Stabilizing Consensus With the Power of Two Choices

Benjamin Doerr[*]
Max Planck Institute for Informatics
Saarbrücken, Germany

Leslie Ann Goldberg[†]
Dept. of Computer Science
University of Liverpool
Liverpool, UK

Lorenz Minder
Computer Science Division
University of California
Berkeley, USA

Thomas Sauerwald[‡]
Max Planck Institute for Informatics
Saarbrücken, Germany

Christian Scheideler[§]
Dept. of Computer Science
University of Paderborn
Paderborn, Germany

ABSTRACT

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* [2]. 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, stable value without necessarily being aware of that. This should work irrespective of the states in which the processes are starting. Our main result is a simple randomized algorithm called *median rule* that, with high probability, just needs $\mathcal{O}(\log m \log\log n + \log n)$ time and work per process to arrive at an almost stable consensus for any set of m legal values as long as an adversary can corrupt the states of at most $\sqrt{n}$ processes at any time. Without adversarial involvement, just $O(\log n)$ time and work is needed for a stable consensus, with high probability. As a by-product, we obtain a simple distributed algorithm for approximating the median of n numbers in time $\mathcal{O}(\log m \log\log n + \log n)$ under adversarial presence.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity

General Terms

Algorithms, Theory

Keywords

distributed consensus, self-stabilization, randomized algorithms

1. INTRODUCTION

Consensus problems occur in many contexts and have therefore been extensively studied in the past (e.g., [8, 32]). Interesting applications are the consolidation of replicated states or information and the synchronization of processes and devices. In the original consensus problem, every process proposes a value, and the goal is to decide on a single value from all those proposed. If all processes are working in a correct and timely manner, the consensus problem is easy to solve by performing a leader election, for example. However, 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 [21]. 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 [20]. However, these two results only apply to deterministic algorithms and only to the case where processes need to commit to a value and this commitment can only be done once.

We are studying a slight variant of the consensus problem called the *stabilizing consensus problem* [2]. In this problem, we do not require that each process irrevocably commits to a final value but that eventually they arrive at a common, stable value without necessarily being aware of that. This should work irrespective of the states in which the processes are starting. In other words, we are searching for a *self-stabilizing* algorithm for the consensus problem. Coming up with such an algorithm is easy without adversarial involvement, but we allow some adversary to continuously corrupt the state of some processes. Despite these corruptions, we would like most of the processes to arrive quickly at a common value that will be preserved for any polynomial in n

[*] Partially supported by ???

[†] Partially supported by EPSRC grant EP/I011528/1

[‡] Email: sauerwal@mpi-inf.mpg.de

[§] Email: scheideler@upb.de. Partially supported by DFG grant SCHE 1592/1-1.

SPAA'11, June 4–6, 2011, San Jose, California, USA.
Copyright 2011 ACM 978-1-4503-0743-7/11/06 ...$10.00.

many steps. Interestingly, we will demonstrate that there is a simple randomized algorithm for this problem that essentially needs logarithmic time and work with high probability to arrive at such a stable value.

1.1 Our approach

We will focus on synchronous message-passing systems with adversarial state corruptions. The time proceeds in synchronized *rounds.* In each round, every process can send out one or more requests, receive replies to its requests, and perform some local computation based on these replies.

We assume that we have a fixed set of n processes that faithfully follow the protocol (based on their current state, which might be corrupted), and every process knows all the other processes in the system (i.e., there are no connectivity constraints). As usual in the literature, the *state* of a process contains all of its variables but does not include its contact information about the other processes and the protocol (which are fixed throughout the lifetime of the processes). The *system state* includes all of the processes' local states. In general, a system is called *self-stabilizing* if in the absence of state corruptions (caused by faults or adversarial behavior) it holds that (1) when started in an arbitrary state, the system eventually reaches a legal state (*convergence*) and (2) given that the system is in a legal state, it will stay in a legal state (*closure*). In the *stabilizing consensus problem* we are more strict as we do not just demand that a consensus is reached and from that point on a consensus is maintained but that we have a *stable* consensus. That is, the n processes may initially have arbitrary states with values $v_1, \dots, v_n$ out of some set S of legal values and the goal is to arrive at a single, stable value among these values. A system state S is called *stable* if in all possible executions starting from S, the values of the processes do not change. If every process has the same value x in a stable system state S, we say that the values *stabilize* to x. A *self-stabilizing consensus protocol* must satisfy the following properties (given that there are no state corruptions):

- **Stabilization**: the protocol eventually reaches a stable state.
- **Validity**: if a process has some value v, then some process must have had v in the previous round.
- **Agreement**: for every reachable stable state, all processes have the same value.

Note that the validity rule prevents the processes from just changing to a default value. Otherwise, the consensus problem would be trivial.

The *runtime* of a self-stabilizing consensus protocol is the number of communication rounds it takes until a stable state is reached. Besides the runtime, we will also consider the *work* of such a protocol, which is the maximum number of messages (i.e., requests and replies) a process has to handle until a stable state is reached. This disqualifies simple strategies like "everybody contacts process 1" as its work would be n while its runtime is 1. For a distributed system to be scalable, both the runtime and the work has to be as low as possible, therefore we are focussing on protocols with low runtime *and* work.

The adversary

We assume that adversarial state corruptions can continuously happen during the self-stabilizing process. Most of the self-stabilizing algorithms proposed in the literature are not guaranteed any more to reach a legal state in this case, so finding algorithms that still converge to a legal state is a non-trivial problem. We assume that there is a *T-bounded adversary* that knows the entire state of the system at the end of each communication round. Based on that information, it may corrupt the state of up to T processes in an arbitrary way before the next round starts.

Of course, under a T-bounded adversary we cannot reach a stable state any more. Therefore, we will only require the system to reach a state S so that for any poly(n) many time steps following S, all but at most $\mathcal{O}(T)$ processes agree on some stable value v (note that these $\mathcal{O}(T)$ processes can be different from round to round). We will call this an *almost stable state.* The goal is to come up with an efficient protocol so that for values of T that are as large as possible, an almost stable state can be reached with a runtime and work that is as low as possible.

