Stabilizing Consensus With the Power of Two Choices

Benjamin Doerr* Leslie Ann Goldberg† Lorenz Minder‡ Thomas Sauerwald§
Christian Scheideler¶

*Max-Planck Institute for Computer Science, Saarbrücken
†Department of Computer Science, University of Liverpool
‡Computer Science Division, University of California, Berkeley
§International Computer Science Institute Berkeley
¶Department of Computer Science, University Paderborn

Abstract

Consensus problems occur in many contexts and have therefore been intensively studied in the past. In the standard consensus problem there are n processes with possibly different input values and the goal is to eventually reach a point at which all processes commit to exactly one of these values. We are studying a slight variant of the consensus problem called the *stabilizing consensus problem* [1]. In this problem, we do not require that each process commits to a final value at some point, but that eventually they arrive at a common value without necessarily being aware of that. This should work irrespective of the states in which the processes are starting. Coming up with a self-stabilizing rule is easy without adversarial involvement, but we allow some T-bounded adversary to manipulate any T processes at any time. In this situation, a perfect consensus is impossible to reach, so we only require that there is a time point t and value v so that at any point after t, all but up to $\mathcal{O}(T)$ processes agree on v, which we call an almost stable consensus. As we will demonstrate, there is a surprisingly simple rule for the standard message passing model that just needs $\mathcal{O}(\log n \log\log n)$ time for any $\sqrt{n}$-bounded adversary and just $\mathcal{O}(\log n)$ time without adversarial involvement, with high probability, to reach an (almost) stable consensus from any initial state. A stable consensus is reached, with high probability, in the absence of adversarial involvement.

1 Introduction

Consensus problems occur in many contexts and have therefore been extensively studied in the past (e.g., [7, 29]). Interesting applications are the consolidation of replicated states or information and the synchronization of processes and devices. In the consensus problem, every process proposes a value, and the goal is to agree on a single value from all those proposed. In the case that all processes are working in a correct and timely manner, the consensus problem is easy to solve by performing a leader election.

If there are faulty or adversarial processes, the consensus problem becomes much harder. In fact, Fischer, Lynch and Paterson have shown that in an asynchronous message passing system, where processes have no common clock and run at arbitrarily varying speeds, the problem is impossible to solve if one process may crash at any time [23]. Also in a synchronous message passing system, where all processes run at the same speed, consensus is impossible if at least a third of the processes can experience Byzantine failures [22]. However, these two results only apply to deterministic algorithms.

Dagstuhl Seminar Proceedings 09371
Algorithmic Methods for Distributed Cooperative Systems
http://drops.dagstuhl.de/opus/volltexte/2010/2429

Randomized algorithms are known that can solve the consensus problem in the asynchronous setting with probability approaching one [2]. The first randomized consensus protocol was given by Ben-Or [10]. Later papers extended Ben-Or's work in two directions: the literature on message-passing consensus protocols has largely concentrated on solving consensus using cryptographic techniques or private channels with a linear bound on the number of faulty processes (including Byzantine faults) while work on shared-memory systems has used the underlying reliability of the shared memory system to solve consensus in the wait-free case, where there is no limit on how many processes may fail but failures are limited to crash failures.

In the context of message-passing systems, Rabin [30] showed that Byzantine agreement can be solved in constant expected time given a shared coin visible to all processes, and described an implementation of such a coin based on digital signatures and a trusted dealer. Feldman and Micali [21] gave a constant-round shared-coin protocol for a synchronous system that uses secret sharing to avoid the need for a trusted dealer. A constant-time shared-coin protocol for an asynchronous system was given later by Canetti and Rabin [12] based in part on further unpublished work by Feldman. For the Canetti and Rabin protocol the cryptographic assumptions can be replaced by the assumption of private channels. Their protocol works as long as fewer than $n/3$ processes are adversarial. That was subsequently improved by Katz and Koo [28] to an expected constant-round consensus protocol that works as long as fewer than $n/2$ processes are adversarial, which is optimal. When assuming a computationally unbounded adversary, one also usually assumes that the adversary knows the entire history as well as the current state of the protocol, so cryptographic primitives and private channels do not exist in this case. For non-adaptive (in the selection of adversarial processes) but otherwise computationally unbounded adversaries, Feige [20] has shown that when broadcasting is allowed, consensus can be achieved in $\mathcal{O}(\log^* n)$ rounds for fewer than $n/(2+\epsilon)$ Byzantine processes, for an arbitrary constant $\epsilon > 0$. For the case that only point-to-point communication is allowed, Ben-Or et al. [11] proved that consensus is possible with an expected running time of $\mathcal{O}(\log n)$ rounds for fewer than $n/(4+\epsilon)$ Byzantine processes. The restriction that the adversary be non-adaptive is essential. Ben-Or and Bar-Joseph [9] have shown that any consensus protocol that tolerates $\Theta(n)$ adaptive fail-stop faults runs for $\tilde{\Omega}(\sqrt{n})$ rounds. All protocols above need $\Omega(n)$ expected individual work per process.

In the context of shared memory systems, the n processes communicate by reading and writing to shared multi-writer multi-reader registers. Each step consists of some local computation and one shared memory event, which is either a read or a write to some register. The interleaving of the processes' events is controlled by an adversary. In the strong adversarial model, the adversary knows the entire history and observes the results of the local coin flips before scheduling the next event (i.e., the order in which the processes proceed). The first randomized consensus protocol to use shared memory was given by Chor, Israeli and Li [13]. Many other randomized consensus algorithms have been proposed since then. Aspnes et al. [3, 4] recently showed that the consensus problem can be solved in this case with $\mathcal{O}(n)$ expected individual work and $\mathcal{O}(n^2)$ expected total work. Attiya and Censor [6] proved that $\Omega(n^2)$ is also a lower bound on the total work. Tight upper and lower bounds for the number of (asynchronous) rounds under a much weaker adversary have been shown by Kapron et al. [26] and Attiya and Censor [5].

1.1 Our approach

We are studying a slight variant of the original consensus problem called the *stabilizing consensus problem* [1]. In this problem, we do not require that each process commits to a final value at some point, but that eventually they arrive at a common value (without necessarily being aware of that) *and stay there*. This should work irrespective of the states in which the processes are starting.

Thus, we are aiming for a self-stabilizing consensus protocol.

In our model we have n processes that are completely interconnected in an anonymous network. That is, no unique process IDs are known, but rather each process has its own, private numbering of the other processes. For the presentation of our protocol and the analysis we assume, though, that the processes are numbered from 1 to n. Each process i initially has some value $v_i \in \mathbb{N}$ (which is assumed to require only $\mathcal{O}(\log n)$ bits for its storage so that values can be efficiently exchanged). The goal is to eventually agree on a common, stable value $v \in \{v_1, \ldots, v_n\}$, that is, afterwards the processes will not deviate from v anymore. We assume that time proceeds in synchronized rounds. In each round, every process can contact at most a logarithmic number of other processes, exchange a logarithmic amount of information with each of them, and perform some local computation based on the received information. A process with more than a logarithmic number of requests directed to it will only receive a logarithmic number of them, possibly selected by an adversary, and the others are dropped.

If no process is ever corrupted, the consensus problem would be easy to solve with the following *minimum rule*: In each round, every process i contacts some random process j in the system and updates its own value to $v_i := \min\{v_i, v_j\}$. In this way, all processes will have the same value within $\mathcal{O}(\log n)$ rounds w.h.p.. This bound is essentially best possible for any protocol within our model.

However, if some processes can be corrupted, finding *and staying* at a consensus is much more tricky. We use the following strong adversarial model. A *T-bounded adversary* is allowed to know the entire history of the protocol. At the beginning of each round, it may decide to change the state of up to T many of the processes in an arbitrary way subject to the constraint that it can only change the value of a process to one out of the initial set of values $\{v_1, \ldots, v_n\}$. This is necessary to stay within the original consensus problem (i.e., the processes agree on one of the initial values) but also reasonable in applications where the values are signed by some outside authority so that the adversary cannot just come up with a value by itself.

Of course, in our adversarial model we cannot enforce any more that eventually *all* processes agree on a stable value. Hence we will only require that there is a round r and value $v \in \{v_1, \ldots, v_n\}$ such that for any round later than r, all but up to $\mathcal{O}(T)$ processes agree on v. We will call this *almost stable consensus*. The goal is to come up with an efficient protocol so that for the number of faulty processes T as high as possible, almost stable consensus can still be achieved (with high probability).

