RELAXED BYZANTINE CONSENSUS

YUZHOU GU, TIANCHENG YU, YUANCHENG YU

1. Introduction

The Byzantine vector consensus problem is a well studied problem in distributed computation, where each process receives a vector in $\mathbb{R}^d$, and the non-faulty processes are required to output a vector lying in the convex hull of the input vectors of all non-faulty processes. In the exact version, all non-faulty processes must output the same vector; while in the approximate version, their outputs must be within ϵ-distance to each other. Vaidya and Garg [VG13] completely solved this problem, showing that

- in the synchronous setting, exact Byzantine vector consensus is solvable if and only if $n \geq \max\{3, d+1\}f + 1$, where n is the number of processes, d is the dimension, and f is the number of Byzantine faults;
- in the asynchronous setting (where exact Byzantine vector consensus is impossible), the approximate version is solvable if and only if $n \geq (d+2)f + 1$.

Since solving the Byzantine vector consensus problem is impossible when $d > n$, which is usually the case when dealing with high-dimensional data, we need to resort to a relaxed guarantee to improve the dependence of n on d.

Two definitions of closeness for relaxed Byzantine vector consensus have been studied by Xiang and Vaidya [XV17]. In k-consensus, any k-dimensional orthogonal projection of the output must be in the convex hull of the same projection of all non-faulty inputs; in (δ, p)-consensus, the output is required to be δ-close in L_p-distance to the convex hull.

For k-consensus, they showed that the dependence of n on d can be eliminated if and only if $k = 1$, where it is trivially achievable by applying Byzantine scalar consensus pointwise. For (δ, p)-consensus, they showed that for $p = 2$ and certain combinations of n, f (namely, $f = 1$ or $n = (d+1)f$), it is possible to achieve a δ which is the product of a function of n, f and the max L_p distance between non-faulty inputs. This adaptive version makes more sense than constant δ since scaling up the inputs would not affect the guarantee of the algorithm. For general n, f, p such that $n \geq 3f + 1$, they conjectured an upper bound for δ which depends on d.

In this paper, we show that for all n, f, p such that $n \geq 3f+1$, the (δ, p)-consensus problem is solvable with δ the product of a function of n, f and the max L_p distance between non-faulty inputs, but independent of d. The corresponding algorithms are surprisingly simple. We also provide lower bounds for δ by constructing a symmetric data set, which match the upper bound up to a factor of 4 for L_1 and L_∞ norm.

1.1. Application.

This problem is motivated by distributed estimation with Byzantine failure [YCKB18]. In a distributed estimation problem, most of the processes

will report their data, which is essentially an empirical distribution of some true distribution μ, while the faulty processes may report arbitrary distributions. As long as the consensus is close to the convex hull of the truthful empirical distributions, it will also converge to the true distribution by Glivenko–Cantelli theorem (or Dvoretzky–Kiefer–Wolfowitz inequality for non-asymptotic version). As a result, it is highly desirable to understand when this task is tractable and the corresponding complexity.

1.2. **Organization.** A summary of our results for the input-dependent relaxed Byzantine vector consensus problem is in Table 1.2.

norm	upper bound	lower bound
L_1	$\frac{2f}{n-f}$	$\frac{f}{2(n-f)}$
L_p, $1 < p < \infty$	$\frac{2f}{n-f}$	$\frac{(n-f)^{\frac{1}{p}-1} f^{\frac{1}{p}}}{2^{\frac{1}{p}} \left[(n-f)^{\frac{1}{p-1}} + f^{\frac{1}{p-1}}\right]^{\frac{p-1}{p}}}$
L_∞	$\frac{f}{n-f}$	$\frac{f}{2(n-f)}$
any norm	$\frac{2f}{n-f}$	none

In Section 2, we prove upper bounds for the input-dependent relaxed Byzantine vector consensus problem. In Section 3, we prove lower bounds for the problem. In Section 4, we prove some other results related to relaxed Byzantine consensus.

2. Input-Dependent Upper Bound

In this section we prove input-dependent upper bounds for relaxed Byzantine consensus problem.

Let us first rigorously define the problem. There are n processes and at most f of them suffer from Byzantine failure. Each process i holds a input $\mathbf{X}_i$, which is a vector in some normed space. Each non-faulty processes i should output a vector, satisfying

- Exact agreement: All non-faulty outputs should be the same.
- Relaxed validity: The non-faulty output should have distance (in the given norm) at most δ to the convex hull of the non-faulty inputs.

