Lower Bounds on Expected Round Complexity of Byzantine Agreement Protocols

Working Draft: Please Do Not Distribute

Ran Cohen*¶ Iftach Haitner†¶ Nikolaos Makriyannis‡¶ Matan Orland§¶

December 17, 2018

Abstract

Keywords: Byzantine agreement; lower bound; round complexity.

*MIT and Northeastern University. E-mail: rancohen@mit.edu.

†School of Computer Science, Tel Aviv University. E-mail: iftachh@cs.tau.ac.il. Member of the Israeli Center of Research Excellence in Algorithms (ICORE) and the Check Point Institute for Information Security.

‡School of Computer Science, Tel Aviv University. E-mail: n.makriyannis@gmail.com.

§School of Computer Science, Tel Aviv University. E-mail: matanorland@mail.tau.ac.il.

¶Research supported by ERC starting grant 638121.

Contents

1 Preliminaries

1.1 Notations

We use calligraphic letters to denote sets, uppercase for random variables, lowercase for values, boldface for vectors, and sans-serif (e.g., A) for algorithms (i.e., Turing Machines). For $n \in \mathbb{N}$, let $[n] = \{1, \cdots, n\}$. Let poly denote the set all positive polynomials and let PPT denote a probabilistic algorithm that runs in *strictly* polynomial time. A function $\nu \colon \mathbb{N} \mapsto [0,1]$ is *negligible*, denoted $\nu(\kappa) = \mathrm{neg}(\kappa)$, if $\nu(\kappa) < 1/p(\kappa)$ for every $p \in \mathrm{poly}$ and large enough κ. The statistical distance between two random variables X and Y over a finite set $\mathcal{U}$, denoted $\mathrm{SD}(X,Y)$, is defined as $\frac{1}{2} \cdot \sum_{u \in \mathcal{U}} |\Pr[X=u] - \Pr[Y=u]|$. Given a random variable X, we write $x \leftarrow X$ to indicate that x is selected according to X.

Two distribution ensembles $X = \{X(a,\kappa)\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ and $Y = \{Y(a,\kappa)\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ are computationally indistinguishable (denoted $X \equiv_c Y$) if for every non-uniform polynomial-time distinguisher A there exists a function $\nu(\kappa) = \mathrm{neg}(\kappa)$, such that for every $a \in \{0,1\}^*$ and $\kappa \in \mathbb{N}$

$$|\Pr[\mathsf{A}(X(a,\kappa), 1^\kappa) = 1] - \Pr[\mathsf{A}(Y(a,\kappa), 1^\kappa) = 1]| \leq \nu(\kappa).$$

The distribution ensembles X and Y are statistically close (denoted $X \equiv_s Y$) if $\mathrm{SD}(X,Y) \leq \nu(\kappa)$ for a negligible function ν.

1.2 Byzantine agreement

Definition 1.1 (Byzantine Agreement)**.** *An (n,t,α,β)-Byzantine Agreement (BA) is an n-party protocol in which each party begins with a single private input bit and must output a single bit under the following conditions:*

- *(t,α)-agreement: In the presence of no more than t corrupt parties all honest parties output the same bit with probability at least $1-\alpha$.*
- *(t,β)-validity: In the presence of no more than t corrupt parties, if all honest parties began with input bit b, then they also output b with probability at least $1-\beta$.*

[**Matan's Note: This should be a definition of its own**] *We say that a BA has (t,δ)-two-round-halting if in the presence of t corrupt parties the probability of all honest parties to halt after two rounds is at least δ.*

2 Lower Bounds for 5-Party 2-Reselient BA

In what follows we will prove some lower bounds on the round complexity of Byzantine agreements, mostly working towards the following conjecture:

Conjecture 2.1. *Let Π be an n-party $2n/5$-resilient Byzantine Agreement. Π's expected round complexity is at least three.*

2.1 Honest Agreement

We start this section with an introduction of a very simple yet surprisingly powerful lemma, which simply states that if a BA protocol Π is *t-resilient* , then in an *honest* execution of Π in which the input vector is unbalanced enough (exactly how much is "enough" will be clearly defined) the output bit can be determined. We define the lemma and summarily prove it:

Lemma 2.2 (Honest Agreement Lemma)**.** *Let* Π *be an* n-party,t-resilient *Byzantine agreement, and let* w *be the hamming weight of the input vector. The following two conditions hold in an* honest *execution of* Π *in which* agreement*is reached:* [**Nikos's Note: w.p.** $1-\beta$]

1. *If* $w \geq n-t$ *then all honest parties output* 1*.*
2. *If* $w \leq t$ *then all honest parties output* 0*.*

Proof. We assume the first case, and the second will follow by symmetry. Let σ_1 be an honest execution of Π, and let $\tilde{P} \subset P$ such that $\left|\tilde{P}\right| \leq t$ be the set of parties with input bit 0. Consider also σ_2, an execution of Π in which all parties have input bits 1, an adversary controls all parties in $\tilde{P}$ and all adversary-controlled parties behave as honest parties with input bit 0. Clearly the views of each honest party in σ_1, σ_2 are equivalent. Thus, the outputs of all honest parties in both executions are equivalent. From *validity*all parties in σ_2 output 1, and so all parties in σ_1 output 1. □

2.2 Public Coin

[**Ran's Note: is the model defined?**]

Theorem 2.3. *Let* Π *be an* $5t$*-party* $2t$*-resilient BA protocol in the public-coin model. Then, the expected round complexity of* Π *is at least* 3*.*

To prove Theorem 2.3 we first prove a simpler lemma.

[**Ran's Note: changed the claim to a lemma**] **Added(Ran):** Write a few words describing what is going to happen. The proof follows the spirit of [1], and changes an honest execution on inputs $(0,0,1,1,1)$, in which the output must be 1, into an honest execution on inputs $(0,0,0,1,1)$, in which the output must be 0. Similarly to [1], this is done by a sequence of hybrids, where P_3 changes its messages one by one, however, special care must be taken... [**Ran's Note: again, I think that a negligible** ϵ **will simplify the statement and the proof**]

Proof. We assume towards a contradiction that $\delta > \epsilon$. We, again, introduce a few different scenarios: **Added(Ran):** We start by considering an execution on inputs $(0,0,1,1,1)$ for the parties where a corrupted party P_3 plays towards P_2 according to $x_3 = 0$ and towards all other parties according to $x_3 = 1$. We show that if P_1 halts after 2 rounds then the common output remains 1.