Why is the stabilizing consensus problem so difficult in our model? As an example, consider the minimum rule introduced above. Consider an initial setting with just two possible values, 1 and 2, and let there be at most T processes initially having the value 1. Then the adversary can change the values of all of these processes to 2. Consequently, no process will change its value afterwards. However, if at any time in the future the adversary changes the value of one of the processes to 1, all other processes will start adopting value 1, and finally consensus on the value 1 will be reached. Since the adversary may delay this arbitrarily long, the minimum rule does not lead to a stable consensus within any time bound. Note that the state in which all processes store the value 2 does not represent an almost stable consensus since eventually none of the processes will stick to this value.

Other basic approaches like leader election also do not work as the adversary can corrupt the state of the leader (or leader group) once it has been elected thereby causing the leader election to fail. Thus, it appears that more advanced protocols (using verifiable secret sharing or other heavy-weight machinery) have to be used, though one would also have to keep in mind to make them self-stabilizing, i.e., to work from *any* initial state. Is there an alternative, *simple* protocol with a bounded running time, even for a large T and *any* initial state? Interestingly, we will show that there is such a protocol. In fact, this protocol does not need any assumptions on cryptographic

input	with adversary	without adversary
worst-case 2 bins	$O(\log n)$	$O(\log n)$
worst-case m bins	$O(\log m \log\log n + \log n)$	$O(\log n)$
average-case m bins	$O(\log m + \log\log n)$ if m is odd, $\Theta(\log n)$ if m is even.	$O(\log m + \log\log n)$ if m is odd, $\Theta(\log n)$ if m is even.

Figure 1: Summary of our results on the time required to reach an almost stable consensus (with adversary) or stable consensus (without adversary) w.h.p.

primitives or private channels to work.

None of the protocols discussed above can be directly applied to our setting as they assume to start in a well-initialized state whereas we require stabilization from *any* state. There is a recent construction for self-stabilizing consensus in the shared memory model [18], but it needs $\Omega(n)$ expected individual work per process. In general, it would be easy to use a universal compiler to turn non-stabilizing deterministic protocols into efficient self-stabilizing ones (e.g., [31]), but finding efficient transformations for randomized protocols appears to be very tricky. This is due to the problem that the standard transformations are based on checking the consistency of logging information. However, besides the challenging problem of checking whether prior actions were based on (sufficiently) random decisions (recall that the nodes may initially be in *any* state), it may turn out that the random process j recorded by process i that i was supposed to communicate with in some prior round is not the correct one. So *checking* whether a state is incorrect as well as *correcting* it may involve global coordination and can therefore be quite expensive. Also, transformations of non-stabilizing into stabilizing protocols usually result in protocols that are far from being light-weight, intuitive and simple.

1.2 Our contributions

We propose a very simple protocol for our model called the *median rule*: In each round, every process i picks two processes j and k uniformly and independently at random among all processes (including itself). It then updates v_i to the *median* of v_i, v_j and v_k. For example, if $v_i = 10$, $v_j = 12$ and $v_k = 100$, then the new value of v_i is 12.

When taking the *mean* of the three values instead of the median, the convergence properties towards a single number have already been formally analyzed [17]. However, with the mean rule we are no longer guaranteed to solve the consensus problem (i.e., one of the initial values wins). Unfortunately, the analysis for the mean rule cannot be adapted to the median rule (also, the model is completely different — in the model of [17], all pairs of processors are allowed to exchange values during every round), so we had to develop completely different proofs.

The median rule works surprisingly well. Classical counter strategies like switching values or hiding values for an unbounded amount of time are ineffective and cannot prevent the median rule from converging quickly. We prove the following results that are also summarized in Figure 1.2.

Theorem 1. *For any initial state it holds that if no process is ever adversarial, then the median rule reaches a stable consensus in $\mathcal{O}(\log n)$ rounds w.h.p.*

Hence, the median rule is as effective as the minimum rule in the non-adversarial case. This still holds for the adversarial case and a constant initial number of different values.

Theorem 2. *For any initial state with a constant number of different values and under any T-bounded adversary with $T \leq \sqrt{n}$, the median rule reaches an almost stable consensus in $\mathcal{O}(\log n)$ rounds w.h.p.*

The bound on T is essentially tight as $T = \tilde{\Omega}(\sqrt{n})$ would not allow the median rule to stabilize any more w.h.p. because the adversary could keep two groups of processes with equal values in perfect balance for at least a polynomially long time. More generally, we can show the following result.

Theorem 3. *For any initial state with m different values and under any T-bounded adversary with $T \leq \sqrt{n}$, the median rule reaches an almost stable consensus in $\mathcal{O}(\log m \cdot \log\log n + \log n)$ rounds w.h.p.*

A further improvement can be obtained in an average-case setting.

Theorem 4. *Let $m \leq n^{1/2-\epsilon}$. Assume that the initial state is chosen uniformly at random from all states having m different values. Then against any T-bounded adversary with $T \leq \sqrt{n}$, the median rule reaches an almost stable consensus within $\Theta(\log n)$ rounds, w.h.p., if m is even, and $\mathcal{O}(\log m + \log\log n)$ rounds, w.h.p., if m is odd.*

With these results, the median rule is yet another demonstration of the *power of two choices* as a deterministic single choice rule would only allow us to implement the minimum or maximum rule, which have an unbounded runtime. However, two other processes are contacted at random in a round, the median rule can be implemented, resulting in a very small runtime as we have seen. This power of two choices has also been demonstrated in many other contexts [27, 16, **?**, 14, 15, 19] (mostly in the balls into bins model, which is why we will use that notation later), but we are not aware of any result using it in the context of consensus.

2 Notation and Preliminaries

2.1 Definitions

We have n balls (processes) numbered from 1 to n, and n bins (values) also numbered from 1 to n. Write $b_{0,1}, \ldots, b_{0,n}$ for the initial assignment of each ball to its bin, where the bins are identified with integers in $\mathbb{N}$. So, the "all-one" assignment that assigns each ball to a different bin, can be achieved by setting $b_{0,i} = i$ for $1 \leq i \leq n$.

In the i-th iteration the algorithm does the following: Let $I_{i,1}, \ldots, I_{i,n}$ and $J_{i,1}, \ldots, J_{i,n}$ be $2n$ independent random variables, chosen uniformly in $[n] = \{1, \ldots, n\}$. Then the bin assignment of the i-th iteration is

$$b_{i,j} = \text{median}(b_{i-1,j}, b_{i-1,I_{i,j}}, b_{i-1,J_{i,j}})$$

for $j = 1, \ldots, n$. The algorithm has reached a fixed point (has stabilized) in the i-th iteration if and only if $b_{i,1} = \cdots = b_{i,n}$.

For each iteration $t \in \mathbb{N}$, we denote by m_t the bin that contains the median-ball, so,

$$|\{j \in [n] : b_{t,j} < m_t\}| \leq n/2,$$

and

$$|\{j \in [n] : b_{t,j} > m_t\}| \leq n/2.$$

2.2 Probabilistic Tools

We write w.h.p. to refer to an event that holds with probability at least $1 - n^{-c}$ for $c > 1$. We will use the following three Chernoff bounds.

Lemma 5 (Chernoff Bound for Bernoulli Variables)**.** *Let $X_1, \dots, X_n$ be independent binary random variables, let $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathrm{E}[X]$. Then it holds for all $\delta > 0$ that*

$$\Pr[X \geq (1+\delta)\mu] \quad \leq \quad \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu} \leq e^{-\min[\delta^2,\, \delta]\cdot\mu/3} .$$

Furthermore, it holds for all $0 < \delta < 1$ that

$$\Pr[X \leq (1-\delta)\mu] \quad \leq \quad \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu} \leq e^{-\delta^2\mu/2} .$$

Lemma 6 (Chernoff Bound for Geometric Variables)**.** *Consider some fixed $0 < \delta < 1$. Suppose that $X_1, \dots, X_n$ are independent geometric random variables on $\mathbb{N}$ with $\Pr[X_i = k] = (1-\delta)^{k-1}\delta$ for every $k \in \mathbb{N}$. Let $X = \sum_{i=1}^{n} X_i$, $\mu = \mathrm{E}[X]$. Then it holds for all $\epsilon > 0$ that*

$$\Pr[X \geq (1+\epsilon)n/\delta] \leq e^{-\epsilon^2 n/2(1+\epsilon)}$$

Lemma 7 (Chernoff Bound for Variables with Exponential Tails)**.** *Suppose that $X_1, \dots, X_n$ are independent random variables on $\mathbb{N}$, such that there is a constant $\gamma > 0$ with $\Pr[X_i = k] \leq \gamma(1-\delta)^{k-1}$ for every $k \in \mathbb{N}$. Let $X = \sum_{i=1}^{n} X_i$, $\mu = \mathrm{E}[X]$. Then it holds for all $\epsilon > 0$ that*

$$\Pr[X \geq (1+\epsilon)\mu + \mathcal{O}(n)] \leq e^{-\epsilon^2 n/2(1+\epsilon)}.$$

2.3 Auxiliary Results for Absorbing Markov Chains

We now prove exponential tail bounds for reaching absorbing states in Markov Chains.

