Tight distortion bounds for distributed metric voting on a line

Alexandros A. Voudouris

School of Computer Science and Electronic Engineering, University of Essex, United Kingdom

A R T I C L E   I N F O

Article history:
Received 5 January 2023
Received in revised form 3 March 2023
Accepted 6 March 2023
Available online 11 March 2023

Keywords:
• Metric voting
• Distortion
• Districts

A B S T R A C T

We consider a voting problem where agents and alternatives are on the line of real numbers, the agents are partitioned into disjoint districts, and the goal is to choose one alternative using a distributed voting mechanism. Such mechanisms select a representative alternative for each district and then choose one of them as the winner. We design simple mechanisms with distortion at most $2 + \sqrt{5}$ for the average-of-max and max-of-average cost objectives, matching the corresponding lower bound shown in previous work.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction and model

We consider the following voting problem. An instance $(I,A)$ consists of a set $N$ of $n$ agents and a set $A$ of $m$ alternatives, all of whom are represented by points on the line of real numbers. For any agent $i \in N$ and alternative $a \in A$, let $\delta(i,a)$ be the distance between $i$ and $a$ on the line (which is equal to the absolute difference between their positions). The agents are also partitioned into a set $D$ of $k$ districts, such that each district contains at least one agent. We denote by $N_d$ the set of agents of district $d \in D$, and by $n_d = |N_d|$ the size of $d$. The goal is to choose an alternative with good social efficiency guarantees using a distributed mechanism that takes as input ordinal information about the instance, such as the ordering of agents and alternatives on the line, and the ordinal preferences of the agents over the alternatives. To be specific, the preference of an agent $i$ over the alternatives is a linear ordering of the alternatives such that alternative $a$ is ranked higher than another alternative $b$ if $\delta(i,a) \leq \delta(i,b)$, breaking ties arbitrarily but consistently.

In general, a distributed mechanism $M$ works as follows:

• Step 1: For each district $d \in D$, given the preferences of the agents in $N_d$ over the alternatives and their relative ordering on the line, $M$ decides a representative alternative $y_d \in A$ for $d$.

• Step 2: Given the representatives of all districts and their relative ordering on the line, $M$ outputs one of them as the overall winner $M(I) \in \bigcup_{d \in D} N_d$.

Clearly, different distributed mechanisms can be designed by changing the method used for deciding the representative alternatives of the districts or the method for choosing a representative as the overall winner.

Interpreting the distance $\delta(i,a)$ as the individual cost of agent $i$ for alternative $a$, we measure the social efficiency of $x$ by some function of the distances of all agents from $x$. There are many well-known social efficiency objectives that have been studied in various social choice problems, such as the social cost (total or average distance of all agents) and the max cost (maximum distance among all agents). In the context of metric distributed voting, Anshelevich et al. [5] considered social objectives that are composed by some function that is applied over the districts and some function that is applied within the districts. In particular, they focused on the following four objectives:

• The average-of-average cost of $x$ is the average over each district of the average individual cost of the agents therein:

$$\text{AVG} \circ \text{AVG}(x) = \frac{1}{k} \sum_{d \in D} \left( \frac{1}{n_d} \sum_{i \in N_d} \delta(i,x) \right)$$

• The max cost of $x$ is the maximal individual cost over all agents:

$$\text{MAX} \circ \text{MAX}(x) = \max_{d \in D} \max_{i \in N_d} \delta(i,x) = \max_{i \in N} \delta(i,x)$$

• The average-of-max cost of $x$ is the average over each district of the max individual cost therein:

$$\text{AVG} \circ \text{MAX}(x) = \frac{1}{k} \sum_{d \in D} \max_{i \in N_d} \delta(i,x)$$
The max-of-average cost of $x$ is the max over each district of the average individual cost therein:

$$(\text{MAX} \circ \text{AVG}) (x) = \max_{d \in D} \left\{ \frac{1}{n_d} \sum_{i \in N_d} \delta(i, x) \right\}.$$ 

We measure the efficiency of a mechanism $M$ with respect to a social objective $f$ (such as the ones defined above) by its distortion, the worst-case ratio (over all possible instances) of the $f$-value of the alternative chosen by the mechanism over the minimum possible $f$-value among all alternatives:

$$\sup_{l = (N, A, D)} \frac{f(M(I))}{\min_{x \in A} f(x)}$$

By definition, the distortion of any mechanism is at least 1. We aim to design distributed mechanisms with an as low distortion as possible.

