

# Social Choice for Eliminating the Worst Candidate under Metric Preferences

November 26, 2020

Jianan Lin (jnl16@fudan.edu.cn),  
Chenhao Wang (chenhwang4-c@my.cityu.edu.hk)

## Abstract

We consider social choice rules under the metric preferences: both voters and candidates are points in a metric space, and each voter prefers candidates that are closer to her to ones that are further away. We study the problem of selecting the least popular candidate to eliminate. Each voter takes the distance to the eliminated one as her utility, and the objective is to maximize the total utility of all voters. The underlying distances can be arbitrary, implicit, and unknown. Based only on the voters' preference rankings that are consistent with the metric distances, a social choice rule may not identify an optimal candidate. The quality of a rule is measured by its distortion, defined as the worst-case ratio between the performance of an optimal candidate and that of the selected one. We present a social choice rule with distortion of 4.236, using a symmetric idea of the Weighted Uncovered rule. Further, we design a new rule with improved distortion of 3.828 for a subclass of instances, including all instances with no more than 6 candidates. Both rules only look at the weighted tournament graph on the candidates. We also contribute to the well-studied problem of selecting a candidate to minimize the total distance to all voters, improving the distortion to 3.828 when there are no more than 6 candidates.

**Keywords:** Social choice; Metric preferences; Distortion

## 1 Introduction

In social choice theory, a *mechanism* (also referred to as a voting rule or a social choice rule) aggregates the preferences of multiple voters over a set of candidates, and returns a candidate as the winner of the election. The preference of a voter is often represented as a linear order over the candidates. There is no consensus on what are the best voting rules, and most of the social choice literature compares different voting rules by defining normative or axiomatic criteria, such as anonymity, neutrality and Pareto efficiency.

An appealing approach to dealing with social choice problems is embedding the input “election” into a *metric space*, *i.e.*, each participant is represented by a point in a metric space, and voters prefer candidates that are closer to them to the ones that are further away. This spatial model has very natural interpretations. For example, in a 2-dimensional Euclidean space, each dimension specifies a political issue (such as military or education), and the position of a voter or candidate

identifies the extent to which the individual supports the issues. Recently, this model has attracted attentions from AI researchers, see, *e.g.*, [2, 11, 1].

As voters' preferences are specified by their distances to candidates, it is natural to quantify the *quality* of the selected candidate by the associated distances. We evaluate the performance of a mechanism in the standard worst-case analysis benchmark (introduced by Procaccia and Rosenschein [19]), which defines the *distortion* of a mechanism to be the worst-case ratio between the quality of an alternative selected by this mechanism and that of the optimal alternative selected by an omniscient mechanism. More generally, the distortion, which stems from an information deficiency, is a measure of how close it always gets to the optimal alternative.

Previous works were mainly concerned about the single-winner election that aims to select one candidate, minimizing the total distance to all voters. In this paper, we focus on the antithesis, the *single-loser* election that aims to select one candidate, maximizing the total distance to all voters. That is, we want to eliminate the least popular candidate. When the context is clear, the selected candidate is also called the winner of the election, even though this is more like a "loser". Every voter wants to stay as far away as possible from this winner. This election is well motivated, for example, some enterprises adopt a *last-out* mechanism in the personnel performance appraisal system, which dismisses the employee with the lowest performance in a department each year. Besides, a voting rule for the single-loser election may be iteratively used as a subroutine for multi-winner elections in TV talent shows, eliminating one candidate at each iteration.

## Our Results and Organizations

We consider the social choice problem under metric preferences, in which all participants are assumed to be points in a metric space, and the preferences are identified by the distances. The commonly studied objective is to minimize the total distance of the selected candidate to all voters, called the social cost objective. In this work, we study the social choice rules for the problem of selecting a candidate with maximum total distance to all voters. As it is only given the ordinal preference orderings, instead of the underlying metric costs which generate these preferences, a rule cannot always select the optimal candidate. We measure the performance of a social choice rule by its distortion, defined as the worst-case ratio between the social value of the optimal candidate and the one returned by the rule.

Section 2 gives the preliminaries. Assume each voter takes the distance to the selected candidate as her utility, and the *social utility* is the sum of all voters' utilities. In this work we consider the problem of maximizing the social utility. Several typical voting rules for the social cost objective in the literatures (*e.g.*, Weighted Uncovered in [17]) are presented: some of them can be applied to our problem with the same distortion, by reversing the input preference orders.

In Section 3, we provide two lemmas that bound the ratio between the social utilities of two candidates. They are the key tools for analysis throughout this paper. Then we adapt the Weighted Uncovered rule for the social utility objective: it selects a candidate  $A$  such that for any other candidate  $B$ , either at least  $z$  fraction of voters prefer  $B$  to  $A$ , or there exists a candidate  $C$  such that at least  $1 - z$  fraction prefers  $C$  to  $A$  and at least  $z$  fraction prefers  $B$  to  $C$ . By setting an appropriate parameter  $z$ , we obtain a voting rule with distortion 4.236 (proved by using the lemmas), the same with that of Weighted Uncovered for the social cost version.

In Section 4, the power of the two lemmas can be used to improve the distortion for a subclass of instances. We propose a voting rule that selects a candidate arbitrarily from a (possibly empty) set, called  $Cov_{su}$ -set, only based on the fraction of voters who prefer  $A$  to  $B$ , for each pair of candidates  $A, B$ . Setting an appropriate parameter, we prove that this rule achieves a distortion

3.828 if the  $Cov_{su}$ -set is non-empty. Although we are not able to identify when this set is empty in general, we prove that this set contains at least one candidate for a subclass of instances, including all instances with no more than 6 candidates. Hence, we improve the distortion from 4.236 to 3.828 for such a subclass, and leave the application range of this rule as an open question.

