bankhamer, Petra Berenbrink, Felix Biermeier, Robert Elsässer, Hamed Hosseinpour, Dominik Kaaser, and Peter Kling. 2022. Fast Consensus via the Unconstrained Undecided State Dynamics. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) (Proceedings)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 3417–3429. doi:10.1137/1.9781611977073.135

[5] Luca Becchetti, Andrea Clementi, and Emanuele Natale. 2020. Consensus Dynamics. *ACM SIGACT News* 51 (2020), 58–104. Issue 1. doi:10.1145/3388392.3388403

[6] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Riccardo Silvestri. 2015. Plurality consensus in the gossip model. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) (SODA '15)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 371–390. https://dl.acm.org/doi/pdf/10.5555/2722129.2722156

[7] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, Riccardo Silvestri, and Luca Trevisan. 2017. Simple Dynamics for Plurality Consensus. *Distributed Computing* 30 (2017), 293–306. Issue 4. doi:10.1007/s00446-016-0289-4

[8] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Luca Trevisan. 2016. Stabilizing consensus with many opinions. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) (SODA '16')*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 620–635. https://dl.acm.org/doi/10.5555/2884435.2884481

[9] Petra Berenbrink, Felix Biermeier, and Christopher Hahn. 2024. *Undecided State Dynamics with Stubborn Agents*. Technical Report 2406.07335. arXiv. doi:10.48550/arXiv.2406.07335

[10] Petra Berenbrink, Andrea Clementi, Robert Elsässer, Peter Kling, Frederik Mallmann-Trenn, and Emanuele Natale. 2017. Ignore or Comply? On Breaking Symmetry in Consensus. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)* (Washington, DC, USA) *(PODC '17)*. Association for Computing Machinery, New York, NY, USA, 335–344. doi:10.1145/3087801.3087817

[11] Petra Berenbrink, Tom Friedetzky, George Giakkoupis, and Peter Kling. 2016. Efficient Plurality Consensus, or: The Benefits of Cleaning Up from Time to Time. In *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP)*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 136:1–136:14. doi:10.4230/LIPICS.ICALP.2016.136

[12] Andrea Clementi, Mohsen Ghaffari, Luciano Gualà, Emanuele Natale, Francesco Pasquale, and Giacomo Scornavacca. 2018. A Tight Analysis of the Parallel Undecided-State Dynamics with Two Colors. In *Proceedings of the International Symposium on Mathematical Foundations of Computer Science (MFCS)*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 28:1–28:15. doi:10.4230/LIPIcs.MFCS.2018.28

[13] Anne Condon, Monir Hajiaghayi, David Kirkpatrick, and Ján Maňuch. 2017. Simplifying Analyses of Chemical Reaction Networks for Approximate Majority. In *DNA Computing and Molecular Programming*. Springer International Publishing, Cham, 188–209. doi:10.1007/978-3-319-66799-7_13

[14] Colin Cooper, Frederik Mallmann-Trenn, Tomasz Radzik, Nobutaka Shimizu, and Takeharu Shiraga. 2025. Asynchronous 3-majority dynamics with many opinions. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 4095–4131. doi:10.1137/1.9781611978322.140

[15] Colin Cooper, Frederik Mallmann-Trenn, Tomasz Radzik, Nobutaka Shimizu, and Takeharu Shiraga. 2026. *Undecided State Dynamics with Many Opinions*. Technical Report 2603.02636. arXiv. doi:10.48550/arXiv.2603.02636 Full version of this paper.

