Yale University

Technical Report

A Population Protocol for Binary Signaling Consensus

Dana Angluin, James Aspnes and Dongqu Chen
Department of Computer Science
Yale University

YALEU/DCS/TR-1527
August 2016

A Population Protocol for Binary Signaling Consensus

Dana Angluin, James Aspnes and Dongqu Chen
Department of Computer Science
Yale University
{angluin, aspnes, chen}@cs.yale.edu

Abstract

We study a simple population protocol for consensus using *binary signaling*, where the communication in each interaction is limited to a single bit transmitted from the initiator to the responder. We prove that with high probability the three-state binary signaling population protocol reaches consensus after $\Theta(n \log n)$ interactions in the worst case, regardless of the initial configuration. Furthermore, we show that this protocol correctly computes the majority value if the initial margin is $\omega(\sqrt{n \log n})$ with high probability, and is able to tolerate $o(\sqrt{n})$ Byzantine agents in the population, making it the minimal protocol for fast and robust approximate majority. In the general case, a continuous-time binary signaling process in the limit will converge within $O(r \log nr)$ time (corresponding to $O(nr \log nr)$ interactions in expectation) if the initial configuration is monotone, where r is the number of confidence levels. In the other direction, we also show a convergence lower bound $\Omega(nr + n \log n)$ on the number of interactions for any $r \geq 2$. Experimental results are presented to support our theoretical results and to provide evidence for some conjectures.

1 Introduction

A population protocol [AAD+06] is where agents may interact in pairs and each individual agent is extremely limited. Then the complex behavior of the system emerges from the rules governing the possible pairwise interactions of the agents. The agents in a population protocol are anonymous, and it is assumed that interactions between agents happen under some kind of a fairness condition.

Angluin et al. [AAE07] introduced a simple population protocol for majority computation. This protocol assigns only three possible states to every agent, including two opposite states and one intermediate state. The essential idea of this protocol is that when two agents with different preferences meet, one drops its preference and enters the intermediate state; an agent at the intermediate state adopts the preference of any biased agent it meets. Nothing happens when two unbiased agents meet. They show that with high probability this protocol reaches convergence within $O(n \log n)$ interactions with a complete interaction graph of n vertices, if the process starts from a biased initial configuration.

Perron et al. [PVV09] analyzed the continuous-time process of Angluin et al.'s three-state protocol by studying the differential equation system modeling the expected change of the protocol. An additional continuous-time three-state protocol is defined where instead of being passive, a blank agent acts as in the two opposite states uniformly at random. The authors gave an elegant upper bound on the convergence time of a differential equation approximation that converges to the behavior of the discrete process for any fixed time in the limit by Kurtz's theorem [Kur81]. They claim the stronger result that this approximation converges for time $\Theta(\log n)$. While this claim may in fact be true, applying Kurtz's theorem in this case requires an unjustified interchange of

limits that gives incorrect results in many cases. To avoid this issue, we employ a potential-function approach similar to that used by Angluin et al. [AAE07]. For more related works and theoretical background about population protocols we refer to the survey of Aspnes and Ruppert [AR09].

In this paper we study a population protocol for binary signaling consensus, where two interacting agents communicate with only one binary bit, without knowing each other's state or identity. One scenario of binary signaling consensus is as a model of the language emergence process in a human society, i.e., the process of how people learn and acquire a language from interactions [RYC14, KDG07, GW94]. Given the connection between population protocols and biological systems [CCN12], more potential applications of binary signaling consensus may be found in biology.

We show that with high probability, the three-state binary signaling protocol converges after $\Theta(n \log n)$ interactions, regardless of the initial configuration. Furthermore, we study the general binary signaling consensus protocol with any resistance $r \geq 2$. We prove that the continuous-time binary signaling process with large r in the limit will reach consensus within $O(r \log nr)$ time (corresponding to $O(nr \log nr)$ interactions in expectation) if the initial configuration is monotone. We also provide a convergence lower bound of $\Omega(nr + n \log n)$ in the general case. Experimental results are presented to support our theoretical results and to provide evidence for some conjectures.

One major interest people have in a population protocol is its capability of majority computation. Angluin et al. prove that the three-state protocol in [AAE07] converges to the initial majority value with high probability if the initial margin is $\omega(\sqrt{n} \log n)$. In addition, this protocol is able to tolerate Byzantine behavior in $o(\sqrt{n})$ of the agents.

We show that not only does the three-state binary signaling population protocol converge quickly, but it also correctly converges to the majority value of population with high probability, and this property is robust against adversarial Byzantine agents. This means we successfully reduce the signaling from ternary to binary, without losing the power on computing approximate majority. Note that we can't reduce further, neither on signaling nor on the state space. Reducing the state space to two makes the population protocol the voter model. The voter model guarantees consensus but doesn't compute majority, since its error probability is a constant away from 1 [HP01]. Reducing signaling to unitary apparently loses the power of majority computation too. Thus this protocol is the minimal population protocol for fast and robust approximate majority.

2 Binary Signaling Consensus

We consider a population consisting of n agents. Define an *interaction graph* $G = (V, E)$ over this population to be a directed graph with $|V| = n$ whose edges indicate the possible interactions that may take place. Each agent $i \in V$ in the crowd has a *confidence level* $cl(i)$ for a preference, which is an integer between 0 and r. We say r is the *resistance*. At each step, an edge (i, j) is chosen uniformly at random from E. The source agent i is the *initiator*, and the sink agent j is the *responder*. The two agents communicate in a way that the initiator sends a binary bit to the responder. With probability $cl(i)/r$ agent i sends a positive bit to agent j and the latter does the update $cl(j) = \min(cl(j) + 1, r)$. Otherwise the initiator sends a negative bit to the responder who updates $cl(j) = \max(cl(j) - 1, 0)$. Starting from an initial configuration, the communication process keeps going until *convergence*, where either all agents are of confidence level r (i.e., *positive convergence*) or all agents are of confidence level 0 (i.e., *negative convergence*).

We inherit the terminology *binary signaling* used by Perron et al. [PVV09], in the sense that the signaling between the agents is binary and the communication between two interacting agents doesn't depend on the knowledge of their states or identities. In this paper, we study the case where the interaction graph G is a complete graph. For algebraic convenience we assume self-loops

are allowed in the interaction graph, while all our results can be easily applied to the setting of no self-loops as the number of agents n goes to infinity.

The parameter r is called the *resistance* or the recalcitrance as the larger r is, the more difficult to persuade an individual of the opposite opinion. A more general model could allow different agents to have different resistance values. In this setting, the range of the confidence level of agent i is from 0 to $r(i)$. Everything remains the same except that the initiator i has probability $cl(i)/r(i)$ of sending a positive bit and the responder updates $cl(j) = \min(cl(j)+1, r(j))$. Although this is a more general setting, it complicates the model and also violates the anonymity condition in population protocols, where the output of the transition function should be independent of the identities of the two involved agents. Hence, we assume all agents are of the same resistance r.

One application of binary signaling consensus is to model the language emergence process in a human society, which is the process of how people learn and acquire a language from interactions [RYC14, KDG07, GW94]. In this scenario we consider a society of population n. Each person has a confidence level for a grammar (or a language). A person with higher confidence level speaks with this grammar with higher probability. Each individual adjusts her opinion of the grammar while interacting with others. Positive convergence of this process means a new grammar eventually emerges while negative convergence indicates extinction of this grammar. More applications of binary signaling consensus can be found in the study of other fields such as rumor spreading, epidemiology and biological systems.

3 Three-State Binary Signaling Consensus

When studying a population protocol, it is common to assume a very large population n and constant parameter, which in our model is the value of resistance r. Since the $r = 0$ case is trivial and the $r = 1$ model doesn't involve probabilistic interactions, the three-state case with $r = 2$ is a reasonable start for us to study this protocol. In this section, we show that starting from any initial configuration, the three-state protocol will converge within $\Theta(n \log n)$ interactions with high probability. Sketches of the proofs are provided in this section, with most technical details deferred to Appendix A.

Let τ_* be the number of interactions until the three-state binary signaling model reaches consensus. The main result of this section is the following theorem. Note that the stated convergence bound is a worst-case bound.

Theorem 1. *With probability* $1 - o(1)$, $\tau_* = \Theta(n \log n)$ *in the worst case. In addition, for any constant* $c > 0$ *we have*

$$\mathbb{P}\left(\tau_* \geq 96930(c+1)n \log n\right) \leq \max\left(9n^{-c}, \frac{c \log n}{\sqrt[3]{n}}\right)$$

The convergence lower bound $\tau_* = \Omega(n \log n)$ can be easily obtained from the well-known coupon collector bound. When the initial configuration is $cl(i)$ being 1 for all $i \in V$, in order to achieve consensus, every agent must participate in at least one interaction, leading to the coupon collector lower bound. However, the upper bound $\tau_* = O(n \log n)$ requires a substantial amount of work. It may be surprising that fast convergence of such a simple consensus process needs such a lengthy proof. Part of the reason is that we want to obtain exact asymptotic bounds with explicit constants that work for arbitrary configurations.

The core of our proof is to construct a supermartingale for each region in the configuration space. This technique is inspired by the proof used by Angluin et al. [AAE07]. Recall that a *supermartingale* is a discrete stochastic process $\{M_t\}$ where M_t satisfies $\mathbb{E}(|M_t|) < +\infty$ and $\mathbb{E}(M_t \mid$

Indicator	Counter
I_t^{g-}: g decreases by 1	$S_t^{g-} = \sum_{i=1}^{t} I_i^{g-}$
I_t^{g+}: g increases by 1	$S_t^{g+} = \sum_{i=1}^{t} I_i^{g+}$
I_t^{sc}: the configuration is changed	$S_t^{sc} = \sum_{i=1}^{t} I_i^{sc}$
I_t^{c}: $\max(\tilde{b}, \tilde{g}, \tilde{w}) < 3/4$	$S_t^{c} = \sum_{i=1}^{t} I_i^{c}$
I_t^{b}: $\tilde{b} \geq 3/4$	$S_t^{b} = \sum_{i=1}^{t} I_i^{b}$
I_t^{g}: $\tilde{g} \geq 3/4$	$S_t^{g} = \sum_{i=1}^{t} I_i^{g}$
I_t^{w}: $\tilde{w} \geq 3/4$	$S_t^{w} = \sum_{i=1}^{t} I_i^{w}$

Table 1: Indicators and Counters

$M_0, \ldots, M_{t-1}) \leq M_{t-1}$. The expected value of each M_t is bounded by the initial value $\mathbb{E}M_t \leq \mathbb{E}M_0$. Supermartingales are commonly studied with a *stopping time*. A stopping time with respect to a stochastic process $\{M_t\}$ is an almost surely finite random variable τ with positive integer values and the property that the event $\tau = t$ depends only on the values of $M_0, M_1, \ldots, M_t$. A supermatingale with a stopping time is still a supermartingale. In this section, we let $\tau = \min(\tau_*, dn \log n)$ for some fixed constant d. Thus τ is a stopping time. This truncation guarantees that τ and quantities defined in terms of it are finite and well-defined, despite the logical possibility that convergence is not achieved and τ_* is ill-defined.

Now that $r = 2$ and an agent has only three possible states, we denote by w (white), g (gray) and b (black) the states with confidence levels 0 (negative), 1 (neutral) and 2 (positive) respectively. For notational convenience we also overload w, g, b to denote the number of each token in a configuration. Meanwhile, let $\tilde{b} = b/n$, $\tilde{g} = g/n$ and $\tilde{w} = w/n$ be the corresponding proportions. Obviously we always have $\tilde{b} + \tilde{g} + \tilde{w} = 1$. Denote $u = b - w$ and $v = b + w$. Note that $-n \leq u \leq n$ and $0 \leq v \leq n$. The point when $|u| = n$ is equivalent to convergence. The change of basis to u and v allows us to take advantage of the symmetry between b and w tokens. Auxiliary 0-1 indicators and counters for the proof are defined in Table 1.

The key to constructing a supermartingale in a region is to design a proper potential function that drops smoothly inside this region and doesn't increase too much elsewhere. Because the behavior of the consensus process is qualitatively different in different regions, we choose a specific potential function for each region of the configuration space. In our proof, we divide the configuration space into four regions:

1. The corner region where at least $3n/4$ agents are of confidence level 0 and $I^w = 1$;
2. The corner region where at least $3n/4$ agents are of confidence level 1 and $I^g = 1$;
3. The corner region where at least $3n/4$ agents are of confidence level 2 and $I^b = 1$;
4. The central region left where the tokens are more evenly balanced and $I^c = 1$.

More concretely, given that the potential function f decreases consistently by $-\Theta(n^{-1})$ in expectation when $I_t^1 = 1$ and increases by a relatively smaller amount in expectation when $I_t^2 = 1$, we are able to construct a stochastic process of the form $\{M_t = \exp\big((c_1 S_t^1 - c_2 S_t^2)/n\big) \cdot f\}$ which is a supermartingale, where I_t^1 and I_t^2 are two different binary indicators, $S_t^1 = \sum_{i=1}^{t} I_t^1$ and $S_t^2 = \sum_{i=1}^{t} I_t^2$ are their counters, and c_1 and c_2 are two carefully chosen positive constants. The supermartingale property $\mathbb{E}M_\tau \leq \mathbb{E}M_0$ together with Markov's inequality then gives us the desired $O(n \log n)$ upper bound for S_τ^1 (depending on S_τ^2). Here we assume either S_τ^2 is already well bounded (Lemma 9, Lemma 10, Lemma 11 and Lemma 12), or there exists some auxiliary inequality relationship between S_τ^1 and S_τ^2 (Lemma 5 and Lemma 7). A formal statement of this proof technique is presented in Lemma 4.

The proof of the upper bound consists of four components. Notice that $t = S_t^c + S_t^b + S_t^g + S_t^w$

for any time t. Thus upper bounds for S_τ^c, S_τ^b, S_τ^g and S_τ^w imply one for τ. We will later find that these four quantities can be bounded using an upper bound on the number of state-changing interactions S_τ^{sc}. Therefore, the proof starts with an $O(n \log n)$ upper bound for S_τ^{sc}.

In three-state binary signaling consensus, every state-changing interaction must increase or decrease the value of g by 1. Hence, we have $S_\tau^{sc} = S_\tau^{g+} + S_\tau^{g-}$. The proof of bounding $S_\tau^{sc} = O(n \log n)$ (Lemma 3) is done case by case. First we show that if the process starts from some point in the region $\{g \leq \min(b, w)/4\}$, then within $O(n \log n)$ state-changing interactions, it will either converge or leave the region (Lemma 5). If the former happens then we are done. Otherwise, we have $g > \min(b, w)/4$ and we prove that within the next $O(n \log n)$ state-changing interactions, either the process will never enter the region $\{g < \min(b, w)/10\}$ again, or it will enter the region $\{\min(b, w) = O(\log n) \bigwedge g = O(\log n)\}$ (Lemma 6 and Corollary 2). In the first case, we show the population protocol will converge within the next $O(n \log n)$ state-changing interactions (Lemma 7). In the latter case, we show the protocol will converge within the next $O(n)$ state-changing interactions (Lemma 8).

Based on the upper bound on state-changing interactions, we are able to construct a family of supermartingales for different regions in the configuration space. To bound the number of interactions S_τ^c in the central region, we prove the stochastic process $C_t = \exp((S_t^c - 9S_t^{sc})/n)$ to be a supermartingale. The key observation is that in the central region where $\max(\tilde{b}, \tilde{g}, \tilde{w}) < 3/4$, we should have either $\tilde{b}$ and $\tilde{w}$ are both $\geq 1/8$, or $\tilde{g} \geq 1/8$. We then show that in both cases we have C_t dropping in expectation, which implies an $O(n \log n)$ upper bound for S_τ^c. For the corner region where $\tilde{g} \geq 3/4$, we choose the potential function to be $f = 1/(2v + 1)$. We show that this potential function drops consistently by $\Theta(-1/n)$ of its current value in expectation in the large-g region, while its rise when $I_t^g = 0$ can be upper-bounded by $O(I_t^{g+}/n)$. With this we then construct a supermartingale in the form $M = \exp(aS/n)f(b, w)$ as described above and achieve the bound $S_\tau^g = O(n \log n)$. For the corner region where $b \geq 3n/4$, the potential function we use is $f = 3w + g + 1$. Similar to the idea of bounding S_τ^g, we bound S_τ^b by showing the value of the potential function decreases by a factor of $\exp(-\Theta(1/n))$ when b is large, and increases otherwise by an amount we can bound using S_t^{g+} and S_t^{g-}. Thus the number of interactions S_τ^b that happen in the large-b region is also $O(n \log n)$. The number of interactions S_τ^w that happen in the large-w region can be bounded in a symmetric way using the potential function $f = 3b + g + 1$. Finally, for $\tau = S_\tau^c + S_\tau^b + S_\tau^g + S_\tau^w$, summing the bounds for all the four regions we will obtain a bound on the total number of interactions. Given a convergence upper bound $O(n \log n)$ with an explicit constant c, we then choose a slightly larger constant $d > c$ to truncate the process and let $\tau = \min(\tau_*, dn \log n)$ to make τ a well defined stopping time. Some readers might think this truncation at $\Theta(n \log n)$ interactions already assumes the correctness of the target statement, but we have proved that the total number of interactions is smaller than $dn \log n$ with high probability so we have the convergence upper bound $\tau_* = O(n \log n)$ as stated in Theorem 1.

4 Fast and Robust Approximate Majority

The ability to compute approximate majority is one important property of a population protocol. In this section, we show that with high probability the three-state binary signaling population protocol correctly computes the majority value when the initial margin is $\omega(\sqrt{n \log n})$, and is able to tolerate $o(\sqrt{n})$ Byzantine agents in the population.

Note that we achieve the same or slightly improved bounds compared with the three-state population protocol presented in [AAE07]. This means we successfully reduce the signaling from ternary to binary, without losing the power on computing approximate majority. Also note that we

can't reduce further, neither on signaling nor on the state space. Reducing the state space to two makes the population protocol the voter model. The voter model guarantees consensus but doesn't compute majority, since its error probability is $(1 - c)$, a constant away from 1 [HP01]. Reducing signaling to unitary apparently loses the power of majority computation too. Thus in this sense our protocol is the minimal population protocol for fast and robust approximate majority.

Theorem 2. *The binary signaling protocol converges to the initial majority value with high probability, if the initial difference between the majority and the minority population is* $\omega(\sqrt{n \log n})$.

The proof is done by constructing a martingale process that provides a lower bound of the difference between the majority population and the minority population. We show that at the point of convergence, this difference is positive with high probability, which means the protocol correctly computes the initial majority value. The detailed proof is deferred to Appendix B.

Theorem 2 shows the correctness of approximate majority, while Theorem 3 and Theorem 4 present the robustness against adversarial Byzantine agents. A Byzantine agent can pretend to be in any normal state in an interaction. The signaling from Byzantine agents is not predictable and can depend on both the global configuration and the identity of the specific agent it interacts with. Due to the adversarial behaviors of Byzantine agents in the worst case, there is always non-zero probability of Byzantine agents pulling the normal population away from convergence. Thus we have to slightly relax the criterion for convergence, by allowing an $O(\sqrt{n})$ gap from the complete converged configuration. We prove that the binary signaling protocol is able to tolerate $o(\sqrt{n})$ Byzantine agents with high probability.

Theorem 3. *Let* τ *be the time when* $b \geq n - \sqrt{n}$, $w \geq n - \sqrt{n}$ *or* $b + w \leq \sqrt{n}$ *first holds. If the number of Byzantine agents in the population is* $o(\sqrt{n})$ *and initially* $b_0 + w_0 \geq \sqrt{n} + c \log_{7/5} n$, *then*

$$\mathbb{P}\left(\tau \geq 96930(c+1)n \log n\right) \leq \max\left(n^{-c+o(1)}, \frac{c \log n}{\sqrt[3]{n}}\right), \text{ and } \mathbb{P}(b_\tau + w_\tau \leq \sqrt{n}) \leq n^{-c}$$

for any constant $c > 0$.

We start from showing that in spite of the existence of Byzantine agents, there is a strong bias pushing the process away from the large-g corner. More concretely, it's unlikely to have $v \leq \sqrt{n}$ within any polynomial number of interactions (Lemma 13). However, once the process is in the large-b corner or the large-w corner, it will remain there within any polynomial number of interactions with high probability. We prove this by showing that $Z_t = \exp((2g_t + 5w_t)/16)$ is a supermartingale process for the large-b case, such that it is difficult to have $2g + 5w \geq 5\sqrt{n}$ within any polynomial time (Lemma 14). The large-w case can be achieved in the symmetric way. Switching to the Byzantine case, we still keep the potential functions used for the non-Byzantine case in Theorem 1, but include an adjustment factor that compensates for increases due to Byzantine interactions. We show that the effect of this adjustment factor is small and the remaining quantities are at most $n^{o(1)}$ times their original values with high probability. Eventually combining all the results gives us Theorem 3.