In Section 5 we extend the results in Section 4 to the well-studied problem for social cost objective, and provide improved distortion 3.828 when there are 5 or 6 candidates (the best known distortion is 4.236 by Weighted Uncovered). Section 6 discusses egalitarian objectives and concludes this paper.

## Related Work

In social choice theory, Procaccia and Rosenschein [19] propose a utilitarian approach – the implicit utilitarian voting – by assuming that voters have latent cardinal utilities and report ordinal preferences induced by them. They measure the performance of popular voting rules by the notion of *distortion*. Subsequently, Caragiannis and Procaccia [8], Oren and Lucier [18], Boutilier *et al.* [7], Bhaskar and Ghosh [6] employ this notion and design selection rules with low distortions.

Anshelevich *et al.* [2] first embed the social choice into a metric space, in which the participants are points, and the costs are driven by the distances. For the social cost objective of selecting a candidate to minimize the total distance to all voters, they show that the uncovered set rules [16] (including the Copeland rule) can guarantee distortion at most 5. They also prove that any deterministic rule has a lower bound for distortion of at least 3. Later, Skowron and Elkind [21] show that the class of scoring rules and STV have super-constant distortion. Goel *et al.* [15] show the ranked pairs rule and the Schulze method working on weighted tournament graph on the candidates both have distortion of at least 5. Recently, Munagala and Wang [17] propose a Weighted Uncovered rule with distortion  $2 + \sqrt{5} \approx 4.236$ . They additionally give a conjecture that the perfect matching rule can always guarantee a distortion at most 3, matching the lower bound. All above mentioned rules are deterministic and work for the social cost objective.

Randomized social choice rules have also been studied in the metric distortion setting, which output a probability distribution over the set of candidates. The lower bound on distortion of randomized rules is 2 [3] [12]. Random dictatorship that randomly selects the top choice of one of the  $m$  voters gets distortion  $3 - 2/m$  [3] [12]. Gross *et al.* [15] propose a simple mechanism that randomly asks voters for their top choice until two voters agree, achieving very low distortion when the number of candidates is not large. However, randomized social choice rules attract less attention than deterministic ones in practise.

Some works focus on other objectives or properties. The work of Anshelevich *et al.* [2] considers the objective of minimizing the median cost of agents, and proves that the Copeland rule achieves the optimal distortion of 5. Fairness property is also considered [13] [17].

*Other problems.* Some works study other social choice problems in the metric distortion setting. When candidates are independently drawn from the population of voters, Cheng *et al.* show that the distortion can be better than 2 [9], and the class of scoring rules with constant expected distortion has a clean characterization [10]. The work of Anshelevich and Zhu [4] studies the social choice and facility location problems in the setting with known locations of candidates. The work of Anshelevich and Zhu [5] proposed a low-distortion algorithm for maximum bipartite matching problem in metric spaces. Very recently, Abramowitz *et al.* [1] consider the assumption that a (very small) amount of information is known about the voter preference strengths, not just about their ordinal preferences.

## 2 Preliminaries

### 2.1 Social Choice Rules

Let  $\mathcal{C} = \{A_1, A_2, \dots, A_n\}$  be the set of  $n$  candidates, and  $\mathcal{V}$  be the set of  $m$  voters (agents). Let  $S = \mathcal{C} \cup \mathcal{V}$  be the set of all  $m + n$  participants in this social choice system. Every voter  $v \in \mathcal{V}$  has a strict *preference ranking*  $\sigma_v$  on the candidates in  $\mathcal{C}$ . By  $X \succ_v Y$  or  $Y \prec_v X$ , we mean that  $X$  is preferred over  $Y$  in the ranking  $\sigma_v$ . We use  $\sigma = \{\sigma_v\}_{v \in \mathcal{V}}$  to denote the *preference profile*. Let  $\mathcal{P}$  be the set of all possible preference profiles.

Given a preference profile  $\sigma \in \mathcal{P}$ , we want to aggregate the preferences of the voters and select a single candidate as the winner of the election. A (deterministic) social choice function  $f : \mathcal{P} \rightarrow \mathcal{C}$  is a mapping from a preference profile to a candidate. It is also referred to as a voting rule or a mechanism.

In this paper, we will typically use upper-case letters  $A, B, C, X, Y$  to refer to candidates, and lower-case letters  $u, v$  to refer to voters. Let  $AB = \{v \in \mathcal{V} \mid A \succ_v B\}$  denote the set of voters who prefer  $A$  to  $B$ . Let  $ABC = \{v \in \mathcal{V} \mid A \succ_v B, B \succ_v C\}$  denote the set of voters that prefer  $A$  to  $B$  and prefer  $B$  to  $C$ . We will use  $|AB|$  to denote the ratio between the size of  $AB$  and the number  $m$  of voters, but not the size of  $AB$ . That is,  $|AB|$  is the fraction of voters who prefer  $A$  to  $B$ , and we have  $|AB| + |BA| = 1$ . Also,  $|ABC|$  is the ratio between the size of  $ABC$  and  $m$ .

### 2.2 Social Utility and Distortion

We assume that all candidates and voters are located in a metric space  $(S, d)$ , where metric  $d : S \times S \rightarrow \mathbb{R}_{\geq 0}$  defines a distance. The voters prefer candidates that are closer to them to the ones that are further away. We say a metric  $d$  is *consistent* with a preference profile  $\sigma$ , if

$$\forall v \in \mathcal{V}, \forall A, B \in \mathcal{C} : A \succ_v B \iff d(A, v) \leq d(B, v).$$

Here we allow a voter to have equal distances to different candidates. Define  $\rho(\sigma)$  to be the set of all metrics consistent with  $\sigma$ .