1.2 Our contributions

We are focussing on stabilizing consensus problems based on an arbitrary (finite or countably infinite) set S of *legal* values with a total order. Classical examples are $S = \{0, 1\}$ and $S = \mathbb{N}$. All initial values of the processes must be from S and also the adversary is restricted to choosing only values in S. (If the adversary chooses a value outside of S in some process p, we may assume that p instantly recognizes that and then switches over to some default value in S.)

If no process is ever corrupted, we can restrict S to be the set of initial values as no new values will ever be introduced by a protocol satisfying the validity rule. In this case, the stabilizing consensus problem could easily be solved with the following *minimum rule*: Suppose that the current value of process i is v_i. In each round, every process i contacts some random process j in the system and updates its value to $v_i := \min\{v_i, v_j\}$. It is easy to see that this rule needs just $\mathcal{O}(\log n)$ time and work with high probability (or short, w.h.p.)[1] until all processes have the same value, namely the minimum of the initial values $v_1, \dots, v_n$. Since they will not deviate from that value any more, we have reached a stable state. However, if some processes can be corrupted, then even for $S = \{0, 1\}$ *no runtime bound* can be given for the minimum rule to reach an (almost) stable state: if all processes start with 1, then the adversary could inject 0 at any time later to cause a change in the consensus. In fact, a 1-bounded adversary would be sufficient for that. Therefore, we are proposing a different rule 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) and requests their values. It then updates v_i to the *median* of v_i, v_j and v_k. Any request sent to process i will be answered with the value that i had *at the beginning* of the current round.

For example, if $v_i = 10$, $v_j = 12$ and $v_k = 130$, then the new value of v_i is 12. When taking the *mean* of a selected group of values instead of the median, the convergence properties towards a single number have already been formally analyzed [16]. However, with the mean rule we are no longer

[1] We write w.h.p. to refer to an event that holds with probability at least $1 - n^{-c}$ for any constant $c > 1$.

guaranteed to solve the stabilizing consensus problem as the validity rule may be violated. Moreover, the approach in [16] is quite different from our approach (as it is based on a repeated all-to-all exchange of values and some filtering mechanism before computing means), so its analysis cannot be adapted to the median rule.

The median rule works surprisingly well. We prove the following results that are also summarized in Figure 1.

Theorem 1.1. *For any initial state it holds that if no process is ever corrupted, then the median rule needs just $\mathcal{O}(\log n)$ time and work to reach a stable consensus w.h.p.*

Hence, the median rule is as effective as the minimum rule in the non-adversarial case. Contrary to the minimum rule, the median rule also works for the adversarial case. Let $m = |S|$ be the number of legal values. Then it holds:

Theorem 1.2. *For any $\sqrt{n}$-bounded adversary, the median rule needs just $\mathcal{O}(\log m \cdot \log\log n + \log n)$ time and work to reach an almost stable consensus w.h.p.*

Of course, $|S|$ may not be finite. In this case, Theorem 1.2 still holds if we define m as the number of legal values between v_ℓ and v_r, where v_ℓ is the $(n/2 - c\sqrt{n \log n})$-smallest and v_r is the $(n/2 + c\sqrt{n \log n})$-smallest value of the initial values for some sufficiently large constant c. As a by-product, the median rule computes a good approximation of the median, even under the presence of an adversary.

Corollary 1.3. *For any $\sqrt{n}$-bounded adversary, the median rule needs just $\mathcal{O}(\log m \cdot \log\log n + \log n)$ time and work to compute an almost stable value that is between the $(n/2 - c\sqrt{n \log n})$-largest value and the $(n/2 + c\sqrt{n \log n})$-largest value of the initial values w.h.p.*

The bound on T is essentially tight as $T = \Omega(\sqrt{n \log n})$ 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. A further improvement of Theorem 1.2 can be obtained in an average-case setting:

Theorem 1.4. *Let $m \le n^{1/2-\epsilon}$. If initially each process chooses one out of m legal values uniformly at random, then for any $\sqrt{n}$-bounded adversary, the median rule needs $\Theta(\log n)$ time and work w.h.p., if m is even, and $\mathcal{O}(\log m + \log\log n)$ time and work w.h.p., if m is odd, to reach an almost stable consensus.*

Finally, if the T-bounded adversary is *static* in a sense that there is a *fixed* set of T faulty processes throughout the execution, then we present a simple extension of the median rule to a so-called *careful median rule* that reaches, within the time bound given in Theorem 1.2, a consensus that is stable for poly(n) many rounds for *all non-faulty processes* w.h.p. This is not possible with the original median rule as with $T = \sqrt{n}$ there is a constant probability that some process contacts two corrupted processes and therefore changes its value to a value selected by the adversary.

With these results, the median rule is yet another demonstration of the *power of two choices* as the time needed by the minimum rule (as well as the maximum rule) can be unbounded even for $S = \{0, 1\}$ and 1-bounded adversaries. This power of two choices has also been demonstrated in many other contexts [26, 15, 9, 13, 14, 18] (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.

1.3 Model discussion

Theorem 1.2 also holds for other adversarial models. We just consider two of them:

Adversarial processes: Suppose that we have a $\sqrt{n}/4$-bounded adversary that can pick any $\sqrt{n}/4$ processes at the beginning of a round that behave in an arbitrary adversarial manner throughout that round. Since these processes will only be contacted by at most $3\sqrt{n}/4$ other processes w.h.p. when using the median rule, we can emulate the effect of $\sqrt{n}/4$ adversarial processes on the system by $\sqrt{n}$ adversarial state corruptions, which is our original model. So our results extend to adversarial processes.

Sleep scheduling: Suppose that we have a $\sqrt{n}/4$-bounded adversary that can just put any $\sqrt{n}/4$ processes to sleep in a round. Also this model can be simulated by our original $\sqrt{n}$-bounded adversary using the same arguments as for adversarial processes. More interestingly, one can already show for the sleep scheduling model that if the adversary can put $\Omega(\sqrt{n \log n})$ processes to sleep in each round, a consensus cannot be reached any more for polynomially many steps w.h.p. even if $|S| = 2$. (The proof as well as the strategy achieving this is simple: given an imbalance of Δ, put 2Δ processes of the majority value to sleep. This increases the presence of the minority value which reduces the imbalance.) This implies that even a slight asynchrony that is under the control of an *adaptive* adversary can lead to a failure of the median rule. However, we note that the protocol by Angluin et al. [1] would suffer from the same problem, so adaptive asynchrony seems to be hard to handle for simple distributed algorithms.