The correctness of Theorem 4 follows Theorem 3. Appendix C presents a complete proof of Theorem 3 and Theorem 4.

Theorem 4. *If the number of Byzantine agents in the population is* $o(\sqrt{n})$ *and initially* $b_0 + w_0 \geq \sqrt{n} + c \log_{7/5} n$, *and the initial difference between the majority and the minority is* $\omega(\sqrt{n \log n})$, *the binary signaling protocol converges to the initial majority value with high probability.*

Another important variant of approximate majority is the model with an epidemic-triggered start. In addition to the confidence level, each agent has an active/inactive bit. Active agents interact as before, while an inactive agent wakes up only after involved into an interaction with an active agent, and all other interactions have no effect. This is important to the application to the register machine simulation, where the signal to start the next operation is broadcast from the leader via an epidemic. We show that the binary signaling protocol maintains the correctness of computing majority with an epidemic-triggered start if there is a large enough initial majority.

Theorem 5. *If the initial difference between the majority and the minority population is* $\omega((n \log n)^{3/4})$ *and there is exactly one initially active agent, then the binary signaling protocol with an epidemic-triggered start converges to the initial majority value with high probability.*

We divide the process into two stages, by the point when there are $\left(n^{3/4}(\log n)^{-1/4}\right)$ active agents in the population. We show that there are at most $O(\sqrt{n \log n})$ active-active interactions in the first stage with high probability, and the majority enjoys an advantage of $\omega(\sqrt{n \log n})$ to start the second stage. Thus we can apply the same analysis from the proof of Theorem 2. A formal proof is shown in Appendix D.

5 Binary Signaling Consensus with $r > 2$

In the previous section we studied the population protocol for binary signaling consensus with $r = 2$. This is a reasonable start for understanding binary signaling consensus process, but to gain an insight into the population protocol in depth, we have to investigate the general binary signaling consensus process with larger r.

In this section we allow the value of resistance r to be arbitrarily large, i.e., not necessarily a small constant. Denote by n_i the number of agents of confidence level i and by $x_i = n_i/n$ the corresponding proportion. Any configuration over the population can be represented as a $(r+1)$-dimensional vector $\vec{x} \in [0,1]^{r+1}$ where $\sum_{i=0}^{r} x_i = 1$. Denote by $p = \sum_{i=0}^{r}(i/r)x_i$. We say an interaction is a *positive interaction* if the initiator sends a positive bit, and is a *negative interaction* otherwise. Then p is the probability of occurrence of a positive interaction. The curve of p serves as a significant indicator of the underlying status of the society. A large p implies the preference is almost accepted and $p = 1$ is equivalent to positive convergence. A small p indicates the preference is close to extinction and $p = 0$ is equivalent to negative convergence. If we expect a positive convergence, then a positive interaction is never harmful while a negative interaction never helps, and vice versa for negative convergence.

Unfortunately, rigorous and comprehensive analysis of large-r case turns out to be rather difficult. This is not surprising given that even the proof for the three-state binary signaling consensus is already very lengthy. The increase of degrees of freedom with large r leads to high dimensionality of the configuration space and makes the process more unpredictable. One path of p could correspond to a large number of possible hidden configuration sequences, which does not permit us to generalize the potential functions in Section 3 to large-r case. In addition, the fact that the corresponding systems of differential equations do not have closed-form solutions (even for the $r = 2$ case) rules out arguments based on techniques involving reduction to a continuous process in the limit. In fact we will see later an essential difference between the $r = 2$ case and the $r > 2$ case. In the $r = 2$ case p is always increasing or always decreasing in the limit, but the curve of p in the $r > 2$ case doesn't have this nice property and is more unpredictable. This intrinsic difference is one indication of that we should expect more difficulties in analyzing the large-r case.

5.1 Continuous-time binary signaling consensus

When the gap between the discrete time steps in the model goes to zero in the limit, the communication process becomes continuous-time. To study this continuous process, we use the asynchronous timing defined by Boyd et al. [BGPS06]. Each agent in the population has a clock which ticks at the times of a Poisson process of rate r. The inter-tick times at each agent are exponentials of rate r, independent across agents and over time. Equivalently, this corresponds to a single clock ticking according to a Poisson process of rate nr at time t_k, $k \geq 1$, where $\{t_{k+1} - t_k\}$ are i.i.d. exponentials of rate nr. At time t_k, an edge (i, j) is chosen uniformly at random from E and the two chosen agents interact as defined in the protocol.

Note that the continuous process can be arbitrarily close to but never reaches complete consensus where $p = 0$ or 1. A direct reason is that the derivative of p goes to 0 as the process approaches to convergence. Therefore, instead of entire convergence, we redefine *consensus* for the continuous process to be the region where $\min(p, 1-p) = O(1/(nr))$, which is the closest point the process can achieve to complete convergence. We say a configuration is *monotone* if it has $x_0 \leq x_1 \leq \ldots \leq x_r$ with at least one $<$ in the middle, or $x_0 \geq x_1 \geq \ldots \geq x_r$ with at least one $>$ in the middle. The set of all monotone configuration is called the monotone region. In this subsection we will show the fast convergence to consensus of the continuous process inside the monotone region.

Theorem 6. *If the initial configuration is monotone, then the continuous binary signaling process will reach consensus within $O(r \log nr)$ time.*

The proof starts with derivation of the corresponding ODE system of the process, which can be inferred by taking the limit of the expectation of the configuration vector. This ODE system provides a mathematical formula of the vector field in the configuration space. We show that the vector field anywhere at the boundary of the monotone region always points inwards into the monotone region, which means the process stays in the monotone region and never leaves. We divide the monotone region into two sub-areas A_+, the region where $x_0 \leq x_1 \leq \ldots \leq x_r$ with at least one $<$ in the middle, and A_-, the region where $x_0 \geq x_1 \geq \ldots \geq x_r$ with at least one $>$ in the middle. The ODE system also gives us the differential equation for p, from which we prove that p is always increasing in A_+ and is always decreasing in A_-. It suffices to show the convergence bound for A_+, as it holds for A_- symmetrically.

The above two facts already tell us that once the process enters A_+, p will keep increasing until convergence. What we need is a positive lower bound for the derivative of p that will lead to the desired convergence time. We need to take care of two cases where $\mathrm{d}p/\mathrm{d}t$ is very small. The first case is when the process is almost at convergence and p is very close to 1. The other is when the configuration vector is almost uniform and p is very close to $1/2$. To do so, we divide the path of p from $1/2 + 1/(nr)$ to $1 - 1/(nr)$ into two corresponding stages: from $2/3$ to $1 - 1/(nr)$ and from $1/2 + 1/(nr)$ to $2/3$. We show the time for the former stage is $O(\log nr)$ and the time for the latter is $O(r \log nr)$. All technical details are deferred to Appendix E.

We have bounded the convergence time for the monotone region. To achieve a complete bound for the whole configuration space, we need either a convergence bound for the non-monotone region separately if the process can stay in the non-monotone region, or to bound the time until the process enters the monotone region and show this always happens. Empirical results presented in Section 6 suggest that the process will eventually enter the monotone region regardless of the initial configuration and that the time needed for this to happen is short (see Conjecture 4), which indicates bounding the convergence time in the monotone region will be essential to the general bound for the entire configuration space. This is why the monotone case is interesting to us.

When the resistance $r = 2$ or $r = O(1)$, the convergence time is $O(\log n)$ and the rate of the

clock is $\Theta(n)$, so the total number of ticks of the clock is $O(n \log n)$, which matches our result for three-state binary signaling consensus in Section 3.

The analysis of the continuous process above gives us the following lemma (see Appendix E for a formal proof).

Lemma 1. *When $r = 2$, p is always increasing when $p > 1/2$ and is always decreasing when $p < 1/2$. This doesn't hold for any $r > 2$.*

Therefore, when $r = 2$ the probability of positive interaction p is always pushed towards convergence in the correct direction, but in the $r > 2$ case the change of p is more unpredictable. This shows an intrinsic difference between the $r = 2$ case and the $r > 2$ case.

5.2 A convergence lower bound

Although convergence upper bounds are our primary interest in the population protocol for binary signaling consensus, in this subsection we study the general protocol in another direction and prove a convergence lower bound on the number of interactions. Recall that for the three-state population protocol, the convergence lower bound $\Omega(n \log n)$ is an immediate result from the well-known coupon collector's bound, because when the initial configuration is $cl(i) = 1$ for all $i \in V$, every agent must participate in at least one interaction in order to achieve consensus. Likewise, to bound the number of interactions for the $r > 2$ case, we consider a generalized version of the coupon collector problem. An r-coupon collector is where instead of collecting at least one copy for each type of coupon, we need to keep drawing coupons until we have collected at least r copies for each type of coupon. The number of steps an $(r/2)$-coupon collector takes gives a convergence lower bound for the general binary signaling consensus process, as every agent must participate in at least $r/2$ interactions before convergence, when the initial configuration of the population is $x_{r/2}$=1 and $x_i = 0$ for all $i \neq r/2$.

Another important reason for us to study the r-coupon collector problem here is the inspiration from the three-state population protocol that the convergence bound of the binary signaling consensus process is exactly the tight bound of coupon collector. This fact leads to our conjecture that this connection also holds for $r > 2$ (see Conjecture 2 in Section 6). To the best of our knowledge, there exists no direct result for this generalized coupon collector problem so we prove the bound here. Since we are only interested in the magnitude, we will consider an r-coupon collector instead of an $(r/2)$-coupon collector, for algebraic convenience.

Theorem 7. *An r-coupon collector needs $\Theta(nr + n \log n)$ steps with high probability.*

To prove this bound, we consider the equivalent balls-in-bins problem: if we keep throwing balls uniformly at random into n bins, how many balls do we need to throw such that every bin has at least r balls with high probability? Let N be the answer to this question. The proof is easy for $r = O(1)$, by doing at most r rounds of classic coupon collector to fill the bins. For $r = \omega(1)$, the proof is done by using Poisson approximation. Let Y be the minimum load among the n bins, which is the minimum among n i.i.d. Poisson random variables with mean N/n in Poisson approximation. We show case by case, depending on the magnitude of r, that we can always find an $N_0 = \Theta(nr + n \log n)$ such that $\mathbb{P}(Y < r)$ goes to zero after throwing N_0 balls. Because $\mathbb{P}(Y < r)$ is monotonically decreasing in N, all $N \geq N_0$ have $\mathbb{P}(Y < r) \to 0$. Therefore, we have $N \leq N_0 = O(nr + n \log n)$. A detailed proof is presented in Appendix F.

An intuitive interpretation of this bound is that we throw the first $\Theta(nr)$ balls to have all the bins almost full, and after that the last stage is to wait for these almost-full bins to be eventually

full, which is a classic coupon collector. The r-coupon collector gives a convergence lower bound for the binary signaling consensus process.

Corollary 1. *With high probability, a binary signaling consensus process needs* $\Omega(nr + n\log n)$ *interactions to converge.*

6 Empirical Results and Conjectures

To support our theoretical results, in this section we present a series of empirical results, based on which we propose several conjectures for different aspects of the binary signaling consensus process. To be more robust against fluctuation from randomness, each test was run for ten times and the medians were taken. All figures for this section are presented in Appendix G.

The experiments start from verifying the fast convergence result for three-state binary signaling consensus process. We study two groups of experiments with different initial configurations. Group 1 is a society starting with everyone in the intermediate state. Group 2 is a society with initial balanced configuration where half of population supports the preference with full confidence while the other half is in the opposite state. These are two worst cases that are expected to have the longest convergence time and are ideal for examining convergence upper bound. We fix the resistance r as 2 and vary the number of agents n. The results are plotted in Figure 1 with the curves of convergence time of the two groups respectively. These two curves indicate the population in group 2 converges slightly slower than the one in group 1. To verify the order of the convergence time and estimate the concrete constant, we divide the number of interactions by $n\log n$ and also show this quotient on the plot. From the results we can see this quotient is stable around 5. This is supportive evidence of our theoretical results on the order of convergence rate. However, the constant we provided in Theorem 1 seems too large, as the experiments suggest this constant be 5, or conservatively speaking, smaller than 10, which leads to our first conjecture.

Conjecture 1. *With high probability, the number of interactions for a population with resistance* 2 *to reach consensus is at most* $10 \cdot n\log n$ *for all sufficiently large* n*.*

What interests us more is the large-r case, for the questions we are not able to answer theoretically. We have noticed that the convergence bound $\Theta(n\log n)$ of the three-state population protocol for binary signaling consensus is exactly the tight bound of the coupon collector problem. This inspires us that the tight bound of the r-coupon collector might also indicate (or at least approximate) the convergence bound of large-r binary signaling consensus. In Section 5.2 we have shown $\Theta(nr + n\log n)$ is a tight bound for the r-coupon collector process. Thus it is reasonable for us to conjecture $\Theta(nr + n\log n)$ as the convergence bound in the large-r case.

Conjecture 2. *With high probability, the number of interactions for a population with resistance* r *to reach consensus in the worst case is* $\Theta(nr + n\log n)$*.*

We seek empirical evidence to support this conjecture. Since the bound $\Theta(nr+n\log n)$ involves both r and n, we conduct two sets of experiments with fixed r (shown in Figure 2) and with fixed n (shown in Figure 3) respectively. With fixed r and varying n, we expect the number of interactions to increase in the order of $\Theta(n\log n)$. In Figure 2, we fix the resistance r as 50 and vary the population n. The four curves are plotted as in the previous experiments with $r = 2$ and have similar shapes. The population protocol converges obviously slower with $r = 50$ than with $r = 2$. As expected, the constant is also larger with larger fixed r. For group 1 the quotient is stable around 28 and for group 2 it is around 33. The process in group 2 is still slower to converge than

the one in group 1 but the difference is now more apparent. Hence, the behaviors of the curves match what our conjecture predicts. Figure 3 shows the curves of convergence time when we fix the population n as 1000 and vary the resistance r. The same four curves are plotted and the only difference is now we divide the number of interactions by r, as we expect the convergence time to be $\Theta(r)$ with fixed n. Group 1 is still faster than group 2 in the sense of convergence rate and also with smaller constant, which is stable around 5000 while the constant of group 2 is about 8200. These large constants are not surprising since all the values of n and r we choose for this set are quite large. Again these results agree with the prediction of our conjecture.

In Section 5.1 we studied the continuous-time process in the limit and showed that the convergence time is $O(r \log nr)$ if it starts from a monotone initial configuration. However, the behavior of the process outside the monotone region is still uncertain. Fortunately, empirical simulation suggests the process will enter the monotone region fast enough and then go to convergence rapidly. We simulate the continuous-time process using the Runge-Kutta method with $n = 100000$. In order to show the process will eventually enter the monotone region with any initial configuration, we conduct more groups of simulations with different types of initial configuration. Figure 4 demonstrates the experimental results in the form of a bar chart to compare the time in the non-monotone region and the time in the monotone region. The initial setup of each group is as follows.

Group 1: 40% of the population at confidence level 0 and 60% at confidence level r;

Group 2: $1/2 - 1/(nr)$ of the population at level 0 and $1/2 + 1/(nr)$ at level r;

Group 3: 0.1% of the population at confidence level 0 and 99.9% at confidence level r;

Group 4: 50% of the population at confidence level 1 and 50% at confidence level r;

Group 5: 40% of the population at confidence level 1 and 60% at confidence level r.

The first group is designed for the majority computation scenario. Group 2 and group 3 are to show the process will enter the monotone region first before convergence regardless of whether the population is almost balanced (group 2) or almost converged (group 3). As expected, group 2 is the slowest to converge while group 3 is the fastest. Group 4 and group 5 are designed to witness the drop of p in the $p > 1/2$ region, which is an essential difference between the $r = 2$ case and the $r > 2$ case (Lemma 1). From these results we propose the following conjecture.

Conjecture 3. *A continuous-time binary signaling process will enter the monotone region before convergence starting from any initial configuration.*

The bar chart also suggests the time needed to enter the monotone region doesn't dominate the whole process. Thus it is reasonable to conjecture that the total convergence time is of the same order as the convergence time inside the monotone region we presented in Theorem 6.

Conjecture 4. *A continuous-time binary signaling process converges within $O(r \log nr)$ time.*

References

[AAD+06] Dana Angluin, James Aspnes, Zoë Diamadi, Michael J Fischer, and René Peralta. Computation in networks of passively mobile finite-state sensors. *Distributed computing*, 18(4):235–253, 2006.

[AAE07] Dana Angluin, James Aspnes, and David Eisenstat. A simple population protocol for fast robust approximate majority. In *Distributed Computing*, pages 20–32. Springer, 2007.

[AAE08] Dana Angluin, James Aspnes, and David Eisenstat. Fast computation by population protocols with a leader. *Distributed Computing*, 21(3):183–199, 2008.

[AR09] James Aspnes and Eric Ruppert. An introduction to population protocols. In *Middleware for Network Eccentric and Mobile Applications*, pages 97–120. Springer, 2009.

[BGPS06] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE/ACM Transactions on Networking (TON)*, 14(SI):2508–2530, 2006.

[CCN12] Luca Cardelli and Attila Csikász-Nagy. The cell cycle switch computes approximate majority. *Scientific reports*, 2, 2012.

[GW94] Edward Gibson and Kenneth Wexler. Triggers. *Linguistic inquiry*, pages 407–454, 1994.

[HP01] Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2001.

[KDG07] Simon Kirby, Mike Dowman, and Thomas L Griffiths. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104(12):5241–5245, 2007.

[Kur81] Thomas G Kurtz. *Approximation of population processes*, volume 36. Society for Industrial and Applied Mathematics, 1981.

[PVV09] Etienne Perron, Dinkar Vasudevan, and Milan Vojnović. Using three states for binary consensus on complete graphs. In *INFOCOM 2009, IEEE*, pages 2527–2535. IEEE, 2009.

[RYC14] Russell Richie, Charles Yang, and Marie Coppola. Modeling the emergence of lexicons in homesign systems. *Topics in cognitive science*, 6(1):183–195, 2014.

Appendix A Proof of Theorem 1

In this section we present a formal proof of the $O(n \log n)$ convergence bound on the three-state binary signaling consensus process.

Theorem 1 (in the main paper) *With probability* $1 - o(1)$, $\tau_* = \Theta(n \log n)$ *in the worst case. In addition, for any constant* $c > 0$ *we have*

$$\mathbb{P}\left(\tau_* \geq 96930(c+1)n \log n\right) \leq \max\left(9n^{-c}, \frac{c \log n}{\sqrt[3]{n}}\right)$$

The convergence lower bound $\tau_* = \Omega(n \log n)$ can be easily obtained from the well-known coupon collector bound. When the initial configuration is $cl(i)$ being 1 for all $i \in V$, in order to achieve consensus, every agent must participate in at least one interaction, leading to the coupon collector lower bound.

Lemma 2. *With probability* $1 - o(1)$, $\tau_* = \Omega(n \log n)$ *in the worst case.*

However, the upper bound $\tau_* = O(n \log n)$ requires a substantial amount of work.

A.1 Bounding $S_\tau^{sc} = O(n \log n)$

In this subsection we show the number of state-changing interactions S_τ^{sc} is at most $O(n \log n)$ with high probability. In the three-state model, every state-changing interaction must increase or decrease the value of g by 1. Hence, we have $S_\tau^{sc} = S_\tau^{g+} + S_\tau^{g-}$ with the following upper bounds.

Lemma 3. *With probability* $1 - o(1)$, $S_\tau^{sc} = O(n \log n)$. *In addition, for any constant* $c > 0$ *we have*

$$\mathbb{P}\left(S_\tau^{sc} \geq 372.72(c+1)n \log n\right) \leq \max\left(5n^{-c}, \frac{c \log n}{\sqrt[3]{n}}\right)$$

$$\mathbb{P}\left(S_\tau^{g+} \geq 186.36(c+1)n \log n\right) \leq \max\left(4n^{-c}, \frac{c \log n}{\sqrt[3]{n}}\right)$$

and

$$\mathbb{P}\left(S_\tau^{g-} \geq 186.36(c+1)n \log n\right) \leq \max\left(4n^{-c}, \frac{c \log n}{\sqrt[3]{n}}\right)$$

The proof of Lemma 3 is done case by case. First we show that if the process starts from some point in the region $\{g \leq \min(b, w)/4\}$, then within $O(n \log n)$ state-changing interactions, it will either converge or leave the region (Lemma 5). If the former happens then we are done. Otherwise, we have $g > \min(b, w)/4$ and we prove that within the next $O(n \log n)$ state-changing interactions, either the process will never enter the region $\{g < \min(b, w)/10\}$ again, or it will enter the region $\{\min(b, w) = O(\log n) \bigwedge g = O(\log n)\}$ (Lemma 6 and Corollary 2). In the first case, we show the population protocol will converge within the next $O(n \log n)$ state-changing interactions (Lemma 7). In the latter case, we show the protocol will converge within the next $O(n)$ state-changing interactions (Lemma 8).