Most literature study the objective of minimizing the *social cost*, in which the social cost of a candidate  $A$  for metric  $d$  is the total distance to all voters:

$$S(A, d) = \sum_{v \in \mathcal{V}} d(v, A).$$

The *distortion* of a social choice rule  $f$  is the worst-case ratio between the social cost induced by  $f$  and the optimum:

$$\Delta_c(f) = \sup_{\sigma} \sup_{d \in \rho(\sigma)} \frac{S(f(\sigma), d)}{\min_{X \in \mathcal{C}} S(X, d)}.$$

Our work focuses on the antithesis, selecting a candidate to maximize the total distance to all voters. This models the scenario that we want to eliminate a single candidate. Each voter wants to stay as far away as possible from the winner, and takes the distance as her utility. The *social utility* of a candidate  $A$  for metric  $d$  is the total utility of all voters, also denoted by  $S(A, d)$ . Under the objective of maximizing the social utility, the distortion of a social choice rule  $f$  is the worst-case ratio between the optimal social utility and that induced by  $f$ :

$$\Delta_u(f) = \sup_{\sigma} \sup_{d \in \rho(\sigma)} \frac{\max_{X \in \mathcal{C}} S(X, d)}{S(f(\sigma), d)}.$$

For the social cost objective, no voting rule has distortion less than 3 (see [2]). This lower bound also holds for the social utility objective, which can be proved by the same analysis: Suppose there are two candidate  $X$  and  $Y$ , and two voters  $u$  and  $v$ ;  $u$  prefers  $X$  and  $v$  prefers  $Y$ . Assume w.l.o.g. that the given rule selects  $Y$ . The underlying metric is Euclidean distance, and  $X, Y, u, v$  are located at points 0, 2, 1, 2 on the real line, respectively. Then  $S(X, d) = 3$  and  $S(Y, d) = 1$ , indicating a distortion at least 3. It is easy to see that the lower bound 3 holds for any  $n$  and even number  $m$ .

## 2.3 Examples of Social Choice Rules

We construct a *tournament graph* on the candidates: let  $\mathcal{C}$  be the set of vertices, and for each pair of  $A, B \in \mathcal{C}$ , there is a directed edge  $(A, B)$  if  $A$  pairwise beats  $B$ :  $|AB| > \frac{m}{2}$ . A voting rule is a *C1 rule* if it makes decisions only based on the tournament graph. We can also construct a *weighted tournament graph* on the candidates, where for each pair of  $A, B \in \mathcal{C}$ , there is a directed edge  $(A, B)$  with weight  $|AB|$ . A voting rule is a *C2 rule* if it makes decisions only based on the weighted tournament graph.

Now we present some voting rules for the social cost objective known in the literature.

*Copeland*. This rule selects the candidate with the largest out-degree in the tournament graph. It is a C1 rule, and has distortion of 5 [2].

*Ranked Pairs*. It is a C2 rule proposed in [22]. Sort all edges in the weighted tournament graph with descending weights. We start from an empty graph and add one edge at each time in the sorted order without creating a cycle. The rule selects the source of the resulted acyclic graph in the end. Generally, the distortion is at least 5 [14]. When the size of the largest cycle in the tournament graph is at most 4, it has distortion of 3 [2].

*Weighted Uncovered*. It is a C2 rule proposed in [17]. It selects a candidate  $A$ , such that for every candidate  $B \neq A$ , either  $|AB| \geq \frac{3-\sqrt{5}}{2}$  or there exists another candidate  $C \neq A, B$  satisfying  $|AC| \geq \frac{3-\sqrt{5}}{2}$  and  $|CB| \geq \frac{\sqrt{5}-1}{2}$ . This rule has the best known distortion of  $2 + \sqrt{5}$ .

Other voting rules for the social cost objective include *Schulze* [20], *Uncovered* [2], *Perfect Matching* [17] and positional scoring rules.

The difference between the problems under the social cost and the social utility objectives lies in that one is a minimization problem and the other is a maximization problem. A natural idea for tackling the social-utility version is taking the *reversed* preference rankings as input, and implementing the existing voting rules for the social-cost version. This approach maintains the distortion of Copeland and Ranked Pairs, but not Weighted Uncovered. We will explore the distortion of C2 rules in this paper.

## 3 Maximizing Social Utility

In this section, we start with two key lemmas, which upper bound the ratio between the social utilities of two candidates, with respect to the edge weights in the weighted tournament graph. These two lemmas play an important role throughout this paper. We present a voting rule for the social utility objective and prove a distortion at most  $2 + \sqrt{5}$ , based on these two lemmas. This rule is an adaptation of the Weighted-Uncovered rule proposed in [17], which works for the single-winner election under the social cost objective (i.e., minimizing the total distance to all voters), and has the same distortion. We give an example to show the tightness of the distortion for our rule.