1.4 Related work

Randomized algorithms are known that can solve the consensus problem with probability approaching one in many different cases ranging from asynchronous message passing models to shared memory models (see, e.g., [34, 12, 19, 10, 17, 11, 30, 25, 4, 5, 27] or [3] for a survey). Most of these algorithms can tolerate a constant fraction of Byzantine fail/stop failures or nodes but at the cost of spending $\Omega(n)$ expected individual work. Also several lower bounds are known. Ben-Or and Bar-Joseph [10] have shown that any consensus protocol that tolerates $\Theta(n)$ adaptive fail-stop faults runs for $\tilde{\Omega}(\sqrt{n})$ rounds. Attiya and Censor [7] proved that $\Omega(n^2)$ is a lower bound on the total work under adaptive adversaries in the shared memory model. The same authors [6] also showed for message passing as well as shared memory systems that for every integer k, the probability that an f-resilient randomized consensus algorithm for n processes does not terminate with agreement within $k(n - f)$ steps is at least $1/c^k$ for some constant c.

Recently, Gilbert and Kowalski [22] presented a randomized consensus algorithm that runs in $\mathcal{O}(\log n)$ time and uses only $\mathcal{O}(n)$ bits in total for all messages. However, the adversary is not fully adaptive; it has to specify the set of faulty processes in advance. In addition, there are some processes which have to send $\Omega(n/\log n)$ bits during the execution of the algorithm. Very recently, King and Saia [29] presented a randomized algorithm for Byzantine agreement that runs in

	with adversary	**without adversary**
worst-case, $m = 2$	$\mathcal{O}(\log n)$	$\mathcal{O}(\log n)$
worst-case, arb. m	$\mathcal{O}(\log m \log\log n + \log n)$	$\mathcal{O}(\log n)$
average-case, arb. m	$\mathcal{O}(\log m + \log\log n)$ if m is odd $\Theta(\log n)$ if m is even	$\mathcal{O}(\log m + \log\log n)$ if m is odd $\Theta(\log n)$ if m is even

Figure 1: Our results on the time and work required to reach an almost stable consensus (with adversary) or stable consensus (without adversary). m **is the number of legal values. By average-case we refer to the case where every initial value is chosen independently and uniformly at random among the** m **legal values. The results for the average case with adversary require that** $m \leq n^{1/2-\epsilon}$ **for some constant** $\epsilon > 0$**.**

polylogarithmic time and only needs $\tilde{\mathcal{O}}(\sqrt{n})$ bits per process against an adaptive adversary.

The consensus problem has also been studied in the context of population protocols, which are protocols for extremely simple, passively mobile systems. Angluin et al. [1] show that with high probability, n agents that meet at random reach consensus in $O(n \log n)$ pairwise interactions and the value chosen is the majority provided that its initial margin is at least $\omega(\sqrt{n \log n})$. This protocol has the additional property of tolerating Byzantine behavior in $o(\sqrt{n})$ of the agents. We can also show these properties for the median rule, but we are more general than Angluin et al. as we allow a set of legal values of arbitrary cardinality whereas Angluin et al. only consider two different values. The result by Angluin et al. can be extended to m different values, but their analysis would only allow one to conclude a runtime of $O(\log m \log n)$ for the non-adversarial as well as the adversarial situation whereas our runtime bounds are $O(\log n)$ and $O(\log m \log\log n + \log n)$ respectively.

Due to the fact that even in the adversarial setting, the median rule stabilizes to a value that is the k-smallest of the initial values for some $k \in [n/2 - c\sqrt{n \log n}, n/2 + c\sqrt{n \log n}]$ w.h.p. and therefore gives a good approximation of the median of the initial values, it is also interesting to compare the median rule with other distributed algorithms for finding the median. Kempe et al. [28] proposed a gossip-based algorithm that computes the median within $O(\log^2 n)$ communication rounds in a complete graph w.h.p. Patt-Shamir [33] showed that the median can be approximated to within ϵn distance from the median with just $O((\log\log n)^3)$ bit transmissions per node if each element can be encoded with $O(\log n)$ bits. Kuhn et al. [31] showed that in networks of diameter D, the median can be found in $O(D \log_D n)$ communication rounds w.h.p. and also prove a matching lower bound holding for a general class of distributed algorithms. The median problem has also been studied in the context of sensors networks (e.g., [35]), but mostly experimentally. However, none of these previous results consider the adversarial case.

2. TWO VALUES WITH ADVERSARY

In this section, we focus on the case that there are only two legal values, x_0 and x_1 with $x_0 < x_1$. Before we analyze the median rule for that case, we propose an alternative notation for our consensus problem based on balls into bins. We have n *balls* representing the processes and 2 *bins* representing the two legal values. In that notation, the state of the system at the beginning of a round is represented by a distribution of the balls among the bins, and a T-bounded adversary may pick up any T balls at the end of each round and throw them into any of the two bins. Even though the two-bin case sounds fairly restrictive, this case turns out to be of general interest, as our analysis for more than two bins will use some results of this section. For the two-bin case, the median rule is equivalent to the *majority rule*, where a ball's next bin is chosen to be the bin that is used by the majority of itself and the two random balls.

THEOREM 2.1. *For* $|S| = 2$ *and any initial distribution of the balls,* $\mathcal{O}(\log n)$ *rounds of the median (majority) rule suffice for any* $\sqrt{n}$*-bounded adversary to reach an almost stable consensus, w.h.p.*

In this theorem as well as the other theorems below, we will just focus on providing time bounds as the work bounds in Section 1.2 immediately follow from the time bounds with the help of standard Chernoff bounds when using the median rule:

LEMMA 2.2 (CHERNOFF BOUNDS). *Let* $X_1, \ldots, X_n$ *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*

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

The rest of this section is dedicated to the proof of this theorem. In the following, let L_t be the number of balls in the left bin at (the end of) 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 focus on the case with even n, since 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 non-negative integer). We will use $\tilde{X}_t, \tilde{\Delta}_t$ to denote the corresponding numbers before the adversary is allowed to change the location of up to T balls at the end of round t. Based on the imbalance Δ_t we distinguish between three cases.