- **Scenario 1**: The input vector for the protocol is $(0,0,1,1,1)$ and P_3^* is corrupt. The adversarial strategy of P_3^* is to play toward P_2 according to the next-message function defined by Π for input bit $x_2 = 0$ [**Ran's Note: this is what we mean by play honestly according to input** $x_2 = 0$], and according to the next-message function for input bit $x_2 = 1$ with all other parties.

- **Scenario 1.1**: The input vector for the protocol is $(0,0,1,1,1)$. **Added(Ran):** Parties P_2^* and P_3^* are corrupt. The adversarial strategy of P_3^* is to play honestly. The adversarial strategy of P_2^* is to play honestly in round one, in round two send P_1 a message corresponding to the next-message function defined in Π had P_3^* sent P_2^* a message corresponding to input bit 0 in round one and send P_4 and P_5 a message corresponding to the next-message function defined in Π had P_3^* sent P_2^* a message corresponding to input bit 1 in round one. P_2^* then halts after round two. We note here that although P_3^*'s strategy [**Ran's Note: try not to use P_3^*'s, writing "the strategy of P_3^*" is easier to read**] is to play honestly, the adversary needs control over P_3^*, since P_2^* must send messages corresponding to different (possibly signed) messages of P_3^*. [**Ran's Note: maybe define $m_{i\to j}(b)$ as the first round message P_i sends to P_j, when P_i plays honestly according to input $x_i = b$?**]

- **Scenario 1.2**: The input vector for the protocol is $(0,0,1,1,1)$. **Added(Ran):** Parties P_1^* and P_2^* are corrupt. Their adversarial strategy is to play *honestly* in the first two rounds and then halt.

Claim 2.4. *In an execution of Scenario 1 in which* agreement*is reached and* P_1 *halts after two rounds of communication all honest parties output* 1*.*

Proof. We first note that in Scenario 1.2 if *agreement*is reached, then all honest parties output 1. This stems from *validity*of Π. Let $\textsc{view}_i^j$ **Added(Ran):** and $\textsc{output}_i^j$ be the random variables corresponding to the view and output of party P_i in Scenario j, respectively, given that P_1 (or P_1^*) halts after two rounds of communication and *agreement*is reached. We first [**Ran's Note: don't overuse "first"**] note that $\textsc{view}_1^1 \equiv \textsc{view}_1^{1.1}$. This can easily be seen by accounting for all incoming messages and randomness of P_1, thus $\textsc{output}_1^1 \equiv \textsc{output}_1^{1.1}$. Given that, by assumption, agreement is reached in both scenarios, we conclude that $\textsc{output}_5^1 \equiv \textsc{output}_5^{1.1}$. Again, by accounting for all incoming messages to P_5 in Scenarios 1.1 and 1.2, we can easily see that $\textsc{view}_5^{1.1} \equiv \textsc{view}_5^{1.2}$. Since $\textsc{output}_5^1 \equiv 1$ [**Ran's Note: $\textsc{output}_5^{1.2}$?**], we have by chain rule [**Ran's Note: ??**] that $\textsc{output}_5^1 \equiv 1$, and by assumption all honest parties in Scenario 1 output 1, which gives us the claim. □

Added(Ran): We continue by showing in the above setting, if corrupted P_3 plays according to $x_3 = 0$ also towards P_4 (in addition to P_2), then the common output remains 1.

- **Scenario 2:** The input vector for the protocol is $(0,0,1,1,1)$ and P_3^* is corrupt. The adversarial strategy of P_3^* is to play according to the next-message function defined by Π for input bit 0 toward P_2 and P_4, and according to the next-message function for input bit 1 with all other parties.

- **Scenario 2.1:** The input vector for the protocol is $(0,0,1,1,1)$. Parties P_3^* and P_4^* are corrupt. The adversarial strategy of P_3^* is to play according to the next-message function defined by Π for input bit 0 toward P_2. The adversarial strategy of P_4^* is to play honestly in round one, and in round two send P_1 a message corresponding to the next-message function defined in Π had P_3^* sent P_4^* a message corresponding to input bit 0 in round one and send P_2 and P_5 a message corresponding to the next-message function defined in Π had P_3^* sent P_4^* a message corresponding to input bit 1 in round one. P_4^* then plays honestly. [**Ran's Note: ?**]

- **Scenario 2.2:** The input vector for the protocol is $(0, 0, 1, 1, 1)$. Parties P_3^* and P_4^* are corrupt. The adversarial strategy of P_3^* is to play according to the next-message function defined by Π for input bit 0 with P_2. The adversarial strategy of P_4^* is to play honestly in round one, and in round two send all honest parties messages corresponding to the next-message function defined in Π had P_3^* sent P_4^* a message corresponding to input bit 1 in round one. P_4^* then plays honestly. [**Ran's Note: ?**]

Claim 2.5. *In an execution of Scenario 2 in which* agreement*is reached and* P_1 *halts after two rounds of communication, and given that on the same randomness* P_1 *halts in an execution of Scenario 1, all honest parties output* 1.