### 3.1 Two Key Lemmas

**Lemma 3.1.** *If  $|AB| > 0$ , then for any metric  $d$ , we have*

$$\frac{S(A, d)}{S(B, d)} \leq \frac{2}{|AB|} - 1.$$

*Proof.* We have

$$\begin{aligned} & \left(\frac{2}{|AB|} - 1\right) \cdot S(B, d) - S(A, d) \\ &= \sum_{v \in AB} \left[\left(\frac{2}{|AB|} - 1\right) \cdot d(v, B) - d(v, A)\right] \\ &+ \sum_{v \in BA} \left[\left(\frac{2}{|AB|} - 1\right) \cdot d(v, B) - d(v, A)\right] \\ &\geq \sum_{v \in AB} \left(\frac{2}{|AB|} - 2\right) \cdot d(v, B) - \sum_{v \in BA} d(A, B) \\ &\geq \sum_{v \in AB} \left(\frac{1}{|AB|} - 1\right) \cdot d(A, B) - \sum_{v \in BA} d(A, B) \\ &= \sum_{v \in AB} \frac{1}{|AB|} \cdot d(A, B) - \sum_{v \in V} d(A, B) \\ &= m \cdot |AB| \cdot \frac{1}{|AB|} \cdot d(A, B) - m \cdot d(A, B) = 0. \end{aligned}$$

□

**Lemma 3.2.** *If  $|AC|, |CB| > 0$ , then for any metric  $d$ , we have*

$$\frac{S(A, d)}{S(B, d)} \leq \max\left\{\frac{2}{|AC|} - 1, \frac{2(1 - |AB|)}{|CB|} + 1\right\}.$$

*Proof Sketch.* Assume for contradiction that  $\frac{S(A, d)}{S(B, d)} > \max\left\{\frac{2}{|AC|} - 1, \frac{2(1 - |AB|)}{|CB|} + 1\right\}$ . We discuss two cases:  $d(A, B) \leq d(B, C)$  and  $d(A, B) > d(B, C)$ .

If  $d(A, B) \leq d(B, C)$ , we have

$$\begin{aligned} & (2 + |CB| - 2|AB|) \cdot S(B, d) - |CB| \cdot S(A, d) \\ &\geq \{|AB| - |AB|^2 - |BA| \cdot (|CB| - |BCA|) \cdot |BA|\} \cdot d(A, B) \\ &\geq 0, \end{aligned}$$

implying  $\frac{S(A, d)}{S(B, d)} \leq \frac{2(1 - |AB|)}{|CB|} + 1$ .

If  $d(A, B) > d(B, C)$ , we have

$$\begin{aligned} & (2 - |AC|) \cdot S(B, d) - |AC| \cdot S(A, d) \\ &\geq \{|AB| - |AC| + |CBA| \cdot (1 - |AC|)\} \cdot d(B, C). \end{aligned}$$

Define  $\Delta = |AB| - |AB|^2 - |CB| \cdot |CBA| + |BAC| \cdot (1 + |CB| - |AB|) - |CB| \cdot |BCA|$ . By the assumption  $\frac{S(A,d)}{S(B,d)} > \frac{2}{|AC|} - 1$ , it derives  $|AB| - |AC| + |CBA| \cdot (1 - |AC|) < 0$  and  $\Delta \geq 0$ . Furthermore, we have

$$\begin{aligned} & (2 + |CB| - 2|AB|) \cdot S(B, d) - |CB| \cdot S(A, d) \\ & \geq \Delta \cdot d(A, B) + \\ & \quad \{|CBA| \cdot (1 - |AB|) - |BAC| \cdot (1 - |AB|)\} \cdot d(B, C). \end{aligned}$$

Since  $\Delta \geq 0$ , it is easy to see that  $(2 + |CB| - 2|AB|) \cdot S(B, d) - |CB| \cdot S(A, d) \geq 0$ , implying  $\frac{S(A,d)}{S(B,d)} \leq \frac{2(1-|AB|)}{|CB|} + 1$  as desired.

Therefore we can confirm that whenever  $\frac{S(A,d)}{S(B,d)} > \frac{2}{|AC|} - 1$ ,  $\frac{S(A,d)}{S(B,d)} \leq \frac{2(1-|AB|)}{|CB|}$  will hold.  $\square$

*Proof.* We perform different analysis for the situation  $d(A, B) \leq d(B, C)$  and  $d(A, B) > d(B, C)$ .

Case(1)  $d(A, B) \leq d(B, C)$ :