Case 1: $\Delta_t \geq n/4$

We will show the following lemma whose proof uses standard Chernoff bounds (see Lemma 2.2).

LEMMA 2.3. *If there is a step* t_0 *with* $\Delta_{t_0} \geq n/4$*, then there is a step* $t_1 = t_0 + \mathcal{O}(\log\log n)$ *at which we reach a stable consensus (if there is no adversary) or an almost stable consensus (for any* $\sqrt{n}$*-bounded adversary) w.h.p.*

PROOF. Note that initially $X_t \leq n/4$ (as by assumption, $\Delta_t \geq n/4$). Assume without loss of generality that the left bin initially has fewer balls, so $L_{t_0} = X_{t_0}$ and $R_{t_0} = Y_{t_0}$, and for simplicity we may assume that $t_0 = 0$. For any step t, set $p_t := L_t/n$.

Without adversary: For any ball i let the binary random variable $L_{t,i}$ be 1 if and only if ball i is in the left bin after t rounds, and 0 otherwise. If ball i was in the left bin in round $t-1$, then writing $p_{t-1} = L_{t-1}/n$, we have $\mathrm{E}[L_{t,i}] = \Pr[L_{t,i} = 1] = 1 - (1 - p_{t-1})^2$. Similarly, if the ball i was in the right bin, we have $\mathrm{E}[L_{t,i}] = p_{t-1}^2$. As $L_t = \sum_i L_{t,i}$ is the total number of balls in the left bin after r rounds, we have

$$\begin{aligned}\mathrm{E}[L_t] &= L_{t-1} \cdot (1 - (1 - p_{t-1})^2) + (n - L_{t-1}) \cdot p_{t-1}^2 \\ &= L_{t-1} p_{t-1} \cdot (3 - 2p_{t-1}) \le L_{t-1}^2/n \cdot 3.\end{aligned}$$

Hence, again by using a Chernoff bound argument, we get

$$\Pr\left[L_t \ge \frac{9L_{t-1}^2}{2n}\right] \le \exp\left(-\frac{9L_{t-1}^2}{4n}\right),$$

which, if $L_{t-1} \ge \sqrt{\epsilon n \log(n)}$, is polynomially small in n (depending on the constant ϵ). To see that this needs $\mathcal{O}(\log\log n)$ steps, note the successive squaring in the mapping $x \mapsto 9x^2/(2n)$. Once we are at a step t with $L_t \le \sqrt{\epsilon n \log n}$, we get

$$\begin{aligned}&\Pr[L_{t+1} \ge 2C \cdot \log n] \\ &\le \binom{\sqrt{\epsilon n \log n}}{C \log n} \cdot \left(1 - \left(1 - \frac{\sqrt{\epsilon n \log n}}{n}\right)^2\right)^{C \log n} \\ &\quad + \binom{n}{C \log n} \cdot \left(\frac{\sqrt{\epsilon n \log n}}{n}\right)^{2C \log n} \\ &\le \left(\frac{e\sqrt{\epsilon n}}{C\sqrt{\log n}}\right)^{C \log n} \cdot \left(\frac{2\sqrt{\epsilon n \log n}}{n}\right)^{C \log n} \\ &\quad + \left(\frac{en}{C \log n}\right)^{C \log n} \cdot \left(\frac{\sqrt{\epsilon n \log n}}{n}\right)^{2C \log n},\end{aligned}$$

which is smaller than n^{-2} for sufficiently large C. If $L_t \le 2C \cdot \log n$ for some step t, then $\mathrm{E}[L_{t+1}] = \mathcal{O}((\log n)^2/n)$ and by Markov's inequality $\Pr[L_{t+1} \ge 1] \le \mathrm{E}[L_{t+1}] = \mathcal{O}((\log n)^2/n)$.

With adversary: As observed earlier, the adversary can only change the location of $T = \sqrt{n}$ balls at the end of each round. Hence we obtain as above that after $\mathcal{O}(\log\log n)$ steps we reach a step t with $L_t \le \sqrt{\epsilon n \log n}$. Then we know from the analysis above, that in the next step we have $L_t \le 2C \cdot \log n + \sqrt{n}$ with high probability and this will hold for polynomially many time steps with high probability. □

Case 2: $c\sqrt{n \ln n} \le \Delta_t < n/4$ *for a sufficiently large constant* c

Here, we will show the following lemma. Again, the proof is elementary and uses standard Chernoff bounds.

LEMMA 2.4. *If there is a step t_0 with $c\sqrt{n \ln n} \le \Delta_{t_0} \le n/4$ for a sufficiently large constant c, then for any $\sqrt{n}$-bounded adversary there is a step $t_1 = t_0 + \mathcal{O}(\log n)$ with $\Delta_{t_1} \ge n/4$ w.h.p.*

PROOF. Recall that $X_t = n/2 - \Delta_t$ is the number of balls in the smaller bin at step t. Furthermore, we define $\delta_t := \Delta_t/n$ and recall that by assumption, $\delta_t \in [c\sqrt{\ln n}/\sqrt{n}, 1/4]$.

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 $(1/2 - \delta_t)^2 = 1/4 - \delta_t + \delta_t^2$. Recall that $\tilde{X}_{t+1}$ is the number of balls in the smaller bin before the action of the adversary at the end of step $t+1$. Linearity of expectation gives

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

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

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

This implies that $\tilde{X}_{t+1} \le (1 - \delta/4)X_t$ w.h.p. Since the adversary can only change the location of at most $T = \sqrt{n}$ balls at the end of round $t+1$, we have w.h.p. that X_{t+1} is at most $(1 - \delta_t/4)X_t + \sqrt{n}$. Hence, w.h.p., $n/2 - \Delta_{t+1} \le (1 - \delta_t/4) \cdot (n/2 - \Delta_t) + \sqrt{n}$, and further rearranging gives that, w.h.p.,

$$\begin{aligned}\Delta_{t+1} &\ge \Delta_t + \delta_t n/8 - \delta_t \Delta_t/4 - \sqrt{n} \\ &\ge \Delta_t + \Delta_t/8 - \Delta_t/16 - \Delta_t/32 \\ &\ge (1 + 1/32)\Delta_t.\end{aligned}$$

Hence, taking the union bound over $\mathcal{O}(\log n)$ rounds, we reach a step with an imbalance of at least $n/4$ w.h.p. □

Case 3: $\Delta_t < c\sqrt{n \ln n}$

In contrast to the previous cases, the imbalance is now rather small which requires a more careful analysis.