In the input-dependent version, we compare δ with e, the maximum distance between any two non-faulty inputs. We aim to minimize $\kappa = \kappa(n, f, ||\cdot||)$ such that $\delta \le \kappa e$ holds.

In Section 2.1, we prove an upper bound for L_∞. In Section 2.2, we prove a (slightly worse) upper bound for general norms. Interestingly, the upper bounds we achieve do not dependent on the dimension d.

2.1. Upper bound for L_∞ distance.

Proposition 1. *For the synchronous relaxed Byzantine consensus problem under L_∞ distance and f failures ($n \ge 3f+1$), there exists an algorithm that can output a vector with at most*

$$\frac{f}{n-f} \max_{i \ne j \text{ non-faulty}} ||\mathbf{X}_i - \mathbf{X}_j||_\infty$$

distance to the convex hull of vectors of non-faulty processes.

Proof. Each process broadcasts $\mathbf{X}_i$ using the standard Byzantine agreement algorithm. This can be done because $n \geq 3f+1$. After this step, all non-faulty process agree on the same sequence $\mathbf{X}_1, \ldots, \mathbf{X}_n$ of vectors.

For each coordinate k, let $y_1 \leq \ldots \leq y_n$ be sorted list of $\mathbf{X}_{1,k}, \ldots, \mathbf{X}_{n,k}$. Every non-faulty process outputs

$$\frac{1}{n-2f} \sum_{f+1 \leq i \leq n-f} y_i$$

for this coordinate. Let $\mathbf{X}^*$ be the output vector. We prove that $\mathbf{X}^*$ satisfies

$$d_{L^\infty}(\mathbf{X}^*, \mathrm{Conv}(\mathbf{X}_S)) \leq \frac{f}{n-f} \max_{i,j \in S} ||\mathbf{X}_i - \mathbf{X}_j||_\infty,$$

where $S \subseteq [n]$ is the set of non-faulty processes.

We have

$$d_{L^\infty}(\mathbf{X}^*, \mathrm{Conv}(\mathbf{X}_S)) \leq ||\mathbf{X}^* - \frac{1}{|S|} \sum_{i \in S} \mathbf{X}_i||_\infty.$$

Fix a coordinate k. WLOG assume that $\mathbf{X}_{1,k} \leq \cdots \leq \mathbf{X}_{n,k}$. So $y_i = \mathbf{X}_{i,k}$ for all $1 \leq i \leq n$. Let $A = S \cap \{1, \ldots, f\}$, and $B = S \cap \{n-f+1, \ldots, n\}$.

Let $e = \max_{i,j \in S} |\mathbf{X}_{i,k} - \mathbf{X}_{j,k}|$. Because $|S| \geq n-f$, $y_{n-f} - y_{f+1} \leq e$. We consider a coupling between $\mathrm{Unif}(S)$ and $\mathrm{Unif}(\{f+1, \ldots, n-f\})$, satisfying: if $i \in A$, $j \in B$, and (i,k) and (j,l) have non-zero weights in the coupling, then $k \leq l$. Such a coupling can easily be constructed. By this coupling, we get

$$|\frac{1}{n-2f} \sum_{f+1 \leq i \leq n-f} y_i - \frac{1}{|S|} \sum_{j \in S} \mathbf{X}_{j,k}| \leq \frac{\max\{|A|, |B|\}}{|S|} e \leq \frac{f}{n-f} e.$$

Applying the above for all coordinates k, we get the desired result. □

2.2. **General upper bound.**

Proposition 2. *For the synchronous relaxed Byzantine consensus problem under arbitrary norm and f failures ($n \geq 3f+1$), there exists an algorithm that can output a vector with at most*

$$\frac{2f}{n-f} \max_{i \neq j \text{ non-faulty}} ||\mathbf{X}_i - \mathbf{X}_j||$$

distance to the convex hull of vectors of non-faulty processes.

Proof. Each process broadcasts $\mathbf{X}_i$ using the standard Byzantine agreement algorithm. This can be done because $n \geq 3f+1$. After this step, all non-faulty process agree on the same sequence $\mathbf{X}_1, \ldots, \mathbf{X}_n$ of vectors. Let $T \subseteq [n]$, $|T| \geq n-f$ be a set with smallest $\max_{i,j \in T} ||\mathbf{X}_i - \mathbf{X}_j||$. All non-faulty processes outputs $\mathbf{X}^* = \frac{1}{|T|} \sum_{i \in T} \mathbf{X}_i$.