The essential idea of our proof is to construct a family of supermartingales for different regions in the configuration space by carefully selecting a series of corresponding potential functions. The following lemma is a general statement of this proof technique.

Lemma 4. *Let f be a potential function and A be a region in the configuration space. If in region A, $\mathbb{E}(\Delta f/f \mid I^1) \leq -k_1/n$ and $\mathbb{E}(\Delta f/f \mid I^2) \leq k_2/n$ where k_1 and k_2 are two constants such that $k_1 > k_2 > 0$, and I^1 and I^2 are two binary indicators such that $I_t^1 \cdot I_t^2 \equiv 0$ at any number of interactions t, then the stochastic process $\{M_t\}$ given by*

$$M_t = \exp\left(\frac{c_1 S_t^1 - c_2 S_t^2}{n}\right) \cdot f_t$$

is a supermartingale in region A, where $S_t^1 = \sum_{i=1}^t I_t^1$ and $S_t^2 = \sum_{i=1}^t I_t^2$, and c_1, c_2 are two constants such that $k_1 > c_1 > c_2 > k_2 > 0$.

In addition, given $f_0/f_t \leq n^{c_3}$ for some positive constant $c_3 > 0$ at any number of interactions t, if the process never leaves region A, we have

$$\mathbb{P}\left(c_1 S_\tau^1 \geq c_2 S_\tau^2 + (c_3 + c_4) n \log n\right) \leq n^{-c_4}$$

for any positive constant $c_4 > 0$.

Proof Given

$$\mathbb{E}\left(\frac{\Delta f}{f} \mid I^1\right) = \mathbb{E}\left(\frac{f_{t+1} - f_t}{f_t} \mid I_{t+1}^1\right) \leq -\frac{k_1}{n}$$

and

$$\mathbb{E}\left(\frac{\Delta f}{f} \mid I^2\right) = \mathbb{E}\left(\frac{f_{t+1} - f_t}{f_t} \mid I_{t+1}^2\right) \leq \frac{k_2}{n}$$

we have

$$\mathbb{E}\left(f_{t+1} \mid I_{t+1}^1\right) \leq \left(1 - \frac{k_1}{n}\right) \cdot f_t \leq \exp\left\{-\frac{c_1}{n}\right\} \cdot f_t$$

and

$$\mathbb{E}\left(f_{t+1} \mid I_{t+1}^2\right) \leq \left(1 + \frac{k_2}{n}\right) \cdot f_t \leq \exp\left\{\frac{c_2}{n}\right\} \cdot f_t$$

Boosting the constants from $-k_1$ to $-c_1$ and from k_2 to c_2 is to absorb the second-order and higher terms in the Taylor series expansion of the exponential.

The expected value of M_{t+1} in each case is as follows.

$$\mathbb{E}\left(M_{t+1} \mid I_{t+1}^1 + I_{t+1}^2 = 0\right) = M_t$$

$$\begin{aligned}
\mathbb{E}\left(M_{t+1} \mid I_{t+1}^1\right) =& \mathbb{E}\left(\exp\left(\frac{c_1(S_t^1 + 1) - c_2 S_t^2}{n}\right) \cdot f_{t+1} \mid I_{t+1}^1\right) \\
=& \mathbb{E}\left(\frac{M_t \cdot f_{t+1} \cdot \exp(c_1/n)}{f_t} \mid I_{t+1}^1\right) \\
=& \exp\left\{\frac{c_1}{n}\right\} \cdot \mathbb{E}\left(f_{t+1} \mid I_{t+1}^1\right) \cdot \frac{M_t}{f_t} \\
\leq& M_t
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left(M_{t+1} \mid I_{t+1}^2\right) =& \mathbb{E}\left(\exp\left(\frac{c_1 S_t^1 - c_2(S_t^2 + 1)}{n}\right) \cdot f_{t+1} \mid I_{t+1}^2\right) \\
=& \mathbb{E}\left(\frac{M_t \cdot f_{t+1} \cdot \exp(-c_2/n)}{f_t} \mid I_{t+1}^2\right) \\
=& \exp\left\{-\frac{c_2}{n}\right\} \cdot \mathbb{E}\left(f_{t+1} \mid I_{t+1}^2\right) \cdot \frac{M_t}{f_t} \\
\leq& M_t
\end{aligned}$$

In any case we always have $\mathbb{E}\left(M_{t+1}\right) \leq M_t$ so the stochastic process $\{M_t\}$ is a supermartingale in region A. If the process never leaves region A, we have $\mathbb{E}(M_\tau) \leq M_0 = f_0$. Given $f_0/f_t \leq n^{c_3}$ at any number of interactions t (including the stopping time $t = \tau$), we have

$$\mathbb{E}(M_\tau) = \mathbb{E}\left(\exp\left(\frac{c_1 S_\tau^1 - c_2 S_\tau^2}{n}\right) \cdot f_\tau\right) \leq M_0 = f_0$$

and

$$\mathbb{E}\left(\exp\left(\frac{c_1 S_\tau^1 - c_2 S_\tau^2}{n}\right)\right) \leq n^{c_3}$$

From Markov's inequality,

$$\mathbb{P}\left(\exp\left(\frac{c_1 S_\tau^1 - c_2 S_\tau^2}{n}\right) \geq n^{c_3+c_4}\right) \leq n^{-c_4}$$

for any positive constant $c_4 > 0$ and then

$$\mathbb{P}\left(c_1 S_\tau^1 - c_2 S_\tau^2 \geq (c_3 + c_4) n \log n\right) \leq n^{-c_4}$$

which completes the proof. ■

Lemma 4 presents the proof technique we use throughout this section. When using this technique, we have either S_τ^2 is already well bounded (Lemma 9, Lemma 10, Lemma 11 and Lemma 12), or there exists some auxiliary inequality relationship between S_τ^1 and S_τ^2 (Lemma 5 and Lemma 7).

Lemma 5. *If the binary signaling consensus process starts with $g \leq \min(b, w)/4$, then for any constant $c > 0$, with probability $1-n^{-c}$ one of the following two events will happen within $O_c(n \log n)$ state-changing interactions:*

1. *$g > \min(b, w)/4$.*
2. *The process converges and*

$$\mathbb{P}\left(S_\tau^{g-} \geq \frac{1000}{7}\left(n \log\left(\frac{2}{5}n + 1\right) + cn \log n\right) + \frac{392}{7}n\right) \leq n^{-c}$$

and

$$\mathbb{P}\left(S_\tau^{g+} \geq \frac{1000}{7}\left(n \log\left(\frac{2}{5}n + 1\right) + cn \log n\right) + \frac{399}{7}n\right) \leq n^{-c}$$

Proof We can prove this fact by showing that if event 1 doesn't happen, then event 2 will surely happen. That is, if we always have $g \leq \min(b, w)/4$ and never have $g > \min(b, w)/4$, then with probability $1 - o(1)$ the process converges after $O(n \log n)$ state-changing interactions. For notational convenience, let value $f = u^2 + 5n/2$ so the potential function is $1/f$. We have

$$\begin{aligned}\Delta f =& (u + \Delta u)^2 + 5n/2 - u^2 - 5n/2 \\ =& u^2 + 2u\Delta u + (\Delta u)^2 - u^2 \\ =& 2u(\Delta u) + (\Delta u)^2\end{aligned}$$

Because $|\Delta u| \leq 1$ and $|\Delta f| \leq 2|u|+1$, we have $|\Delta f/f| \leq (2|u|+1)/(u^2+5n/2) = O(\min(1/|u|, 2|u|/5n))$, which is maximized at $u = \Theta(\sqrt{n})$ so that $|\Delta f/f| = O(1/\sqrt{n})$.

Let I^{bw} be the indicator of the event that neither of the two agents in a state-changing interaction is in state gray. Let I^{gv} be the indicator of the event that the speaker in a state-changing interaction is in gray and the listener is in black or white. Denote by $p = \tilde{b} + \tilde{g}/2$ and by $M = 2bw + \frac{1}{2}gv$. The expected change of value f conditioned on each case of state-changing interactions is as follows.

$$\begin{aligned}\mathbb{E}\left(\Delta f \mid I^{g-}\right) =& p(2u+1) + (1-p)(-2u+1)\\ =& 1 + (2p-1)\cdot 2u\\ =& 1 + 2u\cdot\frac{2b+g-n}{n}\\ =& 1 + \frac{2u^2}{n}\end{aligned}$$

$$\mathbb{E}\left(\Delta f \mid I^{bw}\right) = \frac{1}{2}(2u+1) + \frac{1}{2}(-2u+1) = 1$$

$$\begin{aligned}\mathbb{E}\left(\Delta f \mid I^{gv}\right) =& (2u+1)\frac{w}{v} + (-2u+1)\frac{b}{v}\\ =& 1 + 2u\cdot\frac{w-b}{v}\\ =& 1 - \frac{2u^2}{v}\end{aligned}$$

$$\mathbb{E}\left(\Delta f \mid I^{g+}\right) = \frac{2bw}{M} + \frac{gv}{2M}\left(1 - \frac{2u^2}{v}\right) = 1 - \frac{gu^2}{M}$$

$$\begin{aligned}\mathbb{E}\left((\Delta f)^2 \mid I^{g-}\right) =& p(2u+1)^2 + (1-p)(-2u+1)^2\\ =& p(4u^2+4u+1) + (1-p)(4u^2-4u+1)\\ =& 4u^2 + 1 + (2p-1)\cdot 4u\\ =& 4u^2 + 1 + 4u\cdot\frac{2b+g-n}{n}\\ =& 4u^2 + 1 + \frac{4u^2}{n}\end{aligned}$$

$$\mathbb{E}\left((\Delta f)^2 \mid I^{bw}\right) = \frac{1}{2}(2u+1)^2 + \frac{1}{2}(-2u+1)^2 = 4u^2 + 1$$

$$\begin{aligned}\mathbb{E}\left((\Delta f)^2 \mid I^{gv}\right) =& \frac{w}{v}(2u+1)^2 + \frac{b}{v}(-2u+1)^2\\ =& \frac{w}{v}(4u^2+4u+1) + \frac{b}{v}(4u^2-4u+1)\\ =& 4u^2 + 1 + \frac{w-b}{v}\cdot 4u\\ =& 4u^2 + 1 - \frac{4u^2}{v}\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left((\Delta f)^2 \mid I^{g+}\right) =& \frac{2bw}{M}(4u^2+1) + \frac{gv}{2M}\left(4u^2+1-\frac{4u^2}{v}\right)\\ =& 4u^2+1-\frac{2gu^2}{M}\end{aligned}$$

When $g \le \min(b,w)/4$, we have

$$\begin{aligned}M =& 2bw + \frac{1}{2}gv\\ =& 2\min(b,w)\cdot\max(b,w) + \frac{1}{2}g\cdot(\min(b,w)+\max(b,w))\\ \ge& g\cdot\left(\frac{17}{2}\max(b,w)+\frac{1}{2}\min(b,w)\right)\end{aligned}$$

As $4g = 4(n-\min(b,w)-\max(b,w)) \le \min(b,w) \le \max(b,w)$, we know

$$\frac{4}{5}(n-\max(b,w)) \le \min(b,w) \le \max(b,w)$$

Note that function $\frac{17}{2}x+\frac{1}{2}y$ given $\frac{4}{5}(1-x) \le y \le x$ and $y \ge 0$ and $x \le 1$ is at least 4. Thus we have $M \ge 4gn$. Let $z = u^2/n$.

$$\begin{aligned}&\mathbb{E}\left(\frac{\Delta(1/f)}{1/f} \mid I^{g-}\right)\\ =&\mathbb{E}\left(-\frac{\Delta f}{f}+\left(\frac{\Delta f}{f}\right)^2+O(n^{-3/2}) \mid I^{g-}\right)\\ =&-\frac{1+2u^2/n}{u^2+5n/2}+\frac{4u^2+1+4u^2/n}{(u^2+5n/2)^2}+O(n^{-3/2})\\ =&\left(u^2+5n/2\right)^{-2}\cdot\left(-(1+2u^2/n)(u^2+5n/2)+4u^2+1+4u^2/n\right)+O(n^{-3/2})\\ =&\left(u^2+\frac{5n}{2}\right)^{-2}\left(-u^2-\frac{5n}{2}-\frac{2u^2}{n}\left(u^2+\frac{5n}{2}\right)+4u^2+1+\frac{4u^2}{n}\right)+O(n^{-3/2})\\ =&\left(u^2+\frac{5n}{2}\right)^{-2}\left(3u^2-\frac{5n}{2}+1+\left(-u^2-\frac{5n}{2}+2\right)\cdot\frac{2u^2}{n}\right)+O(n^{-3/2})\\ =&n^{-2}(z+5/2)^{-2}(3zn-5n/2+1+2z(-zn-5n/2+2))+O(n^{-3/2})\\ =&n^{-2}(z+5/2)^{-2}((3z-5/2)n+1-2z(z+5/2)n+4z)+O(n^{-3/2})\\ =&\frac{1}{n}\cdot\frac{-2z^2+(3-5)z-5/2}{(z+5/2)^2}+\frac{1}{n^2}\cdot\frac{1+4z}{(z+5/2)^2}+O(n^{-3/2})\end{aligned}$$

where the first equality is due to $\frac{\Delta(1/f)}{1/f} = \sum_{i=1}^{+\infty}(-\Delta f/f)^i$ for $|\Delta f/f| < 1$. Note that function $\frac{-2x^2-2x-5/2}{(x+5/2)^2}$ given $x \ge 0$ is at most $-2/5$ and function $\frac{1+4x}{(x+5/2)^2}$ given $x \ge 0$ is at most $4/9$. Thus we have

$$\mathbb{E}\left(\frac{\Delta(1/f)}{1/f} \mid I^{g-}\right) \le -\frac{2}{5}n^{-1}+\frac{4}{9}n^{-2}+O(n^{-3/2}) = -\frac{2}{5}n^{-1}+O(n^{-3/2})$$

In the other case when $I^{g+} = 1$, we have

$$\begin{aligned}
&\mathbb{E}\left(\frac{\Delta(1/f)}{1/f} \mid I^{g+}\right)\\
=&\mathbb{E}\left(-\frac{\Delta f}{f} + \left(\frac{\Delta f}{f}\right)^2 + O(n^{-3/2}) \mid I^{g+}\right)\\
=&-\frac{1-gu^2/M}{u^2+5n/2} + \frac{4u^2+1-2gu^2/M}{(u^2+5n/2)^2} + O(n^{-3/2})\\
=&\left(u^2+5n/2\right)^{-2}\cdot\left(-(1-gu^2/M)(u^2+5n/2)+4u^2+1-2gu^2/M\right)+O(n^{-3/2})\\
=&\left(u^2+\frac{5n}{2}\right)^{-2}\left(-u^2-\frac{5n}{2}+\frac{gu^2}{M}\left(u^2+\frac{5n}{2}\right)+4u^2+1-\frac{2gu^2}{M}\right)+O(n^{-3/2})\\
=&\left(u^2+\frac{5n}{2}\right)^{-2}\left(3u^2-\frac{5n}{2}+1+\left(u^2+\frac{5n}{2}-2\right)\cdot\frac{gu^2}{M}\right)+O(n^{-3/2})\\
\le&\left(u^2+\frac{5n}{2}\right)^{-2}\left(3u^2-\frac{5n}{2}+1+\left(u^2+\frac{5n}{2}-2\right)\cdot\frac{u^2}{4n}\right)+O(n^{-3/2})\\
=&n^{-2}(z+5/2)^{-2}(3zn-5n/2+1+z(zn+5n/2-2)/4)+O(n^{-3/2})\\
=&n^{-2}(z+5/2)^{-2}((3z-5/2)n+1+z(z+5/2)n/4-z/2)+O(n^{-3/2})\\
=&\frac{1}{n}\cdot\frac{z^2/4+(3+5/8)z-5/2}{(z+5/2)^2}+\frac{1}{n^2}\cdot\frac{1-z/2}{(z+5/2)^2}+O(n^{-3/2})
\end{aligned}$$

Note that function $\frac{x^2/4+29x/8-5/2}{(x+5/2)^2}$ given $x \ge 0$ is at most $1001/2560$ and function $\frac{1-x/2}{(x+5/2)^2}$ given $x \ge 0$ is at most $4/25$. Thus we have

$$\mathbb{E}\left(\frac{\Delta(1/f)}{1/f} \mid I^{g+}\right) \le \frac{1001}{2560}n^{-1} + \frac{4}{25}n^{-2} + O(n^{-3/2}) = \frac{1001}{2560}n^{-1} + O(n^{-3/2})$$

According to Lemma 4, we can see the stochastic process $\{L_t\}$ given by

$$L_t = \frac{\exp\Big((0.399S_t^{g-} - 0.392S_t^{g+})/n\Big)}{u_t^2 + 5n/2}$$

is a supermartingale if we always have $g \le \min(b,w)/4$. Because $u^2 + 5n/2 \le n^2 + 5n/2$,

$$\mathbb{E}\left(\exp\left(\frac{0.399S_\tau^{g-} - 0.392S_\tau^{g+}}{n}\right)\right) \le \frac{2(n^2+5n/2)}{5n} = \frac{2n}{5} + 1$$

and then

$$\mathbb{P}\left(0.399S_\tau^{g-} - 0.392S_\tau^{g+} \ge n\log(2n/5+1) + cn\log n\right) \le n^{-c}$$

Because at any number of interactions t, the number of gray tokens the process has produced can't be more than the number of gray tokens the process has consumed plus n, we have $S_\tau^{g+} \le S_\tau^{g-} + n$, giving the bound

$$\mathbb{P}\left(0.399S_\tau^{g-} - 0.392(S_\tau^{g-}+n) \ge n\log(2n/5+1) + cn\log n\right) \le n^{-c}$$

and

$$\mathbb{P}\left(S_\tau^{g-} \ge \frac{1000}{7}\left(n\log\left(\frac{2}{5}n+1\right) + cn\log n\right) + \frac{392}{7}n\right) \le n^{-c}$$

interaction	b	w	g	$g-w$
$(b,+,w)$		-1	$+1$	$+2$
$(w,-,b)$	-1		$+1$	$+1$
$(b,+,g)$	$+1$		-1	-1
$(w,-,g)$		$+1$	-1	-2
$(g,+,g)$	$+1$		-1	-1
$(g,-,g)$		$+1$	-1	-2
$(g,+,w)$		-1	$+1$	$+2$
$(g,-,b)$	-1		$+1$	$+1$

Table 2: Changes in $(g-w)$ by state-changing interactions

which implies

$$\mathbb{P}\left(S_\tau^{g+} \geq \frac{1000}{7}\left(n\log\left(\frac{2}{5}n+1\right)+cn\log n\right)+\frac{399}{7}n\right) \leq n^{-c}$$

which completes the proof. ■

Now we have shown that if the population starts from the region $\{g \leq \min(b,w)/4\}$, within $O(n\log n)$ state-changing steps, it will either reach consensus or leave the region with high probability. While once the process leaves the region and has $g > \min(b,w)/4$, we prove that within the next $O(n\log n)$ state-changing interactions, either the population will never enter the region $\{g < \min(b,w)/10\}$, or it will enter the region $\{\min(b,w) = O(\log n) \bigwedge g = O(\log n)\}$ (Lemma 6 and Corollary 2).

Lemma 6. *If the process starts with $g > \min(b,w)/4$, then with probability $1-n^{-\omega(1)}$, for any polynomial $T = poly(n)$, we have either $g_t \geq \min(b_t,w_t)/10$ holds for all $1 \leq t \leq T$ or at some stage $1 \leq t \leq T$, the process reaches $\min(b,w) = O(\log n)$, $g = O(\log n)$ and $\max(b,w) = n - O(\log n)$.*

Proof Again we can show this fact by showing that if latter event doesn't happen, former event will happen. Let's consider how the value of $(g - \min(b,w))$ changes in different state-changing interactions. Without loss of generality, assume that at the current time step $\max(b,w) = b$ and $\min(b,w) = w$. Let $N = ng + 2bw + \frac{1}{2}gv$. Table 2 lists all the cases.