Proof. We first set the randomness for all following discussed executions to one for which in both Scenarios 1 and 2 *agreement*is reached and also P_1 halts after two rounds. A simple observation would be that the view of P_5 in Scenario 2 is identical to the view of P_5 in Scenario 2.1. This is due to the fact that the only message different in the two executions is the one sent from P_4 (P_4^* in Scenario 2.1) to P_1, and by assumption P_1 halts without sending any messages after that message had been received. Thus, in both executions P_5 will output the same bit. We also note that the view of P_1 in an execution of Scenario 2.2 is identical to the view of P_1 in an execution of Scenario 1, given the aforementioned set randomness, since the randomness is identical and all incoming messages are identical as a result of the randomness and adversarial strategy. Thus, from Claim 2.4, and from the assumption about *agreement*[**Ran's Note: this assumption is annoying. If you assume it all the time there is no need for a non-negligible ϵ. Removing this assumption will shorten the proof and simplify it**], P_5 will output 1 in Scenario 2.2. Our final link in the chain would be to notice that, given all assumptions, the view of P_1 in the execution of Scenario 2.2 is identical to its view in the execution of Scenario 2.1. Again, since this means that P_1 will have the same output, and we have the claim. □

Added(Ran): In the next scenario, P_3 plays according to $x_3 = 0$ also towards P_5 (in addition to P_2 and P_2); the common output remains 1.

- **Scenario 3:** The input vector for the protocol is $(0, 0, 1, 1, 1)$ and P_3^* is corrupt. The adversarial strategy of P_3^* is to play according to the next-message function defined by Π for input bit 0 with $\mathsf{P}_2, \mathsf{P}_4$, and P_5, and according to the next-message function for input bit 1 with P_1.

Claim 2.6. *In an execution of Scenario 3 in which* agreement*is reached and* P_1 *halts after two rounds of communication, and given that on the same randomness* P_1 *halts in executions of Scenarios 1 and 2, all honest parties output* 1.

The proof of Claim 2.6 follows in similar lines to proof of Claim 2.5.

- **Scenario 4:** The input vector for the protocol is $(0, 0, 1, 1, 1)$ and P_3^* is corrupt. The adversarial strategy of P_3^* is to play according to the next-message function defined by Π for input bit 0 toward all other parties.

Claim 2.7. *In an execution of Scenario 4 in which* agreement*is reached and* P_5 *halts after two rounds of communication, and given that on the same randomness* P_1 *halts in executions of Scenarios 1, 2, and 3, all honest parties output* 1.

The proof of Claim 2.7 follows in similar lines to proof of Claim 2.5.

Next, we claim that if *agreement*is reached, then the conditions for Claims 2.4 to 2.7 must coexist.

We denote by $\tilde{b} = (b_1, b_2, b_3, b_4)$ the vector of bits used by P_3^* to determine its messages, where b_i is the bit used for the messages to party P_i. In the non-adaptive public-coin model all randomness is revealed as part of a common reference string, and so a corrupt party may only deviate from Π by representing a different input bit to other parties. P_3^*'s first round messages are then determined by this choice of input-vector and the protocol's randomness.

Proof. Let $\mathcal{R}$ be the space of randomness. Let $\mathcal{R}^1, \mathcal{R}^2 \subseteq \mathcal{R}$ be two randomness sub-spaces, and let $\tilde{b}^1, \tilde{b}^2$ be two input-vectors, such that for $i \in \{1, 2\}$ when Π's randomness is $r \in \mathcal{R}^i$ and P_3^* sends messages according to $\tilde{b}^i$, at least one party (any party) does not halt after the second round. We assume towards a contradiction that

$$\frac{|\mathcal{R}^1 \cup \mathcal{R}^2|}{|R|} > 1 - \delta.$$

We describe an adversary controlling P_3^* that upon seeing the CRS r sends messages according to bit vector $\tilde{b}$ according to the following rule:

$$\tilde{b} = \begin{cases} \tilde{b}^1, & r \in \mathcal{R}^1 \\ \tilde{b}^2, & r \notin \mathcal{R}^1. \end{cases}$$

Observe that by assumption, given this adversarial behavior there exists some honest party, not necessarily the same party, which does not halt after two rounds for a subspace of randomness of size at least $|\mathcal{R}^1 \cup \mathcal{R}^2|$. Thus, with probability larger than $1 - \delta$ some honest party continues execution after two rounds of communication. This, in contradiction to the assumption that all honest parties halt after two rounds with probability at least δ, regardless of adversarial strategy. We conclude the claim. □

We can now complete the proof of **??**. Let σ be an *honest* execution of Π with input vector $(0, 0, 0, 1, 1)$ in which all parties halt after two rounds. Since $\delta > \epsilon$ there exists such an execution. Since *agreement*is reached, all honest parties output 0, by Lemma 2.2. By **??**, all honest parties must halt after two rounds on executions of Scenarios 1, 2, 3, and 4 on the same randomness. Thus, for an execution σ' with the same randomness of σ, conditions for Claim 2.7 exist, and P_1 will halt after two rounds and output 0. Accounting for all of P_1's incoming messages, input and randomness, it is easy to see that its view is identical in both σ and σ', yet in those executions it outputs 0 and 1, respectively. This is a contradiction, which gives us **??**.

□

[Ran's Note: please adjust the rest like previous section]

Corollary 2.8. *Let* Π *be a* 5*-party,*2*-resilient Byzantine agreement in the public coin model.* Π*'s expected round complexity is at least three.*

Proof. By definition we have that for Π, $\epsilon = 0$, and from **??** we have that δ, the probability with which all parties halt in Π after one round is smaller or equal to ϵ. We conclude that $\delta = 0$, and the expected round complexity of Π must be at least two. □

We can now prove Theorem 2.3 as a corollary of **??**.

Proof. We assume towards a contradiction that Π has expected round complexity less than three. We construct Π' which is a 5-*party*,2-*resilient* Byzantine agreement, without changing the round complexity, thus reaching a contradiction with Corollary 2.8. In Π' each party, upon receiving input bit b, will simply simulate the behavior of $n/5$ parties in Π, all starting with input bit b. First, we note that an adversary controlling two parties in Π', will control $2n/5$ parties of the simulated version of Π, and since Π is $\lceil 2n/5 \rceil$-*resilient*, *agreement*is guaranteed. Next, we note that if *validity*holds in Π then it also holds in Π' since if all input bits to Π' were identical, so would they be in the simulation of Π by parties of Π', by construction. Since all parties simply execute Π the expected round complexity of Π' is equivalent to that of Π. We have that Π' is a 5-*party*,2-*resilient* Byzantine agreement with expected round complexity less than three, in contradiction to Corollary 2.8. □