We prove that $\mathbf{X}^*$ satisfies

$$d(\mathbf{X}^*, \mathrm{Conv}(\mathbf{X}_S)) \leq \frac{f}{n-f} \max_{i,j \in S} ||\mathbf{X}_i - \mathbf{X}_j||,$$

where $S \subseteq [n]$ is the set of non-faulty processes.

Because $|T|, |S| \geq n-f$ and $n \geq 3f$, we have $T \cap S \neq \emptyset$. If $k \in T \cap S$, then for every $i \in T$, $j \in S$, we have

$$||\mathbf{X}_i - \mathbf{X}_j|| \leq ||\mathbf{X}_i - \mathbf{X}_k|| + ||\mathbf{X}_k - \mathbf{X}_j|| \leq 2 \max_{i,j \in S} ||\mathbf{X}_i - \mathbf{X}_j||.$$

The two distributions $\mathrm{Unif}(S)$ and $\mathrm{Unif}(T)$ has total variation distance at most $\frac{f}{n-f}$. By using a coupling that achieves this total variation distance, we get

$$||\frac{1}{|T|}\sum_{i\in T}\mathbf{X}_i - \frac{1}{|S|}\sum_{i\in S}\mathbf{X}_i|| \leq \frac{2f}{n-f}\max_{i,j\in S}||\mathbf{X}_i - \mathbf{X}_j||.$$

This finishes the proof. □

We give an alternative proof of Proposition 2 using linear metric embedding to L_∞ space. This method uses slightly more machinery and achieves the same bound as Proposition 2, but an improvement to Proposition 1 would imply directly an improved upper bound for general norms.

Alternative proof of Proposition 2. As usual, each process broadcasts $\mathbf{X}_i$ using the standard Byzantine agreement algorithm. This can be done because $n \geq 3f+1$. The output vector is the same as the $\mathbf{X}^*$ in the previous proof. We use a different way to prove that

$$d(\mathbf{X}^*, \mathrm{Conv}(\mathbf{X}_S)) \leq \frac{f}{n-f}\max_{i,j\in S}||\mathbf{X}_i - \mathbf{X}_j||.$$

Note that any finite dimensional normed space embeds isometrically and linearly into L^∞ (with possibly uncountable dimension). Such an embedding could be constructed as the following: Consider the dual space of the original normed space. The codomain space has uncountable dimension, labeled by elements in the unit ball of the dual space. Each vector $\mathbf{X}$ in the original space is mapped to $f(\mathbf{X})$ for coordinate f. One can verify that this is a linear isometric embedding of the original normed space into L^∞.

Let g be the embedding mentioned above. Let $\mathbf{Y}_i = g(\mathbf{X}_i)$. By Proposition 1, there exists $\mathbf{Y}^*$ such that

$$||\mathbf{Y}^* - \frac{1}{|S|}\sum_{i\in S}\mathbf{Y}_i||_\infty \leq \frac{f}{n-f}\max_{i,j\in S}||\mathbf{Y}_i - \mathbf{Y}_j||_\infty$$

for all $S \subseteq [n]$, $|S| \geq n-f$.

So

$$\begin{aligned}
||\mathbf{X}^* - \frac{1}{|S|}\sum_{i\in S}\mathbf{X}_i|| &= ||\frac{1}{|T|}\sum_{i\in T}\mathbf{X}_i - \frac{1}{|S|}\sum_{i\in S}\mathbf{X}_i|| \\
&= ||\frac{1}{|T|}\sum_{i\in T}\mathbf{Y}_i - \frac{1}{|S|}\sum_{i\in S}\mathbf{Y}_i||_\infty \\
&\leq ||\frac{1}{|T|}\sum_{i\in T}\mathbf{Y}_i - \mathbf{Y}^*||_\infty + ||\frac{1}{|S|}\sum_{i\in S}\mathbf{Y}_i - \mathbf{Y}^*||_\infty \\
&\leq \frac{f}{n-f}\max_{i,j\in T}||\mathbf{Y}_i - \mathbf{Y}_j||_\infty + \frac{f}{n-f}\max_{i,j\in S}||\mathbf{Y}_i - \mathbf{Y}_j||_\infty \\
&\leq \frac{2f}{n-f}\max_{i,j\in S}||\mathbf{Y}_i - \mathbf{Y}_j||_\infty \\
&= \frac{2f}{n-f}\max_{i,j\in S}||\mathbf{X}_i - \mathbf{X}_j||.
\end{aligned}$$