Lemma 8. *Consider a Markov Chain $(X_t)_{t=1}^{\infty}$ with state space $\{0, \dots, m\}$ that has the following properties:*

- *there are constants $c_1 > 1$ and $c_2 > 0$, such that for any $t \in \mathbb{N}$,*

$$\Pr[X_{t+1} \geq \min\{m, c_1 X_t\}] \geq 1 - e^{-c_2 X_t},$$

- *$X_t = 0 \Rightarrow X_{t+1} \geq 1$ with probability c_3 that is greater than some constant > 0,*

Then for every constant $c_4 > 0$, there is a time step $t = \mathcal{O}(\log m)$ such that

$$\Pr[X_t \geq c_4 \log m] \geq 1 - m^{-2}.$$

Lemma 9. *Consider a Markov Chain $(X_t)_{t=1}^{\infty}$ with state space $\{0, \dots, m\}$ that has the following properties,*

- *there are constants $c_1 > 1$ and $c_2 > 0$, such that for any $t \in \mathbb{N}$,*

$$\Pr[X_{t+1} \geq \min\{m, c_1 X_t\}] \geq 1 - e^{-c_2 X_t},$$

- *$X_t = 0 \Rightarrow X_{t+1} = 0$ with probability 1,*
- *There is a number $m \in \mathbb{N}$ such that $X_t = m \Rightarrow X_{t+1} = m$ with probability 1.*

Then it holds for $t = \mathcal{O}(\log m)$ that

$$\Pr[X_t \in \{0\} \cup \{m\}] \geq 1 - m^{-2}.$$

The proof is almost the same as the proof of Lemma 8. The only difference is that we stop the process if we have reached the state $\{0\}$ and that we run the process until we have reached state m (instead a state that is at least $c_4 \log m$).

3 Two Bins with Adversary

In this section, we consider the case with only two bins. This special case is of more general interest, as the later analysis with more than two bins will use some results of this section.

We also consider an adversary with the following abilities. In each round after all n balls have made their random choices, the adversary is allowed to change the choices of at most $\sqrt{n}$ balls. We prove that despite the adversary, we reach a stable state where one bin contains at least $n - \sqrt{n}$ balls w. h. p.

Note that for the two bin-case, the median rule coincides with the *majority rule*, where a balls next bin is chosen to be the majoritary bin of itself and two random balls.

Theorem 10. *For any initial distribution of the balls, $\mathcal{O}(\log n)$ rounds of the majority rule suffice to get $n - \mathcal{O}(\sqrt{n})$ balls into the same bin, w.h.p. This holds even if an adversary is allowed to manipulate the random choices of at most $\mathcal{O}(\sqrt{n})$ balls.*

In the following, let L_t be the number of balls in the left bin at step t and let R_t be the number of balls in the right bin at step t. Let $X_t = \min(L_t, R_t)$ and let $Y_t = \max(L_t, R_t)$. For simplicity, we will assume that n is even. The proof for odd n follows along the same lines. The *imbalance* at a step t is given by $\Delta_t = (Y_t - X_t)/2$ (which is a positive integer if n is even). The *labelled imbalance* is given by $\Psi_t = (R_t - L_t)/2$. Based on the imbalance Δ_t we distinguish between three cases.

Case 1: $\Delta_0 \geq n/3$

We will show the following lemma.

Lemma 11. *If $\Delta_0 \geq cn$, where $c > 0$ is any constant, then $\mathcal{O}(\log \log n)$ rounds of the majority rule suffice to reach a stable consensus, w.h.p.*

Case 2: $c\sqrt{n \ln n} \leq \Delta_0 < n/3$ for a sufficiently large c

Here, we will show the following lemma.

Lemma 12. *If there is a step t with $c\sqrt{n \ln n} \leq \Delta_t < n/3$ for a sufficiently large constant c, then $\mathcal{O}(\log n)$ additional rounds of the majority rule suffice to arrive at case 1 w.h.p.*

Case 3: $\Delta_0 \leq c\sqrt{n \ln n}$

We say that a random variable Y *stochastically dominates* a random variable Z, and write $Y \succeq Z$, if $\Pr[Y \geq x] \geq \Pr[Z \geq x]$ for any x.

Lemma 13. *Assume no adversary is present and $L_t \leq R_t$. For any two labelled imbalances Ψ_t and Ψ'_t with $\Psi_t > \Psi'_t$ it holds that $\Psi_{t+1} \succeq \Psi'_{t+1}$.*

In the next lemma, we use the Central Limit Theorem to prove that with constant probability, we have a sufficiently large imbalance regardless of the previous imbalance.

Lemma 14. *Assume no adversary is present. Let $\varepsilon > 0$ be an arbitrary constant, and let c be the constant from Lemma 12. For any $\Psi_t \geq 0$, $\Pr[\Psi_{t+1} \geq c\sqrt{n}] \geq \frac{1}{\sqrt{2\pi}(1+4c/\sqrt{3})} e^{-8c^2/3} - \varepsilon$, provided n is large enough.*

Finally, we prove that there is a strong drift to increase the imbalance by a constant factor.

Lemma 15. *Assume no adversary is present. If $\Delta_t \geq c\sqrt{n}$, where c is the constant from Lemma 12, then*

$$\Pr[\Delta_{t+1} \geq (4/3)\Delta_t] \geq 1 - \exp\left(-\Theta(\Delta_t^2/n)\right).$$

We are now ready to finish the case 3.

Lemma 16. *If initially, we have $\Delta_0 < c\sqrt{n \ln n}$ for the value of c needed by Lemma 12, then after $\mathcal{O}(\log n)$ steps, we have $t \geq c\sqrt{n \ln n}$ w. h. p.*

4 More than Two Bins

In this section we consider the more challenging case when we have up to n different bins. We analyze the models with and without adversary separately.

In order to be able to reduce our problem with n bins to the one with 2 bins, we first establish a partial order of *fineness* on the distribution of balls into bins.

4.1 Partial Order on the Balls-and-Bins-Distribution

We say that an initial assignment with k_i balls in bin i ($i \in [n]$) is *finer* than an assignment $(\tilde{k}_i)_{i\in[n]}$, if there is a monotonic function $f : [n] \to [n]$ mapping bins to bins such that

$$\tilde{k}_i = \sum_{j \in f^{-1}(i)} k_j$$

for each $i \in [n]$. Obviously, the "is finer"-relation induces a partial order on the set of all balls-and-bins-distributions.

The all-one assignment $k_1 = \cdots = k_n = 1$ is finer than every other assignment $(\tilde{k}_i)_{i\in[n]}$: The corresponding mapping can be constructed by setting $f(1) = \cdots = f(\tilde{k}_1) = 1$, $f(\tilde{k}_1 + 1) = \cdots = f(\tilde{k}_1 + \tilde{k}_2) = 2$, and so on. (Since $n = \sum k_i = \sum \tilde{k}_i$, this mapping is well defined.)

Lemma 17. *If the assignment $(k_i)_{i\in[n]}$ is finer than the assignment $(\tilde{k}_i)_{i\in[n]}$, then the number of iterations of the algorithm needed for convergence with the initial assignment $(k_i)_{i\in[n]}$ upper bounds the number of iterations needed with the initial assignment $(\tilde{k}_i)_{i\in[n]}$. This is true point wise in the probability space and hence applies to both average and high probability running times.*

4.2 More than Two Bins without Adversary

In this section, we prove Theorem 1. Our proof proceed as follows. We show that after $\mathcal{O}(\log n)$ rounds, we end up with at most 2 non-empty bins. Then we can directly use our result for two bins to conclude that after additional $\mathcal{O}(\log n)$ rounds, the process stabilizes.

Assume that the bins and balls are numbered from 1 to n such that all balls with higher numbers are in higher bins and balls in the same bin form a consecutive interval. As an example, $(1, 2, 3 \,|\, 4, 5 \,|\, |\, 6 \,|\, 7, 8)$ describes a distribution of 8 balls into 5 bins, where the first bin holds 3 balls, the second bin 2, the third bin zero and so on.

We associate with each ball $i \in [n]$ a value $g(i)$ called *gravity*, which is the expected number of balls that choose i as their median for the next step (when considering the ball ordering). The

gravity $g(i)$ can be estimated as

$$g(i) = (n-i-1)\cdot\frac{2(i-1)}{n^2} + (i-1)\cdot\frac{2(n-i-1)}{n^2} + 1\cdot\frac{2(n-i-1)(i-1)}{n^2} + \mathcal{O}\left(\frac{1}{n}\right)$$
$$= 6\frac{(n-i)i}{n^2} + \mathcal{O}\left(\frac{1}{n}\right). \quad (1)$$

Note that the gravity of a ball i is maximized for the median-ball, which has number $\lceil n/2 \rceil$ according to our ordering.