1.1. Our contribution

In previous work, Anshelevich et al. [5] considered the problem of distributed metric voting where agents and alternatives are in some arbitrary metric space. By carefully composing centralized voting mechanisms for making decisions within and over the districts, they designed distributed mechanisms with distortion guarantees for general metric spaces and with respect to the aforementioned objectives as well as more general ones. For the special case of a line metric (which is our focus here), they showed that the distortion of distributed mechanisms that use ordinal information is exactly 7 with respect to the average-of-average cost, exactly 3 with respect to the max cost, and in the interval $[2 + \sqrt{5}, 5]$ with respect to the average-of-max or the max-of-average-cost objectives.

Inspired by the recent work of Filos-Ratsikas et al. [15] who designed distributed mechanisms with tight distortion bounds for the continuous distributed facility location problem (where each point on the line of real numbers can be considered as an alternative), we resolve the distortion of distributed mechanisms for voting on a line with respect to the average-of-max and the max-of-average costs. In particular, we design two essentially symmetric distributed mechanisms that achieve a tight distortion bound of $2 + \sqrt{5}$ for the average-of-max cost and the max-of-average cost. For the average-of-max cost, our distributed mechanism chooses the favorite alternative of the rightmost agent in each district as the line representative, and then the $(\alpha \cdot k)$-th leftmost alternative as the overall winner. For the max-of-average cost, our mechanism chooses the favorite alternative of the $(\alpha \cdot n_d)$-th agent in each district $d$ as the district representative, and then the rightmost alternative as the overall winner. We show that both mechanisms achieve distortion at most $\max \left\{ \frac{2 - \alpha}{2}, \frac{\sqrt{5}}{2} - 1 \right\}$ for the corresponding objectives; for $\alpha = \frac{1}{2 - \sqrt{5}}$, the two terms in the bound balance out to $2 + \sqrt{5}$.

1.2. Other related work

Since its definition by Procaccia and Rosenschein [18], the distortion of voting mechanisms has been studied extensively for several settings under different assumptions about the preferences of the agents. The most well-studied setting is that of single-winner voting that has been considered under the premise of normalized agent valuations [1,7,11] as well as metric preferences [3,16,17]. Several other models have also been considered, such as multi-winner voting [8,9], participatory budgeting [6], and matching [2,13].

The distortion of distributed voting was first considered by Filos-Ratsikas et al. [14] who focused on bounding the distortion of max-weight mechanisms with respect to the social welfare when agents have normalized valuations for the alternatives. Filos-Ratsikas and Voudouris [12] then studied the distortion of mechanisms for a distributed facility location setting where the agents are positioned on a line and the goal is to choose a single location from a set of alternative ones, which can be discrete (finite) or continuous (infinite). The discrete variant was studied further and generalized to arbitrary metric spaces by Anshelevich et al. [5], who also introduced and studied the average-of-max and max-of-average costs for the first time. The distortion of distributed mechanisms for the continuous variant on a line was recently resolved by Filos-Ratsikas et al. [15]. We refer the reader to the recent survey of Anshelevich et al. [4] for more details on the distortion of voting mechanisms in different models, and to the survey of Chan et al. [10] for details on related facility location models.

2. Average-of-max cost

Let $\alpha \in [0, 1]$ be a parameter. We consider the $\alpha$-LEFTMOST-OF-RIGHTMOST mechanism, which works as follows. For each district $d \in D$, the mechanism chooses the favorite alternative of the rightmost agent in $d$ as the representative $y_d$. Afterwards, it chooses the $(\alpha \cdot k)$-th leftmost alternative as the overall winner. See Mechanism 1.

**Mechanism 1: $\alpha$-LEFTMOST-OF-RIGHTMOST.**

```
for each district $d \in D$ do
    $y_d :=$ favorite alternative of the rightmost agent in $N_d$;
return $w := (\alpha \cdot k)$-th leftmost representative;
```

**Theorem 2.1.** For average-of-max, the distortion of the $\alpha$-LEFTMOST-OF-RIGHTMOST mechanism is at most $\max \left\{ \frac{2 - \alpha}{2}, \frac{\sqrt{5}}{2} - 1 \right\}$.

**Proof.** Let $w$ be the alternative chosen by the mechanism when given as input an arbitrary instance, and $o$ the optimal alternative. For each district $d$, let $i_d$ be the most distant agent from $w$, and $i_d'$ the most distant agent from $o$. So, $(\text{AVG} \circ \text{MAX})(w) = \frac{1}{k} \sum_{d \in D} \delta(i_d, w)$, and $(\text{AVG} \circ \text{MAX})(o) = \frac{1}{k} \sum_{d \in D} \delta(i_d', o)$. Also, let $k_d$ and $r_d$ denote the leftmost and rightmost agents in $d$, respectively. Note that $i_d, i_d' \in \{k_d, r_d\}$. We consider the following two cases depending on the relative positions of $w$ and $o$.