$$\begin{aligned}
& (2 + |CB| - 2|AB|) \cdot S(B, d) - |CB| \cdot S(A, d) \\
&= \sum_{v \in AB} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} + \sum_{v \in CBA} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} \\
&+ \sum_{v \in BAC} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} + \sum_{v \in BCA} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} \\
&= \sum_{v \in AB} \{(2 - 2|AB|) \cdot d(v, B) + |CB| \cdot [d(v, B) - d(v, A)]\} \\
&+ \sum_{v \in CBA} \{(2 - 2|AB|) \cdot d(v, B) + |CB| \cdot [d(v, B) - d(v, A)]\} \\
&+ \sum_{v \in BAC} \{(2 - 2|AB|) \cdot d(v, B) + |CB| \cdot [d(v, B) - d(v, A)]\} \\
&+ \sum_{v \in BCA} \{(2 - 2|AB|) \cdot d(v, B) + |CB| \cdot [d(v, B) - d(v, A)]\} \\
&\geq \sum_{v \in AB} \{(2 - 2|AB|) \cdot d(v, B)\} + \sum_{v \in CBA} \{(2 - 2|AB|) \cdot d(v, B) - |CB| \cdot d(A, B)\} \\
&+ \sum_{v \in BAC} \{|CB| \cdot [d(v, B) - d(v, A)]\} + \sum_{v \in BCA} \{(2 - 2|AB|) \cdot d(v, B) - |CB| \cdot d(A, B)\} \\
&\geq \sum_{v \in AB} \{(1 - |AB|) \cdot d(A, B)\} + \sum_{v \in CBA} \{(1 - |AB|) \cdot d(B, C) - |CB| \cdot d(A, B)\} \\
&- \sum_{v \in BAC} \{|CB| \cdot d(A, B)\} + \sum_{v \in BCA} \{(1 - |AB|) \cdot [(d(v, B) + d(v, C)) + (d(v, B) - d(v, A))] - |CB| \cdot d(A, B)\} \\
&\geq \sum_{v \in AB} \{(1 - |AB|) \cdot d(A, B)\} + \sum_{v \in CBA} \{(1 - |AB|) \cdot d(B, C) - |CB| \cdot d(A, B)\} \\
&- \sum_{v \in BAC} \{|CB| \cdot d(A, B)\} + \sum_{v \in BCA} \{(1 - |AB|) \cdot [d(B, C) - d(A, B)] - |CB| \cdot d(A, B)\} \\
&= |AB| \cdot \{(1 - |AB|) \cdot d(A, B)\} + |CBA| \cdot \{(1 - |AB|) \cdot d(B, C) - |CB| \cdot d(A, B)\} \\
&- |BAC| \cdot \{|CB| \cdot d(A, B)\} + |BCA| \cdot \{(1 - |AB|) \cdot [d(B, C) - d(A, B)] - |CB| \cdot d(A, B)\} \\
&= \{|AB| - |AB|^2 - (|CBA| + |BAC| + |BCA|) \cdot |CB| + |BCA| \cdot (|AB| - 1)\} \cdot d(A, B) \\
&+ (1 - |AB|) \cdot (|CBA| + |BCA|) \cdot d(B, C) \\
&\geq \{|AB| - |AB|^2 - |BA| \cdot |CB| - |BCA| \cdot |BA|\} \cdot d(A, B) + |BA| \cdot (|CBA| + |BCA|) \cdot d(A, B) \\
&\geq [|AB| - |AB|^2 - |BA| \cdot (|CB| - |CBA|)] \cdot d(A, B) \\
&\geq [|AB| - |AB|^2 - (1 - |AB|) \cdot |AB|] \cdot d(A, B) \\
&= 0.
\end{aligned}$$

Case(2)  $d(A, B) > d(B, C)$ :



First we have:

$$\begin{aligned}
& (2 + |CB| - 2|AB|) \cdot S(B, d) - |CB| \cdot S(A, d) \\
&= \sum_{v \in AB} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} + \sum_{v \in CBA} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} \\
&+ \sum_{v \in BAC} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} + \sum_{v \in BCA} \{(2 + |CB| - 2|AB|) \cdot d(v, B) - |CB| \cdot d(v, A)\} \\
&= \sum_{v \in AB} \{(2 - 2|AB|) \cdot d(v, B) + |CB| \cdot [d(v, B) - d(v, A)]\} \\
&+ \sum_{v \in CBA} \{(2 - 2|AB|) \cdot d(v, B) + |CB| \cdot [d(v, B) - d(v, A)]\} \\
&+ \sum_{v \in BAC} \{(2 + |CB| - 2|AB|) \cdot [d(v, B) + d(v, A)] - (2 - 2|AB|) \cdot d(v, A)\} \\
&+ \sum_{v \in BCA} \{(2 - 2|AB|) \cdot d(v, B) + |CB| \cdot [d(v, B) - d(v, A)]\} \\
&\geq \sum_{v \in AB} \{(2 - 2|AB|) \cdot d(v, B)\} + \sum_{v \in CBA} \{(2 - 2|AB|) \cdot d(v, B) - |CB| \cdot d(A, B)\} \\
&+ \sum_{v \in BAC} \{(2 + |CB| - 2|AB|) \cdot d(A, B) - (2 - 2|AB|) \cdot [d(A, C) - d(v, C)]\} + \sum_{v \in BCA} \{|CB| \cdot [d(v, B) - d(v, A)]\} \\
&\geq \sum_{v \in AB} \{(1 - |AB|) \cdot d(A, B)\} + \sum_{v \in CBA} \{(1 - |AB|) \cdot d(B, C) - |CB| \cdot d(A, B)\} \\
&+ \sum_{v \in BAC} \{(2 + |CB| - 2|AB|) \cdot d(A, B) - (1 - |AB|) \cdot d(A, C)\} - \sum_{v \in BCA} \{|CB| \cdot d(A, B)\} \\
&\geq \sum_{v \in AB} \{(1 - |AB|) \cdot d(A, B)\} + \sum_{v \in CBA} \{(1 - |AB|) \cdot d(B, C) - |CB| \cdot d(A, B)\} \\
&+ \sum_{v \in BAC} \{(2 + |CB| - 2|AB|) \cdot d(A, B) - (1 - |AB|) \cdot [d(A, B) + d(B, C)]\} - \sum_{v \in BCA} \{|CB| \cdot d(A, B)\} \\
&= |AB| \cdot \{(1 - |AB|) \cdot d(A, B)\} + |CBA| \cdot \{(1 - |AB|) \cdot d(B, C) - |CB| \cdot d(A, B)\} \\
&+ |BAC| \cdot \{(1 + |CB| - |AB|) \cdot d(A, B) - (1 - |AB|) \cdot d(B, C)\} - |BCA| \cdot \{|CB| \cdot d(A, B)\} \\
&= \{|AB| - |AB|^2 - |CB| \cdot |CBA| + |BAC| \cdot (1 + |CB| - |AB|) - |CB| \cdot |BCA|\} \cdot d(A, B) \\
&+ \{|CBA| \cdot (1 - |AB|) - |BAC| \cdot (1 - |AB|)\} \cdot d(B, C)
\end{aligned}$$