□

3. Input-Dependent Lower Bound

In this section, we construct input-dependent lower bounds for relaxed Byzantine consensus problem. We begin with the general construction in Section 3.1 and then specialize the result for different norms (L_1 in Section 3.2, L_∞ in Section 3.3 and general L_p in Section 3.4.)

3.1. **General construction.** Since we are interested in dimension-free bounds, i.e., dimension d does not appear in the bound of distance explicitly, let's consider the case when d is large. In following, we will construct the lower bound when $d = n!$.

Fix n scalars $\{a_i\}_{i=1}^n$, whose value will be chosen later in the subsequent sections to cater specific norms. To simplify the exposition, we will use any permutation $\sigma \in S_n$ to denote a number in $[n!]$ when there is no confusion. Here, S_n is the set of all the permutations on $[n]$. Now let's define the input $\{\mathbf{X}_i\}_{i=1}^n$ based on $\{a_i\}_{i=1}^n$:

$$X_{i,\sigma} = a_{\sigma(i)}$$

Intuitively, if we consider $\{\mathbf{X}_i\}_{i=1}^n$ as columns of a matrix $\mathbf{X} \in \mathbb{R}^{n\times n!}$,

$$\mathbf{X} = \left[\mathbf{X}_1, \mathbf{X}_2, \quad \cdots, \quad \mathbf{X}_n\right],$$

then the σ-th row of $\mathbf{X}$ is the vector we get by applying σ to the coordinates of $\mathbf{a} = [a_1, a_2, \cdots, a_n]$.

This symmetric construction indicates a symmetric property in different coordinates. Define the permutation version $\mathbf{Y}^\sigma$ by $Y^\sigma_\tau = Y_{\sigma\circ\tau}$, then as a result, $\mathbf{Y}^\sigma$ is "as good as" $\mathbf{Y}$. Mathematically, let $\mathcal{T} \subseteq [n]$ be the set of indices of non-faulty processes, then

$$d\left(\mathbf{Y}, \mathrm{Conv}\left(\mathbf{X}_\mathcal{T}\right)\right) = d\left(\mathbf{Y}^\sigma, \mathrm{Conv}\left(\mathbf{X}_{\sigma(\mathcal{T})}\right)\right).$$

So in worst case the distance is no worse then that of $\mathbf{Y}$. Choose

$$\mathcal{T}^* = \arg\max_{\mathcal{T}} d\left(\mathbf{Y}, \mathrm{Conv}\left(\mathbf{X}_\mathcal{T}\right)\right),$$

then by convexity of the distance function we have

$$\begin{aligned} d\left(\frac{1}{n!}\sum_{\sigma\in S_n}\mathbf{Y}^\sigma, \mathrm{Conv}\left(\mathbf{X}_\mathcal{T}\right)\right) \leq & \frac{1}{n!}\sum_{\sigma\in S_n} d\left(\mathbf{Y}^\sigma, \mathrm{Conv}\left(\mathbf{X}_\mathcal{T}\right)\right) \\ \leq & \frac{1}{n!}\sum_{\sigma\in S_n} d\left(\mathbf{Y}, \mathrm{Conv}\left(\mathbf{X}_{\mathcal{T}^*}\right)\right) = d\left(\mathbf{Y}, \mathrm{Conv}\left(\mathbf{X}_{\mathcal{T}^*}\right)\right). \end{aligned}$$

As a result, we only need to consider outputs in the form $\frac{1}{n!}\sum_{\sigma\in S_n}\mathbf{Y}^\sigma$, which has the same value in different coordinates. From now on, we only consider $\mathbf{Y}$ s.t. for any $\sigma \in S_n$, $Y_\sigma = b$. Clearly $\mathbf{Y}^\sigma = \mathbf{Y}$ for any $\sigma \in S_n$.