Thus

$$\mathbb{P}(\Delta(g-w) = +1 \mid I^{sc}) = \left(bw + \frac{1}{2}gb\right)/N = \frac{n^2}{N}(1-\tilde{g}-\tilde{w})\left(\tilde{w}+\frac{1}{2}\tilde{g}\right)$$

$$\begin{aligned}\mathbb{P}(\Delta(g-w) = -1 \mid I^{sc}) &= \left(bg + \frac{1}{2}g^2\right)/N \\ &= \frac{n^2}{N}\left((1-\tilde{w}-\tilde{g})\tilde{g} + \frac{1}{2}\tilde{g}^2\right) \\ &= \frac{n^2}{N}\left((1-\tilde{w})\tilde{g} - \frac{1}{2}\tilde{g}^2\right)\end{aligned}$$

$$\begin{aligned}\mathbb{P}(\Delta(g-w)=+2 \mid I^{sc}) =& \left(bw+\frac{1}{2}gw\right)/N \\ =& \frac{n^2}{N}\left(\tilde{w}\left(1-\tilde{g}-\tilde{w}+\frac{1}{2}\tilde{g}\right)\right) \\ =& \frac{n^2}{N}\left(\tilde{w}\left(1-\tilde{w}-\frac{1}{2}\tilde{g}\right)\right)\end{aligned}$$

$$\mathbb{P}(\Delta(g-w)=-2 \mid I^{sc}) = \frac{n^2}{N}\left(\tilde{w}\tilde{g}+\frac{1}{2}\tilde{g}^2\right)$$

Note that if the process enters the region $\{g < \min(b,w)/10\}$ from the initial region $\{g > \min(b,w)/4\}$, it must pass through the region $\{\min(b,w)/10 \le g \le \min(b,w)/4\}$. We show that even passing through this intermediate region already requires strictly more than a polynomial number of state-changing interactions, let alone the whole fleeing path.

Note that function $\frac{(1-x-y)(x+y/2)}{(1-x)y-y^2/2}$ conditioned on $0 \le x \le (1-y)/2$ and $x/10 \le y \le x/4$ is always ≥ 4. Also, function $\frac{x(1-x-y/2)}{xy+y^2/2}$ conditioned on $0 \le x \le (1-y)/2$ and $x/10 \le y \le x/4$ is always ≥ 4. Thus when $w/10 \le g \le w/4$, we always have

$$\frac{\mathbb{P}(\Delta(g-w)=+1 \mid I^{sc})}{\mathbb{P}(\Delta(g-w)=-1 \mid I^{sc})} \ge 4 \text{ and } \frac{\mathbb{P}(\Delta(g-w)=+2 \mid I^{sc})}{\mathbb{P}(\Delta(g-w)=-2 \mid I^{sc})} \ge 4$$

The value of $(g-w)$ never stays put with $I^{sc}=1$.

When $\min(b,w)=\omega(\log n)$, the length of this gap $\min(b,w)/4-\min(b,w)/10 = 3\min(b,w)/20$ is also $\omega(\log n)$. Let length $\ell=\omega(\log n)$. Consider the following one-dimensional random walk on integers from 0 to ℓ. State 0 is a reflecting barrier always pushing the walk back to state 1. At any state $1 \le i \le \ell-1$, the forward probability is $1/5$ and the backward probability is $4/5$. The walk starts at state 1 and we are interested in the first hitting time of state ℓ.

The number of steps until this random walk first hits state ℓ provides an upper bound on the number of interactions needed by the process to flee from the region $\{g > \min(b,w)/4\}$ and enter the region $\{g < \min(b,w)/10\}$, conditioned on the event that $\min(b,w)=\omega(\log n)$ always holds. Now we show it needs strictly more than a polynomial number of steps with high probability.

Note that every time that the walk hits state 0, the reflecting barrier "resets" it to state 1. Everything the walk does between two consecutive "resets" can be viewed as a Bernoulli trial. And we shall show with probability $1-o(1)$, this Bernoulli process needs strictly more than polynomially many trials to succeed. Here for each trial, hitting 0 before hitting ℓ is a failure and otherwise it succeeds.

Denote by $\beta_i = \mathbb{P}(\text{hitting state 0 before hitting state } \ell \mid \text{ starting at state } i)$. Then $\beta_0=1$ and $\beta_\ell=0$. The probability of failure is β_1.

For any $1 \le i \le \ell-1$, $\beta_i = \beta_{i+1}/5 + 4\beta_{i-1}/5$. Define

$$\begin{aligned}\Delta\beta_i =& \beta_i - \beta_{i+1} \\ =& \beta_i - \frac{1}{5}\beta_{i+2} - \frac{4}{5}\beta_i \\ =& \frac{1}{5}(\beta_i - \beta_{i+2}) \\ =& \frac{1}{5}(\beta_i - \beta_{i+1} + \beta_{i+1} - \beta_{i+2}) \\ =& \frac{1}{5}\Delta\beta_i + \frac{1}{5}\Delta\beta_{i+1}\end{aligned}$$

which implies $\frac{4}{5}\Delta\beta_i = \frac{1}{5}\Delta\beta_{i+1}$ or $\Delta\beta_{i+1} = 4\Delta\beta_i$. Thus $\Delta\beta_i = 4^i\Delta\beta_0 = 4^i(1-\beta_1)$.

Note that

$$\sum_{i=0}^{\ell-1}\Delta\beta_i = \beta_0 - \beta_1 + \beta_1 - \beta_2 + \ldots + \beta_{\ell-1} - \beta_\ell = \beta_0 - \beta_\ell = 1$$

Then

$$\sum_{i=0}^{\ell-1}\Delta\beta_i = (1-\beta_1)\sum_{i=0}^{\ell-1}4^i = (1-\beta_1)\cdot\frac{4^\ell-1}{3} = 1$$

Therefore, $(1-\beta_1)(4^\ell-1) = 3$ and $\beta_1 = 1 - 3/(4^\ell-1)$.

Let c be any arbitrarily large constant. The probability that all the first n^c trials fail is

$$\beta_1^{n^c} = \left(1 - \frac{3}{4^\ell - 1}\right)^{n^c} = \left(1 - \frac{1}{n^{\omega(1)}}\right)^{n^c} = \exp\left(-n^{c-\omega(1)}\right) \sim 1 - n^{c-\omega(1)}$$

The probability goes to 1 in order $n^{\omega(1)-c}$. Thus with probability $1 - O(n^{-\omega(1)})$ the random walk won't hit state ℓ within a polynomial number of steps.

Therefore, we can have $g < \min(b,w)/10$ only when $\min(b,w) = O(\log n)$ happens. In this case $g < \min(b,w)/10 = O(\log n)$ too so the other event happens. ■

Because in this problem we are only interested in the next $O(n\log n)$ state-changing interactions, we have

Corollary 2. *If the process starts with $g > \min(b,w)/4$, then with probability $1 - n^{-\omega(1)}$, for any $T = O(n\log n)$, we have either $g_t \geq \min(b_t, w_t)/10$ holds for all $1 \leq t \leq T$ or at some stage $1 \leq t \leq T$, the process reaches $\min(b,w) = O(\log n)$, $g = O(\log n)$ and $\max(b,w) = n - O(\log n)$.*

Next we will show the process also converges fast within the region $\{g \geq \min(b,w)\cdot 1/10\}$.

Lemma 7. *If $g \geq \min(b,w)/10$ holds for a polynomial number of state-changing interactions, then for any constant $c > 0$, with probability $1 - n^{-c}$, after $O_c(n\log n)$ state-changing interactions, the process will converge and we have*

$$\mathbb{P}\left(S_\tau^{sc} \geq 87n\log\left(\frac{1}{64}n + 1\right) + 87cn\log n\right) \leq n^{-c}$$

$$\mathbb{P}\left(S_\tau^{g+} \geq \frac{87}{2}n\log\left(\frac{1}{64}n + 1\right) + \frac{87}{2}cn\log n + \frac{1}{2}n\right) \leq n^{-c}$$

and

$$\mathbb{P}\left(S_\tau^{g-} \geq \frac{87}{2}n\log\left(\frac{1}{64}n + 1\right) + \frac{87}{2}cn\log n + \frac{1}{2}n\right) \leq n^{-c}$$

Proof In this proof we use the potential function $1/(u^2+64n)$ and denote by $f = u^2 + 64n$. Similarly we have $\Delta f = 2u(\Delta u) + (\Delta u)^2$ and $|\Delta f/f| = O(1/\sqrt{n})$. Recall that $N = ng + 2bw + \frac{1}{2}gv$. We have

$$\begin{aligned}\mathbb{E}\left(\Delta f \mid I^{sc}\right) =& \frac{ng}{N}\left(1 + \frac{2u^2}{n}\right) + \frac{2bw}{N} + \frac{gv}{2N}\left(1 - \frac{2u^2}{v}\right)\\ =& 1 + 2u^2\cdot\left(\frac{ng}{N}\cdot\frac{1}{n} - \frac{gv}{2N}\cdot\frac{1}{v}\right)\\ =& 1 + \frac{gu^2}{N}\end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}\left((\Delta f)^2 \mid I^{sc}\right)\\
=&\frac{ng}{N}\left(4u^2+1+\frac{4u^2}{n}\right)+\frac{2bw}{N}(4u^2+1)+\frac{gv}{2N}\left(4u^2+1-\frac{4u^2}{v}\right)\\
=&4u^2+1+4u^2\left(\frac{ng}{N}\cdot\frac{1}{n}-\frac{gv}{2N}\cdot\frac{1}{v}\right)\\
=&4u^2+1+\frac{2gu^2}{N}
\end{aligned}
$$

When $g \geq \min(b,w)/10$, we have $bw = \min(b,w)\cdot\max(b,w) \leq \min(b,w)\cdot n \leq 10bn$ and $N = ng + 2bw + gv/2 \leq gn + 20gn + gn/2 = 43gn/2$. Again let $z = u^2/n$. We have

$$
\begin{aligned}
&\mathbb{E}\left(\frac{\Delta(1/f)}{1/f} \mid I^{sc}\right)\\
=&\mathbb{E}\left(-\frac{\Delta f}{f}+\left(\frac{\Delta f}{f}\right)^2+O(n^{-3/2}) \mid I^{sc}\right)\\
=&-\frac{1+gu^2/N}{u^2+64n}+\frac{4u^2+1+2gu^2/N}{(u^2+64n)^2}+O(n^{-3/2})\\
=&\left(u^2+64n\right)^{-2}\cdot\left(-(1+gu^2/N)(u^2+64n)+4u^2+1+2gu^2/N\right)+O(n^{-3/2})\\
=&\left(u^2+64n\right)^{-2}\left(-u^2-64n-\frac{gu^2}{N}\left(u^2+64n\right)+4u^2+1+\frac{2gu^2}{N}\right)+O(n^{-3/2})\\
=&\left(u^2+64n\right)^{-2}\left(3u^2-64n+1+\left(-u^2-64n+2\right)\cdot\frac{gu^2}{N}\right)+O(n^{-3/2})\\
\leq&\left(u^2+64n\right)^{-2}\left(3u^2-64n+1+\left(-u^2-64n+2\right)\cdot\frac{2u^2}{43n}\right)+O(n^{-3/2})\\
=&n^{-2}(z+64)^{-2}(3zn-64n+1+2z(-zn-64n+2)/43)+O(n^{-3/2})\\
=&n^{-2}(z+64)^{-2}((3z-64)n+1-2z(z+64)n/43+4z/43)+O(n^{-3/2})\\
=&\frac{1}{n}\cdot\frac{-2z^2/43+(3-128/43)z-64}{(z+64)^2}+\frac{1}{n^2}\cdot\frac{1+4z/43}{(z+64)^2}+O(n^{-3/2})
\end{aligned}
$$

Note that function $\frac{-2x^2/43+x/43-64}{(x+64)^2}$ given $x \geq 0$ is at most $-22015/1893376 < -1/87$ and function $\frac{1+4x/43}{(x+64)^2}$ given $x \geq 0$ is at most $4/9159$. Thus from Lemma 4 we have the stochastic process $\{K_t\}$ given by

$$K_t = \frac{\exp(S_t^{sc}/(87n))}{u_t^2+64n}$$

is a supermartingale if we always have $g \geq \min(b,w)/10$. This gives us

$$\mathbb{E}\left(\exp(S_\tau^{sc}/(87n))\right) \leq (n^2+64n)/(64n) = n/64+1$$

For Markov's inequality

$$\mathbb{P}\left(\exp(S_\tau^{sc}/(87n)) \geq n^c(n/64+1)\right) \leq n^{-c}$$

and

$$\mathbb{P}\left(S_\tau^{sc}/87 \geq n\log(n/64+1)+cn\log n\right) \leq n^{-c}$$

$$\mathbb{P}\left(S_\tau^{sc} \geq 87n\log\left(\frac{1}{64}n+1\right)+87cn\log n\right) \leq n^{-c}$$

Since $S_\tau^{sc} = S_\tau^{g+} + S_\tau^{g-}$, $S_\tau^{g+} \le S_\tau^{g-} + n$ and $S_\tau^{g-} \le S_\tau^{g+} + n$, we have

$$\mathbb{P}\left(S_\tau^{g+} \ge \frac{87}{2} n \log\left(\frac{1}{64}n + 1\right) + \frac{87}{2} cn \log n + \frac{1}{2}n\right) \le n^{-c}$$

and

$$\mathbb{P}\left(S_\tau^{g-} \ge \frac{87}{2} n \log\left(\frac{1}{64}n + 1\right) + \frac{87}{2} cn \log n + \frac{1}{2}n\right) \le n^{-c}$$

which completes the proof. ■

The only case left is when the protocol enters the region $\{\min(b, w) = O(\log n) \bigwedge g = O(\log n)\}$. Recall that $p = \tilde{b} + \tilde{g}/2$. Once it enters this region, we will have $p = O(\log n/n)$ or $1-p = O(\log n/n)$.

Lemma 8. *If the process starts with $p = O(\log n/n)$ or $1 - p = O(\log n/n)$, then with probability $1 - O\left(\frac{\log n}{\sqrt[3]{n}}\right)$ the population will reach consensus within $O(n)$ state-changing interactions.*

Proof The proof is completed by worst-case analyses. Without loss of generality, assume $1 - p = O(\log n/n)$ and we will show with high probability p will converge to 1 within $O(n)$ state-changing interactions. In this case, we have

$$\mathbb{P}\left(p_{t+1} = p_t + \frac{1}{2n} \mid I^{sc}\right) = \frac{p_t(1 - \tilde{x}_t)}{p_t(1 - \tilde{x}_t) + (1 - p_t)(1 - \tilde{y}_t)}$$

and

$$\mathbb{P}\left(p_{t+1} = p_t - \frac{1}{2n} \mid I^{sc}\right) = \frac{(1 - p_t)(1 - \tilde{y}_t)}{p_t(1 - \tilde{x}_t) + (1 - p_t)(1 - \tilde{y}_t)}$$

Note that $x_t \le p_t$. We have

$$\frac{\mathbb{P}\left(p_{t+1} = p_t + 1/2n \mid I^{sc}\right)}{\mathbb{P}\left(p_{t+1} = p_t - 1/2n \mid I^{sc}\right)} \ge \frac{p_t(1 - p_t)}{1 - p_t} = p_t$$

To provide an upper bound on the moves of p in the region $\{1 - 2\sqrt[3]{n}/(2n) \le p \le 1\}$, consider the following one-dimensional random walk on integers from 0 to $2\sqrt[3]{n}$. (Obviously each state i corresponds to the configuration $p = 1 - (2\sqrt[3]{n} - i)/(2n)$.) State 0 is a reflecting barrier always pushing the walk back to state 1. At any state $1 \le i \le 2\sqrt[3]{n} - 1$, the forward probability is $p/(p+1) = \frac{1-(2\sqrt[3]{n}-i)/(2n)}{2-(2\sqrt[3]{n}-i)/(2n)}$ and the backward probability is $1/(p+1) = \frac{1}{2-(2\sqrt[3]{n}-i)/(2n)}$. The walk starts at some state $k = 2\sqrt[3]{n} - O(\log n)$. Denote by t_i the number of steps needed to first hit state $2\sqrt[3]{n}$, starting at state i. And let h_i be the number times hitting state 0 before reaching state $2\sqrt[3]{n}$, starting at state i. Then the total number of state-changing interactions for the process starting from this region to entirely converge is at most $h_k \cdot O(n \log n) + t_k$.

Though this walk is already simple, we can further simplify it to the same walk with fixed forward probability $q_+ = \frac{1-n^{-2/3}}{2-n^{-2/3}}$ and backward probability $q_- = \frac{1}{2-n^{-2/3}}$, which also provides an upper bound, because in the region $\{1 - 2\sqrt[3]{n}/(2n) \le p \le 1\}$ we always have $p \ge 1 - n^{-2/3}$. We overload the notation t_i and h_i for this simpler walk. Denote by $\bar{t}_i = \mathbb{E}t_i$ and let $\Delta\bar{t}_i$ be the expected number of steps the walk takes from state $i-1$ to state i. Then $\Delta\bar{t}_1 = 1$ due to the reflecting barrier at state 0. For $i \ge 2$, we have

$$\begin{aligned}\Delta\bar{t}_i =& 1 + q_+ \cdot 0 + q_- \cdot \mathbb{E}(\text{number of steps from state } i-2 \text{ to state } i)\\ =& 1 + q_-(\Delta\bar{t}_{i-1} + \Delta\bar{t}_i)\end{aligned}$$

which implies

$$\begin{aligned}
\Delta\bar{t}_i =& \frac{1}{q_+} + \frac{q_-}{q_+}\Delta\bar{t}_{i-1} \\
=& \frac{1}{q_+} + \frac{q_-}{q_+}\left(\frac{1}{q_+} + \frac{q_-}{q_+}\Delta\bar{t}_{i-2}\right) \\
=& \frac{1}{q_+} + \frac{q_-}{q_+}\left(\frac{1}{q_+} + \frac{q_-}{q_+}\left(\frac{1}{q_+} + \frac{q_-}{q_+}\Delta\bar{t}_{i-3}\right)\right) \\
=& \dots \\
=& \frac{1}{q_+} + \frac{q_-}{q_+^2} + \frac{q_-^2}{q_+^3} + \dots + \frac{q_-^{i-2}}{q_+^{i-1}} + \frac{q_-^{i-1}}{q_+^{i-1}} \\
=& \frac{1}{q_+}\left(1 + \frac{q_-}{q_+} + \dots + \frac{q_-^{i-2}}{q_+^{i-2}}\right) + \left(\frac{q_-}{q_+}\right)^{i-1}
\end{aligned}$$

Note that for any $0 \le i \le 2\sqrt[3]{n}$, we have

$$\left(1 - n^{-\frac{2}{3}}\right)^i \ge \left(1 - n^{-\frac{2}{3}}\right)^{2\sqrt[3]{n}} = \left(\left(1 - n^{-\frac{2}{3}}\right)^{n^{\frac{2}{3}}}\right)^{2n^{-\frac{1}{3}}} = \exp\left(-2/\sqrt[3]{n}\right) \to 1$$

Thus all $\left(\frac{q_-}{q_+}\right)^i \to 1$ for large n. Then we have $\Delta\bar{t}_i = 2(i-1)+1 = 2i-1$. Hence, $\bar{t}_k = \sum_{i=k+1}^{\sqrt[3]{n}} \Delta\bar{t}_i = \Theta(\sqrt[3]{n}\log n)$. Markov's inequality gives

$$\mathbb{P}(t_k \ge n) \le \frac{\bar{t}_k}{n} = \Theta\left(\frac{\log n}{n^{2/3}}\right)$$

Now if we can show with high probability $h_k = O(1)$ then we are done. But in fact we can do much better: with high probability $h_k = 0$. Denote by $\gamma_i = \mathbb{P}(h_i = 0)$. Then for $1 \le i \le 2\sqrt[3]{n} - 1$, $\gamma_i = q_+\gamma_{i+1} + q_-\gamma_{i-1}$. Define

$$\begin{aligned}
\Delta\gamma_i =& \gamma_{i+1} - \gamma_i \\
=& q_+\gamma_{i+2} + q_-\gamma_i - \gamma_i \\
=& q_+(\gamma_{i+2} - \gamma_i) \\
=& q_+(\gamma_{i+2} - \gamma_{i+1} + \gamma_{i+1} - \gamma_i) \\
=& q_+\Delta\gamma_{i+1} + q_+\Delta\gamma_i
\end{aligned}$$

which implies $\Delta\gamma_{i+1} = \frac{q_-}{q_+}\Delta\gamma_i$ and $\Delta\gamma_i = \left(\frac{q_-}{q_+}\right)^i \Delta\gamma_0$. Note that $\gamma_{2\sqrt[3]{n}} = 1$ and $\gamma_0 = 0$.

$$\sum_{i=0}^{2\sqrt[3]{n}-1} \Delta\gamma_i = \gamma_1 - \gamma_0 + \gamma_2 - \gamma_1 + \dots + \gamma_{2\sqrt[3]{n}} - \gamma_{2\sqrt[3]{n}-1} = \gamma_{2\sqrt[3]{n}} - \gamma_0 = 1$$

Then $\sum_{i=0}^{2\sqrt[3]{n}-1} \Delta\gamma_i = 2\sqrt[3]{n}\Delta\gamma_0 = 1$ so $\Delta\gamma_0 = 1/(2\sqrt[3]{n})$. And we have

$$\gamma_k = \sum_{i=0}^{k-1} \Delta\gamma_i = \frac{k}{2\sqrt[3]{n}} = \frac{2\sqrt[3]{n} - O(\log n)}{2\sqrt[3]{n}} = 1 - O\left(\frac{\log n}{\sqrt[3]{n}}\right)$$

Thus with probability $1 - O\left(\frac{\log n}{\sqrt[3]{n}}\right)$, we have $h_k = 0$, which means with high probability, a population starting from region $\{\min(b, w) = O(\log n) \bigwedge g = O(\log n)\}$ will reach consensus after $O(n)$ state-changing interactions without leaving this region. ■

Combining all the lemmas we have so far yields Lemma 3. Note that the error bounds in these lemmas are all at most n^{-c} except in Lemma 8 where the error bound is $O\left(\frac{\log n}{\sqrt[3]{n}}\right)$, which dominates the other error terms (with a tiny increase in the constant) when c is large. Also note that the constants in the O's in Lemma 6 ("$\min(b, w) = O(\log n)$ and $g = O(\log n)$") and Lemma 8 ("$\min(p, 1-p) = O(\log n/n)$") can be chosen arbitrarily. We simply make them consistent, and choose a proper value to have the error bound $\frac{c \log n}{\sqrt[3]{n}}$ as claimed in Lemma 3. Eventually, we have Lemma 3. ■

A.2 Bounding $S_\tau^c = O(n \log n)$

In this subsection we bound the number of interactions S_τ^c in the central region where $\max(\tilde{b}, \tilde{g}, \tilde{w}) < 3/4$, using the total number of state-changing interactions S_τ^{sc}.