If  $\{|AB| - |AB|^2 - |CB| \cdot |CBA| + |BAC| \cdot (1 + |CB| - |AB|) - |CB| \cdot |BCA|\} \geq 0$ , then

$$\begin{aligned}
& (2 + |CB| - 2|AB|) \cdot S(B, d) - |CB| \cdot S(A, d) \\
&\geq \{|AB| - |AB|^2 - |CB| \cdot |BA| + |BAC| \cdot (1 + 2|CB| - |AB|)\} \cdot d(B, C) \\
&+ \{(1 - |AB|) \cdot (|CBA| - |BAC|)\} \cdot d(B, C) \\
&= \{(1 - |AB|) \cdot (|AB| + |CBA|) + |CB| \cdot (2|BAC| - |BA|)\} \cdot d(B, C) \\
&\geq \{(1 - |AB|) \cdot |CB| + |CB| \cdot (2|BAC| - |BA|)\} \cdot d(B, C) \\
&= (1 - |AB| - |BA|) \cdot |CB| + 2|CB| \cdot |BAC| \\
&= 2|CB| \cdot |BAC| \\
&\geq 0
\end{aligned}$$

Second we have:

$$\begin{aligned}
& (2 - |AC|) \cdot S(B, d) - |AC| \cdot S(A, d) \\
&= \sum_{v \in AB} \{(2 - |AC|) \cdot d(v, B) - |AC| \cdot d(v, A)\} + \sum_{v \in CBA} \{(2 - |AC|) \cdot d(v, B) - |AC| \cdot d(v, A)\} \\
&+ \sum_{v \in BAC} \{(2 - |AC|) \cdot d(v, B) - |AC| \cdot d(v, A)\} + \sum_{v \in BCA} \{(2 - |AC|) \cdot d(v, B) - |AC| \cdot d(v, A)\} \\
&= \sum_{v \in AB} \{(2 - 2|AC|) \cdot d(v, B) + |AC| \cdot [d(v, B) - d(v, A)]\} \\
&+ \sum_{v \in CBA} \{(2 - 2|AC|) \cdot d(v, B) + |AC| \cdot [d(v, B) - d(v, A)]\} \\
&+ \sum_{v \in BAC} \{(2 - |AC|) \cdot [d(v, B) + d(v, A)] \cdot 2 \cdot d(v, A)\} \\
&+ \sum_{v \in BCA} \{(2 - 2|AC|) \cdot d(v, B) + |AC| \cdot [d(v, B) - d(v, A)]\} \\
&\geq \sum_{v \in AB} \{(2 - 2|AC|) \cdot d(v, B)\} + \sum_{v \in CBA} \{(2 - 2|AC|) \cdot d(v, B) - |AC| \cdot d(A, B)\} \\
&+ \sum_{v \in BAC} \{(2 - |AC|) \cdot d(A, B) - 2 \cdot [d(A, C) - d(v, C)]\} + \sum_{v \in BCA} \{|AC| \cdot [d(v, B) - d(v, A)]\} \\
&\geq \sum_{v \in AB} \{(1 - |AC|) \cdot d(A, B)\} + \sum_{v \in CBA} \{(1 - |AC|) \cdot d(B, C) - |AC| \cdot d(A, B)\} \\
&+ \sum_{v \in BAC} \{(2 - |AC|) \cdot d(A, B) - d(A, C)\} - \sum_{v \in BCA} \{|AC| \cdot d(A, B)\} \\
&\geq \sum_{v \in AB} \{(1 - |AC|) \cdot d(A, B)\} + \sum_{v \in CBA} \{(1 - |AC|) \cdot d(B, C) - |AC| \cdot d(A, B)\} \\
&+ \sum_{v \in BAC} \{(2 - |AC|) \cdot d(A, B) - [d(A, B) + d(B, C)]\} - \sum_{v \in BCA} \{|AC| \cdot d(A, B)\} \\
&\geq |AB| \{(1 - |AC|) \cdot d(A, B)\} + |CBA| \{(1 - |AC|) \cdot d(B, C) - |AC| \cdot d(A, B)\} \\
&+ |BAC| \{(2 - |AC|) \cdot d(A, B) - [d(A, B) + d(B, C)]\} - |BCA| \{|AC| \cdot d(A, B)\} \\
&= \{|AB| - |AB| \cdot |AC| - |AC| \cdot |CBA| + |BAC| \cdot (1 - |AC|) - |AC| \cdot |BCA|\} \cdot d(A, B) \\
&+ \{|CBA| \cdot (1 - |AC|) - |BAC|\} \cdot d(B, C)
\end{aligned}$$

Because we know that  $|AB| - |AC| + |BAC| \geq 0$ , and we can get that

$$\{|AB| - |AB| \cdot |AC| - |AC| \cdot |CBA| + |BAC| \cdot (1 - |AC|) - |AC| \cdot |BCA|\} \geq 0,$$

thus

$$\begin{aligned}
& (2 - |AC|) \cdot S(B, d) - |AC| \cdot S(A, d) \\
&\geq \{|AB| - |AB| \cdot |AC| - |AC| \cdot |CBA| + |BAC| \cdot (1 - |AC|) - |AC| \cdot |BCA|\} \cdot d(B, C) \\
&+ \{|CBA| \cdot (1 - |AC|) - |BAC|\} \cdot d(B, C) \\
&= \{|AB| - |AB| \cdot |AC| - |AC| \cdot |BA| + |BAC| + |CBA| \cdot (1 - |AC|) - |BAC|\} \cdot d(B, C) \\
&= \{|AB| - |AC| + |CBA| \cdot (1 - |AC|)\} \cdot d(B, C)
\end{aligned}$$

If  $|AB| - |AC| + |CBA| \cdot (1 - |AC|) \geq 0$ , then  $(2 - |AC|) \cdot S(B, d) - |AC| \cdot S(A, d) \geq 0$ . Thus if  $(2 - |AC|) \cdot S(B, d) - |AC| \cdot S(A, d) < 0$ , then  $|AB| - |AC| + |CBA| \cdot |CA| < 0$ .