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 2.5. *Assume no adversary is present. Let $\gamma > 0$ be any constant. Then for any $\Delta_t \ge 0$, $\Pr[\Delta_{t+1} \ge \gamma\sqrt{n}] \ge \frac{1}{\sqrt{4\pi}(1+4\gamma/\sqrt{3})} e^{-8\gamma^2/3}$, provided n is large enough.*

PROOF. For the proof of Lemma 2.5, we need the following notation. We say that a random variable Y *stochastically dominates* a random variable Z, and write $Y \succeq Z$, if $\Pr[Y \ge x] \ge \Pr[Z \ge x]$ for any x. Finally, we define the *labeled imbalance* by $\Psi_t = (R_t - L_t)/2$.

CLAIM 2.6. *For any two labeled imbalances Ψ_t and Ψ'_t with $\Psi_t \ge \Psi'_t \ge 0$ it holds that $\Psi_{t+1} \succeq \Psi'_{t+1}$.*

PROOF. We show stochastic domination for any two labeled 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'_t, 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 therefore, $R_{t+1} - L_{t+1}$ dominates $R'_{t+1} - L'_{t+1}$, which proves the claim. □

We now return to the proof of Lemma 2.5. We only need to prove it for $\Psi_t = 0$ because the general case follows from stochastic domination (see Claim 2.6). 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 $Z_1, \ldots, Z_{n/2} \in \{0,1\}$ be random variables defined as follows:

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

Then the Z_i are independent Bernoulli variables with $\Pr[Z_i = 1] = 1/4$. Analogously, for balls $n/2+1$ to n, we define the random variables $Z_{n/2+1}$ to Z_n to be 1 if the corresponding ball moves from the right bin to the left bin. We again have $\Pr[Z_i = 1] = 1/4$ for these variables. Then

$$\Psi_{t+1} = \sum_{i=1}^{n/2} Z_i - \sum_{i=n/2+1}^{n} Z_i.$$

Let $\Psi^{(1)}_{t+1} = \sum_{i=1}^{n/2} Z_i$ and $\Psi^{(2)}_{t+1} = \sum_{i=n/2+1}^{n} Z_i$. Each $\Psi^{(j)}_{t+1}$ is binomially distributed with parameters $n/2$ and $p = 1/4$. Thus, $\mathrm{E}[\Psi^{(j)}_{t+1}] = p \cdot n/2 = n/8$ and $\mathrm{V}[\Psi^{(j)}_{t+1}] = p(1-p) \cdot n/2 = (3/4) \cdot n/8$. Since $\Psi^{(1)}_{t+1}$ and $\Psi^{(2)}_{t+1}$ are stochastically independent, it holds that $\mathrm{E}[\Psi_{t+1}] = \mathrm{E}[\Psi^{(1)}_{t+1}] - \mathrm{E}[\Psi^{(2)}_{t+1}] = 0$ and $\mathrm{V}[\Psi_{t+1}] = \mathrm{V}[\Psi^{(1)}_{t+1}] - \mathrm{V}[\Psi^{(2)}_{t+1}] = 3n/16$. Let $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du$ and let $X = \sum_{i=1}^{n} X_i$ be a sum of independent random variables X_i with finite $\mu = \mathrm{E}[X]$ and $\nu = \mathrm{V}[X]$. From the Central Limit Theorem it follows for $n \to \infty$ that for any $a < b$,

$$\lim_{n \to \infty} \Pr\left[a < \frac{X - \mu}{\sqrt{\nu}} < b\right] = \Phi(b) - \Phi(a)$$

Thus, it holds for any $\gamma > 0$ that

$$\Pr[\Psi_{t+1} \geq \gamma\sqrt{n}] \geq 1 - \Phi(\sqrt{16/3}\gamma) - \varepsilon$$

where $\varepsilon \to 0$ as $n \to \infty$. For $x \geq 0$, the value of $\Phi(x)$ can be bounded as follows (see e.g., [23] p. 17 and [24] 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

$$\frac{1}{\sqrt{2\pi}(1+4\gamma/\sqrt{3})} e^{-\frac{8\gamma^2}{3}} - \varepsilon \geq \frac{1}{\sqrt{4\pi}(1+4\gamma/\sqrt{3})} e^{-\frac{8\gamma^2}{3}}$$

if n is sufficiently large, which finishes the proof. □

Finally, we prove via standard Chernoff bounds that beyond an imbalance of about $\sqrt{n}$ there is a strong drift to increase the imbalance by a constant factor.

Lemma 2.7. *If $6\sqrt{n} \leq \Delta_t \leq c\sqrt{n \log n}$, then for any $\sqrt{n}$-bounded adversary,*

$$\Pr[\Delta_{t+1} \geq (7/6)\Delta_t] \geq 1 - \exp\left(-\Theta(\Delta_t^2/n)\right).$$

Proof. In the proof of Lemma 2.4 (Equation 1) we showed that $\mathrm{E}[\tilde{X}_{t+1}] = n/2 - (3/2)\Delta_t + 2\delta^2\Delta_t$ which implies that $\mathrm{E}[\tilde{\Delta}_{t+1}] = (3/2 - 2\delta^2)\Delta_t$. As $\delta = \Delta_t/n = o(1)$ by our upper bound on Δ_t, it follows from the Chernoff bounds that

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

As $\Delta_{t+1} \geq \tilde{\Delta}_{t+1} - \sqrt{n}$, we note that $\tilde{\Delta}_{t+1} \geq (4/3)\Delta_t$ implies that $\Delta_{t+1} \geq (4/3)\Delta_t - \sqrt{n} \geq (4/3)\Delta_t - (1/6)\Delta_t \geq (7/6)\Delta_t$. □

Now we can finish the case 3.