Lemma 9. *With probability $1 - o(1)$, $S_\tau^c = O(n \log n)$. In addition, for any constant $c > 0$, we have*

$$\mathbb{P}(S_\tau^c \geq 9S_\tau^{sc} + cn \log n) \leq n^{-c}$$

Proof We show that the stochastic process $\{C_t\}$ given by

$$C_t = \exp((S_t^c - 9S_t^{sc})/n)$$

is a supermartingale. When $I_t^c = 0$, the value of C_t cannot increase so obviously $\mathbb{E}(C_t \mid \mathcal{F}_{t-1} \bigwedge I_t^c = 0) \leq C_{t-1}$. When $I_t^c = 1$, i.e., $\max(\tilde{b}, \tilde{g}, \tilde{w}) < 3/4$, at least two of $\tilde{w}, \tilde{b}$ and $\tilde{g}$ must be at least $1/8$:

- If $\tilde{b}$ and $\tilde{w}$ are both $\geq 1/8$, then $\mathbb{P}(I^{g+} = 1) = (2bw + gv/2)/n^2 \geq 1/8$, because function $2xy + (x+y)(1-x-y)/2$ given $1/8 \leq x, y \leq 3/4$ and $x + y \leq 1$ is at least $1/8$. Then the probability of the event that the current interaction increases S_t^c but not S_t^{g-} or S_t^{g+} and multiplies C_t by $\exp(1/n)$ is at most $7/8$. The probability of the event that the current interaction increases both S_t^c and S_t^{g+} but not S_t^{g-} and multiplies C_t by $\exp(-8/n)$ is at least $1/8$.
- If $\tilde{g} \geq 1/8$, we have $\mathbb{P}(I^{g-} = 1) = ng/n^2 \geq 1/8$. Then the probability of the event that the current interaction increases S_t^c but not S_t^{g-} or S_t^{g+} and multiplies C_t by $\exp(1/n)$ is at most $7/8$. The probability of the event that the current interaction increases both S_t^c and S_t^{g-} but not S_t^{g+} and multiplies C_t by $\exp(-8/n)$ is at least $1/8$.

This gives the bound

$$\begin{aligned}
\mathbb{E}(C_t \mid \mathcal{F}_{t-1} \bigwedge I_t^c) \leq & C_{t-1}\left(\frac{7}{8}\exp\left(\frac{1}{n}\right) + \frac{1}{8}\exp\left(-\frac{8}{n}\right)\right) \\
= & \left(\frac{7}{8}\left(1 + \frac{1}{n}\right) + \frac{1}{8}\left(1 - \frac{8}{n}\right) + O(n^{-2})\right) \\
= & C_{t-1}\left(1 - \frac{1}{8}n^{-1} + O(n^{-2})\right) \\
< & C_{t-1}
\end{aligned}$$

where the first equality is due to the Taylor expansion of the exponential function. Thus from Lemma 4 $\{C_t\}$ is a supermartingale and

$$\mathbb{P}(S_\tau^c \geq 9S_\tau^{sc} + cn\log n) \leq n^{-c}$$

which completes the proof. ■

A.3 Bounding $S_\tau^g = O(n\log n)$

In this subsection we bound the number of interactions S_τ^g in the corner region where $\tilde{g} \geq 3/4$, using the total number of g-decreasing interactions S_τ^{g-}.

Lemma 10. *With probability $1 - o(1)$, $S_\tau^g = O(n\log n)$. In addition, for any constant $c > 0$, we have*

$$\mathbb{P}\left(S_\tau^g \geq 26S_\tau^{g-} + 6cn\log n + 6n\log(2n+1) + \frac{45}{2}n\right) \leq n^{-c}$$

Proof In the large-g region, we choose the potential function $1/(2v+1)$ and let $f = 2v+1$. When $v = 0$, the whole population is at state g and the next interaction is surely (g, g). Then

$$\frac{\Delta(1/f)}{1/f} = 1 \cdot \left(\frac{1}{2\times 1+1} - \frac{1}{2\times 0+1}\right) = -\frac{2}{3}$$

When $v \geq 1$, we have expectation

$$\begin{aligned}
&\mathbb{E}\left(\frac{\Delta(1/f)}{1/f}\right)\\
=&\mathbb{E}\left((2v+1)\left(I^{g-}\left(\frac{1}{2v+3} - \frac{1}{2v+1}\right) + I^{g+}\left(\frac{1}{2v-1} - \frac{1}{2v+1}\right)\right)\right)\\
=&\mathbb{E}\left(-\frac{2I^{g-}}{2v+3} + \frac{2I^{g+}}{2v-1}\right)\\
=&n^{-2}\left(-\frac{2ng}{2v+3} + \frac{2(2bw+gv/2)}{2v-1}\right)\\
=&(n^2(2v+3)(2v-1))^{-1}\cdot(-2ng(2v-1) + (4bw+gv)(2v+3))\\
=&(n^2(4v^2+4v-3))^{-1}\cdot(-4ngv + 2ng + 8bwv + 12bw + 2gv^2 + 3gv)\\
\leq&(n^2(4v^2+4v-3))^{-1}\cdot(-4ngv + 2ng + 8v\cdot v^2/4 + 12v^2/4 + 2gv^2 + 3gv)\\
=&(n^2(4v^2+4v-3))^{-1}\cdot(-4ngv + 2ng + 2v^3 + 3v^2 + 2gv^2 + 3gv)\\
=&\frac{-4n(n-v)v + 2n(n-v) + 2v^3 + 3v^2 + 2(n-v)v^2 + 3(n-v)v}{n^2(4v^2+4v-3)}\\
=&\frac{-4n^2v + 4nv^2 + 2n^2 - 2nv + 2v^3 + 3v^2 + 2nv^2 - 2v^3 + 3nv - 3v^2}{n^2(4v^2+4v-3)}\\
=&\frac{-4n^2v + 6nv^2 + 2n^2 + nv}{n^2(4v^2+4v-3)}\\
=&\frac{-4nv + 6v^2 + 2n + v}{n(4v^2+4v-3)}\\
=&\frac{(2-4v)n + 6v^2 + v}{n(4v^2+4v-3)}
\end{aligned}$$

When $v \geq 1$, we have $2 - 4v < 0$. Since $v = n - g \leq n/4$, we have $n \geq 4v$ and

$$\mathbb{E}\left(\frac{\Delta(1/f)}{1/f}\right) \leq \frac{(2-4v)\cdot 4v + 6v^2 + v}{n(4v^2+4v-3)} \leq -\frac{1}{5}n^{-1}$$

This is because function $(-10x^2+9x)/(4x^2+4x-3)$ given $x \geq 1$ is at most $-1/5$.

When the population is not in the large-g region, i.e., $\tilde{g} < 3/4$, we have

$$\frac{\Delta(1/f)}{1/f} = -\frac{2I^{g-}}{2v+3} + \frac{2I^{g+}}{2v-1} \leq -\frac{2I^{g-}}{2n+3} + \frac{2I^{g+}}{2n/4-1}$$

Hence, from Lemma 4 the stochastic process $\{G_t\}$ given by

$$G_t = \frac{\exp\left(\left(\frac{1}{6}S_t^g + \sum_{i=1}^t \left(\frac{2}{3}I_i^{g-} - 5I_i^{g+}\right)(1-I_i^g)\right)/n\right)}{2v_t+1}$$

is a supermartingale process. This gives us the bound

$$\mathbb{E}\left(\frac{\exp\left(\left(\frac{1}{6}S_\tau^g + \sum_{i=1}^\tau \left(\frac{2}{3}I_i^{g-} - 5I_i^{g+}\right)(1-I_i^g)\right)/n\right)}{2n+1}\right) \leq \mathbb{E}G_\tau \leq G_0 \leq 1$$

Again for Markov's inequality,

$$\mathbb{P}\left(\frac{1}{6}S_\tau^g + \sum_{i=1}^\tau \left(\frac{2}{3}I_i^{g-} - 5I_i^{g+}\right)(1-I_i^g) \geq cn\log n + n\log(2n+1)\right) \leq n^{-c}$$

Note that $\sum_{i=1}^t I_i^{g-}(1-I_i^g)$ is the number of I^{g-} interactions that occur in the region $\{g < 3n/4\}$ and $\sum_{i=1}^t I_i^{g+}(1-I_i^g)$ is the number of I^{g+} interactions that occur in the region $\{g < 3n/4\}$. If the process never leaves the region after entering it, we have $\sum_{i=1}^t I_i^{g+}(1-I_i^g) \leq \sum_{i=1}^t I_i^{g-}(1-I_i^g)+3n/4$. If it passes the boundary of the region more than once, because every time that the process leaves the region it must have $g = 3n/4-1$ and every time that it returns to the region it must have $g = 3n/4-1$ too, we still have $\sum_{i=1}^t I_i^{g+}(1-I_i^g) \leq \sum_{i=1}^t I_i^{g-}(1-I_i^g) + 3n/4$. In addition, $\sum_{i=1}^t I_i^{g-}(1-I_i^g) \leq S_t^{g-}$ so we have

$$\mathbb{P}\left(\frac{1}{6}S_\tau^g + \sum_{i=1}^\tau \left(-\frac{13}{3}I_i^{g-}\right)(1-I_i^g) - \frac{15}{4}n \geq cn\log n + n\log(2n+1)\right) \leq n^{-c}$$

$$\mathbb{P}\left(\frac{1}{6}S_\tau^g - \frac{13}{3}S_\tau^{g-} \geq cn\log n + n\log(2n+1) + \frac{15}{4}n\right) \leq n^{-c}$$

and

$$\mathbb{P}\left(S_\tau^g \geq 26S_\tau^{g-} + 6cn\log n + 6n\log(2n+1) + \frac{45}{2}n\right) \leq n^{-c}$$

which completes the proof. ■

interaction	b	w	g	$3w+g+1$
$(b,+,w)$		-1	$+1$	-2
$(w,-,b)$	-1		$+1$	$+1$
$(b,+,g)$	$+1$		-1	-1
$(w,-,g)$		$+1$	-1	$+2$
$(g,+,g)$	$+1$		-1	-1
$(g,-,g)$		$+1$	-1	$+2$
$(g,+,w)$		-1	$+1$	-2
$(g,-,b)$	-1		$+1$	$+1$
Others				0

Table 3: Changes in $(3w+g+1)$

A.4 Bounding $S_\tau^b = O(n\log n)$ and $S_\tau^w = O(n\log n)$

We first bound the number of interactions S_τ^b in the corner region where $\tilde{b} \geq 3/4$. Then the upper bound for the number of interactions S_τ^w in the other corner region where $\tilde{w} \geq 3/4$ follows in a symmetric way.

Lemma 11. *With probability $1-o(1)$, $S_\tau^b = O(n\log n)$. In addition, for any constant $c>0$, we have*

$$\mathbb{P}\left(S_\tau^b \geq 153S_\tau^{g-} + 85S_\tau^{g+} + 17cn\log n + 17n\log(3n+1)\right) \leq n^{-c}$$

Proof In the large-b region, we choose the potential function $f = 3w+g+1$. Table 3 lists the changes in f by different types of interactions. Suppose $3/4 \leq \tilde{b} < 1$ so $\max(g,w) \geq 1$. (The case when $\tilde{b}=1$ is convergence.) Again we need to bound the expectation $\mathbb{E}(\Delta f/f)$:

$$\begin{aligned}
\mathbb{E}\left(\frac{\Delta f}{f}\right) =& (n^2(3w+g+1))^{-1}\cdot\left(-2bw+bw-bg+2gw-\frac{1}{2}g^2+g^2-gw+\frac{1}{2}bg\right)\\
=& (n^2(3w+g+1))^{-1}\cdot\left(-bw-\frac{1}{2}bg+gw+\frac{1}{2}g^2\right)\\
=& (n^2(3w+g+1))^{-1}\cdot\left(-\frac{1}{2}g(2w+g)+\frac{1}{2}g(2w+g)\right)\\
\leq& -\frac{b}{2n^2}\cdot\frac{2w+g}{3w+g+1}+\frac{g(2w+g)}{2n^2(3w+g)}\\
=& \frac{1}{2n}\left(-\tilde{b}\cdot\frac{2w+g}{3w+g+1}+\frac{\tilde{g}(2\tilde{w}+\tilde{g})}{3\tilde{w}+\tilde{g}}\right)\\
\leq& \frac{1}{2n}\left(-\frac{3}{4}\cdot\frac{1}{2}+\frac{1}{4}\right)\\
=& -\frac{1}{16n}
\end{aligned}$$

where the last inequality comes from the facts that function $(2y+x)/(3y+x+1)$ given $x,y\geq 0$ and $\max(x,y)\geq 1$ is at least $1/2$, and that function $(x(2y+x))/(3y+x)$ given $x,y\geq 0$ and $x+y\leq 1/4$ is at most $1/4$.

When the population is not in the large-b region, i.e., $\tilde{b} < 3/4$, we have $w + g \geq n/4$ and

$$\frac{\Delta f}{f} \leq \frac{2I^{g-} + I^{g+}}{3w + g + 1} \leq \frac{2I^{g-} + I^{g+}}{n/4} = (8I^{g-} + 4I^{g+})n^{-1}$$

Hence, from Lemma 4 the stochastic process $\{B_t\}$ given by

$$B_t = (3w_t + g_t + 1) \cdot \exp\left(\left(\frac{1}{17}S_t^b - \sum_{i=1}^{t}\left(9I_i^{g-} + 5I_i^{g+}\right)(1 - I_i^b)\right)/n\right)$$

is a supermartingale. This gives us the bound

$$\mathbb{E}\left(\exp\left(\left(\frac{1}{17}S_\tau^b - \sum_{i=1}^{\tau}\left(9I_i^{g-} + 5I_i^{g+}\right)(1 - I_i^b)\right)/n\right)\right) \leq \mathbb{E}B_\tau \leq B_0 \leq 3n + 1$$

Again for Markov's inequality and $\sum_{i=1}^{t} I_i^{g-}(1 - I_i^b) \leq S_t^{g-}$ and $\sum_{i=1}^{t} I_i^{g+}(1 - I_i^b) \leq S_t^{g+}$, we have

$$\mathbb{P}\left(\frac{1}{17}S_\tau^b - 9S_\tau^{g-} - 5S_\tau^{g+} \geq cn\log n + n\log(3n+1)\right) \leq n^{-c}$$

and

$$\mathbb{P}\left(S_\tau^b \geq 153S_\tau^{g-} + 85S_\tau^{g+} + 17cn\log n + 17n\log(3n+1)\right) \leq n^{-c}$$

which completes the proof. ■

Then the number of interactions S_τ^w in the other region where $\tilde{w} \geq 3/4$ can be bounded in a symmetric way using the potential function $f = 3b + g + 1$.

Lemma 12. *With probability* $1 - o(1)$, $S_\tau^w = O(n\log n)$. *In addition, for any constant* $c > 0$, *we have*

$$\mathbb{P}\left(S_\tau^w \geq 153S_\tau^{g-} + 85S_\tau^{g+} + 17cn\log n + 17n\log(3n+1)\right) \leq n^{-c}$$

Finally, combining all the above lemmas implies a bound on $\tau = S_\tau^c + S_\tau^b + S_\tau^g + S_\tau^w$ that

$$\mathbb{P}\left(\tau \geq 96930(c+1)n\log n\right) \leq \max\left(9n^{-c}, \frac{c\log n}{\sqrt[3]{n}}\right)$$

As we explained in Section 2, let $\tau = \min(\tau_*, 10^5(c+1)n\log n)$. This makes τ a well-defined stopping time and eventually, we have Theorem 1. Again the $O\left(\frac{\log n}{\sqrt[3]{n}}\right)$ error term is because all the lemmas give error bounds at most $O(n^{-c})$ except Lemma 8, which gives an $O\left(\frac{\log n}{\sqrt[3]{n}}\right)$ error bound and dominates the other error terms (with a tiny increase in the constant) when c is large. ■

Appendix B Proof of Theorem 2

This section provides our proof of Theorem 2.

The proof is done by constructing a martingale process that provides a lower bound of the difference between the majority population and the minority population. We show that at the

point of convergence, this difference is positive with high probability, which means the population protocol correctly computes the initial majority value.

Theorem 2 (in the main paper) *The binary signaling protocol converges to the initial majority value with high probability, if the initial difference between the majority and the minority population is* $\omega(\sqrt{n \log n})$.

Proof Let $u_t = b_t - w_t$ being the difference between black and white population at time t. Denote by $p_t = (\frac{1}{2}g_t + b_t)(n - b_t)/n^2$ the probability of $\Delta u_t = 1$ and by $q_t = (\frac{1}{2}g_t + w_t)(n - w_t)/n^2$ the probability of $\Delta u_t = -1$. Without loss of generality, we assume that black is the initial majority. We construct a martingale process u'_t as a lower bound of u_t by defining the joint distribution of $(\Delta u_t, \Delta u'_t)$ as below:

$$(\Delta u_t, \Delta u'_t) = \begin{cases} (0,0) & \text{with probability } 1 - p_t - q_t \\ (1,-1) & \text{with probability } p_t - \frac{1}{2}(p_t + q_t) \\ (1,1) & \text{with probability } \frac{1}{2}(p_t + q_t) \\ (-1,-1) & \text{with probability } q_t \end{cases}$$

The margin distribution of Δu_t matches the behavior of u_t and from the margin distribution of $\Delta u'_t$, we have $u't$ being a martingale process $\mathbb{E}(\Delta u'_t) = 0$ and $|\Delta u'_t| \leq 1$. Note that the probability $\mathbb{P}(\Delta u_t = 1, \Delta u'_t = -1) = p_t - \frac{1}{2}(p_t + q_t)$ is well defined if $p_t \geq q_t$.

$$\begin{aligned} p_t - q_t =& \frac{1}{n^2}\left(\left(\frac{1}{2}g_t + b_t\right)(n - b_t) - \left(\frac{1}{2}g_t + w_t\right)(n - w_t)\right) \\ =& \frac{1}{n^2}\left(\frac{1}{2}g_t(w_t - b_t) + n(b_t - w_t) - (b_t^2 - w_t^2)\right) \\ =& \frac{1}{n^2}\left(\frac{1}{2}g_t(w_t - b_t) + (b_t - w_t)(n - (b_t + w_t))\right) \\ =& \frac{1}{n^2}\left(-\frac{1}{2}g_t(b_t - w_t) + g_t(b_t - w_t)\right) \\ =& \frac{1}{2n^2} \cdot g_t(b_t - w_t) \\ =& \frac{1}{2n^2} \cdot g_t u_t \end{aligned}$$

So $p_t \geq q_t$ if $u_t \geq 0$. Thus when u_t drops to zero, we end the process. Letting $u'0 = u_0$, then from the coupling definition of $(\Delta u_t, \Delta u'_t)$, we always have $u_t \geq u'_t$ before the process ends.