2.3 Extending to Adaptive Public Coins

We next extend our model to have some more randomness in it. Specifically, part of the CRS is devoted to selecting a party which can flip extra coins in round two and distribute them. We call this model the CRS-MC (MC for Magic Coin) model. We call the party who is chosen to flip coins in round two the leader. We denote the CRS space as $(\mathcal{R}, i)$, where $\mathcal{R}$ is some random string and $i \in [n]$. We further denote by $\mathcal{C}$ the space of second round leader coins. We give another caveat in the form of restrictions on the halting decision of honest parties in Π. [**Matan's Note: Need to add restrictions on halting function**]

Theorem 2.9. *Let Π be an n-party $2n/5$-resilient Byzantine agreement in the CRS-MC model. Then other than with negligible probability Π's expected round complexity is at least three.*

Proof. We prove Theorem 2.9 in much the same way we proved Theorem 2.3, as first establishing a lemma about the simpler 5-*party*,2-*resilient* Byzantine agreement case.

Proving **??** is very similar to proving **??**, in that the witness chain created in Claims 2.4 to 2.7 holds for this model as well, as the fact that the protocol was in the CRS (or public coin) model did not play a part in the statement or proof of these claims. The only claim left to prove is a version of **??**, with its model adjusted to the CRS-MC model. This results in a constant fraction loss in probability, which is nullified in $o(1)$-sound protocols.

Proof. We first note that if the second round coins do not influence honest-party halting, then there is no difference between the CRS-MC model and the CRS model in regards to this claim. Let $\mathcal{R}$ be the space of randomness. We now claim that for any randomness $r \in \mathcal{R}$ and adversarial strategy for which an honest party does not halt after two rounds with positive probability there exists $c \in \mathcal{C}$ for which at least one honest party does not halt after two rounds. If this were not true, then simply setting the randomness to r and the coins to c would result in a deterministic two-round BA. Let $\mathcal{R}^1, \mathcal{R}^2 \subseteq \mathcal{R}$ be two randomness sub-spaces, and let $(\tilde{b}^1, c_1), (\tilde{b}^2, c_2)$ be two input-vector and second-round randomness string pairs, such that for $j \in \{1, 2\}$ when Π's randomness is $r \in \mathcal{R}^i$ and P_3^* sends messages according to $(\tilde{b}^i, c_i)$, at least one party (any party) does not halt after the second round. We assume towards a contradiction that

$$\frac{|\mathcal{R}^1 \cup \mathcal{R}^2|}{|R|} > 1 - \delta.$$

We describe an adversary controlling P_3^* that upon seeing the CRS r sends messages according to bit vector $\tilde{b}$ and round-two randomness c according to the following rule:

$$(\tilde{b}, c) = \begin{cases} (\tilde{b}^1, c_1), & r \in \mathcal{R}^1 \\ (\tilde{b}^2, c_2) & r \notin \mathcal{R}^1. \end{cases}$$

Observe that by assumption, given this adversarial behavior there exists some honest party, not necessarily the same party, which does not halt after two rounds for a subspace of randomness of size at least $|\mathcal{R}^1 \cup \mathcal{R}^2|$. Thus, with probability larger than $1 - \delta$ some honest party continues execution after two rounds of communication. This, in contradiction to the assumption that all honest parties halt after two rounds with probability at least δ, regardless of adversarial strategy. We conclude the claim. □

We can now prove **??**, by assuming towards a contradiction that $\delta > 5\epsilon$. On the surface it would appear as though we cannot make the same claims made for the proof of **??**, as the δ-fraction of coins for which all parties halt after two rounds may be disjoint from the set of coins for which P_3^* is the leader. Still we can claim that since the CRS is input-independent, and thus there must exist a party such that when it is the leader, for at least a δ-fraction of coins all parties halt after two rounds, in which case we simply substitute the input-vector and corrupt parties in all of our previous claims to maximize the adversary's probability of attack. We have that for a at least a $\delta/5$ fraction of coins all parties halt after two rounds, and since by assumption there exist such coins for which *agreement*holds, and we conclude the same contradiction as in **??**. □

[**Matan's Note: maybe define good and bad events**]

Proof. The proof of Theorem 2.9 follows the same lines as the proof of Theorem 2.3 with the adage that in a simulation of a larger (as in more participants) protocol by five parties, there is still a 0.2 probability of the leader being under the control of an adversary. □

We can now extend this case to have more coins. In the CRS-PC model the adversary first selects t parties to corrupt, a CRS is then revealed, and in the second round all parties have fresh randomness which they must send with their second round messages. **??** naturally extends to this model, if we consider halting decision which is input-independent. Since each honest party must halt on at least a δ-fraction of randomness for each input, there must exist a party which can influence halting for at least a $\delta/5$-fraction of second-round coins. We can simply take this party to be the corrupt party, and adjust our input vector accordingly, as the halting decision is independent of the input vector.

[**Matan's Note: There is a problem with extending this to Theorem 2.9 - in a simulation now there must exist a party, all of whose simulated parties have influence - this is strange.**]

2.4 Deterministic First Round, MC second round

We next consider the following model. An adversary first chooses t parties to corrupt, then the protocol is run, with the first round being deterministic, and in the second round each party can flip a single coin, and send it (along with a second round message) to all other parties. [**Ran's Note: parties also have CRS and signing oracle, right?**]

Theorem 2.10. *Let* Π *be a* $5n$*-party* $2n$*-resilient BA in the above model. Except for negligible probability,* Π*'s expected round complexity is* 3*.*

[**Ran's Note: I'm not sure the multiparty case works, let's start with 5-party protocols**]

We prove Theorem 2.10 by proving a stronger lemma first. We show that given a non-negligible probability for a 5-party 2-resilient Byzantine agreement to complete after two rounds of communication, there must be non-negligible probability of disagreement. [**Matan's Note: reformulate**] [**Ran's Note: I don't follow...**]

Lemma 2.11. [**Ran's Note: missing lemma?**]