We can further simplify the distance by symmetry. There exist $\{\rho_i\}_{i\in\mathcal{T}}$ s.t. $\rho_i \geq 0$,$\sum_{i\in\mathcal{T}}\rho_i = 1$ and

$$d\left(\mathbf{Y}, \mathrm{Conv}\left(\mathbf{X}_\mathcal{T}\right)\right) = d\left(\mathbf{Y}, \sum_{i\in\mathcal{T}}\rho_i\mathbf{X}_i\right)$$

Now for any permutation $\tau \in S_{\mathcal{T}}$,

$$d\left(\mathbf{Y}, \sum_{i\in\mathcal{T}} \rho_i \mathbf{X}_i\right) = d\left(\mathbf{Y}^\tau, \sum_{i\in\mathcal{T}} \rho_{\tau(i)} \mathbf{X}_i\right) = d\left(\mathbf{Y}, \sum_{i\in\mathcal{T}} \rho_{\tau(i)} \mathbf{X}_i\right).$$

Put everything together,

$$\begin{aligned} d\left(\mathbf{Y}, \mathrm{Conv}\left(\mathbf{X}_{\mathcal{T}}\right)\right) =& \frac{1}{|\mathcal{T}|!} \sum_{\tau\in S_{\mathcal{T}}} d\left(\mathbf{Y}, \sum_{i\in\mathcal{T}} \rho_{\tau(i)} \mathbf{X}_i\right) \\ \geq& d\left(\mathbf{Y}, \frac{1}{|\mathcal{T}|!} \sum_{\tau\in S_{\mathcal{T}}} \sum_{i\in\mathcal{T}} \rho_{\tau(i)} \mathbf{X}_i\right) = d\left(\mathbf{Y}, \frac{1}{|\mathcal{T}|} \sum_{i\in\mathcal{T}} \mathbf{X}_i\right). \end{aligned}$$

Therefore we only need to lower bound $d\left(\mathbf{Y}, \frac{1}{|\mathcal{T}|}\sum_{i\in\mathcal{T}} \mathbf{X}_i\right)$.

The last step is to simplify $\max\limits_{i,j\in\mathcal{T}} \|\mathbf{X}_i - \mathbf{X}_j\|$, which is just $d\left(\mathbf{X}_1, \mathbf{X}_2\right)$ because the distance between different $\mathbf{X}_i$s are same. To see this, consider the permutation $(i,k)(j,l)$ to prove $d\left(\mathbf{X}_i, \mathbf{X}_j\right) = d\left(\mathbf{X}_k, \mathbf{X}_l\right)$.

Finally, given $\mathbf{a}$, choose $\mathcal{T} = [n-f]$, we just need to compute

$$\frac{\min\limits_{\mathbf{Y}=b\mathbf{1}} d\left(\mathbf{Y}, \frac{1}{(n-f)}\sum_{i\in[n-f]} \mathbf{X}_i\right)}{d\left(\mathbf{X}_1, \mathbf{X}_2\right)}.$$

3.2. **Lower bound for L_1 distance.** Here we choose $a_1 = 1$ and $a_i = 0$ for $i > 1$. Now coordinates of $\sum_{i\in[n-f]} \mathbf{X}_i$ can take only two values: 0 and 1. $(n-f)\,(n-1)!$ of them are 1 and $f\,(n-1)!$ of them are 0. Therefore

$$d\left(\mathbf{Y}, \frac{1}{n-f} \sum_{i\in[n-f]} \mathbf{X}_i\right) = (n-1)! \left[(n-f)\left|b - \frac{1}{n-f}\right| + f\,|b|\right]$$

which is minimized when $b = \frac{1}{n-f}$.

We also have $d\left(\mathbf{X}_1, \mathbf{X}_2\right) = 2\,(n-1)!$. Putting everything together, we have

Proposition 3. *For the synchronous relaxed Byzantine consensus problem under L_1 distance with n processes and f failures ($n \geq 3f+1$). There exist an instance s.t. it is impossible to output a vector with less than $\frac{f}{2(n-f)} \max_{i\neq j \text{ non-faulty}} ||\mathbf{X}_i - \mathbf{X}_j||_1$ distance to convex hull of vectors of non-faulty processes.*

3.3. **Lower bound for L_∞ distance.** Here we set $a_i = 1$ for $1 \leq i \leq f$ and $a_i = 0$ for $i > f$. Since coordinates of $\sum_{i\in[n-f]} \mathbf{X}_i$ lie within $[0, f/(n-f)]$, we can choose $b = f/2(n-f)$ to guarantee

$$d\left(\mathbf{Y}, \frac{1}{n-f} \sum_{i\in[n-f]} \mathbf{X}_i\right) \leq \frac{f}{2\,(n-f)}.$$