Fix a bin j. By linearity of the expectation, the expected load of j at a time $t+1$ is equal to the sum of gravities of the balls in bin j at time t.

For each bin $j \in [n]$ at step t, we define a set of heavy balls $\mathcal{H}_{t,j}$ which is defined as the subset of the $\Phi = C\sqrt{n \log n}$ balls in bin j with largest gravity. $C > 0$ is a sufficiently large constant. Note that by definition, $0 \leq |\mathcal{H}_{t,j}| \leq \Phi$. We first prove the following:

Lemma 18. *If there is a ball $i \in \mathcal{H}_{t,j}$ with $g(i) < 4/3$, then at step $t+1$ either there is a ball $l \in \mathcal{H}_{t+1,j}$ with $g(l) < 4/3$ or bin j is empty w. h. p.*

Lemma 19. *Consider an arbitrary bin $j \in [n]$. Then there is a round $t_1 = \mathcal{O}(\log n)$ such that at least one of the following holds w.h.p.:*

1. *at least one ball $i \in \mathcal{H}_{t,j}$ satisfies $g(i) < 4/3$ (or $\mathcal{H}_{t,j}$ is empty), or*
2. *$|\mathcal{H}_{t,j}| = \Phi$.*

We are now ready to prove the main result of this section.

Proof of Theorem 1. Since Lemma 19 holds w.h.p., there is a round $t_1 = \mathcal{O}(\log n)$ where Lemma 19 holds for *every* bin w.h.p. For this t_1 let $j_{\min}, j_{\max}$ be the leftmost and rightmost balls in bin m_t, respectively. Note that by Lemma 19, the load of bin m_t is at least Φ. We proceed by a case distinction on the position of m_t.

1. $n/2 - j_{\min} \leq \Phi/2$. Let $m_t - 1$ be the left bin from bin m_t. Equation 1 implies that all heavy balls in bin $m_t - 1$ have gravity at least $4/3$. So, Lemma 19 implies that the load of bin $m_t - 1$ is at least Φ. Consolidate all bins from $1, \ldots, m_t - 2$ and all bins from $m_t + 1, \ldots, n$ into two superbins A and B, respectively. By our arguments above, both superbins have a load of at most $n/2 - \Phi/2$. Therefore for C large enough, Lemmas 11 and 12 imply that both superbins will die out within the next $\mathcal{O}(\log n)$ steps w.h.p. After this has happened, we only end up with two bins $m_t - 1$ and m_t. A final application of our two bin-analysis (Theorem 10) reduces the number of bins from 2 to 1 within additional $\mathcal{O}(\log n)$ rounds, and our theorem follows.

2. $j_{max} - n/2 \leq \Phi/2$. This case is handled exactly as before.

3. $j_{\max} - j_{\min} > \Phi$. In this case, it follows as in the previous cases that by Lemmas 11 and 12 that all bins except bin m_t will vanish after the next $\mathcal{O}(\log n)$ rounds.

□

4.3 More than Two Bins with Adversary

The following theorem extends our two-bin result to the many-bin case in the presence of an adversary.

Theorem 20. *For any $\sqrt{n}$-bounded adversary and m bins, it will take at most $\mathcal{O}(\log m \log\log n + \log n)$ rounds until the median rule reaches an almost stable consensus.*

The proof considers $O(\log m)$ phases. The expected number of rounds taken by each phase is $O(\log\log n)$. In each phase, we show that there is a small set of non-empty bins such that bins outside of this set have too few balls, and will therefore eventually become empty. The size of the "small set" halves with each new phase. The analysis of each phase applies the two-bin analysis to two "meta-bins", one of which consists of the left half of the small set of bins, and all bins to the left of that — the other meta-bin consists of the remaining bins. After $O(\log m)$ phases, at most three actual bins remain non-empty so the two-bin analysis again applies, giving an almost stable consensus in $O(\log n)$ more rounds.

5 Average Case Analysis

In this section, we investigate the case where all n balls are initially put independently and randomly into m bins.

5.1 The Uniform Random Case without Adversary

Without the adversary, we get the following result.

Theorem 21. *Assume that each of the n balls is initially assigned uniformly at random to one of the m bins. Then the median rule reaches a stable consensus in the following number of rounds w.h.p.:*

$$\begin{array}{ll} \mathcal{O}(\log m + \log\log n) & \text{if } m \text{ is odd,} \\ \Theta(\log n) & \text{if } m \text{ is even.} \end{array}$$

Proof. If m is even, we can use Lemma 17 to reduce the problem to that of two superbins with $m/2$ bins each. The Central Limit Theorem implies that these superbins will initially have at most $n/2 + \sqrt{n}/8$ balls with constant probability. Given that this bound is true in some round t, it will also be true in round $t+1$ with constant probability, so $\Omega(\log n)$ rounds are needed to make sure that the number of balls in a superbin is more than $n/2 + \sqrt{n}/8$ in at least one round w.h.p. This implies the lower bound on the runtime for even m. The upper bound follows from Theorem 1.

We now study the case where m is odd, and we assume w. l. o. g. that $m = \mathcal{O}(\sqrt[3]{n})$ (otherwise we can use Theorem 1). We merge the $(m-1)/2$ leftmost bins into a single superbin, and we do the same with the $(m-1)/2$ rightmost bins. So we have three superbins (we simply call bins) left. The left and the right ones contain $n/2 - n/(2m) + \mathcal{O}(\sqrt{n}\log n)$ balls each w.h.p., and the middle bin contains $n/m + \mathcal{O}(\sqrt{n}\log n)$ balls w.h.p.

Consider now any round t in which the left bin has at most $n/2 - \Delta$ balls, where $n/(3m) \leq \Delta < n/3$ (which is initially true in our case). Then it follows from the analysis of Lemma 12 that the expected number of balls in the left bin in round $t+1$ is at most $n/2 - (5/4)\Delta$. Since $\Delta \geq n/(3m) = \Omega(n^{2/3})$, the Chernoff bounds imply that this is also at most $n/2 - (6/5)\Delta$ w.h.p. Hence, after k many rounds with $(n/(3m))(6/5)^k \geq n/3$, we reach a round with $\Delta > n/3$ w.h.p. This is true for $k = O(\log m)$. But then Lemma 11 shows that from that point it takes only $O(\log\log n)$ further steps to make the left bin empty w.h.p. The same argument also applies to

the right bin. Hence, after $O(\log m + \log\log n)$ many rounds, the middle bin will have all balls w.h.p. □

5.2 The Uniform Random Case with Adversary

Corollary 22. *Consider any $\sqrt{n}$-bounded adversary and suppose that $m \leq n^{1/2-\epsilon}$ for some constant $\epsilon > 0$. Then the median rule reaches an almost stable consensus w.h.p. after the following number of rounds:*

$$\begin{array}{ll} \mathcal{O}(\log m + \log\log n) & \text{if } m \text{ is odd and} \\ \Theta(\log m) & \text{if } m \text{ is even.} \end{array}$$

Proof. If m is odd, then the proof follows along the same lines as the previous proof as initially $\Delta \geq n/(3m) \geq n^{1/2+\epsilon}/3$, which is much larger than the adversarial influence on the balls. Hence, w.h.p. the middle bin will keep drawing balls into it until there are only $O(\sqrt{n})$ outside balls left.

If m is even, we consider the two middle bins as a single bin to reduce this case to an odd m. In this case, $\mathcal{O}(\log m + \log\log n)$ rounds suffice w.h.p. to get all but $O(\sqrt{n})$ balls into the two middle bins. Afterwards, we are left with the 2-bin case which leads to an almost stable consensus in at most $O(\log n)$ further steps w.h.p. □

6 Conclusions

In this paper we presented a surprisingly simply and effective synchronization mechanism. Unfortunately, we were only able to rigorously analyze its performance for the one-dimensional case. It would be very interesting though probably very challenging to prove a time bound of $\mathcal{O}(\log n)$ also for higher dimensions. Also, the robustness of the protocol deserves further studies.

Acknowledgments

We would like to thank Michael Bender for suggesting a related problem that initiated this research and Goran Konjevod, Andrea Richa and Donglin Xia for helpful discussions on early versions of the paper.

References

[1] D. Angluin, M. Fischer, and H. Jiang. Stabilizing consensus in mobile networks. In *Proc. of the Intl. Conference on Distributed Computing in Sensor Networks (DCOSS)*, pages 37–50, 2006.

[2] J. Aspnes. Randomized protocols for aynchronous consensus. *Distributed Computing*, 16(2-3):165–176, 2003.