For Theorem 1, the binary signaling process reaches convergence within $O(n \log n)$ interactions with high probability. Let $\tau = O(n \log n)$ be the time when the process converges. Given that u'_t is a martingale process, by Azuma's inequality, we have

$$\mathbb{P}(|u'_\tau - u'_0| \geq \delta) \leq 2\exp\left(\frac{-\delta^2}{2\tau \cdot 1^2}\right) = 2\exp\left(\frac{-\delta^2}{O(n \log n)}\right)$$

Thus for any $\delta = \omega(\sqrt{n \log n})$, we have $|u'_\tau - u'_0| < \delta$ with high probability. if initially $u'_0 = u_0 = \omega(\sqrt{n \log n})$, then we have $u_t \geq u'_t > 0$ before the process ends with high probability, which also indicates $u_\tau > 0$ at the convergence. Hence, the binary signaling protocol converges to the initial majority value correctly with high probability. ■

Appendix C Proof of Theorem 3

In this section we present a formal proof of Theorem 3. The correctness of Theorem 4 follows Theorem 3.

Denote by $z = o(\sqrt{n})$ the number of Byzantine agents in the population. We start from showing that in spite of the existence of Byzantine agents, there is a strong bias pushing the process away from the large-g corner. More concretely, it's unlikely to have $v \leq \sqrt{n}$ within any polynomial number of interactions (Lemma 13). However, once the process is in the large-b corner or the large-w corner, it will remain there within any polynomial number of interactions with high probability. We prove this by showing that $Z_t = \exp((2g_t + 5w_t)/16)$ is a supermartingale process for the large-b case, such that it is difficult to have $2g + 5w \geq 5\sqrt{n}$ within any polynomial time (Lemma 14). The large-w case can be achieved in the symmetric way. Switching to the Byzantine case, we still keep the potential functions used for the non-Byzantine case in Theorem 1, but include an adjustment factor that compensates for increases due to Byzantine interactions. We show that the effect of this adjustment factor is small and the remaining quantities are at most $n^{o(1)}$ times their original values with high probability. Eventually combining all the results gives us Theorem 3.

Lemma 13. *If the process starts with $v_0 \geq \sqrt{n} + c\log_{7/5} n$ and $z = o(\sqrt{n})$, then with probability $1 - n^{-c}$, we have $v_t > \sqrt{n}$ holds for all $1 \leq t \leq T$, for any $T = poly(n)$ and any constant $c > 0$.*

Proof Note that in the worst case against this statement, all the Byzantine agents will attempt to decrease v. We have $\mathbb{P}(\Delta v = 1) = (vg+g^2)/(z+n)^2$ and $\mathbb{P}(\Delta v = -1) = (2bw+zv+\frac{1}{2}gv)/(z+n)^2$. Suppose the process is at the large-g corner, i.e., $g \geq (3/4)n$ and $\sqrt{n} \leq v \leq n/4$. Then

$$\begin{aligned}
\mathbb{P}(\Delta v = 1 | \Delta v \neq 0) =& \frac{vg + g^2}{2bw + zv + \frac{1}{2}gv + vg + g^2} \\
>& \frac{vg}{2bw + zv + \frac{1}{2}gv + vg} \\
>& \frac{vg}{\frac{1}{2}v^2 + zv + \frac{1}{2}gv + vg} \\
=& \frac{g}{\frac{1}{2}v + z + \frac{3}{2}g} \\
>& \frac{g}{\frac{1}{2}v + \frac{1}{14}(3g - 7v) + \frac{3}{2}g} \\
=& \frac{7}{12}
\end{aligned}$$

This is due to $g \geq 3v$ and $z = o(\sqrt{n}) < \frac{1}{14}(3g - 7v) = \Theta(n)$. This implies $\mathbb{P}(\Delta v = -1 | \Delta v \neq 0) < 5/12$, so the bias of v being increasing is strong, in spite of the existence of Byzantine agents.

To prove this lemma, consider a one-dimensional random walk on integers from 0 to ℓ with forward probability 7/12 and backward probability 5/12. Each state i corresponds to the configuration with $v = \sqrt{n} + i$. In our case we let $\ell = n/4$.

Denote by $\beta_i = \mathbb{P}(\text{hitting state } \ell \text{ before hitting state } 0 \mid \text{ starting at state } i)$. Then $\beta_0 = 0$ and $\beta_\ell = 1$.

For any $1 \le i \le \ell - 1$, $\beta_i = \frac{7}{12}\beta_{i+1} + \frac{5}{12}\beta_{i-1}$. Define

$$\begin{aligned}\Delta\beta_i =& \beta_i - \beta_{i-1}\\ =& \beta_i - \frac{7}{12}\beta_i - \frac{5}{12}\beta_{i-2}\\ =& \frac{5}{12}(\beta_i - \beta_{i-2})\\ =& \frac{5}{12}(\beta_i - \beta_{i-1} + \beta_{i-1} - \beta_{i-2})\\ =& \frac{5}{12}\Delta\beta_i + \frac{5}{12}\Delta\beta_{i-1}\end{aligned}$$

which implies $\frac{7}{12}\Delta\beta_i = \frac{5}{12}\Delta\beta_{i-1}$ or $\Delta\beta_i = \frac{5}{7}\Delta\beta_{i-1}$. Thus $\Delta\beta_i = (5/7)^{i-1}\Delta\beta_1 = (5/7)^{i-1}\beta_1$.

Note that

$$\sum_{i=1}^{\ell} \Delta\beta_i = \beta_1 - \beta_0 + \beta_2 - \beta_1 + \ldots + \beta_\ell - \beta_{\ell-1} = \beta_\ell - \beta_0 = 1$$

Then

$$\sum_{i=1}^{\ell} \Delta\beta_i = \beta_1 \sum_{i=1}^{\ell} (5/7)^{i-1} = \beta_1 \cdot \frac{1 - (5/7)^\ell}{2/7} = 1$$

Therefore, $\beta_1 = \frac{2/7}{1-(5/7)^\ell}$ and for $k = c\log_{7/5} n$,

$$\begin{aligned}\beta_k =& \sum_{i=1}^{k} \Delta\beta_i = \beta_1 \sum_{i=1}^{k} (5/7)^{i-1}\\ =& \frac{2/7}{1-(5/7)^\ell} \cdot \frac{1-(5/7)^k}{2/7}\\ =& \frac{1-(5/7)^k}{1-(5/7)^\ell}\\ =& 1 - \frac{(5/7)^k - (5/7)^\ell}{1-(5/7)^\ell}\\ =& 1 - \frac{(7/5)^{\ell-k} - 1}{(7/5)^\ell - 1}\\ =& 1 - n^{-c}\end{aligned}$$

This means if the process starts with $v_0 \ge \sqrt{n} + c\log_{7/5} n$, with probability $1 - n^{-c}$, it will reach $v = n/4$ first, instead of reaching $v = \sqrt{n}$.

Once the process has $v \ge n/4$, we can show that the probability of the process dropping to $v = \sqrt{n}$ within polynomial steps is exponentially small. We have

$$\beta_{\ell-1} = 1 - \frac{(7/5) - 1}{(7/5)^\ell - 1} = 1 - \Theta\left(\left(\frac{5}{7}\right)^n\right)$$

Every time when v drops to $n/4 - 1 = \ell - 1$, consider it as a Bernoulli trial with success probability $1 - \beta_{\ell-1}$. For any polynomial $T = poly(n) = n^d$ for some constant $d > 0$. The probability of at least one success in the first T trials is at most

$$T \cdot (1 - \beta_{\ell-1}) = n^d \cdot \Theta\left(\left(\frac{5}{7}\right)^n\right) = \Theta\left(\left(\frac{5}{7}\right)^{n - d\log n}\right)$$

which is also exponentially small and is absorbed in n^{-c}. Therefore, the probability of the process dropping to $v \le \sqrt{n}$ is less than n^{-c}. ■

Lemma 13 tells us it is difficult to reach a configuration with a low value of v, while the following lemma shows that once the protocol reaches a configuration with a high value of b, it will remain in that corner.

Lemma 14. *If the process starts with $2g_0 + 5w_0 \le 3\sqrt{n}$ and $z = o(\sqrt{n})$, then with probability $1 - \exp(-\sqrt{n}/9)$, we have $2g_t + 5w_t < 5\sqrt{n}$ holds for all $1 \le t \le T$, for any $T = poly(n)$.*

Proof We show that the stochastic process $\{Z_t\}$ given by

$$Z_t = \exp((2g_t + 5w_t)/16)$$

is a supermartingale. Consider the worst case where all the Byzantine agents attempt to increase $2g + 5w$, we have

$$\Delta(2g+5w) = \begin{cases} +3 & \text{with probability } ((w+z)g + \frac{1}{2}g^2)/(n+z)^2 \\ +2 & \text{with probability } ((w+z)b + \frac{1}{2}gb)/(n+z)^2 \\ -2 & \text{with probability } (bg + \frac{1}{2}g^2)/(n+z)^2 \\ -3 & \text{with probability } (bw + \frac{1}{2}gw)/(n+z)^2 \end{cases}$$

We divide all the possibilities into three types of events: event A where the interaction involves a black token as either initiator or responder, event B where $\Delta(2g+5w) = 3$, and event C where $\Delta(2g+5w) < 0$ with a gray initiator. Note that $A \cap B \cap C = \emptyset$.

Suppose $\sqrt{n} \le 2g + 5w \le 5\sqrt{n}$, so $b \ge n - O(\sqrt{n})$ and $z = o(\sqrt{n}) < \frac{1}{4}w + \frac{3}{4}g$. We have

$$\begin{aligned}
\mathbb{E}(\exp(\Delta((2g+5w)/16))|A) =& \frac{\left(\frac{1}{2}g + w + z\right)\exp\left(\frac{1}{8}\right) + g\exp\left(-\frac{1}{8}\right) + w\exp\left(-\frac{3}{16}\right)}{2w + \frac{3}{2}g + z} \\
\le& \frac{\left(\frac{1}{2}g + w + z\right)\left(1 + \frac{1}{8} + \frac{1}{64}\right) + g\left(1 - \frac{1}{8} + \frac{1}{64}\right) + w\left(1 - \frac{3}{16} + \frac{9}{156}\right)}{2w + \frac{3}{2}g + z} \\
=& \frac{\left(2w + \frac{3}{2}g + z\right) + \frac{1}{16}\cdot(z - g - w) + \frac{1}{256}\cdot(6g + 13w + 4z)}{2w + \frac{3}{2}g + z} \\
=& 1 - \frac{1}{16}\cdot\frac{(1-o(1))(g+w)}{2w + \frac{3}{2}g + z} + \frac{1}{256}\cdot\frac{13w + 6g + 4z}{2w + \frac{3}{2}g + z} \\
<& 1 - \frac{1}{16}\cdot\frac{(1-o(1))(g+w)}{2w + \frac{3}{2}g + \frac{1}{4}w + \frac{3}{4}g} + \frac{1}{256}\cdot\frac{13w + 6g + 4z + \frac{15}{4}g + \frac{5}{2}z}{2w + \frac{3}{2}g + z} \\
=& 1 - \frac{1}{16}\cdot\frac{4}{9} + \frac{1}{256}\cdot\frac{13}{2} \\
=& 1 - \frac{11}{4608}
\end{aligned}$$

The probability of event B is proportional to $g\left(w + z + \frac{1}{2}g\right)$ and the probability of event C is proportional to $\frac{1}{2}g(g+w)$, while the probability of event A is proportional to $b\left(2w + \frac{3}{2}g + z\right)$. Recall that $b = O(n)$, $g+w = O(\sqrt{n})$ and $z = o(\sqrt{n})$, we have $\mathbb{P}(A) = O(\sqrt{n}\cdot\mathbb{P}(B)) = O(\sqrt{n}\cdot\mathbb{P}(C))$.

Therefore, we have

$$\begin{aligned}
&\mathbb{E}(\exp(\Delta((2g+5w)/16))|A\cup B\cup C)\\
<&\left(1-\frac{11}{4608}\right)\cdot\mathbb{P}(A|A\cup B\cup C)+\exp\left(\frac{3}{16}\right)\cdot\mathbb{P}(B|A\cup B\cup C)+\exp\left(-\frac{1}{8}\right)\cdot\mathbb{P}(C|A\cup B\cup C)\\
<&1-\frac{11}{4608}+O\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

which is strictly less than 1.

If none of $A\cup B\cup C$ happens, $Z_t = Z_{t-1}$. Thus we have $\mathbb{E}(Z_t)\le Z_{t-1}$, and that $\{Z_t\}$ is a supermartingale.

Assume that initially $\sqrt{n} < 2g+5w \le 3\sqrt{n}$. Let τ be the stopping time when $2g+5w$ first reaches either $\sqrt{n}$ or $5\sqrt{n}$. Given that $\{Z_t\}$ is a supermartingale, we have

$$\mathbb{P}(2g_\tau+5w_\tau=5\sqrt{n})\cdot\exp\left(\frac{5\sqrt{n}}{16}\right)\le\mathbb{E}(Z_\tau)\le Z_0<\exp\left(\frac{3\sqrt{n}}{16}\right)$$

Thus, we have $\mathbb{P}(2g_\tau+5w_\tau=5\sqrt{n})<\exp\left(-\frac{\sqrt{n}}{8}\right)$.

Similar to the proof of Lemma 13, every time when the process drops into $\sqrt{n}<2g+5w\le 3\sqrt{n}$, consider it as a Bernoulli trial with success probability $\exp\left(-\frac{\sqrt{n}}{8}\right)$. For any polynomial $T=poly(n)=n^c$ for some constant $c>0$. The probability of at least one success in the first T trials is at most

$$T\cdot\exp\left(-\frac{\sqrt{n}}{8}\right)=n^c\cdot\exp\left(-\frac{\sqrt{n}}{8}\right)=\exp\left(-\frac{\sqrt{n}}{8}+c\log n\right)<\exp\left(-\frac{\sqrt{n}}{9}\right)$$

which is exponentially small. Therefore, the probability of having $2g+5w>5\sqrt{n}$ is less than $\exp(-\sqrt{n}/9)$. ■

Same argument for the large-w case can be achieved in the symmetric way.

Lemma 15. *If the process starts with $2g_0+5b_0\le 3\sqrt{n}$ and $z=o(\sqrt{n})$, then with probability $1-n^{-c}$, we have $2g_t+5b_t<5\sqrt{n}$ holds for all $1\le t\le T$, for any $T=poly(n)$ and any constant $c>0$.*

To continue our proof of Theorem 3, we need a lemma from [AAE07]. Denote by S_t^z the number of interactions with a Byzantine initiator by time t, we have

Lemma 16 ([AAE07])**.** *Let f be a function of the states of the non-Byzantine agents and their interaction history in some one-way population protocol, such that f_t is a supermartingale in the absence of Byzantine agents. Let $(f+\Delta f)/f\le m$ for all transitions involving a Byzantine initiator starting in some subset D of the configuration space. Then $f'_t=f_t m^{-S_t^z}$ is a supermartingale for all times t less than the first time at which the protocol leaves D.*

Based on the above lemmas, now we are able to prove Theorem 3.

Theorem 3 (in the main paper) *Let τ be the time when $b\ge n-\sqrt{n}$, $w\ge n-\sqrt{n}$ or $b+w\le\sqrt{n}$ first holds. If the number of Byzantine agents in the population is $o(\sqrt{n})$ and initially $b_0+w_0\ge\sqrt{n}+c\log_{7/5}n$, then*

$$\mathbb{P}\left(\tau\ge 96930(c+1)n\log n\right)\le\max\left(n^{-c+o(1)},\frac{c\log n}{\sqrt[3]{n}}\right)$$

and

$$\mathbb{P}(b_\tau + w_\tau \le \sqrt{n}) \le n^{-c}$$

for any constant $c > 0$.

Proof Recall that in our proof of Theorem 1 for the non-Byzantine case, we construct five potential functions. Now we apply Lemma 16 to each potential function.

For $f = 1/(u^2 + 5n/2)$, let D be the entire space. We have

$$\frac{f + \Delta f}{f} \le \frac{u^2 + 5n/2}{(u-1)^2 + 5n/2} = 1 + \frac{2(u-1)+1}{(u-1)^2 + 5n/2} \le 1 + \frac{1}{\sqrt{5n/2}} < 1 + \frac{3}{\sqrt{n}}$$

For $f = 1/(u^2 + 64n)$, let D be the entire space. We have

$$\frac{f + \Delta f}{f} \le \frac{u^2 + 64n}{(u-1)^2 + 64n} = 1 + \frac{2(u-1)+1}{(u-1)^2 + 64n} \le 1 + \frac{1}{8\sqrt{n}} < 1 + \frac{3}{\sqrt{n}}$$

For $f = 1/(2v+1)$, let D be all points with $v \ge \sqrt{n}$. We have

$$\frac{f + \Delta f}{f} \le \frac{2v+1}{2(v-1)+1} = 1 + \frac{2}{2v-1} \le 1 + \frac{2}{2\sqrt{n}-1} < 1 + \frac{3}{\sqrt{n}}$$

For $f = 3w + g + 1$, let D be all points with $3w + g + 1 \ge \sqrt{n}$. We have

$$\frac{f + \Delta f}{f} \le \frac{3(w+1) + g + 1}{3w + g + 1} = 1 + \frac{3}{3w + g + 1} \le 1 + \frac{3}{\sqrt{n}}$$

Similarly, for $f = 3b + g + 1$, let D be all points with $3b + g + 1 \ge \sqrt{n}$. We have $(f + \Delta f)/f \le 1 + 3/\sqrt{n}$. Therefor, we set $m = 1 - 3/\sqrt{n}$ in Lemma 16 and let D be where $v \ge \sqrt{n}$, $b \le n - \sqrt{n}$ and $w \le n - \sqrt{n}$ all hold. Note that D lies in the interaction of the above five domains, since we have $3w + g + 1 > n - b + 1 \ge \sqrt{n}$ and $3b + g + 1 > n - w + 1 \ge \sqrt{n}$.

Let $\tau' = O(n \log n)$ be a stopping time. With $z = o(\sqrt{n})$, we have $\mathbb{E}((\sqrt{n}/n)(n \log n)) = o(\sqrt{n} \log n)$. Chernoff bounds show that this bound holds without the expectation with probability at least $1 - n^{-c}$ for any constant $c > 0$. Hence, we have

$$m^{S^z_{\tau'}} \le \left(1 + \frac{3}{\sqrt{n}}\right)^{o(\sqrt{n} \log n)} \le \exp\left(\frac{3}{\sqrt{n}} \cdot o(\sqrt{n} \log n)\right) = \exp(o(\log n)) = n^{o(1)}$$

which implies $\mathbb{P}(m^{S^z_{\tau'}} > n^o(1)) \le n^{-c}$.

Recall that all the bounds in our proof of Theorem 1 are obtained by applying Markov's inequality to the ratio of two potential function values. Any terms in these potential functions involving only indicator variables are unaffected by Byzantine interactions. Now applying Lemma 16, it follows that the remaining quantities are at most $n^{o(1)}$ times their original values, with probability at least $1 - n^{-c}$. This increases the probability of failure obtained from Markov's inequality from n^{-c} to at most $n^{-c+o(1)}$. Combining the results in this section with the bounds in Theorem 1 gives the bound on τ. Lemma 13 provides the bound on $v_\tau = b_\tau + w_\tau$ and completes the proof. ■

The correctness of Theorem 4 follows Theorem 3.

Theorem 4 (in the main paper) *If the number of Byzantine agents in the population is* $o(\sqrt{n})$ *and initially* $b_0 + w_0 \ge \sqrt{n} + c \log_{7/5} n$*, and the initial difference between the majority and the*

minority population is $\Omega(\sqrt{n}\log n)$, the binary signaling protocol converges to the initial majority value with high probability.

Proof We apply the same analysis in the proof of Theorem 2 to the Byzantine case, based on the result of Theorem 3. For $z = o(\sqrt{n})$, with high probability, there are at most $o(\sqrt{n}) \cdot O(n \log n) = o(\sqrt{n}\log n)$ Byzantine interactions within $O(n \log n)$ interactions. In the worst case the Byzantine interactions will affect at most $o(\sqrt{n}\log n)$. Thus we lift the requirement on the initial difference between the majority and the minority from $\omega(\sqrt{n \log n})$ in Theorem 2 to $\Omega(\sqrt{n}\log n)$. This gives us Theorem 4. ■

Appendix D Proof of Theorem 5

In this section we prove Theorem 5.

We divide the process into two stages, by the point when there are $\left(n^{3/4}(\log n)^{-1/4}\right)$ active agents in the population. We show that there are at most $O(\sqrt{n \log n})$ active-active interactions in the first stage with high probability, and the majority enjoys an advantage of $\omega(\sqrt{n \log n})$ to start the second stage. Thus we can apply the same analysis from the proof of Theorem 2.