We also have $d\left(\mathbf{X}_1, \mathbf{X}_2\right) = 1$. Putting them together we have

Proposition 4. *For the synchronous relaxed Byzantine consensus problem under L_∞ distance with n processes and f failures ($n \geq 3f+1$). There exist an instance s.t. it is impossible to output a vector with less than $\frac{f}{2(n-f)} \max_{i\neq j \text{ non-faulty}} ||\mathbf{X}_i - \mathbf{X}_j||_\infty$ distance to convex hull of vectors of non-faulty processes.*

3.4. **Lower bound for L_p distance.** Similar to the construction in Section 3.2, We also give lower bounds for L_p distance with generic $p > 0$. However, unlike the results above, the lower bound here does not match the upper bound we prove.

We also choose $a_1 = 1$ and $a_i = 0$ for $i > 1$. Now coordinates of $\sum_{i\in[n-f]} \mathbf{X}_i$ can take only two values: 0 and 1. $(n-f)(n-1)!$ of them are 1 and $f(n-1)!$ of them are 0. Therefore

$$d\left(\mathbf{Y}, \frac{1}{n-f}\sum_{i\in[n-f]}\mathbf{X}_i\right) = \sqrt[p]{(n-1)!\left[(n-f)\left|b-\frac{1}{n-f}\right|^p + f|b|^p\right]}$$

which is minimized when

- $b = \frac{1}{n-f}$ if $0 < p \leq 1$ and
- $b = \frac{1}{n-f}\frac{(n-f)^{\frac{1}{p-1}}}{(n-f)^{\frac{1}{p-1}}+f^{\frac{1}{p-1}}}$ if $p > 1$.

We also have $d(\mathbf{X}_1, \mathbf{X}_2) = \sqrt[p]{2(n-1)!}$. Putting everything together, we have

Proposition 5. *For the synchronous relaxed Byzantine consensus problem under L_p distance with n processes and f failures ($n \geq 3f+1$). There exist an instance s.t. it is impossible to output a vector within*

- $\frac{1}{n-f}\sqrt[p]{\frac{f}{2}}\max_{i\neq j \text{ non-faulty}}||\mathbf{X}_i - \mathbf{X}_j||_p$ *if* $0 < p \leq 1$,
- $\frac{(n-f)^{\frac{1}{p}-1}f^{\frac{1}{p}}}{2^{\frac{1}{p}}\left[(n-f)^{\frac{1}{p-1}}+f^{\frac{1}{p-1}}\right]^{\frac{p-1}{p}}}\max_{i\neq j \text{ non-faulty}}||\mathbf{X}_i - \mathbf{X}_j||_p$ *if* $p > 1$

distance to convex hull of vectors of non-faulty processes.

For the special case $p = 2$, one can verify that it is not possible to achieve a lower bound better than $\sqrt{\frac{f}{2n(n-f)}}$ by using a different sequence $\{a_i\}_{i=1}^n$.

4. Other Results

In this section we prove some sporadic results related to the relaxed Byzantine consensus problem.

4.1. **Byzantine distribution consensus.** We consider the Byzantine distribution consensus problem, where the input are probability distributions on the compact interval $[0,1]$, and the non-faulty processes must output a distribution (resp. measure) on the interval that is close in total variation distance (resp. in L_1 distance) to the convex hull of distributions of the non-faulty processes. For this problem, we can prove matching lower and upper bounds.

Proposition 6. *For the synchronous relaxed Byzantine distribution consensus problem under total variation distance (resp. L_1 distance) with f failures ($n \geq 3f+1$) where each $\mathbf{X}_i$ is a probability distribution on $[0,1]$, the following holds.*

1. *There is an algorithm that outputs a distribution with $\leq \frac{f}{n}$ TV-distance to the convex hull of distributions of the non-faulty processes.*
2. *There is no algorithm that outputs a distribution with $< \frac{f}{n}$ TV-distance to the convex hull.*
3. *There is an algorithm that outputs a measure with $\leq \frac{f}{n-f}$ L_1-distance to the convex hull.*

(4) *There is no algorithm that outputs a measure with* $< \frac{f}{n-f}$ L_1*-distance to the convex hull.*