**Case 1: $o < w$.** By the triangle inequality, and since $\delta(i_d, o) \leq \delta(i_d', o)$, we have

$$(\text{AVG} \circ \text{MAX})(w) = \frac{1}{k} \sum_{d \in D} \delta(i_d, w) \leq \frac{1}{k} \sum_{d \in D} \left( \delta(i_d, o) + \delta(w, o) \right) \leq (\text{AVG} \circ \text{MAX})(o) + \delta(w, o).$$

Now, let $S$ be the set of representatives that are to the right of $w$. Since $w$ is by definition the $\alpha$-leftmost representative, we have that $|S| \geq (1 - \alpha) \cdot k$. For every $d$ such that $y_d \in S$, since $o < w \leq y_d$, agent $r_d$ is closer to $w$ than to $o$, and thus $\delta(r_d, o) \geq \frac{1}{2} \delta(w, o)$. Hence,
\[ (\text{AVG} \circ \text{MAX})(o) \geq \frac{1}{k} \sum_{d \in D} \delta(r_d, o) \geq \frac{1}{k} \cdot |S| \cdot \frac{\delta(w, o)}{2} \geq \frac{1 - \alpha}{2} \cdot \delta(w, o), \]

or, equivalently,

\[ \delta(w, o) \leq \frac{2}{1 - \alpha} \cdot (\text{AVG} \circ \text{MAX})(o). \] (2)

Hence, by (1) and (2), we obtain

\[ (\text{AVG} \circ \text{MAX})(w) \leq \left( 1 + \frac{2}{1 - \alpha} \right) \cdot (\text{AVG} \circ \text{MAX})(o) = \frac{3 - \alpha}{1 - \alpha} \cdot (\text{AVG} \circ \text{MAX})(o). \]

**Case 2:** \( w < o \). Let \( L \) be a set of \( \alpha \cdot k \) districts from the one with the leftmost representative until the one with the \((\alpha \cdot k)\)-th representative (which is \( w \)); denote by \( R \) the set of the remaining \((1 - \alpha) \cdot k \) districts. Observe that:

- For every \( d \in L \), since \( y_d \) is the alternative that is closest to \( r_d \) and \( y_d \leq w < o \), both \( r_d \) and \( d \) are closer to \( w \) than to \( o \). So, \( \delta(i_d, w) \leq \delta(i_d^*, o) \).
- For every \( d \in R \), since \( \delta(i_d, o) \leq \delta(i_d^*, o) \) by the definition of \( i_d^* \), using the triangle inequality, we obtain \( \delta(i_d, w) \leq \delta(i_d, o) + \delta(w, o) \leq \delta(i_d^*, o) + \delta(w, o) \).

Hence,

\[
(\text{AVG} \circ \text{MAX})(w) = \frac{1}{k} \sum_{d \in D} \delta(i_d, w) \\
= \frac{1}{k} \sum_{d \in L} \delta(i_d, w) + \frac{1}{k} \sum_{d \in R} \delta(i_d, w) \\
\leq \frac{1}{k} \sum_{d \in L} \delta(i_d^*, o) + \frac{1}{k} \sum_{d \in R} \left( \delta(i_d^*, o) + \delta(w, o) \right) \\
= (\text{AVG} \circ \text{MAX})(o) + \frac{|R|}{k} \cdot \delta(w, o). \] (3)

Since \( r_d \) is closer to \( w \) than to \( o \) for every \( d \in L \), we also have that

\[
(\text{AVG} \circ \text{MAX})(o) \geq \frac{1}{k} \sum_{d \in L} \delta(r_d, o) \geq \frac{|L|}{2k} \delta(w, o),
\]

or, equivalently,

\[
\delta(w, o) \leq \frac{2k}{|L|} \cdot (\text{AVG} \circ \text{MAX})(o). \] (4)

Therefore, by (3) and (4), we obtain

\[
(\text{AVG} \circ \text{MAX})(w) \leq (\text{AVG} \circ \text{MAX})(o) + \frac{|R|}{k} \cdot (\text{AVG} \circ \text{MAX})(o) = \left( \frac{2}{1 - \alpha} - 1 \right) \cdot (\text{AVG} \circ \text{MAX})(o).
\]

Putting everything together, we obtain an upper bound of \( \max \left\{ \frac{3 - \alpha}{1 - \alpha}, \frac{2}{1 - \alpha} - 1 \right\} \).