Let  $|ABC| = a, |BCA| = b, |CAB| = c, |ACB| = d, |CBA| = e, |BAC| = f$ , and we can conclude that  $|CAB| - |BAC| + |CBA| \cdot |CA| < 0$ , which means that  $c - f + e \cdot (b + c + e) < 0$ . So  $be + e^2 < f - c - ce$ . At this time, thus according to the first situation in case(2), we have

$$\begin{aligned}
& (2 + |CB| - 2|AB|) \cdot S(B, d) - |CB| \cdot S(A, d) \\
&= |AB| - |AB|^2 - |CB| \cdot |CBA| + |BAC| \cdot (1 + |CB| - |AB|) - |CB| \cdot |BCA| \\
&= |AB| \cdot |BA| - |CB| \cdot |BA| + |BAC| \cdot (1 + 2|CB| - |AB|) \\
&= |BA| \cdot (|ABC| - |CBA|) + |BAC| \cdot (1 + |CB| + |CBA| - |ABC|) \\
&= (b + e + f) \cdot (a - e) + f \cdot (1 + c + d + 2e - a) \\
&= (ab + ae + af + cf + df + 2ef + f) - (af + be + e^2 + ef) \\
&= (ab + ae + cf + df + ef + f) - (be + e^2) \\
&\geq (ab + ae + cf + df + ef + f) - (f - c - ce) \\
&= ab + ae + c + ce + cf + df + ef \\
&\geq 0
\end{aligned}$$

And now, we prove that:

$$\frac{S(A, d)}{S(B, d)} \leq \max\left\{\frac{2}{|AC|} - 1, \frac{2(1 - |AB|)}{|CB|} + 1\right\}.$$

□

### 3.2 A C2 Voting Rule

The work of Munagala and Wang [17] proposes the Weighted Uncovered rule for the single-winner election under the social cost objective, aiming to minimize the total distance to voters. This rule selects a candidate  $A$ , such that for any other candidate  $B$ , the winner  $A$  covers  $B$  directly, or indirectly via some intermediary candidate. We adopt a dual idea for our problem, and prove the distortion based on the above upper bounds on the ratio of the social utilities of two candidates.

**Mechanism 1.** (ADAPTED WEIGHTED UNCOVERED.) Let  $z \in [0, 0.5]$  be a constant. Select a candidate  $A \in \mathcal{C}$  such that, for any candidate  $B \neq A$ , we have

1. either  $|BA| \geq z$ ,
2. or there exists an intermediary candidate  $C \neq A, B$  such that  $|CA| \geq 1 - z$  and  $|BC| \geq z$ .

**Lemma 3.3.** *There exists a candidate satisfying the conditions in Mechanism 1, for any  $z \in [0, 0.5]$ .*

*Proof.* Suppose there is no such candidate. Then for every candidate  $A$ , there exists a candidate  $B \neq A$  such that  $|BA| < z$  and for any candidate  $C$ , either  $|CA| < 1 - z$  or  $|BC| < z$ ; it is referred to as  $A \triangleright B$ . There must be some voters  $A_1, A_2, \dots, A_k, A_{k+1} = A_1$ , satisfying  $A_i \triangleright A_{i+1}$  ( $i = 1, 2, \dots, k$ ).

We have  $|A_2 A_1| < z$  and  $|A_k A_1| > 1 - z \geq z$ . Then there is some index  $j$  such that  $|A_{j-1} A_1| < z$  and  $|A_j A_1| > z$ . Hence, candidate  $A_1$  can be an intermediary for  $A_{j-1}$  and  $A_j$ , which contradicts with the assumption  $A_{j-1} \triangleright A_j$ . □

Now we show the distortion of Adapted Weighted Uncovered with appropriate parameter  $z$ .

**Theorem 3.4.** *The Adapted Weighted Uncovered rule with  $z = \frac{3-\sqrt{5}}{2}$  has distortion of at most  $2 + \sqrt{5} \approx 4.236$  under the social utility objective.*

*Proof.* Let  $A \in \mathcal{C}$  be the selected candidate, where the existence is guaranteed by Lemma 3.3. Consider an arbitrary candidate  $B \neq A$ . If the first condition in Mechanism 1 (i.e.,  $|BA| \geq z$ ) is satisfied, by Lemma 3.1, we have  $\frac{S(B,d)}{S(A,d)} \leq \frac{2}{z} - 1 = 2 + \sqrt{5}$ .

Otherwise, there is an intermediary  $C$  such that the second condition (i.e.,  $|CA| \geq 1 - z$  and  $|BC| \geq z$ ) is satisfied. By Lemma 3.2, we have

$$\begin{aligned} \frac{S(B,d)}{S(A,d)} &\leq \max\left\{\frac{2}{|BC|} - 1, \frac{2(1 - |BA|)}{|CA|} + 1\right\} \\ &\leq \max\left\{\frac{2}{z} - 1, \frac{2}{1 - z} + 1\right\} = 2 + \sqrt{5}, \end{aligned}$$

which completes the proof.  $\square$

The following example shows that the distortion  $2 + \sqrt{5}$  for Adapted Weighted Uncovered with  $z = \frac{3-\sqrt{5}}{2}$  is tight.