Without loss of generality we will assume throughout the proof of Lemma 2.11, that, as a corrupt party, P_5^* can choose the best second-round randomness it wants. [**Ran's Note: we're jumping into the deep water. The should be an explanation of what the attack is all about. Start with** (00111) **and change to** (00110) **one message at a time, and so on**] Since in any case, at least one coin is better for it (be its goal attacking correctness or round complexity), it can simply choose that coin. Thus, we can define a second round randomness vector received by party P_i as $\boldsymbol{r}^i = (r_1^i, r_2^i, r_3^i, r_4^i)$, with r_5^i being a constant, and therefore ignored. We will also denote by $\boldsymbol{b}_i \in \{0,1\}^4$ the vector of first round messages a corrupt P_5^* sends out during the first round, with b_i^j being the bit it chooses to represent towards P_j. We define $\boldsymbol{b}_i \in \{0,1\}^4$ such that the first i bits of $\boldsymbol{b}_i$ are 1, and the rest are 0 (e.g., $\boldsymbol{b}_3 = (1110)$). [**Ran's Note: be consistent between $\boldsymbol{b}_i$ and $\boldsymbol{r}^i$**] We will restrict our adversaries to dishonest behavior only in the first two rounds, in which only P_5^* is dishonest in round one, and another party is dishonest in round two in respect to its randomness. Thus, after two rounds of communication, the view of an honest party P_i is well defined by a pair $(\boldsymbol{b}_i, \boldsymbol{r}_i)$. [**Ran's Note: change of notation for $\boldsymbol{r}_i$**] [**Ran's Note: what about the second round message of the second corrupted party? We can assume parties echo their view from round 1**] Thus, we will refer to $\Pi(\boldsymbol{b}_i, \boldsymbol{r}_i)$ as an execution with input vector $(0,0,1,1,*)$ with the appropriate $\boldsymbol{b}_i, \boldsymbol{r}_i$ pair. We will denote with $\text{OUT}(\Pi(\boldsymbol{b}_i, \boldsymbol{r}_i))$ and $\text{VIEW}_{\mathsf{P}_i}(\Pi(\boldsymbol{b}_i, \boldsymbol{r}_i))$ the output of all honest parties and the view of P_i in an execution of Π with $\boldsymbol{b}_i, \boldsymbol{r}_i$, respectively. [**Ran's Note: over the random coins of honest parties**] We now define the concepts of *robustness* and *double robustness* which will be used throughout this proof.

Definition 2.12. *(Robustness) A pair* $(\boldsymbol{b}, \boldsymbol{r})$ *is* robust *for* P_j *if the following conditions hold:*

1. *All honest parties halt after two rounds in* $\Pi(\boldsymbol{b}, \boldsymbol{r})$*.*
2. *All honest parties halt after two rounds in* $\Pi(\boldsymbol{b}, \boldsymbol{r} \oplus \boldsymbol{e}_j)$*.*

where $\boldsymbol{e}_j \in \{0,1\}^4$ *is the vector with* 1 *in the* j*'th entry and* 0 *in all other entries.* [**Ran's Note: this should move to preliminaries**]

Definition 2.13. *(Doubly-Robustness) A pair* $(\boldsymbol{b}, \boldsymbol{r})$ *is* doubly robust *for* P_j *if the following conditions hold:*

1. *All honest parties halt after two rounds in* $\Pi(\boldsymbol{b}, \boldsymbol{r})$*.*
2. *All honest parties halt after two rounds in* $\Pi(\boldsymbol{b}, \boldsymbol{r} \oplus \boldsymbol{e}_j)$*.*
3. *All honest parties halt after two rounds in* $\Pi(\boldsymbol{b} \oplus \boldsymbol{e}_j, \boldsymbol{r})$*.*

4. *All honest parties halt after two rounds in* $\Pi(\boldsymbol{b} \oplus \boldsymbol{e}_j, \boldsymbol{r} \oplus \boldsymbol{e}_j)$.

We make three claims regarding the existence of doubly-robust [**Ran's Note: no hyphen**] pairs, and their expected output.

Claim 2.14. *If a pair* $(\boldsymbol{b}, \boldsymbol{r})$ *is doubly robust with respect to* P_j*, then* [**Ran's Note: these are distribution, they should be identically distributed**]

$$\text{OUT}(\Pi(\boldsymbol{b}, \boldsymbol{r})) = \text{OUT}(\Pi(\boldsymbol{b}, \boldsymbol{r} \oplus \boldsymbol{e}_j)) = \text{OUT}(\Pi(\boldsymbol{b} \oplus \boldsymbol{e}_j, \boldsymbol{r})) = \text{OUT}(\Pi(\boldsymbol{b} \oplus \boldsymbol{e}_j, \boldsymbol{r} \oplus \boldsymbol{e}_j)).$$

Proof. Suppose this is not the case, then a corrupt P_j can simply choose to represent different pairs to different parties. By the doubly robust property, all honest parties will halt after two rounds, and they will output different values, in contradiction to agreement. □

Claim 2.15. *Let* Π *be a* 5*-party* 2*-resilient BA such that for every adversary controlling no more than* 2 *parties, all honest parties halt with non-negligible probability after two rounds of communication. Then, for every* $i \in \{0,1,2,3,4\}$ *there exist* $\boldsymbol{r}_i \in \{0,1\}^4$ *and* P_j *with* $j \neq i$ [**Matan's Note: Change** i, j **notation to simplify (this is currently mismatching)**]*, such that* $(\boldsymbol{b}_i, \boldsymbol{r}_i)$ *is doubly robust for* P_j*.*