Theorem 5 (in the main paper) *If the initial difference between the majority and the minority population is $\omega((n \log n)^{3/4})$ and there is exactly one initially active agent, then the binary signaling protocol with an epidemic-triggered start converges to the initial majority value with high probability.*

Proof We consider the whole process as two stages. The first stage is from the initial state with exactly one active agent to the time when there are $\left(n^{3/4}(\log n)^{-1/4}\right)$ active agents in the population. The second stage is from the end of the first stage to convergence. Without of loss of generality, let black be the initial majority value. Denote by b_t^{act} the number of active black agents at time t and by w_t^{act} the number of active white agents. Let $u_t^{act} = b_t^{act} - w_t^{act}$. Denote by b_t^{init} the number of active initially black agents at time t and by w_t^{init} the number of active initially white agents. Let $u_t^{init} = b_t^{init} - w_t^{init}$.

Angluin et al. [AAE08] show that an initial configuration with at least one active agent will reach a configuration where all gents are active within $O(n \log n)$ interactions with high probability. Hence, the first stage will end within $O(n \log n)$ interactions. The value of u_t^{act} can be changed by either a biased agent getting active, or an interaction between two active agents. During the first stage there are at most $\left(n^{3/4}(\log n)^{-1/4}\right)$ active agents, so the probability of an active-active interaction is at most

$$\mathbb{P}(\text{active-active interaction} \mid \text{stage 1}) \leq \left(\frac{n^{\frac{3}{4}}}{n \cdot (\log n)^{\frac{1}{4}}}\right)^2 = \frac{1}{\sqrt{n \log n}}$$

Let S_1^{act} be the number of active-active interactions in stage 1. By Markov's inequality,

$$\mathbb{P}\left(S_1^{act} \geq \delta\right) \leq \frac{\mathbb{E}(S_1^{act})}{\delta} \leq \frac{1}{\delta} \cdot O(n \log n) \cdot \frac{1}{\sqrt{n \log n}} = \frac{O(\sqrt{n \log n})}{\delta}$$

Therefore, $\mathbb{P}\left(S_1^{act} \geq \delta\right) = o(1)$ for any $\delta = \omega(\sqrt{n \log n})$ and we have $S_1^{act} = O(\sqrt{n \log n})$ with high probability.

At the start of the second stage, there are $\left(n^{3/4}(\log n)^{-1/4}\right)$ active agents, so the expectation of u_t^{init} is

$$\mathbb{E}(u_t^{init}|t = \text{ start of stage 2}) = \frac{n^{\frac{3}{4}}}{n \cdot (\log n)^{\frac{1}{4}}} \cdot \omega\left((n \log n)^{\frac{3}{4}}\right) = \omega(\sqrt{n \log n})$$

Let $X_1, X_2, \ldots, X_t$ be the responders of the first t interactions. These are independent random variables from the same space (i.e., the population). Also note that the colors of the initiators and the signals from them don't affect the value of u_t^{init}. Therefore, we view $u_t^{init} = f(X_1, X_2, \ldots, X_t)$ as a function of $X_1, X_2, \ldots, X_t$. We can see function f is Lipschitz with respect to Hamming distance on its domain, with bound 2. By McDiarmids inequality,

$$\mathbb{P}(|u_t^{init} - \mathbb{E}(u_t^{init})| \geq \delta) \leq 2\exp\left(-\frac{2\delta^2}{t \cdot 2^2}\right) = 2\exp\left(-\frac{\delta^2}{2 \cdot O(n \log n)}\right)$$

For any $\delta = \omega(\sqrt{n \log n})$, this probability goes to 0, so $|u_t^{init} - \mathbb{E}(u_t^{init})| = O(\sqrt{n \log n})$ with high probability. Thus the value of u_t^{init} is $\omega(\sqrt{n \log n})$ at the start of stage 2 with high probability. Recall that there are $O(\sqrt{n \log n})$ active-active interactions in the first stage with high probability, which can change u_t^{act} from u_t^{init} by at most $O(\sqrt{n \log n})$. Therefore, with high probability, $u_t^{act} = \omega(\sqrt{n \log n})$ at the start of stage 2. Since the majority enjoys an advantage of $\omega(\sqrt{n \log n})$ to start the second stage, we can apply the same analysis from the proof of Theorem 2 and complete this proof. ■

Appendix E Proof of Theorem 6

This section provides a complete proof of Theorem 6.

Theorem 6 (in the main paper) *If the initial configuration is monotone, then the continuous process will reach consensus within $O(r \log nr)$ time.*

The proof starts with derivation of the corresponding ODE system of the process, which can be inferred by taking the limit on the expectation of the configuration vector. This ODE system provides a mathematical formula of the vector field in the configuration space. We show that the vector field anywhere at the boundary of the monotone region always points inwards into the monotone region, which means the process stays in the monotone region and never leaves. We divide the monotone region into two sub-areas A_+, the region where $x_0 \leq x_1 \leq \ldots \leq x_r$ with at least one $<$ in the middle, and A_-, the region where $x_0 \geq x_1 \geq \ldots \geq x_r$ with at least one $>$ in the middle. The ODE system also gives us the differential equation for p, from which we prove that p is always increasing in A_+ and is always decreasing in A_-. It suffices to show the convergence bound for A_+ and it will holds for A_- in a symmetric way.

The above two facts already tell us that once the process enters A_+, p will keep increasing until convergence. What we need is a positive lower bound for the derivative of p that will lead to the desired convergence time. We need to take care of two cases where $\mathrm{d}p/\mathrm{d}t$ is very small. The first case is when the process is almost at convergence and p is very close to 1. The other is when the configuration vector is almost uniform and p is very close to $1/2$. To do so, we divide the path of p from $1/2 + 1/(nr)$ to $1 - 1/(nr)$ into two corresponding stages: from $2/3$ to $1 - 1/(nr)$ and from $1/2 + 1/(nr)$ to $2/3$. We show the time for the former stage is $O(\log nr)$ and the time for the latter is $O(r \log nr)$.

Lemma 17. *Once the process enters the monotone region, it never leaves.*

Proof The corresponding systems of differential equations of the process can be inferred by taking the limit of the expectation of the configuration vector. For the change of $\{x_i\}$ and p from time tick t_k to the next time tick t_{k+1}, we have

$$\begin{cases} x_0(t_{k+1}) = x_0(t_k) + (1-p(t_k)) \cdot x_1(t_k) \cdot \frac{1}{n} - p(t_k) \cdot x_0(t_k) \cdot \frac{1}{n} \\ x_i(t_{k+1}) = x_i(t_k) + p(t_k) \cdot x_{i-1}(t_k) \cdot \frac{1}{n} + (1-p(t_k)) \cdot x_{i+1}(t_k) \cdot \frac{1}{n} - x_i(t_k) \cdot \frac{1}{n} \\ x_r(t_{k+1}) = x_r(t_k) - (1-p(t_k)) \cdot x_r(t_k) \cdot \frac{1}{n} + p(t_k) \cdot x_{r-1}(t_k) \cdot \frac{1}{n} \end{cases}$$

where the second equation is for $1 \leq i \leq r-1$. Dividing both sides of the equations by the infinitesimal $1/(nr)$ gives

$$\begin{cases} \dfrac{x_0(t_{k+1}) - x_0(t_k)}{1/(nr)} = r \cdot ((1-p(t_k)) \cdot x_1(t_k) - p(t_k) \cdot x_0(t_k)) \\ \dfrac{x_i(t_{k+1}) - x_i(t_k)}{1/(nr)} = r \cdot (p(t_k) \cdot x_{i-1}(t_k) + (1-p(t_k)) \cdot x_{i+1}(t_k) - x_i(t_k)) \\ \dfrac{x_r(t_{k+1}) - x_r(t_k)}{1/(nr)} = r \cdot (p(t_k) \cdot x_{r-1}(t_k) - (1-p(t_k)) \cdot x_r(t_k)) \end{cases}$$

Since nr is the rate of the Poisson process and we let $1/(nr)$ go to 0, we have the ODE system

$$\begin{cases} \dfrac{\mathrm{d}x_0}{\mathrm{d}t} = r \cdot ((1-p) \cdot x_1 - p \cdot x_0) \\ \dfrac{\mathrm{d}x_i}{\mathrm{d}t} = r \cdot (p \cdot x_{i-1} + (1-p) \cdot x_{i+1} - x_i) \\ \dfrac{\mathrm{d}x_r}{\mathrm{d}t} = r \cdot (p \cdot x_{r-1} - (1-p) \cdot x_r) \end{cases}$$

This ODE system provides a mathematical formula for the vector field in the configuration space. To complete the proof, we need to show that the vector field anywhere at the boundary of the monotone region always points inwards into the monotone region. Let A_+ be the region where $x_0 \leq x_1 \leq \ldots \leq x_r$ with at least one $<$ in the middle and A_- be the region where $x_0 \geq x_1 \geq \ldots \geq x_r$ with at least one $>$ in the middle. It is sufficient to prove the lemma for A_+ and the proof for A_- will hold in a symmetric way.

For some $1 \leq i \leq r-2$, when the process is close to a point where $x_{i+1} - x_i = 0$, given that the process is in region A_+, we have

$$\begin{aligned} \frac{\mathrm{d}(x_{i+1} - x_i)}{\mathrm{d}t} =& r \cdot (px_i + (1-p)x_{i+2} - x_{i+1} - px_{i-1} - (1-p)x_{i+1} + x_i) \\ =& r \cdot (p(x_i - x_{i-1}) + (1-p)(x_{i+2} - x_{i+1})) > 0 \end{aligned}$$

which pushes the system back to the area with $x_{i+1} - x_i > 0$.

Notice that the probability of positive interaction $p = \sum_{i=0}^{r} (i/r) x_i$ is always greater than $1/2$ in region A_+. When the process is close to a point where $x_1 - x_0 = 0$ or $x_r - x_{r-1} = 0$, we have

$$\begin{aligned} \frac{\mathrm{d}(x_1 - x_0)}{\mathrm{d}t} =& r \cdot (px_0 + (1-p)x_2 - x_1 - (1-p)x_1 + px_0) \\ =& r \cdot (2px_0 - x_1 + (1-p)(x_2 - x_1)) > 0 \end{aligned}$$

and

$$\begin{aligned}\frac{\mathrm{d}(x_r - x_{r-})}{\mathrm{d}t} =& r \cdot (px_{r-1} - (1-p)x_r - px_{r-2} - (1-p)x_r + x_{r-1}) \\ =& r \cdot (p(x_{r-1} - x_{r-2}) - 2(1-p)x_r + x_{r-1}) > 0\end{aligned}$$

which pushes the system back to the area with $x_1 - x_0 > 0$ and $x_r - x_{r-1} > 0$ respectively. Therefore, the continuous process will never escape from region A_+ once it is inside. We can similarly show symmetric results in the other region A_-. ■

Lemma 18. *The derivative of p is always positive in region A_+ and always negative in region A_-.*

Proof From the above ODE system and $p = \sum_{i=0}^{r}(i/r)x_i$ and $\sum_{i=0}^{r} x_i = 1$, we can infer the differential equation for p, which turns out to be very neat.

$$\frac{\mathrm{d}p}{\mathrm{d}t} = (1 - x_r)p - (1 - x_0)(1 - p)$$

We can interpret the differential equation in a very simple way. With probability p a positive interaction occurs. This is a stay-put interaction if and only if the responder is already fully confident. Thus with conditional probability $(1-x_r)$ it is a state-changing interaction and increases p by $1/(nr)$. Likewise, with probability $(1 - x_0)(1 - p)$, we will have a negative interaction that decreases p by $1/(nr)$.

Again, it suffices to prove the statement for region A_+. We have already shown the process stays in A_+ once it is inside. Denote by $B = \sum_{i=1}^{r-1}(i/r)x_i$, which is the contribution of x_1 to $x_r - 1$ to the probability p. Then $p = B + x_r$ and

$$\begin{aligned}\frac{\mathrm{d}p}{\mathrm{d}t} =& (1 - x_r)p - (1 - x_0)(1 - p) \\ =& (B + x_r)(1 - x_r) - (1 - B - x_r)(1 - x_0) \\ =& (2 - x_r - x_0)B + (1 - x_r)(x_r - 1 + x_0)\end{aligned}$$

Since $2 - x_r - x_0$ and $B \geq \sum_{i=1}^{r-1}(i/r) \cdot (1 - x_0 - x_r)/(r-1) = (1 - x_0 - x_r)/2$ in region A_+, we have

$$\frac{\mathrm{d}p}{\mathrm{d}t} \geq (2 - x_r - x_0) \cdot \frac{1}{2}(1 - x_0 - x_r) + (1 - x_r)(x_r - 1 + x_0)$$

Region A_+ also gives $1 = \sum_{i=0}^{r} x_i \geq x_0 + (r-1)x_0 + x_r$ and $0 \leq x_0 \leq (1 - x_r)/r$. Note that function $(2 - x - y)(1 - x - y)/2 + (1 - x)(x + y - 1)$ for $x \geq y$, $x + y < 1$ and $0 \leq y \leq (1 - x)/2$ is always non-negative. The only case where this function is 0 is when $x = y$ but we always have $x_0 < x_r$ in A_+. For any $r \geq 2$, $x_0 \leq (1 - x_r)/r \leq (1 - x_r)/2$. Thus, we always have $\mathrm{d}p/\mathrm{d}t > 0$ inside region A_+. In a symmetric way, we know $\mathrm{d}p/\mathrm{d}t < 0$ inside region A_-. ■

Proof (of Theorem 6) Again without loss of generality, we only study the region A_+. We know the probability of positive interaction p is always greater than $1/2$ inside A_+ and we expect a positive convergence $p \to 1$. The above two lemmas already tell us that once the process enters A_+, p will keep increasing until convergence. What we need is a positive lower bound for $\mathrm{d}p/\mathrm{d}t$ that will lead to the desired convergence time. Let $\varepsilon = p - 1/2$ and $\delta = 1 - p$. There are two cases where $\mathrm{d}p/\mathrm{d}t$ is very small.

1. The process is almost at convergence and p is very close to 1 with a very small δ;
2. The configuration vector is almost uniform and p is very close to $1/2$ with a very small ε.

To do so, we divide the path of p from $1/2 + 1/(nr)$ to $1 - 1/(nr)$ into two corresponding stages:

1. p goes from $\frac{2}{3}$ to $1 - \frac{1}{nr}$;
2. p goes from $\frac{1}{2} + \frac{1}{nr}$ to $\frac{2}{3}$.

For Stage 1, we have $2/3 \le p = \sum_{i=0}^{r}(i/r)x_i \le \sum_{i=0}^{r}(i/r)x_r = (1+r)x_r/2$ and $x_r \ge 4/(3(r+1))$. Note that function $(2-x-y)(1-x-y)/2+(1-x)(x+y-1)$ for $x \ge y$, $x+y<1$, $0 \le y \le (1-x)/r$ and $4/(3(r+1)) \le x \le 1-\delta$ is minimized at $(y = \delta/r, x = 1-\delta)$. Thus

$$\begin{aligned}
\frac{\mathrm{d}p}{\mathrm{d}t} \ge & (2 - x_r - x_0) \cdot \frac{1}{2}(1 - x_0 - x_r) + (1 - x_r)(x_r - 1 + x_0) \\
\ge & \left(2 - 1 + \delta - \frac{\delta}{r}\right) \cdot \frac{1}{2}\left(1 - 1 + \delta - \frac{\delta}{r}\right) + \delta\left(1 - \delta - 1 + \frac{\delta}{r}\right) \\
= & \left(1 + \left(1 - \frac{1}{r}\right)\delta\right) \cdot \frac{1}{2}\left(1 - \frac{1}{r}\right)\delta - \delta\left(1 - \frac{1}{r}\right)\delta \\
= & \frac{1}{2}\left(1 - \frac{1}{r}\right)\delta \cdot \left(1 + \left(1 - \frac{1}{r}\right)\delta - 2\delta\right) \\
= & \frac{1}{2}\left(1 - \frac{1}{r}\right)\delta \cdot \left(1 - \left(1 + \frac{1}{r}\right)\delta\right)
\end{aligned}$$

As $\delta < \frac{1}{2}$, we have

$$\begin{aligned}
\frac{\mathrm{d}p}{\mathrm{d}t} > & \frac{1}{2}\left(1 - \frac{1}{r}\right)\delta \cdot \left(1 - \left(1 + \frac{1}{r}\right)\frac{1}{2}\right) \\
= & \frac{1}{2}\left(1 - \frac{1}{r}\right)\delta \cdot \left(1 - \frac{1}{2} - \frac{1}{2r}\right) \\
= & \left[\frac{1}{2}\left(1 - \frac{1}{r}\right)\right]^2 \cdot \delta
\end{aligned}$$

We let $c = \left[\frac{1}{2}\left(1 - \frac{1}{r}\right)\right]^2 > 0$, which doesn't change with time. Now the ODE becomes simply $\mathrm{d}p/\mathrm{d}t > c(1-p)$ which is easy to solve. Let $p(t_1) = 1/2+1/(nr)$, $p(t_2) = 2/3$ and $p(t_3) = 1-1/(nr)$. We have

$$c(t_3 - t_2) < -\log(1 - p(t_3)) + \log(1 - p(t_2)) = \log nr - \log 3$$

Hence, $t_3 - t_2 < (\log nr - \log 3)/c = O(\log nr)$ time since $1/16 \le c < 1/4$ for $r \ge 2$.

For Stage 2 from t_1 to t_2, we have $\frac{1}{2} + \varepsilon \le (1+r)x_r/2$ and $x_r \ge (1+2\varepsilon)/(r+1)$. Note that function $(2-x-y)(1-x-y)/2 + (1-x)(x+y-1)$ for $x \ge y$, $x+y<1$, $0 \le y \le (1-x)/r$ and $(1+2\varepsilon)/(r+1) \le x \le 1-\delta$ is minimized at $(y = (1-x)/r, x = (1+2\varepsilon)/(r+1))$. Thus letting z

be $(1+2\varepsilon)/(r+1)$,

$$\begin{aligned}
\frac{\mathrm{d}p}{\mathrm{d}t} \geq & (2-x_r-x_0)\cdot\frac{1}{2}(1-x_0-x_r)+(1-x_r)(x_r-1+x_0)\\
\geq & \left(2-z-\frac{1-z}{r}\right)\cdot\frac{1}{2}\left(1-z-\frac{1-z}{r}\right)+(1-z)\left(z-1+\frac{1-z}{r}\right)\\
= & \frac{1-z}{2}\left(2-z-\frac{1-z}{r}\right)\left(1-\frac{1}{r}\right)+2\left(z-1+\frac{1-z}{r}\right)\\
= & \frac{1-z}{2r^2}\left(2r-rz-1+z\right)(r-1)+2r\left(rz-r+1-z\right)\\
= & \frac{(1-z)(r-1)}{2r^2}\left[2r-1-(r-1)z-2r\left(1-z\right)\right]\\
= & \frac{(1-z)(r-1)}{2r^2}\left[(r+1)z-1\right]\\
= & \frac{r-1}{2r^2}\cdot 2\varepsilon\cdot\frac{r-2\varepsilon}{r+1}\\
= & \frac{(r-1)(r-2\varepsilon)\varepsilon}{(r+1)r^2}
\end{aligned}$$

As $\varepsilon<\frac{1}{2}$, we have

$$\frac{\mathrm{d}p}{\mathrm{d}t}\geq\frac{(r-1)(r-2\varepsilon)\varepsilon}{(r+1)r^2}>\frac{(r-1)^2}{(r+1)r^2}\cdot\varepsilon$$

Again we let $g=(r-1)^2/((r+1)r^2)>0$, which doesn't change with time. Solving the simple ODE $\mathrm{d}p/\mathrm{d}t>c(p-1/2)$ gives

$$g(t_2-t_1)<\log(2p(t_2)-1)-\log(2p(t_1)-1)=\log\frac{1}{3}-\log\frac{2}{nr}$$

and $t_2-t_1<(\log nr-\log 6)/g=O(r\log(nr))$ time. Thus the total time from $p=1/2+1/(nr)$ to $p=1-1/(nr)$ is $t_3-t_1=(t_3-t_2)+(t_2-t_1)=O(r\log nr)$. The same statement can be proved for the other region A_- in a symmetric way. ■

The analysis of the continuous process above gives us the following lemma.