Lemma 2.8. *If at a round t_0 we have $\Delta_{t_0} < c\sqrt{n \ln n}$ for the value of c needed by Lemma 2.4, then for any $\sqrt{n}$-bounded adversary there is a round $t_1 = t_0 + \mathcal{O}(\log n)$ with $\Delta_{t_1} \geq c\sqrt{n \ln n}$ w.h.p.*

Proof. Lemma 2.5 implies that the expected number of steps until we are in the hypothesis of Lemma 2.7 is $\mathcal{O}(1)$. That is, $\Delta_t \geq c\sqrt{n}$. Now let $\Upsilon_\tau = \lfloor \Delta_{t+\tau-1}/(c\sqrt{n}) \rfloor$ and let $q = \lfloor (n/2)/(c\sqrt{n}) \rfloor$ denote maximum value of Υ_τ, that is, the possible values of Υ_τ are in $\{0, \ldots, q\}$. To continue, we need the following technical result. Its proof is omitted due to space constraints.

Claim 2.9. *Let $(X_t)_{t=1}^{\infty}$ be a Markov Chain with state space $\{0, \ldots, q\}$ that has the following properties:*

- *there are constants $c_1 > 1$ and $c_2 > 0$, such that for any $t \in \mathbb{N}$, $\Pr[X_{t+1} \geq \min\{c_1 X_t, q\}] \geq 1 - e^{-c_2 X_t}$,*
- *$X_t = 0 \Rightarrow X_{t+1} \geq 1$ with probability c_3 which is a constant greater than 0,*

Let $c_4 > 0$ be an arbitrary constant and $T := \min\{t \in \mathbb{N} \colon X_t \geq c_4 \log q\}$. Then for every constant $c_6 > 0$ there is a constant $c_5 = c_5(c_4, c_6) > 0$ such that

$$\Pr[T \leq c_5 \cdot \log q + \log_{c_1}(c_4 \log q)] \geq 1 - q^{-c_6}.$$

By this claim, $O(\log q)$ rounds suffice to achieve $\Upsilon_\tau \geq c_4 \log q$, or $\Delta_{t+\tau-1} \geq c\sqrt{n} \cdot c_4 \log q$, w.h.p. for any constant $c_4 > 0$, which finishes the proof. □

3. MORE THAN TWO VALUES

In this section we consider the more challenging case that $|S| > 2$. We analyze the models with and without adversary separately.

3.1 Convergence without Adversary

In this section, we prove Theorem 1.1. Our proof proceeds as follows. Initially, we may have up to n non-empty bins as there can be up to n different initial values, but after just $\mathcal{O}(\log n)$ 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 median rule 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 consecutive numbers. 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 with numbers 1, 2 and 3, the second bin 2 balls, the third bin none 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 computed as follows. Ball i may either be chosen twice by a ball, or ball $j \in \{i+1,\ldots,n\}$ chooses ball i and a ball $i' \in \{1,\ldots,i-1\}$, or ball $j \in \{1,\ldots,i-1\}$ chooses ball i and a ball $i' \in \{i+1,\ldots,n\}$, or ball i chooses one ball out of $\{1,\ldots,i-1\}$ and the other out of $\{i+1,\ldots,n\}$, or ball i chooses itself and some ball $j \neq i$. This gives

$$\begin{aligned} g(i) &= n\cdot\frac{1}{n^2} + (n-i)\cdot\frac{2(i-1)}{n^2} + (i-1)\cdot\frac{2(n-i)}{n^2} + \\ &\quad 1\cdot\frac{2(n-i)(i-1)}{n^2} + 1\cdot\frac{2(n-1)}{n^2}\end{aligned}$$

Simplifying this expression gives

$$g(i) = 6\frac{(n-i)(i-1)}{n^2} + \frac{3n-2}{n^2}. \qquad (2)$$

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 and the definition of gravity, 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, $0 \le |\mathcal{H}_{t,j}| \le \Phi$. We first prove the following:

Lemma 3.1. *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.*

Proof. Assume w.l.o.g. that $j \le m_t$ (the case $j \ge m_t$ follows with identical arguments), where m_t is the median ball at round t. 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 (2), we get

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

which readily implies that $i \le n/3 + \mathcal{O}(1)$. 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 $L_{t,A} \le n/3 + \Phi + \mathcal{O}(1)$. Using the arguments from the two-bin case (Lemmas 2.3 and 2.4) we conclude that ,w.h.p., $L_{t+1,A} \le n/(3+\epsilon)$, for a constant $\epsilon > 0$. Hence by (2), every ball $l \in \mathcal{H}_{t+1,j}$ in bin j satisfies $g(l) < 4/3$ w.h.p. (provided that $\mathcal{H}_{t+1,j} \neq \emptyset$). □

On the other hand, it holds:

Lemma 3.2. *If $|\mathcal{H}_{t,j}| = \Phi$ and there is no ball in $\mathcal{H}_{t,j}$ with $g(i) < 4/3$, then $|\mathcal{H}_{t+1,j}| = \Phi$ w.h.p.*

Proof. Suppose that $\mathcal{H}_{t,j} \ge \Phi$ and there is no ball $l \in \mathcal{H}_{t+1,j}$ with $g(l) < 4/3$. Then it follows from the definition of the gravity that the expected number of balls in bin j at step $t+1$ is at least $(4/3)\cdot\Phi$. Thus, the Chernoff bounds imply that the number of balls in bin j at step $t+1$ is at least Φ w.h.p., and therefore, $\mathcal{H}_{t+1,j} \ge \Phi$. □

Using Lemma 3.1 and Lemma 3.2, we can now prove the following.

Lemma 3.3. *For any initial configuration it takes at most $\mathcal{O}(\log n)$ rounds until at least one of the following two cases holds for all bins j 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. *$|\mathcal{H}_{t,j}| = \Phi$.*

Proof. Consider an arbitrary but fixed round t. Our goal is to apply the following technical result. Its proof is similar to Claim 2.9 and omitted due to space constraints.