Proof. (1). Each process broadcasts $\mathbf{X}_i$. Each process outputs $\mathbf{X}^* := \frac{1}{n}\sum \mathbf{X}_i$.
Let $S \subseteq [n]$ be the set of non-faulty processes. Then

$$d_{\mathrm{TV}}(\mathbf{X}^*, \mathrm{Conv}(\mathbf{X}_S)) \leq \mathrm{TV}(\mathbf{X}^*, \frac{1}{|S|}\sum_{i\in S}\mathbf{X}_i) \leq \frac{n-|S|}{n} \leq \frac{f}{n}.$$

(3). Each process broadcasts $\mathbf{X}_i$. Each process outputs $\mathbf{X}^* := \frac{1}{n-f}\sum \mathbf{X}_i$.
Let $S \subseteq [n]$ be the set of non-faulty processes. Then

$$d_{L_1}(\mathbf{X}^*, \mathrm{Conv}(\mathbf{X}_S)) \leq d_{L_1}(X^*, \frac{1}{|S|}\sum_{i\in S}\mathbf{X}_i) \leq \frac{f}{n-f}.$$

(2)(4). The proof is very similar to Section 3. Suppose each $\mathbf{X}_i$ is a point distribution δ_{x_i} with x_is distinct. Suppose in an algorithm, all non-faulty processes output distribution/measure $\mathbf{X}$. By symmetry, for any permutation $\sigma \in S_n$, $\sigma(\mathbf{X})$ is also a valid output. By convexity of the distance function, $\frac{1}{|S_n|}\sum_{\sigma\in S_n}\sigma(\mathbf{X})$ is also a valid output. Therefore there exists an algorithm not worse than the original one, that outputs a distribution/measure that has the same weight on all x_i's. We can compute the optimal one among such distributions/measures and conclude. □

4.2. **Weaker input-dependence.** In the input dependent version, we compare the distance to convex hull to the maximum distance between inputs of the non-faulty processes. One may wonder what happens if we compare to the maximum distance between all X_is.

For the upper bound we can get a better algorithm.

Proposition 7. *For the synchronous relaxed Byzantine consensus problem under arbitrary norm and* f *failures* ($n \geq 3f+1$)*, there exists an algorithm that can output a vector with at most*

$$\frac{f}{n}\max_{i\neq j}||\mathbf{X}_i - \mathbf{X}_j||$$

distance to the convex hull of vectors of non-faulty processes.

Proof. Each process broadcasts $\mathbf{X}_i$. Each process outputs $\mathbf{X}^* := \frac{1}{n}\sum \mathbf{X}_i$.
Let $S \subseteq [n]$ be the set of non-faulty processes. Then

$$\begin{aligned} d(\mathbf{X}^*, \mathrm{Conv}(\mathbf{X}_S)) &\leq ||\mathbf{X}^* - \frac{1}{|S|}\sum_{i\in S}\mathbf{X}_i|| \\ &= ||\frac{1}{n}\sum_{i\notin S}\mathbf{X}_i - (\frac{1}{|S|}-\frac{1}{n})\sum_{i\in S}\mathbf{X}_i|| \\ &\leq \frac{n-|S|}{n}\max_{i\notin S, j\in S}||\mathbf{X}_i - \mathbf{X}_j|| \\ &\leq \frac{f}{n}\max_{i\neq j}||\mathbf{X}_i - \mathbf{X}_j||. \end{aligned}$$

□

For the lower bound, it is easy to see that all lower bounds in Section 3 still hold.

References

[VG13] Nitin H. Vaidya and Vijay K. Garg. Byzantine vector consensus in complete graphs. In *Proceedings of the 2013 ACM Symposium on Principles of Distributed Computing*, PODC '13, page 65–73, New York, NY, USA, 2013. Association for Computing Machinery.

[XV17] Zhuolun Xiang and Nitin H. Vaidya. Relaxed Byzantine Vector Consensus. In Panagiota Fatourou, Ernesto Jiménez, and Fernando Pedone, editors, *20th International Conference on Principles of Distributed Systems (OPODIS 2016)*, volume 70 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:15, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[YCKB18] Dong Yin, Yudong Chen, Ramchandran Kannan, and Peter Bartlett. Byzantine-robust distributed learning: Towards optimal statistical rates. In *International Conference on Machine Learning*, pages 5650–5659, 2018.