Proof. Assume towards a contradiction that there exists $i \in \{0,1,2,3,4\}$ such that for every $\boldsymbol{r}_i \in \{0,1\}^4$ and every party P_j, where $j \neq i$, it holds that $(\boldsymbol{b}_i, \boldsymbol{r}_i)$ is *not* doubly robust for P_j. Stated differently, party P_j [**Matan's Note: exists k such that party k does not halt**] does not halt after two rounds of communication when it's view is distributed as one of $\text{VIEW}_{\mathsf{P}_j}(\Pi(\boldsymbol{b}_i, \boldsymbol{r}_i))$, $\text{VIEW}_{\mathsf{P}_j}(\Pi(\boldsymbol{b}_i \oplus \boldsymbol{e}_i, \boldsymbol{r}_i))$, $\text{VIEW}_{\mathsf{P}_j}(\Pi(\boldsymbol{b}_i, \boldsymbol{r}_i \oplus \boldsymbol{e}_i))$, or $\text{VIEW}_{\mathsf{P}_j}(\Pi(\boldsymbol{b}_i \oplus \boldsymbol{e}_i, \boldsymbol{r}_i \oplus \boldsymbol{e}_i))$. [**Ran's Note: The proof probably needs adjustments**]

We can now simply define the following adversarial strategy. The adversary corrupts party P_5 and P_k. [**Matan's Note: k?!?!?!?!**] Party P_5^* sends the message vector according to $\boldsymbol{b}_i$. Given second round coins from all honest parties, the adversary now chooses which of the four options for $(\boldsymbol{b}, \boldsymbol{r})$ to present to P_j [**Matan's Note: should be** P_{k_j} **or** $k = k(j)$] as a second round message. By assumption, P_j will continue execution. Thus, the protocol will never halt after two rounds, in contradiction to what is assumed in the claim. □

Notice that by definition if $(\boldsymbol{b}_i, \boldsymbol{r}_i)$ is doubly-robust [**Ran's Note: what is** $\boldsymbol{r}_i$**?**] with respect to P_j then also $(\boldsymbol{b}_i \oplus e_j, \boldsymbol{r}_i)$ is doubly-robust with respect to P_j.

Claim 2.16. *Let* r_1, r_2 *be two random strings and* $j_1 \neq j_2 \in [4]$ *two indices such that* $(b, r_1), (b, r_2)$ *are doubly robust with respect to parties* $\mathsf{P}_{j_1}, \mathsf{P}_{j_2}$*, respectively. Then* $\text{OUT}(\Pi(b, r_1)) = \text{OUT}(\Pi(b, r_2))$*.*

[**Ran's Note: I stopped here**]

Proof. For two strings r_1, r_2 we will use the notation $d(r_1, r_2)$ to signify the number of coordinates in which the strings differ. We have two cases:

- $d(r_1, r_2) = 4$ - Without loss of generality assume that $r_1 = 0000, r_2 = 1111$, $(b, r_1), (b, r_2)$ are doubly-robust pairs with respect to $\mathsf{P}_1, \mathsf{P}_4$, respectively. Then, by definition the pairs $(b, r_3), (b, r_4)$, where $r_3 = 1000$, $r_4 = 1110$ are robust with respect to $\mathsf{P}_1, \mathsf{P}_4$, respectively. Observe an execution of Π in which P_5^* sends b in the first round, with $r = 1100$ being the second round randomness. Consider execution σ_0 in which an adversary has corrupted

P_1, such that it halts after the second round. All honest parties must reach agreement on an output bit. Without loss of generality that output bit is 0, and assume towards a contradiction that $\text{OUT}(\Pi(b, r_3)) = 1$. We reach a contradiction in agreement with an adversary corrupting P_5^*, which sends vector b in the first round, and P_2^* which sends random bit 0 to P_1 and random bit 1 to all other parties. P_1's view is that of an honest execution of $\Pi(b, r_3)$. Thus, it halts after one round, and by assumption outputs 1. All other honest parties have the same view as in σ_0, since their incoming messages in the first round, and second round randomness are the same, and P_1 halts after two rounds. Thus, by assumption they all output 0. This, in contradiction to Π reaching agreement. The same holds for an execution of $\Pi(b, r_4)$. Thus $\text{OUT}(\Pi(b, r_3)) = \text{OUT}(\Pi(b, r_4))$. It is easy to see that $\text{OUT}(\Pi(b, r_1)) = \text{OUT}(\Pi(b, r_3))$ and $\text{OUT}(\Pi(b, r_2)) = \text{OUT}(\Pi(b, r_4))$, since assuming otherwise, by definition, an adversary corrupting P_1 (resp. P_4) can make honest parties output different bits by simply sending some parties randomness bit 0, and others randomness bit 1, in $\Pi(b, r_1)$ (resp. $\Pi(b, r_2)$). Since (b, r_1) (resp. (b, r_2)) is robust for P_1 (resp. P_4), all honest parties will halt after two rounds and output different bits. [**Matan's Note: needs to be done nicer**]

- $d(r_1, r_2) < 4$ - If $d(r_1, r_2) = 1$, then the output is trivially equal. Without loss of generality $r_1 = 0000, r_2 = 0001$, and an adversary corrupting P_4 can simply send randomness bit 0 to some parties, and 1 to other parties, making all parties halt with different output values.
  If $d(r_1, r_2) = 2$ or $d(r_1, r_2) = 3$, then the previous argument involving $\Pi(b, r_3), \Pi(b, r_4)$ holds for this case as well.

We conclude the claim. □

We can now use Claims 2.15 and 2.16 to prove Lemma 2.11.

Proof of Lemma 2.11. [**Matan's Note: need to insert probability δ.**] From Claim 2.15 we know that $\forall b_i \exists r_i$ such that the pair (b_i, r_i) is doubly-robust with respect to P_i. From definition we know that the pair $(b_i \oplus e_i, r_i)$ is also doubly robust with respect to P_i. But $b_i \oplus e_i = b_{i-1}$. Thus, $\text{OUT}(\Pi(b_{i-1}, r_i)) = \text{OUT}(\Pi(b_i, r_i))$. From the same claim, we also have a random string r_{i+1} such that the pair (b_{i+1}, r_{i+1}) is doubly-robust with respect to P_{i+1}, and by a similar argument we have that (b_i, r_{i+1}) is doubly-robust with respect to P_{i+1}. Since both (b_i, r_i) and (b_i, r_{i+1}) are doubly-robust with respect to different parties, by Claim 2.16 $\text{OUT}(\Pi(b_i, r_i)) = \text{OUT}(\Pi(b_i, r_{i+1}))$. Since this is true for every i, we deduce that there exists two randomness strings $r_0, r_4 \in \{0,1\}^4$ such that $\text{OUT}(\Pi(b_0, r_0)) = \text{OUT}(\Pi(b_4, r_4))$. Finally, we note that $\Pi(b_0, r_0)$ is an honest execution of Π with input vector $(0, 0, 1, 1, 0)$, and $\Pi(b_4, r_4)$ is an honest execution of Π with input vector $(0, 0, 1, 1, 1)$. By Lemma 2.2 $\text{OUT}(\Pi(b_0, r_0)) = 0$ and $\text{OUT}(\Pi(b_4, r_4)) = 1$, in contradiction. We conclude the lemma. □