Observe that the bound \( \max \left\{ \frac{3 - \alpha}{1 - \alpha}, \frac{2}{1 - \alpha} - 1 \right\} \) consists of two functions of \( \alpha \), one non-decreasing and one non-increasing in \( \alpha \). To minimize the maximum between the two, we need to find the value of \( \alpha \) for which the two functions intersect. So, we need to solve the equation

\[
\frac{3 - \alpha}{1 - \alpha} = \frac{2}{1 - \alpha} - 1 \Leftrightarrow \alpha^2 - 3\alpha + 1 = 0.
\]

Since \( \alpha < 1 \), its solution is \( \alpha = \frac{3 - \sqrt{5}}{2} \). For this value of \( \alpha \), both functions have value \( \frac{2}{1 - \alpha^2} - 1 = 2 + \sqrt{5} \), and we obtain the following corollary.

**Corollary 2.2.** For average-of-max, the distortion of the \( \frac{3 - \sqrt{5}}{2} \)-leftmost-of-rightmost mechanism is at most \( 2 + \sqrt{5} \).

### 3. Max-of-average cost

Let \( \alpha \in [0, 1] \) be a parameter. We consider the RIGHTMOST-OF-\( \alpha \)-LEFTMOST mechanism, which works as follows. For each district \( d \in D \), the mechanism chooses the favorite alternative of the \((\alpha \cdot n_d)\)-th agent in \( d \) as the representative \( y_d \). Afterwards, it chooses the rightmost alternative as the overall winner. See Mechanism 2.

**Mechanism 2:** **RIGHTMOST-OF-\( \alpha \)-LEFTMOST.

for each district \( d \in D \) do
\[
y_d := \text{favorite alternative of the } (\alpha \cdot n_d)\text{-th leftmost agent in } N_d;
\]
return \( w := \text{rightmost representative}; \)

**Theorem 3.1.** For max-of-average, the distortion of the RIGHTMOST-OF-\( \alpha \)-LEFTMOST mechanism is at most \( \max \left\{ \frac{3 - \alpha}{1 - \alpha}, \frac{2}{1 - \alpha} - 1 \right\} \).

**Proof.** Let \( w \) be the alternative chosen be the mechanism when given as input an arbitrary instance, and \( o \) the optimal alternative. For any district \( d \) and alternative \( x \), let \( \text{AVG}_d(x) = \frac{1}{n_d} \sum_{i \in N_d} \delta(i, x) \) be the total average distance of the agents in \( d \) for alternative \( x \). So, \( \text{AVG}_d(o) \leq (\text{MAX} \circ \text{AVG})(o) \) for every district \( d \). Denote by \( d^* \) a district that gives the max average cost for \( w \), such that \( (\text{MAX} \circ \text{AVG})(w) = \text{AVG}_{d^*}(w) \). Also, let \( d_w \) be a district represented by \( w \). In addition, let \( i^o \) and \( i_w \) be the \((\alpha \cdot n_d)\)-th leftmost agents in districts \( d^* \) and \( d_w \), respectively. We now switch between the following two cases.

**Case 1:** \( o < w \). By the definition of \( d^* \) and the triangle inequality, we have

\[
(\text{MAX} \circ \text{AVG})(w) = \frac{1}{n_d} \sum_{i \in N_{d^*}} \delta(i, w) \leq \frac{1}{n_d} \sum_{i \in N_{d^*}} \left( \delta(i, o) + \delta(o, w) \right) \leq (\text{MAX} \circ \text{AVG})(o) + \delta(o, w). \] (5)

Denote by \( S \) the set of agents that are positioned weakly to the right of \( i_w \) in \( d_w \). By the definition of \( i_w \), \( |S| \leq (1 - \alpha)n_{d_w} \). Since \( o < w \) and \( w \) is the favorite alternative of \( i_w \), all agents in \( S \) are closer to \( w \) than to \( o \), and thus \( \delta(i, o) \geq \frac{1}{2} \delta(w, o) \) for any \( i \in S \). Using all these, we obtain:

\[
\text{AVG}_{d_w}(o) = \frac{1}{n_{d_w}} \sum_{i \in N_{d_w}} \delta(i, o) \geq \frac{1}{n_{d_w}} \sum_{i \in S} \delta(i, o) \geq \frac{1}{n_{d_w}} |S| \cdot \delta(w, o) \geq \frac{1 - \alpha}{2} \cdot \delta(w, o),
\]