[3] J. Aspnes, H. Attiya, and K. Censor. Randomized consensus in expected $o(n \log n)$ individual work. In *Proc. of the 27th ACM Symp. on Principles of Distributed Computing (PODC)*, pages 325–333, 2008.

[4] J. Aspnes and K. Censor. Approximate shared-memory counting despite a strong adversary. In *Proc. of the 20th ACM Symp. on Discrete Algorithms (SODA)*, pages 441–450, 2009.

[5] H. Attiya and K. Censor. Lower bounds for randomized consensus under a weak adversary. In *Proc. of the 27th ACM Symp. on Principles of Distributed Computing (PODC)*, pages 315–324, 2008.

[6] H. Attiya and K. Censor. Tight bounds for asynchronous randomized consensus. *Journal of the ACM*, 55(5), 2008.

[7] H. Attiya and J. Welch. *Distributed Computing: Fundamentals, Simulations, and Advanced Topics (2nd Edition)*. John Wiley and Sons, 2004.

[8] Y. Azar, A. Broder, A. Karlin, and E. Upfal. Balanced allocation. In *Proc. of the 26th ACM Symp. on Theory of Computing (STOC)*, pages 593–602, 1994.

[9] Z. Bar Joseph and M. Ben-Or. A tight lower bound for randomized synchronous consensus. In *Proc. of the 17th ACM Symp. on Principles of Distributed Computing (PODC)*, pages 193–199, 1998.

[10] M. Ben-Or. Another advantage of free choice: completely asynchronous agreement protocols. In *Proc. of the 2ndpodc*, pages 27–30, 1983.

[11] M. Ben-Or, E. Pavlov, and V. Vaikuntanathan. Byzantine agreement in the full-information model in $o(\log n)$ rounds. In *Proc. of the 38th ACM Symp. on Theory of Computing (STOC)*, pages 179–186, 2006.

[12] R. Canetti and T. Rabin. Fast asynchronous Byzantine agreement with optimal resilience. In *Proc. of the 25thstoc*, pages 42–51, 1993.

[13] B. Chor, A. Israeli, and M. Li. Wait-free consensus using asynchronous hardware.

[14] R. Cole, A. Frieze, B.M. Maggs, M. Mitzenmacher, A.W. Richa, R.K. Sitaraman, and E. Upfal. On balls and bins with deletions. In *Proc. of the 2nd Intl. Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM)*, 1998.

[15] R. Cole, B.M. Maggs, F. Meyer auf der Heide, M. Mitzenmacher, A.W. Richa, K. Schröder, R.K. Sitaraman, and B. Vöcking. Randomized protocols for low-congestion circuit routing in multistage interconnection networks. In *Proc. of the 29th ACM Symp. on Theory of Computing (STOC)*, pages 378–388, 1998.

[16] M. Dietzfelbinger and F. Meyer auf der Heide. Simple, efficient shared memory simulations. In *Proc. of the 10th ACM Symp. on Parallel Algorithms and Architectures (SPAA)*, pages 110–119, 1993.

[17] D. Dolev, N. Lynch, S. Pinter, E. Stark, and W. Weihl. Reaching approximate agreement in the presence of faults. *Journal of the ACM*, 33(3):499–516, 1986.

[18] S. Dolev, R. Kat, and E. Schiller. When consensus meets self-stabilization. In *In OPODIS*, pages 45–63, 2006.

[19] R. Elsäesser and T. Sauerwald. The power of memory in randomized broadcasting. In *Proc. of the 19th ACM Symp. on Discrete Algorithms (SODA)*, pages 773–781, 2008.

[20] U. Feige. Noncryptographic selection protocols. In *Proc. of the 40th IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 142–153, 1999.

[21] P. Feldman and S. Micali. An optimal probabilistic protocol for synchronous Byzantine agreement. *SIAM Journal on Computing*, 26(4):873–933, 1997.

[22] M. Fischer, N. Lynch, and M. Merritt. Easy impossibility proofs for distributed consensus problems. *Distributed Computing*, 1(1):26–39, 1986.

[23] M. Fischer, N. Lynch, and M. Paterson. Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(2):374–382, 1985.

[24] K. Ito and H.P. McKean. *Diffusion Processes and their Sample Paths*. Springer Verlag, Heidelberg, 1974.

[25] N.L. Johnson and S. Kotz. *Encyclopedia of Statistical Sciences*. John Wiley, New York, 1982.

[26] B. Kapron, D. Kempe, V. King, J. Saia, and V. Sanwalani. Fast asynchronous Byzantine agreement and leader election with full information. In *Proc. of the 19th ACM Symp. on Discrete Algorithms (SODA)*, pages 1038–1047, 2008.

[27] R. Karp, M. Luby, and F. Meyer auf der Heide. Efficient PRAM simulation on a distributed memory machine. In *Proc. of the 24th ACM Symp. on Theory of Computing (STOC)*, pages 318–326, 1992.

[28] J. Katz and C.-Y. Koo. On expected constant-round protocols for Byzantine agreement. *Journal of Computer and System Sciences*, 75(2):91–112, 2009.

[29] N. Lynch. *Distributed Algorithms*. Morgan Kaufmann Publishers, 1996.

[30] M. Rabin. Randomized Byzantine generals. In *Proc. of the 24th IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 403–409, 1983.

[31] G. Varghese. *Self-Stabilization by local checking and correction*. PhD thesis, PhD Thesis. MIT, 1992.

Appendix

Proof of Lemma 6. Consider transforming every $X_i = k$ into a binary string $B_k = (000...01)$ with $(k-1)$ zeroes. Then the series of $X_1 = k_1, X_2 = k_2, X_3 = k_3, ...$ can be represented by a string B of the form of $\{B_{k_1}B_{k_2}B_{k_3}...\} = \{00...100...100...100...1\}$. Note that B contains n many 1's and the total number of positions (0 or 1) in B is $K = \sum_i k_i$.

Now, consider instead an infinite set of binary random variables $Y_1, Y_2, Y_3, \ldots$ with $\Pr[Y_i = 1] = \delta$. Viewing Y_i as representing the i-th position in B, it is not difficult to check that

$$\Pr\left[\sum_{i=1}^{n} X_i \geq k\right] = \Pr\left[\sum_{j=1}^{k} Y_j \leq n\right]$$

Let $k = (1+\epsilon)\mu$ with $\mu = \mathrm{E}[X]$, $Y = \sum_{j=1}^{(1+\epsilon)\mu} Y_j$, and $\mu' = \mathrm{E}[Y]$. It follows from the Chernoff bounds that for any $\epsilon' \in [0, 1)$,

$$\Pr[Y \leq (1-\epsilon')\mu'] \leq e^{-(\epsilon')^2\mu'/2}$$

Since $\mathrm{E}[X_i] = 1/\delta$ and therefore $\mu = n/\delta$, it follows that $\mu' = (1+\epsilon)\mu \cdot \delta = (1+\epsilon)n$. Hence, if we set $n = (1-\epsilon')\mu'$, we get that $(1-\epsilon')(1+\epsilon) = 1$ and therefore $\epsilon' = \epsilon/(1+\epsilon)$. Plugging this into the equations above, we obtain

$$\Pr\left[\sum_{i=1}^{n} X_i \geq (1+\epsilon)n/\delta\right] \leq e^{-\epsilon^2 n/2(1+\epsilon)}$$

which proves the lemma. □

Proof of Lemma 7. By our assumption on X_i, we have for any integer $k \in \mathbb{N}$ that

$$\begin{aligned}\Pr[X_i \geq k] &\leq \gamma \sum_{l=k}^{\infty}(1-\delta)^l = \gamma \cdot \left(\frac{1}{\delta} - \sum_{l=1}^{k-1}(1-\delta)^l\right) \\ &\leq \gamma \cdot \left(\frac{1}{\delta} - \frac{1-(1-\delta)^k}{\delta}\right) = \frac{\gamma(1-\delta)}{\delta} \cdot (1-\delta)^{k-1}.\end{aligned}$$

Note that if Y is a geometric random variable with parameter δ, we have for any integer $k \in \mathbb{N}$, $\Pr[Y \geq k] = (1-\delta)^{k-1}$. Hence the tails of X_i and Y differ only by a factor of $\frac{\gamma(1-\delta)}{\delta}$ and our strategy is to relate X_i and Y by a domination argument.