Claim 3.4. *Let $(X_t)_{t=1}^{\infty}$ be a Markov Chain with state space $\{0,\ldots,q\}$ that has the following properties,*

- *there are constants $c_1 > 1$ and $c_2 > 0$, such that for any $t \in \mathbb{N}$, $\Pr[X_{t+1} \ge \min\{c_1 X_t, q\}] \ge 1 - e^{-c_2 X_t}$,*
- *$X_t = 0 \Rightarrow X_{t+1} = 0$ with probability 1,*
- *$X_t = q \Rightarrow X_{t+1} = q$ with probability 1.*

Let $c_4 > 0$ be an arbitrary constant and $T := \min\{t \in \mathbb{N}\colon X_t \in \{0\} \cup \{q\}\}$. Then for every constant $c_6 > 0$ there is a constant $c_5 > 0$ such that $\Pr[T \le c_5 \log q] \ge 1 - q^{-c_6}$.

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 3.1 implies that $\mathcal{H}_{t+1,j}$ contains at least one ball l with $g(l) < 4/3$, or $\mathcal{H}_{t+1,j}$ is empty, w.h.p.
2. $|\mathcal{H}_{t,j}| = \Phi$ and all balls $i \in \mathcal{H}_{t,j}$ satisfy $g(i) \ge 4/3$. Then it follows from Lemma 3.2 that $|\mathcal{H}_{t+1,j}| = \Phi$ w.h.p.

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, the Chernoff bounds imply

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

Thus, applying Claim 3.4, we conclude that one of the two absorbing states is reached within $t_1 = \mathcal{O}(\log n)$ rounds w.h.p. □

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

Proof of Theorem 1.1: Let $t_1 = \mathcal{O}(\log n)$ be the round that satisfies Lemma 3.3. For this t_1 let $j_{\min}$ and $j_{\max}$ be the positions (w.r.t. our unique ball ordering) of the leftmost and rightmost balls in bin b_t, where bin b_t is the one containing the median ball. Note that by Lemma 3.3, the load of bin b_t is at least Φ. We proceed by a case distinction on the positions of $j_{\min}$ and $j_{\max}$.

1. $n/2 - j_{\min} \leq \Phi/2$. Let $b_t - 1$ be the left bin of bin b_t. Equation 2 implies that all heavy balls in bin $b_t - 1$ have gravity at least $4/3$. So, Lemma 3.3 implies that the load of bin $b_t - 1$ is at least Φ. Consolidate all bins from $1, \ldots, b_t - 2$ and all bins from $b_t + 1, \ldots, n$ 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 2.3 and 2.4 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, $b_t - 1$ and b_t. A final application of our two-bin analysis (Theorem 2.1) 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 in the same way as before.

3. $n/2 - j_{\min} > \Phi/2$ and $j_{max} - n/2 > \Phi/2$: In this case, it follows as in the previous cases that by Lemmas 2.3 and 2.4, all bins except bin b_t will vanish after the next $\mathcal{O}(\log n)$ rounds.

□

3.2 Convergence with Adversary

In this section we prove Theorem 1.2. First, let $m = |S|$ and assume that m is finite.

THEOREM 3.5. *For any $\sqrt{n}$-bounded adversary, it will take at most $\mathcal{O}(\log m \log\log n + \log n)$ time w.h.p. until the median rule reaches an almost stable consensus.*

PROOF. We shall use the following Chernoff bound which can be easily derived from the standard Chernoff bound for binomial random variables.

LEMMA 3.6. *Consider some fixed $0 < \delta < 1$. Suppose that $X_1, \ldots, X_n$ 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*

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

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 $\log m + 1$. Each phase i, $1 \leq i \leq \log m$ takes only $\mathcal{O}(\log\log n)$ steps in expectation, while the last phase will take $\mathcal{O}(\log n)$ steps. 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 set $S_i \subseteq \{1, \ldots, m\}$ of consecutive bins of size $|S_i| \leq m/2^i + 1$ that satisfies

$$\min\{R(S_i), L(S_i)\} \geq \frac{n}{2} + C\sqrt{n\log n}, \tag{3}$$

where $R(S_i)$ (resp. $L(S_i)$) denotes the total load of all bins that are in the set S_i or located right (resp. 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 at the end (up to $\mathcal{O}(\sqrt{n})$ balls due to the adversary) is located in S_i (which follows from applying the two-bin analysis to the sets $R(S_i)$ and $\{1, \ldots, m\} \setminus R(S_i)$ as well as $L(S_i)$ and $\{1, \ldots, m\} \setminus L(S_i)$).

Let us now prove (3) 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 2.5 and Lemma 2.7 from the two-bin analysis. 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 2.5 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 (w.h.p.), we have $\Delta_{t+1} \geq \sqrt{n}$ with constant probability. Then we apply Lemma 2.7 to obtain

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

which is at least a constant greater than zero. As this holds for any imbalance Δ_t, 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, $R(S_i) \geq \frac{n}{2} + C\sqrt{n\log n}$. Also, the proof of Lemma 2.4 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 (3).

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 3.6 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 $S = S_{\log m} = \{j, j+1\}$ with

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

Applying Lemma 2.3 and Lemma 2.4 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 2.1, we conclude that after additional $\mathcal{O}(\log n)$ rounds, we will have reached an almost stable consensus. □

As an alternative definition of m, we can also define m to be the number of legal values between v_ℓ and v_r, where v_ℓ is the $(n/2 - c\sqrt{n \log n})$-smallest and v_r is the $(n/2 + c\sqrt{n \log n})$-smallest value of the initial values for some sufficiently large constant c. Let us throw all values $v < v_\ell$ into one superbin A and all values $v > v_r$ into one superbin B. Then it follows from Lemmas 2.3 and 2.4 that after $\mathcal{O}(\log n)$ rounds, the superbins A and B run empty except for $\mathcal{O}(\sqrt{n})$ many balls, w.h.p., which implies the following lemma.

Lemma 3.7. *For any $\sqrt{n}$-bounded adversary it holds that after $\mathcal{O}(\log n)$ rounds, all processes apart from $\mathcal{O}(\sqrt{n})$ have values between v_ℓ and v_r w.h.p.*

Using the outcome of this lemma as the starting point in the analysis of Theorem 3.5, one can easily check that Theorem 3.5 is still valid when defining m to be the number of legal values between v_ℓ and v_r.