Observe that the bound \( \max \left\{ \frac{3 - \alpha}{1 - \alpha}, \frac{2}{1 - \alpha} - 1 \right\} \) consists of two functions of \( \alpha \), one non-decreasing and one non-increasing in \( \alpha \). To minimize the maximum between the two, we need to find the value of \( \alpha \) for which the two functions intersect. So, we need to solve the equation

\[
\frac{3 - \alpha}{1 - \alpha} = \frac{2}{1 - \alpha} - 1 \Leftrightarrow \alpha^2 - 3\alpha + 1 = 0.
\]

Since \( \alpha < 1 \), its solution is \( \alpha = \frac{3 - \sqrt{5}}{2} \). For this value of \( \alpha \), both functions have value \( \frac{2}{1 - \alpha^2} - 1 = 2 + \sqrt{5} \), and we obtain the following corollary.

**Corollary 2.2.** For average-of-max, the distortion of the \( \frac{3 - \sqrt{5}}{2} \)-leftmost-of-rightmost mechanism is at most \( 2 + \sqrt{5} \).
or, equivalently,
\[
\delta(w, o) \leq \frac{2}{1 - \alpha} \cdot \text{AVG}_{w}(o) \leq \frac{2}{1 - \alpha} \cdot (\text{MAX} \circ \text{AVG})(o).
\]

Therefore, by (5) and (6), we obtain
\[
(\text{MAX} \circ \text{AVG})(w) \leq \frac{3 - \alpha}{1 - \alpha} \cdot (\text{MAX} \circ \text{AVG})(o).
\]

**Case 2:** If \( w < o \), let \( L \) be the set of the first \( \alpha \cdot n_{d^*} \) agents of \( d^* \) (from the leftmost agent to \( * \)), and \( R \) be the set of the remaining \((1 - \alpha) n_{d^*} \) agents. As \( w \) is the rightmost representative, \( y_{d^*} \leq w < o \). Since \( y_{d'} \) is the favorite alternative of \( * \), every agent \( i \in L \) prefers \( w \) over \( o \), and thus \( \delta(i, w) = \delta(i, o) \). Using this in combination with the triangle inequality for every agent of \( R \), we have
\[
(\text{MAX} \circ \text{AVG})(w) = \frac{1}{n_{d^*}} \sum_{i \in L} \delta(i, w) + \frac{1}{n_{d^*}} \sum_{i \in R} \delta(i, w) \leq \frac{1}{n_{d^*}} \sum_{i \in L} \delta(i, o) + \frac{1}{n_{d^*}} \sum_{i \in R} \delta(i, o) = \frac{|R|}{n_{d^*}} \cdot \delta(w, o)
\]

Since each agent \( i \in L \) prefers \( w \) over \( o \), we also have that \( \delta(i, w) \geq \frac{1}{2} \delta(w, o) \), and thus
\[
(\text{MAX} \circ \text{AVG})(o) \geq \text{AVG}_{d^*}(o) = \frac{1}{n_{d^*}} \sum_{i \in L} \delta(i, o) \geq \frac{1}{n_{d^*}} \sum_{i \in L} \delta(i, o) \geq \frac{|L|}{2n_{d^*}} \cdot \delta(w, o).
\]

or, equivalently,
\[
\delta(w, o) \leq \frac{2n_{d^*}}{|L|} \cdot (\text{MAX} \circ \text{AVG})(o).
\]

Therefore, by (7) and (8), we obtain
\[
(\text{MAX} \circ \text{AVG})(w) \leq \frac{3 - \alpha}{1 - \alpha} \cdot (\text{MAX} \circ \text{AVG})(o) = \left( \frac{2\alpha - 1}{\alpha} \right) \cdot (\text{MAX} \circ \text{AVG})(o).
\]

**Corollary 3.2.** For max-of-average, the distortion of the Rightmost-of-\( \frac{3 - \alpha}{1 - \alpha} \)-Leftmost mechanism is at most \( 2 + \sqrt{3} \).

**4. Open questions**

In this paper, we showed a tight distortion bound of \( 2 + \sqrt{3} \) with respect to the average-of-max and max-of-average cost functions for the single-winner distributed single-winner voting problem, thus completing the distortion picture with respect to the four basic objectives considered by Anshelevich et al. [5] for the line metric. The most interesting direction for future work is to prove tight distortion bounds for general metric spaces, and also consider other social objectives or settings beyond single-winner voting.

**References**