Note first that if $\frac{\gamma(1-\delta)}{\delta} < 1$, then X_i is stochastically smaller than Y. Now if not, each X_i is stochastically smaller than $Y + \lceil \log_{1-\delta} \frac{\delta}{\gamma(1-\delta)} \rceil$, since for any integer $k \in \mathbb{N}$,

$$\begin{aligned}\Pr\left[Y + \left\lceil \log_{1-\delta} \frac{\delta}{\gamma(1-\delta)} \right\rceil \geq k\right] &= \Pr\left[Y \geq k - \left\lceil \log_{1-\delta} \frac{\delta}{\gamma(1-\delta)} \right\rceil\right] \\ &\geq (1-\delta)^{k - \log_{1-\delta} \frac{\delta}{\gamma(1-\delta)} - 1} = \frac{\gamma(1-\delta)}{\delta} \cdot (1-\delta)^{k-1}.\end{aligned}$$

Hence with Lemma 6 we have

$$\begin{aligned}\Pr\left[\sum_{i=1}^{n} X_i \geq (1+\epsilon)\mu + n\left(\left\lceil \log_{1-\delta} \frac{\delta}{\gamma(1-\delta)} \right\rceil\right)\right] &\leq \Pr\left[\sum_{i=1}^{n} Y_i \geq (1+\epsilon)\mu\right] \\ &\leq e^{-\epsilon^2 n/2(1+\epsilon)}.\end{aligned}$$

□

Proof of Lemma 8. A round t is called *successful* if

$$X_{t+1} \geq \begin{cases} 1 & \text{if } X_t = 0 \\ \min\{c_1 X_t, m\} & \text{otherwise} \end{cases}.$$

Let the random variable $Y \in \mathbb{N}_0 \cup \{\infty\}$ denote the number of successful rounds when starting with a round t_0 with $X_{t_0} = 0$ until we reach a round in which we fail. Certainly, $\Pr[Y = 0] \leq 1 - c_3$ and for any $k \geq 1$,

$$\Pr[Y = k] \leq c_3 \cdot \prod_{i=1}^{k-1} \left(1 - e^{-c_2 c_1^i}\right) \cdot e^{-c_2 c_1^k} \leq c_5 \cdot e^{-c_2 c_1^k}$$

for some constant $0 < c_5 < 1$. Hence, there are constants $0 < \gamma, \delta < 1$ with $\Pr[Y = k] \leq \gamma \cdot \delta^k$ for any k. Since $\Pr[Y < \infty] \geq 1 - c_3 + c_3 e^{-c_2}$, which is at least some constant $\gamma' > 0$, it holds that

$$\Pr[Y = k \mid Y < \infty] = \frac{\Pr[Y = k \ \wedge Y < \infty]}{\Pr[Y < \infty]} \leq (\gamma/\gamma')\delta^k = \gamma'' \delta^k$$

for some constant $\gamma'' > 0$. Thus, $\mathrm{E}[Y \mid Y < \infty] = \Theta(1)$.

So far, we only bounded $\Pr[Y = k \mid X_{t_0} = 0]$. However, in a similar way it also follows that for any $x_0 \geq 0$, $\Pr[Y = k \mid X_{t_0} = x_0] \leq \gamma \cdot \delta^k$ for some constants $0 < \gamma, \delta < 1$, so irrespective of X_{t_0}, $\Pr[Y = k] \leq \gamma \cdot \delta^k$ for some constants $0 < \gamma, \delta < 1$. Now, suppose our process had $f = \Theta(\log m)$ failures. We can upper bound the time it consumes for that by the sum $Z = \sum_{i=1}^{f} Z_i$ of independent random variables Z_i with $\Pr[Z_i = k] = \gamma\delta^k$ (so that they stochastically dominate Y for any X_{t_0}). Then Lemma 7 implies that Z is at most $\mathcal{O}(\log m)$ w.h.p., so at most $\mathcal{O}(\log m)$ rounds are needed for f failures.

On the other hand, it holds that for any X_{t_0} that

$$\Pr[Y \geq \log_{c_1} c_4 \log m] \geq c_3 \cdot \prod_{i=1}^{c_4 \log m} \left(1 - \exp(-c_2 c_1^i)\right) \geq c_6,$$

where $c_6 > 0$ is a constant that can be chosen independently of c_4. Hence, the probability that we have more than $2\log_{1/(1-c_6)} m$ attempts without achieving a value at least $X_t \geq c_4 \log m$ is smaller than $1 - m^{-2}$. So with probability at least $1 - m^{-2}$ we need at most $2\log_{1/(1-c_6)} m = \Theta(\log m)$ attempts and with high probability, from above, these $\Theta(\log m)$ attemts take $\Theta(\log m)$ rounds in total. □

Proof of Lemma 11. First we show that after a constant number t_1 of steps, we will have $X_t \leq n/4$ with high probability.

With high probability, the identity of the bin associated with X_t will not change during this process. Assume without loss of generality that the left bin initially has fewer balls so $L_0 = X_0$ and $R_0 = Y_0$.

For any ball i let $L_{j,i}$ be the random variables which equals 1 if ball i is in the left bin after j rounds, and 0 otherwise. If the ball i was in the left bin in round $j-1$, then writing $p_{j-1} = L_{j-1}/n$, we have $\mathrm{E}[L_{j,i}] = \Pr(L_{j,i} = 1) = 1 - (1 - p_{j-1})^2$. Similarly, if the ball i was in the right bin, we have $\mathrm{E}[L_{j,i}] = p_{j-1}^2$. Then $L_j = \sum_i L_{j,i}$ is the total number of balls in the left bin after j rounds. Then we have

$$\mathrm{E}[L_j] = L_{j-1}p_{j-1}(3 - 2p_{j-1}) = n(p_{j-1})^2(3 - 2p_{j-1}).$$

Let $\hat{C}$ be a constant greater than 1 such that

$$p^2(3 - 2p) \leq 1/2 - \hat{C}(1/2 - p)$$

for every $p \in [1/4, 1/2)$. Such a constant exists since the function $p \mapsto p^2(3-2p)$ is convex in this interval. Then

$$\mathrm{E}[L_1] \leq n \left[\frac{1-\hat{C}}{2} + \hat{C}p_0 \right] = L_0 \underbrace{\left[1 - \frac{(1-2p_0)(\hat{C}-1)}{2p_0} \right]}_{*}$$

provided $p_0 \geq 1/4$. Now, note that the gap $*$ is less than 1 for every fixed $p_0 < 1/2$, and so a Chernoff bound argument can be used to show that the probability of L_j, and hence p_j, not decreasing by a constant factor is $\exp(-\Omega(n))$. The gap $*$ increases with decreasing p_0, and hence, after a constant number of t_1 iterations, we have $p_{t_1} \leq 1/4$ w. h. p. (Note that t_1 depends on p_0, i.e., the constant c, but not on n.)

The number of steps for this first part remains the same if we allow for $\mathcal{O}(\sqrt{n})$ balls being in an adversarial matter, because the contraction of the smaller bin by a constant nontrivial factor is still possible.

Next, we now show that if $p_{t_1} \leq 1/4$, in another $t_2 = \mathcal{O}(\log\log n)$ rounds, we get

$$L_{t_1+t_2} \leq \sqrt{4n\log n}$$

with probability at least $1 - \mathcal{O}(\log\log(n)/n)$. We have $\mathrm{E}[L_j] \leq 3L_{j-1}^2/n$. Hence, again by using a Chernoff bound argument, we get

$$\Pr\left(L_j \geq \frac{9L_{j-1}^2}{2n} \right) \leq \exp\left(-\frac{9L_{j-1}^2}{4n} \right),$$

which, if $L_{j-1} \geq \sqrt{4n\log(n)}$, is bounded from above by $\mathcal{O}(n^{-1})$. To see that this needs $\mathcal{O}(\log\log n)$ steps, note the successive squaring in the mapping $x \mapsto 9x^2/(2n)$.

Again, the adversarial choice of $\mathcal{O}(\sqrt{n})$ balls in each round works the same, since there is an arbitrarily small $\varepsilon > 0$, for which $\mathrm{E}[L_j] + \mathcal{O}(\sqrt{n}) \leq (3+\varepsilon)L_{j-1}^2/n$. This is true because $L_{j-1}^2 \in \Omega(n)$ at this stage.

Once we are at $L_{t_1+t_2} \leq \sqrt{4n\log n}$, we have, by yet another application of Chernoff's inequality $L_{t_1+t_2+1} = \mathcal{O}(\log(n))$ with probability at least $1 - \mathcal{O}(n^{-4})$, and hence

$$\mathrm{E}[L_{t_1+t_2+2}] = \mathcal{O}(\log(n)^2/n).$$

Therefore, we are done with high probability after $t_1 + t_2 + 2$ steps.

In the adversarial case, we terminate as follows instead: We know that using Chernoff's inequality $L_{t_1+t_2+1} = \mathcal{O}(\sqrt{n})$ w. h. p, finishing the proof.