2.5 Committing to Randomness

2.5.1 Model, Protocols and Adversaries

2.5.2 Single Coin Randomness

We further extend our model to one in which parties can no longer choose their second-round randomness. Each party has access to an oracle which outputs one bit in the second round. The only option a party has is to either let the receiving party see this bit, or to abort it. The rest of the model is similar to the previous one.

Theorem 2.17. *Let Π be an n-party,t-resilient Byzantine agreement in the above model, where t is any constant fraction of n, and let δ be a lower-bound for the probability that for any input all honest parties halt after two rounds, regardless of adversarial behavior. Then $\delta \in o(1)$.*

We prove the theorem for the case of honest majority where $n = 4\ell + 1$ and $t = 2\ell$ and discuss later on how to convert the proof to such that t is any constant fraction of n. We first prove the following combinatorial lemma.

Lemma 2.18. *Let $\gamma > 0$, $s \in \Theta(n)$ and $C \subset \{0,1\}^n$ such that for every $S \in \binom{[n]}{s}$ it holds that $|C|_S| < \gamma \cdot 2^s$. Then, $|C| \in o(1)$.*

Proof. KKL! □

Definition 2.19. *Let $r \in \{0,1\}^n$, $b = (b_1, \ldots, b_{n-1}) \in \{0,1\}^{n-1}$ and $S \subset [n]$. Denote by $\Pi(b, r^{\star S})$ an execution of Π such that:*

- *Party P_n sends b_i to P_i.*
- *r denotes the randomness in the second round.*
- *All parties in S abort prior to sending their second round messages.*

To prove Theorem 2.17 we make a series of claims, showing that either $\delta \in o(1)$, or there is an adversarial behavior which makes different parties output different values. We split our input parties into 5 distinct sets, each of size ℓ.

Claim 2.20. *For every $S \in \binom{[n]}{s}$ there exists $R_0^S \subseteq \mathcal{R}$ of size $\left|R_0^S\right| \geq (\delta - \beta - 2\alpha) \cdot 2^n$ such that for every $r \in R_0^S$ it holds that:*

1. *$\Pi(b_0, r^{\star S})$ halts after two rounds and all honest parties output 0.*
2. *$\Pi(b_1, r^{\star S})$ halts after two rounds and all honest parties output 0.*

Proof. Define R_0^S to be the set of r's such that $\Pi(b_c, r), \Pi(b_c, r^{\star S})$ halt in two rounds for $c \in \{0,1\}$. We claim that $\left|R_0^S\right| \geq \delta \cdot 2^n$. Otherwise, by corrupting P_n, the first ℓ parties and the parties in S, the adversary can prevent some honest parties from halting with probability greater than $1 - \delta$, in contradiction to the assumption. Namely, a rushing adversary having seen $r \notin R_0^S$ can instruct the first ℓ parties to propagate either b_0 or b_1 and the parties in S to abort accordingly, thus preventing halting. Next, we show that honest parties in $\Pi(b_0, r^{\star S})$ and $\Pi(b_1, r^{\star S})$ output 0, for every $r \in R_0^S$. We claim that the honest parties in an execution of $\Pi(b_1, r^{\star S})$ must output 0, for $r \in \tilde{R}_0^S$ of relative size $\delta - \beta - 2\alpha$. First, we have $\left|\{r \in R_0^S \mid \Pi(b_0, r^{\star S}) = 0\}\right| \geq (\delta - \beta - \alpha) \cdot 2^n$. This stems from Lemma 2.2 as with probability at least $1 - \beta$, $\Pi(b_0, r) = 0$, and from (t, α)-agreement as with probability at least $1 - \alpha$ $\Pi(b_0, r) = \Pi(b_0, r^{\star S})$. To conclude, $\left|\{r \in R_0^S \mid \Pi(b_0, r^{\star S}) = 0 \wedge \Pi(b_1, r^{\star S}) = 0\}\right| \geq (\delta - \beta - 2\alpha) \cdot 2^n$ again by (t, α)-agreement. □

Claim 2.21. *For every $S \in \binom{[n]}{s}$ there exists $R_1^S \subseteq \mathcal{R}$ of size $\left|R_1^S\right| \geq (\delta - \beta - 2\alpha) \cdot 2^n$ such that for every $r \in R_1^S$ it holds that:*

1. *$\Pi(b_3, r^{\star S})$ halts after two rounds and all honest parties output 1.*

2. $\Pi(b_4, r^{\star S})$ *halts after two rounds and all honest parties output* 1.

Proof. The proof of this claim is analogous to the proof of Claim 2.20. □

Claim 2.22. *For every* $S \in \binom{[n]}{s}$ *there exists* $R_0^S \subseteq \mathcal{R}$ *of size* $\left|R_0^S\right| \geq ((\delta - 3\alpha)/2) \cdot 2^n$ *such that for every* $r \in R_0^S$ *it holds that:*

1. $\Pi(b_1, r^{\star S})$ *halts after two rounds and all honest parties output* 0.
2. $\Pi(b_2, r^{\star S})$ *halts after two rounds and all honest parties output* 0.