[16] Colin Cooper and Nicolás Rivera. 2016. The Linear Voting Model. In *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP)*, Vol. 55. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 144:1–144:12. doi:10.4230/LIPIcs.ICALP.2016.144

[17] Francesco d'Amore, Andrea Clementi, and Emanuele Natale. 2020. Phase Transition of a Non-linear Opinion Dynamics with Noisy Interactions. In *Proceedings of the International Colloquium on Structural Information and Communication Complexity (SIROCCO)*. Springer International Publishing, Cham, 255–272. doi:10.1007/978-3-030-54921-3_15

[18] Benjamin Doerr, Leslie Ann Goldberg, Lorenz Minder, Thomas Sauerwald, and Christian Scheideler. 2011. Stabilizing consensus with the power of two choices. In *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)* (San Jose, California, USA). Association for Computing Machinery, New York, NY, USA, 149–158. doi:10.1145/1989493.1989516

[19] Sharmila Duppala, George Z. Li, Juan Luque, Aravind Srinivasan, and Renata Valieva. 2025. Concentration of Submodular Functions and Read-k Families Under Negative Dependence. In *Proceedings of the Innovations in Theoretical Computer Science Conference (ITCS) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 325)*, Raghu Meka (Ed.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 47:1–47:16. doi:10.4230/LIPIcs.ITCS.2025.47

[20] Antoine El-Hayek, Robert Elsässer, and Stefan Schmid. 2025. An almost tight lower bound for plurality consensus with undecided state dynamics in the population protocol model. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*. ACM, New York, NY, USA, 532–540. doi:10.1145/3732772.3733505

[21] Robert Elsässer, Tom Friedetzky, Dominik Kaaser, Frederik Mallmann-Trenn, and Horst Trinker. 2017. Brief Announcement: Rapid Asynchronous Plurality Consensus. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)* (Washington, DC, USA) *(PODC '17)*. Association for Computing Machinery, New York, NY, USA, 363–365. doi:10.1145/3087801.3087860

[22] Marcelo Matheus Gauy, Anna Abramishvili, Eduardo Colli, Tiago Madeira, Frederik Mallmann-Trenn, Vinícius Franco Vasconcelos, and David Kohan Marzagão. 2025. Voter Model Meets Rumour Spreading: A Study of Consensus Protocols on Graphs with Agnostic Nodes. In *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. Association for Computing Machinery, Richland, SC, 848–857.

[23] Dmitry Gavinsky, Shachar Lovett, Michael Saks, and Srikanth Srinivasan. 2015. A tail bound for read-k families of functions. *Random Structures & Algorithms* 47, 1 (Aug. 2015), 99–108. doi:10.1002/rsa.20532

[24] Mohsen Ghaffari and Johannes Lengler. 2018. Nearly-Tight Analysis for 2-Choice and 3-Majority Consensus Dynamics. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*. ACM, New York, NY, USA, 305–313. doi:10.1145/3212734.3212738

[25] Mohsen Ghaffari and Merav Parter. 2016. A Polylogarithmic Gossip Algorithm for Plurality Consensus. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*. ACM, New York, NY, USA, 117–126. doi:10.1145/2933057.2933097

[26] Yehuda Hassin and David Peleg. 2001. Distributed Probabilistic Polling and Applications to Proportionate Agreement. *Information and Computation* 171, 2 (2001), 248–268. doi:10.1006/inco.2001.3088

[27] Dominik Kaaser, Frederik Mallmann-Trenn, and Emanuele Natale. 2015. On the Voting Time of the Deterministic Majority Process. arXiv:1508.03519 [cs.DC] https://arxiv.org/abs/1508.03519

[28] Richard Karp, Christian Schindelhauer, Scott Shenker, and Berthold Vöcking. 2000. Randomized Rumor Spreading. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS) (FOCS '00')*. IEEE Computer Society, USA, 565. doi:10.1109/SFCS.2000.892324

[29] Johannes Lengler. 2020. *Drift Analysis*. Springer International Publishing, Cham, 89–131. doi:10.1007/978-3-030-29414-4_2

[30] Nobutaka Shimizu and Takeharu Shiraga. 2025. 3-Majority and 2-Choices with Many Opinions. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*. ACM, New York, NY, USA, 207–217. doi:10.1145/3732772.3733511