Lemma 1 (in the main paper) *When $r=2$, p is always increasing when $p>1/2$ and is always decreasing when $p<1/2$. This doesn't hold for any $r>2$.*

Proof When $r=2$, we have $p=x_2+(1-x_2-x_0)/2=(1+x_2-x_0)/2$ and

$$\begin{aligned}
\frac{\mathrm{d}p}{\mathrm{d}t} = & (1-x_2)p-(1-x_0)(1-p)\\
= & (1-x_2)\cdot\frac{1+x_2-x_0}{2}-(1-x_0)\cdot\left(1-\frac{1+x_2-x_0}{2}\right)\\
= & \frac{1}{2}\left((1-x_2)(1+x_2)-(1-x_2)x_0-(1-x_0)(1+x_0)+(1-x_0)x_2\right)\\
= & \frac{1}{2}(1-x_2^2-x_0+x_0x_2-1+x_0^2+x_2-x_0x_2)\\
= & \frac{1}{2}((x_2-x_2^2)-(x_0-x_0^2))
\end{aligned}$$

Note that $p > 1/2$ is equivalent to $x_0 < x_2$. Since $x_0 + x_2 \leq 1$ and $0 \leq x_0 < x_2$, we have either $0 \leq x_0 < x_2 < 1/2$ or $0 \leq x_0 < 1/2 \leq x_2 < 1$. In the former case we have $(x_2 - x_2^2) - (x_0 - x_0^2) > 0$ and $\mathrm{d}p/\mathrm{d}t > 0$. In the latter case we have $x_0 \leq 1 - x_2$ and because function $x - x^2$ is symmetric with respect to line $x = 1/2$, we also have $\mathrm{d}p/\mathrm{d}t > 0$. Likewise we have $\mathrm{d}p/\mathrm{d}t < 0$ when $p < 1/2$ when $r = 2$.

When $r \geq 3$, we have $\mathrm{d}p/\mathrm{d}t = (1 - x_r)p - (1 - x_0)(1 - p)$. We consider a case in which $x_0 = 0$ and $x_1 + x_r = 1$. Then we have $p = x_r + (1 - x_r)/r$ and this becomes

$$\begin{aligned}
\frac{\mathrm{d}p}{\mathrm{d}t} =& (1 - x_r)p - (1 - x_0)(1 - p) \\
=& (1 - x_r)p - (1 - p) \\
=& (2 - x_r)p - 1 \\
=& (2 - x_r) \cdot \left(x_r + \frac{1 - x_r}{r} \right) - 1 \\
=& \frac{1}{r} \cdot ((2 - x_r)((r - 1)x_r + 1) - r) \\
=& \frac{1}{r} \cdot ((2r - 3)x_r - (r - 1)x_r^2 + 2 - r)
\end{aligned}$$

Thus $\mathrm{d}p/\mathrm{d}t$ is negative when $x_r < (r - 2)/(r - 1)$. When $r \geq 3$ we know $(r - 2)/(r - 1) \geq 1/2$. We now let $x_r = (r - 2)/(r - 1) - 1/5$ where $\mathrm{d}p/\mathrm{d}t$ is surely negative. Then

$$\begin{aligned}
p =& x_r + \frac{1 - x_r}{r} = \frac{1}{r} + \left(1 - \frac{1}{r} \right) x_r \\
=& \frac{1}{r} + \frac{r - 1}{r} \cdot \left(\frac{r - 2}{r - 1} - \frac{1}{5} \right) \\
=& \frac{1}{r} + 1 - \frac{2}{r} - \frac{1}{5} \left(1 - \frac{1}{r} \right) \\
=& \frac{4}{5} \cdot \left(1 - \frac{1}{r} \right) \geq \frac{4}{5} \cdot \left(1 - \frac{1}{3} \right) = \frac{8}{15} > \frac{1}{2}
\end{aligned}$$

which disproves the statement for any $r > 2$. ■

Appendix F Proof of Theorem 7

In this section we prove a $\Theta(nr + n \log n)$ bound for an r-coupon collector.

Theorem 7 (in the main paper) *An r-coupon collector needs $\Theta(nr + n \log n)$ time with high probability.*

Proof Consider the equivalent balls-in-bins problem: if we keep throwing balls uniformly at random into n bins, how many balls do we need to throw such that every bin has at least r balls with high probability? Let N be the answer to this question. When $r = O(1)$ is a constant, we have $N = \Omega(n \log n)$ from the classic coupon collector's bound. We also have $N = O(n \log n)$ because at most r rounds of coupon collector are enough to fill the bins. Thus $N = \Theta(n \log n) = \Theta(nr + n \log n)$ is a tight bound for $r = O(1)$.

When $r = \omega(1)$, the lower bound $N = \Omega(nr + n\log n)$ is also easy to see. We must throw at least nr balls to fill the bins and the addend $\Omega(n\log n)$ is again from classic coupon collector. To prove the upper bound $N = O(nr + n\log n)$, by using Poisson approximation, we know the joint distribution of the number of balls in all the bins is well approximated by assuming the load at each bin is an independent Poisson random variable with mean $\lambda = N/n$ after we have thrown N balls in total. More concretely, if the probability of an event is either monotonically increasing or monotonically decreasing in the number of balls, then if this event has probability q in Poisson approximation, it has probability at most $2q$ in the exact balls-in-bins case. Let Y be the minimum load among the n bins. Then the probability of $Y < r$ is monotonically decreasing in the number of balls and satisfies the condition of Poisson approximation. If $\mathbb{P}(Y < r) \to 0$ holds in Poisson approximation, $\mathbb{P}(Y < r)$ also goes to zero in the exact balls-in-bins case (or equivalently the r-coupon collector problem).

In Poisson approximation, Y is the minimum among n i.i.d. Poisson random variables with mean N/n. We have

$$\mathbb{P}(Y < r) = \mathbb{P}(Y \le r-1) = 1 - \left(1 - \frac{\Gamma(r,\lambda)}{(r-1)!}\right)^n$$

where $\lambda = N/n$ and $\Gamma(\cdot,\cdot)$ is the incomplete Gamma function. An asymptotic representation for $\Gamma(\cdot,\cdot)$ is $\Gamma(r,\lambda) = \lambda^{r-1}e^{-\lambda} + o(1)$ when $\lambda \to +\infty$. When $N = \Omega(nr + n\log n)$, $\lambda = N/n = \Omega(r + \log n) = \omega(1)$ and this asymptotic representation is applicable. Letting $s = r - 1$, we have

$$\begin{aligned}
\frac{s!}{n\cdot\Gamma(s+1,\lambda)} =& O(1)\cdot\frac{\sqrt{s}\cdot s^s}{n\cdot e^s\lambda^s e^{-\lambda}} \\
=& O(1)\cdot\exp\left(\frac{1}{2}\log s + s\log s + \frac{N}{n} - s\log\frac{N}{n} - s - \log n\right) \\
=& O(1)\cdot\exp\left(s\left(\frac{N}{ns} - \log\frac{N}{ns} - 1 - \frac{1}{s}\log\frac{n}{\sqrt{s}}\right)\right)
\end{aligned}$$

Denote by f the exponent in this expression. The sign and magnitude of f are crucial for the convergence bound. When $f \to -\infty$, we have $s!/\Gamma(s+1,\lambda) = o(n)$ and $\mathbb{P}(Y \ge r) \to 0$; When $f = O(1)$ is a constant, we have $s!/\Gamma(s+1,\lambda) = \Theta(n)$ and $\mathbb{P}(Y \ge r)$ is a constant between 0 and 1; When $f \to +\infty$, we have $s!/\Gamma(s+1,\lambda) = \omega(n)$ and $\mathbb{P}(Y < r) \to 0$.

When $r = o(\log n)$, we have $\Theta(nr + n\log n) = \Theta(n\log n)$. Choose $N_1 = 2n\log n$ and then

$$\begin{aligned}
&\frac{N_1}{ns} - \log\frac{N_1}{ns} - 1 - \frac{1}{s}\log\frac{n}{\sqrt{s}} \\
=& 2\cdot\frac{\log n}{s} - \log\frac{2\log n}{s} - 1 - \frac{1}{s}\log\frac{n}{\sqrt{s}} \\
>& \frac{\log n}{s} - \frac{1}{s}\log\frac{n}{\sqrt{s}} = \frac{1}{2s}\log s
\end{aligned}$$

Since $r = \omega(1)$, s is also $\omega(1)$. Thus $f > s\cdot\log s/(2s) = \log s/2 = \omega(1)$ and $\mathbb{P}(Y < r) \to 0$. Because $\mathbb{P}(Y < r)$ is monotonically decreasing in N, all $N \ge N_1$ have $\mathbb{P}(Y < r) \to 0$. Therefore, we have $N \le N_1 = O(nr + n\log n)$.

When $r = \omega(\log n)$, we choose $N_2 = 2ns + n\log n$ and then

$$\begin{aligned}
&\frac{N_2}{ns} - \log\frac{N_2}{ns} - 1 - \frac{1}{s}\log\frac{n}{\sqrt{s}} \\
=&2 + \frac{\log n}{s} - \log\left(2 + \frac{\log n}{s}\right) - 1 - \frac{1}{s}\log n + \frac{1}{2s}\log s \\
=&1 - \log\left(2 + \frac{\log n}{s}\right) + \frac{1}{2s}\log s > \frac{1}{2s}\log s
\end{aligned}$$

Thus $f > s\cdot\log s/(2s) = \log s/2 = \omega(1)$ and $\mathbb{P}(Y < r) \to 0$. Hence, all $N \geq N_2$ have $\mathbb{P}(Y < r) \to 0$. Therefore, we have $N \leq N_2 = O(nr + n\log n)$.

The only case left is when $r = \Theta(\log n)$ and we need to take care of the constant. When $\lim \frac{\log n}{s} \leq 2$, we choose $N_3 = 3ns + n\log n$ and then

$$\begin{aligned}
&\frac{N_3}{ns} - \log\frac{N_3}{ns} - 1 - \frac{1}{s}\log\frac{n}{\sqrt{s}} \\
=&3 + \frac{\log n}{s} - \log\left(3 + \frac{\log n}{s}\right) - 1 - \frac{1}{s}\log n + \frac{1}{2s}\log s \\
=&2 - \log\left(3 + \frac{\log n}{s}\right) + \frac{1}{2s}\log s \\
\geq&2 - \log 5 + \frac{1}{2s}\log s > \frac{1}{2s}\log s
\end{aligned}$$

Thus $f = \omega(1)$ and $\mathbb{P}(Y < r) \to 0$. Hence, all $N \geq N_3$ have $\mathbb{P}(Y < r) \to 0$. Therefore, we have $N \leq N_3 = O(nr + n\log n)$.

When $\lim \frac{\log n}{s} > 2$, we choose $N_4 = ns + 2n\log n$. Notice that function $x - \log(1 + 2x)$ is always greater than 0.2 for all $x > 2$.

$$\begin{aligned}
&\frac{N_4}{ns} - \log\frac{N_4}{ns} - 1 - \frac{1}{s}\log\frac{n}{\sqrt{s}} \\
=&1 + 2\cdot\frac{\log n}{s} - \log\left(1 + 2\cdot\frac{\log n}{s}\right) - 1 - \frac{1}{s}\log n + \frac{1}{2s}\log s \\
=&\frac{\log n}{s} - \log\left(1 + 2\cdot\frac{\log n}{s}\right) + \frac{1}{2s}\log s \\
>&0.2 + \frac{1}{2s}\log s > \frac{1}{2s}\log s
\end{aligned}$$

Thus $f = \omega(1)$ and $\mathbb{P}(Y < r) \to 0$. Hence, all $N \geq N_4$ have $\mathbb{P}(Y < r) \to 0$. Therefore, we have $N \leq N_4 = O(nr + n\log n)$.

Combining with the lower bound $N = \Omega(nr + n\log n)$ we have $N = \Theta(nr + n\log n)$ and complete the proof. ■

Appendix G Figures of Experimental Results

In this section we present figures of the empirical results described in Section 6.

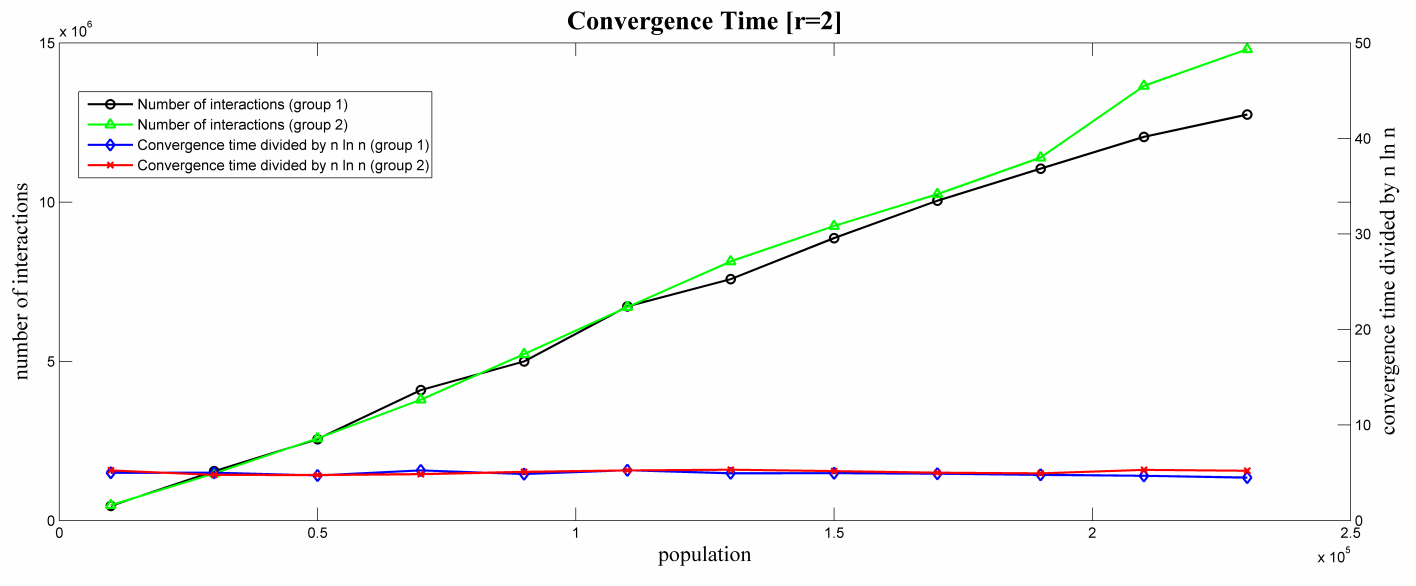

Figure 1: Convergence time with fixed resistance 2 and varying population

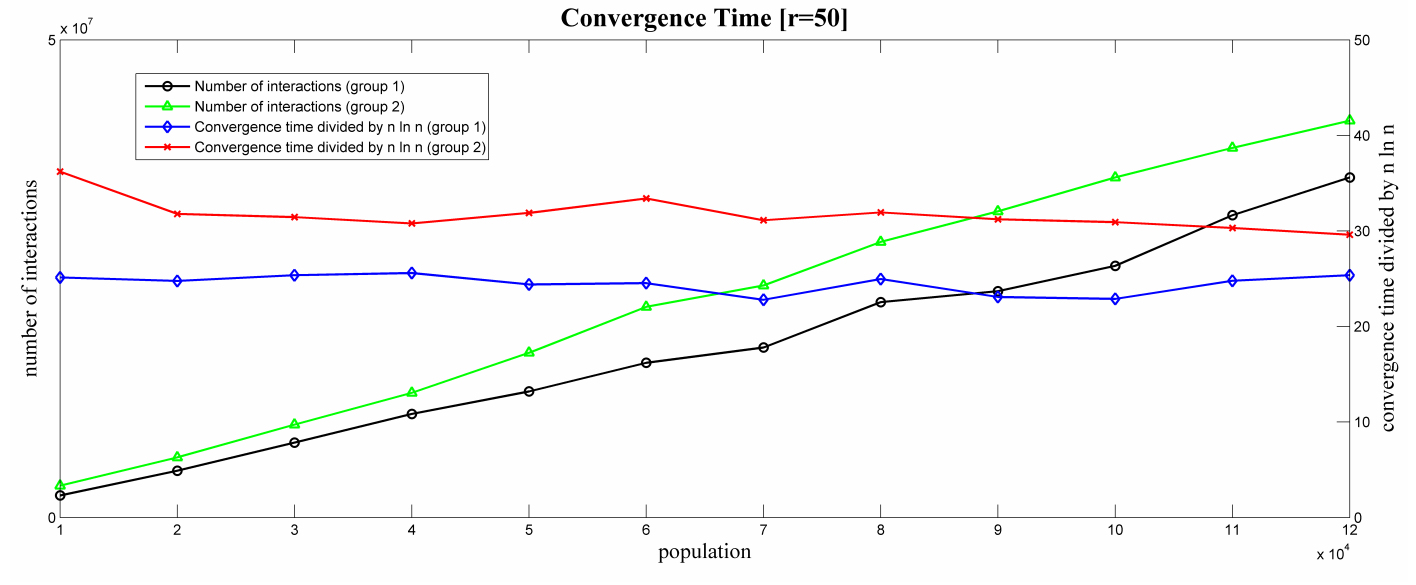

Figure 2: Convergence time with fixed resistance 50 and varying population

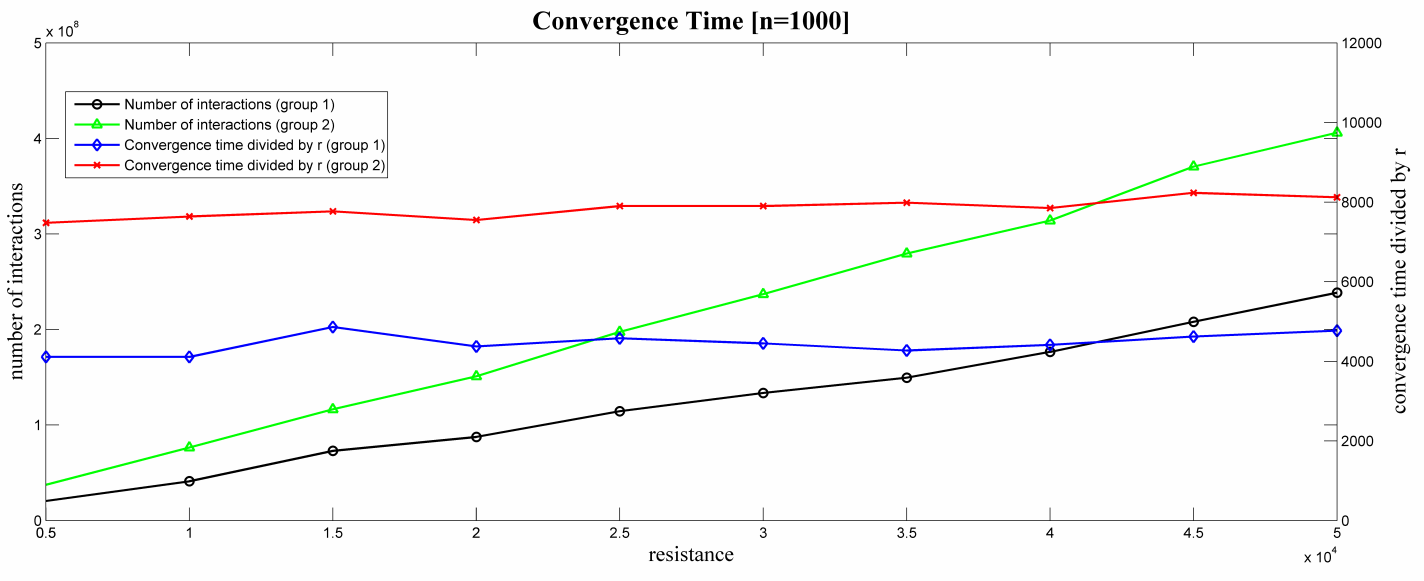

Figure 3: Convergence time with fixed population 1000 and varying resistance

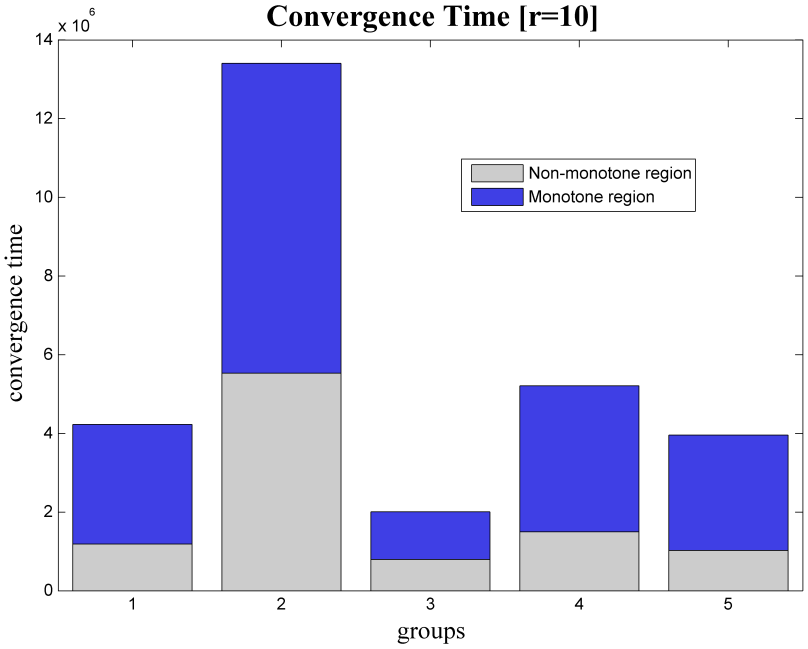

Figure 4: Convergence time comparison for continuous process