Proof. Fix $S \in \binom{[n]}{s}$. Let $R^S \subseteq \mathcal{R}$ such that $\Pi(b_1, r^{\star S}), \Pi(b_2, r^{\star S})$ halt after two rounds, for every $r \in R^S$. Similarly to Claim 2.20, $\left|R^S\right| \geq \delta \cdot 2^n$. Let $R_0^S, R_1^S \subseteq R^S$ such that $\Pi(b_1, r^{\star S}) = c$, for every $r \in R_c^S$. Let $c(S) = \text{argmax}_{c \in \{0,1\}}(\left|R_c^S\right|)$ and observe that $\left|R_{c(S)}^S\right| \geq ((\delta - \alpha)/2) \cdot 2^n$. Similarly to Claim 2.20, There exists $\tilde{R}_{c(S)}^S$ of size $\left|\tilde{R}_{c(S)}^S\right| \geq ((\delta - 3\alpha)/2) \cdot 2^n$ such that $\Pi(b_2, r^{\star S})$ halts after two rounds and all honest parties output $c(S)$ for all $r \in \tilde{R}_{c(S)}^S$. If not, then similarly to the proof of Claim 2.20, a rushing adversary having seen $r \in \tilde{R}_{c(S)}^S$ can instruct parties $\ell+1, \ldots, 2\ell$ to propagate either b_1 or b_2 and the parties in S to abort accordingly, thus violating *agreement*. It remains to show that $c(S) = 0$ for every S.

For every S' define $\tilde{R}_0^{S'} = \{r \in \mathcal{R} \mid \Pi(b_1, r^{\star S}) = 0\}$. By the previous claims, we know that $\left|\tilde{R}_0^{S'}\right| \geq (\delta - \beta - 2\alpha) \cdot 2^n$, for every S'. We claim that $\left|R_1^S|_{\overline{S}'} \cap \tilde{R}_0^{S'}|_{\overline{S}'}\right| \leq \alpha \cdot 2^{|\overline{S}'|}$. To see this, we note that $\tilde{R}_c^{S'} \equiv \{0,1\}^s \times \tilde{R}_c^{S'}|_{\overline{S}'}$, for every S' and for every $c \in \{0,1\}$. This is due to the definition of $\tilde{R}_c^{S'}$, as for every $r \in \tilde{R}_c^{S'}$ and $r' \in \mathcal{R}$, if $r|_{\overline{S}'} = r'|_{\overline{S}'}$ then $r' \in \tilde{R}_c^{S'}$. Thus, we have $\left|\tilde{R}_1^S|_{\overline{S \cup S'}} \cap \tilde{R}_0^{S'}|_{\overline{S \cup S'}}\right| \leq \alpha \cdot 2^{|\overline{S \cup S'}|}$, since otherwise an adversary can corrupt both S and S', and cause disagreement with probability larger than α. This also holds for projections on subsets containing $\overline{S \cup S'}$, namely $\overline{S}'$. By definition, $\left|\tilde{R}_0^{S'}|_{\overline{S}'}\right| \geq (\delta - \beta - 2\alpha) \cdot 2^{|\overline{S}'|}$ and we conclude that for every S', $\left|\tilde{R}_1^S|_{\overline{S}'}\right| < (1 - \delta + \beta + 3\alpha) \cdot 2^{|\overline{S}'|}$. Thus, by Lemma 2.24, $\left|\tilde{R}_1^S\right| \in o(1) \cdot 2^n$ and the proof is concluded. [**Nikos's Note: Parameters come from here**] □

Claim 2.23. *For every* $S \in \binom{[n]}{s}$ *there exists* $R_1^S \subseteq \mathcal{R}$ *of size* $\left|R_1^S\right| \geq ((\delta - 3\alpha)/2) \cdot 2^n$ *such that for every* $r \in R_1^S$ *it holds that:*

1. $\Pi(b_2, r^{\star S})$ *halts after two rounds and all honest parties output* 1.
2. $\Pi(b_3, r^{\star S})$ *halts after two rounds and all honest parties output* 1.

Proof. The proof of this claim is analogous to the proof of Claim 2.21. □

Proof of Theorem 2.17. By Claims 2.22 and 2.23, for every $S \in \binom{[n]}{s}$ there exist $R_1^S, R_0^S \in \mathcal{R}$ of size $\left|R_1^S\right|, \left|R_0^S\right| \geq ((\delta - 3\alpha)/2) \cdot 2^n$ such that:

- $\Pi(b_2, r_0^{\star S})$ halts after two rounds and all honest parties output 0, for every $r_0 \in R_0^S$.
- $\Pi(b_2, r_1^{\star S})$ halts after two rounds and all honest parties output 1, for every $r_1 \in R_1^S$.

Fix $S \in \binom{[n]}{s}$. By the same argument as in the proof of Claim 2.22, $\left|R_1^S|_{\overline{S}'} \cap \tilde{R}_0^{S'}|_{\overline{S}'}\right| \leq \alpha \cdot 2^{\left|\overline{S}'\right|}$, for every $S' \in \binom{[n]}{s}$. Therefore, $\left|R_1^S|_{\overline{S}'}\right| < (1 - \frac{\delta - 5\alpha}{2}) \cdot 2^{\left|\overline{S}'\right|}$, for every S'. By Lemma 2.24 $\left|R_1^S\right| \in o(1) \cdot 2^n$ and the proof is concluded.

□

2.5.3 Generalizing to Any n

2.5.4 What If Not All Sets Have a Small Projection?

Lemma 2.24. *Let $\gamma > 0$, $s \in \Theta(n)$ and $C \subset \{0,1\}^n$ such that for all but a vanishing fraction of $S \in \binom{[n]}{s}$ it holds that $|C|_S| < \gamma \cdot 2^s$. Then, $|C| \in o(1)$.*

Proof. Alex! □

2.5.5 Committed Randomness is Now a String

2.6 Second Round is Private Coin

3 Reformulation in Terms of Computing Probabilities

In this section we formulate an attack on Byzantine Agreement protocols in which there is a constant probability of halting after two rounds. We later show that in certain cases this probability is either computable or can be approximated to attack the correctness of the protocol. (I.e that with some probability honest parties will output differing bits)

References

[1] M. J. Fischer and N. A. Lynch. A lower bound for the time to assure interactive consistency. *Inf. Process. Lett.*, 14(4):183–186, 1982.