□

Proof of Lemma 12. Recall that $X_t = n/2 - \Delta_t$ is the number of balls in the smaller bin at step t. Furthermore, we define $\delta_t := \Delta_t/n$. The probability that a ball that is in the smaller bin at step t chooses its new median also in the same bin at step $t+1$ is

$$1 - (1/2 + \delta_t)^2 = 3/4 - \delta_t - \delta_t^2.$$

Similarly, the probability that a ball in the larger bin at step t chooses its new median in the other bin is

$$(1/2 - \delta_t)^2 = 1/4 - \delta_t + \delta_t^2.$$

Now denote by $\tilde{X}_{t+1}$ the number of balls that choose its new median in step $t+1$ in the bin corresponding to X_t, in the absence of an adversary (at step $t+1$). Linearity of expectation gives

$$\begin{aligned}
\mathrm{E}[\tilde{X}_{t+1}] &= (1/2-\delta_t)n(3/4-\delta_t-\delta_t^2) + (1/2+\delta_t)n(1/4-\delta_t+\delta_t^2) \\
&= (1/2-(3/2)\delta_t+2\delta_t^3)n \\
&= n/2-\Delta_t-\left((1/2)\delta_t-2\delta_t^3\right)n \\
&\leq n/2-\Delta_t-(1/4)\delta_t n && \text{(using } \delta_t < 1/3\text{)} \\
&\leq X_t-(\delta_t/2)X_t && \text{(using } X_t \leq n/2\text{)} \\
&= (1-\delta_t/2)X_t.
\end{aligned}$$

Since the choices of the balls are independent, it follows from the Chernoff bounds that for $\epsilon = \delta_t/4$,

$$\begin{aligned}
\Pr[\tilde{X}_{t+1} \geq (1-\delta_t/4)X_t] &\leq \Pr[\tilde{X}_{t+1} \geq (1+\epsilon)\mathrm{E}[\tilde{X}_{t+1}]] \leq e^{-\epsilon^2\mathrm{E}[\tilde{X}_{t+1}]/3} \\
&\leq e^{-(\delta_t/4)^2(1-\delta_t/2)(n/2)/3} \leq e^{-(c^2\ln n/n)n/96} = n^{-c^2/96}.
\end{aligned}$$

This implies that $\tilde{X}_{t+1} \leq (1-\delta/4)X_t$ w.h.p. If X_t is the left bin then $L_{t+1} \leq \tilde{X}_{t+1}+\mathcal{O}(\sqrt{n})$, which shows that w.h.p.

$$\begin{aligned}
L_{t+1} &\leq \tilde{X}_{t+1}+\mathcal{O}(\sqrt{n}) \\
&\leq (1-\delta_t/4)X_t+\mathcal{O}(\sqrt{n}) \\
&\leq (1-\delta_t/8)X_t && \text{(since } \Delta_t = \Omega(\sqrt{n\log n})\text{)},
\end{aligned}$$

which in turn implies that $\Delta_{t+1} \geq (10/9)\,\Delta_t$. Hence taking the union bound over $\mathcal{O}(\log n)$ rounds, we reach a step with an imbalance of at least $n/3$ w. h. p. □

Proof of Lemma 13. We show stochastic domination for any two labelled imbalances Ψ_t and $\Psi'_t = \Psi_t - 1$. The rest follows by induction. Let $z = n/2 - \Psi'_t$. Without loss of generality, we assume that balls 1 to $z-1$ are in the left bin in both Ψ_t and Ψ'_2 and balls $z+1,\ldots,n$ are in the right bin in both Ψ_t and Ψ'_t. Ball z is in the right bin in Ψ_t and in the left bin in Ψ'_t.

Let Ω be the space of all possible outcomes of the random experiment in which every ball chooses two balls independently and uniformly at random. Consider any such outcome $w \in \Omega$.

Any ball $b \neq z$ that does not choose ball z in w goes to the same bin in both scenarios.

If ball z goes to the right bin in the Ψ'_t scenario (in which it started in the left bin) then it will also go the right bin in the Ψ_t scenario (in which it started in the right bin).

Finally, consider a ball $b \neq z$ that chooses ball z as one (or both) of its choices in w. If it goes to the right bin in the Ψ'_t scenario (in which the z ball is dragging it left) it will also go to the right bin in the Ψ_t scenario.

So R_{t+1} dominates R'_{t+1} and L'_{t+1} dominates L_{t+1} and $R_{t+1} - L_{t+1}$ dominates $R'_{t+1} - L'_{t+1}$. □

Proof of Lemma 14. We only need to prove the claim for $\Psi_t = 0$ because the general case follows from stochastic domination (see Lemma 13).

Assume that at step t balls 1 to $n/2$ reside in the left bin and balls $n/2+1$ to n reside in the right bin. Let $X_1,\ldots,X_{n/2} \in \{0,1\}$ be random variables defined as follows:

$$X_i = \begin{cases} 1 & \text{if ball } i \text{ moves to the right bin,} \\ 0 & \text{otherwise.} \end{cases}$$

Then the X_i are independent Bernoulli variables with $\Pr(X_i = 1) = 1/4$. Analogously, for balls $n/2+1$ to n, we define the random variables $X_{n/2+1}$ to X_n which are equal to 1 if the balls move from the left bin to the right bin. We again have $\Pr(X_i = 1) = 1/4$ for these variables. Then Ψ_{t+1} is

$$Y = \sum_{i=1}^{n/2} X_i - \sum_{i=n/2+1}^{n} X_i.$$

Now by the Central Limit Theorem, as $n \to \infty$, the random variable $\sqrt{2/n}\sum_{i=1}^{n/2}(X_i - 1/4)$ converges in distribution to a normal variable of mean 0 and variance 3/16. So, for large enough n, we have

$$\begin{aligned}
\Pr(Y \geq c\sqrt{n}) &= \Pr\left(\sqrt{\frac{2}{n}}Y \geq c\sqrt{2}\right) \\
&= \Pr\Big(\underbrace{\sqrt{2/n}\Big(\sum_{i=1}^{n/2}\Big(X_i - \frac{1}{4}\Big) - \sum_{i=n/2+1}^{n}\Big(X_i - \frac{1}{4}\Big)\Big)}_{\text{asympt. normal with } \mu = 0,\ \sigma^2 = 3/8} \geq c\sqrt{2}\Big) \\
&\geq 1 - \Phi(c\sqrt{16/3}) - \varepsilon,
\end{aligned}$$

where Φ denotes the cumulative distribution function of a standard-normal random variable.

For $x \geq 0$, the value of $\Phi(x)$ can be bounded as follows (see e.g., [24] p. 17 and [25] p. 505):

$$\frac{1}{\sqrt{2\pi}(1+x)} \cdot e^{-x^2/2} \leq 1 - \Phi(x) \leq \frac{1}{\sqrt{\pi}(1+x)} \cdot e^{-x^2/2},$$

therefore we can lower bound the above probability by

$$\frac{1}{\sqrt{2\pi}(1 + 4c/\sqrt{3})} \exp\left(-\frac{8c^2}{3}\right) - \varepsilon,$$

finishing the proof.

□

Proof of Lemma 15. As in Lemma 12, a simple calculation reveals that $\mathrm{E}[\Delta_{t+1}] \geq (3/2)\Delta_t$. Using Hoeffdings bound, we have w. h. p.

$$\Pr[\Delta_{t+1} \leq (4/3)\Delta_t] \leq \exp(-\Theta(\Delta_t)^2/n).$$

□

Proof of Lemma 16. If there is no adversary, then Lemma 14 states that after $\mathcal{O}(1)$ steps, we are in the hypothesis of Lemma 15. That is, $\Delta_t \leq c\sqrt{n}$. Now let $\Upsilon_\tau = \lfloor \Delta_{t+\tau-1}/(c\sqrt{n} \rfloor$ so $\Upsilon_1 \in \{0, 1\}$. Let $m = \lfloor (n/2)/(c\sqrt{n} \rfloor$ so the possible values of Υ_τ are $\{0, \ldots, m\}$. By Lemma 8, $O(\log m)$ rounds suffice to achieve $\Upsilon_\tau \geq c_4 \log m$ so $\Delta_{t+\tau-1} \geq c\sqrt{n}c_4 \log m$ which is sufficient.

If there is an adversary, he can reduce the drift by at most $\sqrt{n}/2$ per round. Hence, choosing the constant c large enough, we can still arrange for a drift factor > 1 in Lemma 15. □

Proof of Lemma 17. If $b_{0,1}, \ldots, b_{0,n}$ is a configuration corresponding to the initial assignment $(k_i)_{i\in[n]}$, then $\tilde{b}_{0,i} := f(b_{0,i})$ is a configuration corresponding to the assignment $\tilde{k}_i$.