3.3 Static Adversary

Suppose that the adversary has to choose a fixed set of $\sqrt{n}$ corrupted balls throughout the execution. Since this is a special case of our T-bounded adversary, Theorem 3.5 still holds. However, the expected number of non-corrupted balls leaving the stable bin is at least $(n-\sqrt{n}) \cdot (1/\sqrt{n})^2 = 1 - o(1)$ in each round, and it can be up to $\Theta(\log n)$ with probability at least $1/n$. Thus, they still have to do some update work. To prevent any update work (i.e., *all* non-corrupted balls stay at the stable bin for at least a polynomial number of rounds w.h.p.), a simple extension of the median rule, called the *careful median rule*, suffices:

Each process i executes the median rule as before but in addition to this maintains a *stable value* sv_i and keeps track of the last k outcomes of the median rule for some constant $k \geq 3$. Whenever the majority of the last k outcomes agrees on a single value, say v, sv_i is set to v.

Our goal is to reach a consensus for the sv_i values that holds as long as possible.

Theorem 3.8. *For any static $\sqrt{n}$-bounded adversary, the careful median rule needs at most $\mathcal{O}(\log m \, \log\log n + \log n)$ rounds to reach a stable consensus for all non-corrupted processes that holds for poly(n) many steps w.h.p.*

Proof. Theorem 3.5 implies that after $\mathcal{O}(\log m \log\log n + \log n)$ rounds the (standard median rule) values v_i of the honest processes form an almost stable consensus w.h.p., say, on value v. Now, focus on any honest process i. Once an almost stable consensus is reached, the probability that v_i deviates from v at the end of round t is bounded by $\mathcal{O}(T/n)$, which is the probability that i contacts one of the at most $\mathcal{O}(T)$ processes that deviate from the consensus, even if i itself deviates from the consensus at the beginning of t. Since this upper bound on the probability holds independently for each round, it follows that the probability that at least $k/2$ of the k last values of i deviate from v is at most

$$\binom{k}{k/2} \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)^{k/2} = 2^k \cdot \mathcal{O}\left(\frac{1}{n}\right)^{k/4} = \mathcal{O}\left(\frac{1}{n}\right)^{k/4}$$

for a constant k which implies the theorem. □

4. AVERAGE CASE ANALYSIS

In this section, we investigate the case where all n balls are initially put independently and uniformly at random into m bins. Without the adversary, we get the following result.

Theorem 4.1. *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 w.h.p. after the following time:*

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

Intuitively, the reason for this dichotomy is that for odd m there is already a large imbalance at the beginning when we consider all balls that are in bins left to the middle bin versus all the remaining balls. Hence we reach in just $\mathcal{O}(\log m)$ an imbalance of $\Omega(n)$, for which we have shown in Lemma 2.3 that already $\mathcal{O}(\log\log n)$ further steps are enough reach a stable consensus. However, if m is even, then there is only a relatively small imbalance at the beginning and it takes $\Omega(\log n)$ rounds to reach a sufficiently large imbalance.

Let us now analyze the adversarial model.

Theorem 4.2. *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 time:*

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

5. CONCLUSIONS

In this paper we presented a surprisingly simple, efficient and robust consensus mechanism demonstrating the power of two choices. While we were able to prove a tight time bound for this algorithm in the non-adversarial case, the time bound for the adversarial case is not optimal yet (it is $\mathcal{O}(\log n \log\log n)$ instead of the suspected $\mathcal{O}(\log n)$), so further work is needed. Also, it is open whether lightweight self-stabilizing consensus mechanisms exist beyond $\mathcal{O}(\sqrt{n \log n})$ adversarial processes.

Acknowledgments

We would like to thank Michael Bender for suggesting a related problem that initiated this research.

6. REFERENCES

[1] D. Angluin, J. Aspnes, and D. Eisenstat. A simple population protocol for fast robust approximate majority. In *Proc. of the 21st Int. Symposium on Distributed Computing (DISC)*, pages 20–32, 2007.

[2] 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.

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

[4] 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.

[5] 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.

[6] 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.

[7] H. Attiya and K. Censor. Tight bounds for asynchronous randomized consensus. *Journal of the ACM*, 55(5):1–26, 2008.

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

[9] Y. Azar, A. Broder, A. Karlin, and E. Upfal. Balanced allocations. *SIAM Journal on Computing*, 29(1):180–200, 1999.

[10] 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.

[11] M. Ben-Or, E. Pavlov, and V. Vaikuntanathan. Byzantine agreement in the full-information model in $\mathcal{O}(\log n)$ 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 25th ACM Symp. on Theory of Computing (STOC)*, pages 42–51, 1993.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] S. Dolev, R. Kat, and E. Schiller. When consensus meets self-stabilization. In *Proc. of the 10th International Conference on Principle of Distributed Systems (OPODIS)*, pages 45–63, 2006.

[18] R. Elsässer 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.

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

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

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

[22] S. Gilbert and D. Kowalski. Distributed agreement with optimal communication complexity. In *Proc. of the 21st ACM Symp. on Discrete Algorithms (SODA)*, pages 965–977, 2010.

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

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

[25] 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.

[26] 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.

[27] 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.

[28] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proc. of the 44 IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 482–491, 2003.

[29] V. King and J. Saia. Breaking the $O(n^2)$ bit barrier: Scalable Byzantine agreement with an adaptive adversary. In *Proc. of the 29th ACM Symp. on Principles of Distributed Computing (PODC)*, pages 420–429, 2010.

[30] V. King, J. Saia, V. Sanwalani, and E. Vee. Towards secure and scalable computation in peer-to-peer networks. In *Proc. of the 47th IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 87–98, 2006.

[31] F. Kuhn, T. Locher, and R. Wattenhofer. Tight bounds for distributed selection. In *Proc. of the 19th ACM Symp. on Parallel Algorithms and Architectures (SPAA)*, pages 145–153, 2007.

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

[33] B. Patt-Shamir. A note on efficient aggregate queries in sensor networks. *Theoretical Computer Science*, 370(1–3):254–264, 2007.

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

[35] N. Shrivastava, C. Buragohain, D. Agrawal, and S. Suri. Medians and beyond: new aggregation techniques for sensor networks. In *Proc. of the 2nd Intl. Conference on Embedded Networked Sensor Systems (SenSys)*, pages 239–249, 2004.