Since f is monotonic, it commutes with the median function, i.e.,

$$\text{median}(f(a), f(b), f(c)) = f(\text{median}(a, b, c)).$$

Therefore we have by induction that $\tilde{b}_{t,j} = f(b_{t,j})$ for any $t \in \mathbb{N}$ and $j \in [n]$, if we run both instances with the same random values $I_{t,j}$ and $J_{t,j}$.

Now if the $(b_{0,j})_{j\in[n]}$ instance has converged in the t-th iteration, we have $b_{t,1} = \cdots = b_{t,n}$, and hence also

$$f(b_{t,1}) = \cdots = f(b_{t,n}),$$

and so the $(\tilde{b}_{0,j})_{j\in[n]}$ instance has also converged in this iteration. □

Proof of Lemma 18. Assume w.l.o.g. that $j \leq m_t$ (the case $j \geq m_t$ follows with identical arguments). Let i be the number of a ball in $\mathcal{H}_{t,j}$ with gravity $g(i) < 4/3$. When plugging $g(i) < 4/3$ into Equation (1), we get

$$\frac{4}{3} > 6\frac{(n-i)i}{n^2} + \mathcal{O}\left(\frac{1}{n}\right),$$

which readily implies that $i \leq n/3 + \mathcal{O}(1)$. Hence, there are at most $n/3 + \Phi + \mathcal{O}(1)$ balls in the bins 1 to j. Then consolidate all bins from 1 to j into a superbin A, and all other bins into a superbin B. Let $L_{t,A}$ be the load of superbin A in step t, so $L_{t,A} \leq n/3 + \Phi + \mathcal{O}(1)$. By Lemma 17, we can using the arguments from the two-bin case (Lemmas 11 and 12) to conclude that w. h. p.

$$L_{t+1,A} \leq \frac{n}{3+\epsilon},$$

for a constant $\epsilon > 0$. Hence by (1), every ball $l \in \mathcal{H}_{t+1,j}$ in bin j satisfies $g(l) < 4/3$ w.h.p. (provided that $\mathcal{H}_{t+1,j} \neq \emptyset$). □

Proof of Lemma 19. Consider an arbitrary round t. Our goal is to apply Lemma 9. We first identify two absorbing states concerning $\mathcal{H}_{t,j}$:

1. There is a ball $i \in \mathcal{H}_{t,j}$ with $g(i) < 4/3$. Then Lemma 18 implies that at step $t+1$, $\mathcal{H}_{t+1,j}$ contains at least one ball l with $g(l) < 4/3$, or $\mathcal{H}_{t+1,j}$ is empty.
2. $|\mathcal{H}_{t,j}| = \Phi$ and all balls $i \in \mathcal{H}_{t,j}$ satisfy $g(i) \geq 4/3$. Then it follows by a Chernoff bound that $|\mathcal{H}_{t+1,j}| = \Phi$ with probability at least $1 - n^{-2}$.

If $\mathcal{H}_{t,j}$ does not fulfill one of these conditions, all balls in $\mathcal{H}_{t,j}$ have a gravity of at least 4/3. In this case, the expected number of balls in bin j at step $t+1$ would be at least $(4/3)|\mathcal{H}_{t,j}|$. Since the median rule is applied independently at random to each ball, a Chernoff bound implies

$$\Pr\left[|\mathcal{H}_{t+1,j}| \geq \min\{\Phi, \frac{5}{4}|\mathcal{H}_{t,j}|\}\right] \geq 1 - \exp(-\Theta(|\mathcal{H}_{t,j}|)).$$

Thus, applying Lemma 9, we conclude that one of the two absorbing states is reached within $t_1 = \mathcal{O}(\log n)$ rounds w. h. p. □

Proof of Theorem 20. Let the set of non-empty bins be $\{1, \ldots, m\}$ at the beginning. We divide the time into $\log m + 1$ phases, numbered from 1 to $\log m + 1$. For each phase i with $1 \leq i \leq \log m$, we shall prove by induction that at the end of the phase, there is a subset $S_i \subseteq \{1, \ldots, m\}$ of size $|S_i| \leq m/2^i + 1$ that satisfies

$$\min\big\{R(S_i), L(S_i)\big\} \geq \frac{n}{2} + C\sqrt{n \log n}, \tag{2}$$

where $R(S_i)$ ($L(S_i)$) denotes the total load of all bins that are in the set S_i or located right (left) from S_i, respectively. The idea behind the definition is that at the end of each phase i, we know that the bin that gets all balls (up to $\mathcal{O}(\sqrt{n})$ balls) at the end is located in S_i.

Let us now prove (2) by induction. For the induction base, cut the set of all bins into two equally-sized, consecutive sets of bins $S_1^{\text{left}} := \{1, \ldots, \lfloor m/2 \rfloor\}$ and $S_1^{\text{right}} := \{\lfloor m/2 \rfloor + 1, \ldots, m\}$. Now regard S_1^{left} and S_1^{right} as two bins. Our aim is to prove that after $\mathcal{O}(\log \log n)$ steps, one of the two bins will have at least $\frac{n}{2} + C\sqrt{n \log n}$ balls. To show this, we apply the Lemma 14 and Lemma 15 from the two bin-anlysis. Let t be the first time step of phase i and recall that Δ_t is the imbalance at time t.

First we apply Lemma 14 to get that with constant probability > 0, $\Delta_{t+1} \geq 5\sqrt{n}$ holds (if there is no adversary). Since the adversary can influence at most $4\sqrt{n}$ balls, we have $\Delta_{t+1} \geq \sqrt{n}$ with constant probability. Then we apply Lemma 15 to get

$$\Pr[\Delta_{t+\mathcal{O}(\log \log n)} \geq C\sqrt{n \log n}] \geq \prod_{k=1}^{\mathcal{O}(\log \log n)} \left(1 - \exp(-\Theta((4/3)^k)\right),$$

which is at least a constant greater than zero. Hence the expected time to reach a step t_0 with $\Delta_{t_0} \geq C\sqrt{n \log n}$ is $\mathcal{O}(\log \log n)$, which completes the induction base.

Assume now more generally, that at the end of phase i, a set S_i of size at most $m/2^i + 1$ exists with

$$\min\{R(S_i), L(S_i)\} \geq \frac{n}{2} + C\sqrt{n \log n}.$$

Again, we divide S_i into two consecutive sets of bins S_i^{left} and S_i^{right}, each of size at most $m/2^{i+1}+1$. Now regard S_i^{left} together with all bins left from it and S_i^{right} together will all bins right from it as two separate bins, $L(S_i^{\text{left}})$ and $R(S_i^{\text{right}})$. Applying the same arguments as from the induction base, we obtain that after expected $\mathcal{O}(\log \log n)$ steps, the imbalance between $L(S_i^{\text{left}})$ and $R(S_i^{\text{right}})$ is at least $C\sqrt{n \log n}$. Assume w.l.o.g. that $L(S_i^{\text{left}}) \geq \frac{n}{2} + C\sqrt{n \log n}$. Then we set $S_{i+1} := S_i^{\text{left}}$ and note that by assumption,

$$L(S_{i+1}) = L(S_i^{\text{left}}) \geq \frac{n}{2} + C\sqrt{n \log n}.$$

Moreover, we know from the induction hypothesis, that at the end of the previous phase,

$$R(S_i) \geq \frac{n}{2} + C\sqrt{n \log n}.$$

Moreover, the proof of Lemma 12 implies that if the load of any set of balls is above $n/2+C\sqrt{n \log n}$, it never decreases in any following round with high probability. Hence using that the leftmost bin in S_i^{left} is also the leftmost bin in S_i,

$$R(S_{i+1}) = R(S_i^{\text{left}}) \geq R(S_i) \geq \frac{n}{2} + C\sqrt{n \log n}.$$

This completes the induction and proves (2).

So we have shown that the time to reach the end of phase $\log m$ can be bounded by the sum of $\log m$ independent geometric random variables, each with mean $\mathcal{O}(\log \log n)$. Hence Lemma 6 implies that after $\mathcal{O}(\log m \log \log n + \log n)$ steps, we have completed phase $\log m$ with high probability.

Now at the end phase of $\log m$, there is a set of two bins $S = S_{\log m} = \{j, j+1\}$

$$\min\{R(S), L(S)\} \geq \frac{n}{2} + C\sqrt{n \log n}.$$

Applying Lemma 11 and Lemma 12 to $R(S)$ and $L(S)$, we obtain that $R(S)$ and $L(S)$ are both larger than $n - (C/2)\sqrt{n \log n}$ after additional $\mathcal{O}(\log n)$ rounds with high probability. Since the intersection of bins in $R(S)$ and $L(S)$ is at most two, we conclude that there is a set of at most two bins that contains $n - C\sqrt{n \log n}$ balls with high probability. Applying Theorem 10, we conclude that after additional $\mathcal{O}(\log n)$ rounds, we will have reached an almost stable consensus.

□