*Example 3.5.* As shown in Figure 1, there are three candidates  $A, B, C$  located on the real line. The voters belong to either  $ABC$  or  $CAB$ , and  $|ABC| = \frac{3-\sqrt{5}}{2} = z$ ,  $|CAB| = \frac{\sqrt{5}-1}{2} = 1 - z$ . Since  $|CA| \geq 1 - z > z$  and  $|BC| \geq z$ , candidate  $A$  can be selected as the winner by Mechanism 1, while  $B$  is the optimal candidate. Then we have  $\frac{S(B,d)}{S(A,d)} = 2 + \sqrt{5}$ , indicating the distortion at least  $2 + \sqrt{5}$ .

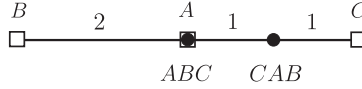


Figure 1: Tight example for the Adapted Weighted Uncovered rule. Candidates and voters are denoted by squares and circles, respectively.

## 4 Improving the Distortion

The tools of Lemma 3.1 and 3.2 allow us to design more involved mechanisms. In this section, we propose a mechanism which selects a candidate from a subset of candidates, if this set is not empty. In particular, when the number of candidates is at most 6, the mechanism can always successfully return a candidate, and the distortion is at most  $2\sqrt{2} + 1 \approx 3.828$ , improving the distortion 4.236 of Adapted Weighted Uncovered.

**Mechanism 2.** ( $Cov_{su}(y)$  RULE.) Let  $y \geq 1$  be a constant. The  $Cov_{su}(y)$ -set is the set of candidates  $A \in \mathcal{C}$  such that for any candidate  $B \neq A$ , we have

1. either  $|BA| \geq \frac{2}{y+1}$ ,
2. or there exists an intermediary candidate  $C \neq A, B$  such that  $|BA| + \frac{y-1}{2}|CA| \geq 1$  and  $|BC| \geq \frac{2}{y+1}$ .

The rule selects an arbitrary candidate in the  $Cov_{su}(y)$ -set if possible.

**Theorem 4.1.** *If the  $Cov_{su}(y)$ -set is non-empty, then  $Cov_{su}(y)$  rule has distortion of at most  $y$ .*

*Proof.* Let  $A \in \mathcal{C}$  be a candidate in the  $Cov_{su}(y)$ -set. Consider an arbitrary candidate  $B \neq A$ . If  $|BA| \geq \frac{2}{y+1}$ , by Lemma 3.1, we have

$$\frac{S(B, d)}{S(A, d)} \leq \frac{2}{|BA|} - 1 \leq \frac{2}{\frac{2}{y+1}} - 1 = y.$$

If  $|BC| \geq \frac{2}{y+1}$  and  $|BA| + \frac{y-1}{2}|CA| \geq 1$ , by Lemma 3.2, we also have

$$\begin{aligned} \frac{S(B, d)}{S(A, d)} &\leq \max\left\{\frac{2}{|BC|} - 1, \frac{2(1 - |BA|)}{|CA|} + 1\right\} \\ &\leq \max\left\{\frac{2}{\frac{2}{y+1}} - 1, \frac{2(1 - |BA|)}{\frac{2}{y-1}(1 - |BA|)} + 1\right\} = y. \end{aligned}$$

□

An immediate question is, when the  $Cov_{su}(y)$ -set is non-empty? Before that, we give some necessary definitions.

**Definition 4.2.** 1. For two candidates  $A$  and  $B$ , we say  $A \times B$ , if  $|BA| < \frac{2}{y+1}$  and for any  $C \neq A, B$ , either  $|BA| + \frac{y-1}{2}|CA| < 1$  or  $|BC| < \frac{2}{y+1}$ .

2. A  $y$ -covered-cycle in the weighted tournament graph is a vertex sequence  $(A_1, A_2, \dots, A_k, A_{k+1} = A_1)$ , such that  $A_i \times A_{i+1}$  for any  $i = 1, 2, \dots, k$ , in which all vertices are distinct except for  $A_1$  and  $A_{k+1}$ .

Note that  $A \times B$  implies candidate  $B$  does not satisfy the conditions in Mechanism 2, and thus  $Cov_{su}(y)$  rule selects a candidate  $A$  such that for any other candidate  $B$ ,  $A \times B$  does not hold. The length of such a  $y$ -covered-cycle is the number of distinct vertices. Let  $L$  be the length of the longest  $y$ -covered-cycle in the weighted tournament graph.

**Lemma 4.3.** *When  $n = 3$ , the  $Cov_{su}(3)$ -set is non-empty, and the distortion of  $Cov_{su}(3)$  rule is at most 3.*

*Proof.* Assume for contradiction that the  $Cov_{su}(3)$ -set is empty. Let  $A, B, C$  be the three candidates in this election. W.l.o.g. we can assume  $A \times B, B \times C, C \times A$ , indicating  $|BA|, |CB|, |AC| < 0.5$ . Then we have  $|AB|, |BC|, |CA| \geq 0.5$ . Hence, by the definition, there must be  $|BA| + |CA| < 1$ ,  $|CB| + |AB| < 1$  and  $|AC| + |BC| < 1$ . Summing up these three inequalities, it derives  $3 < 3$ , a contradiction. □

Before analyzing the case with more candidates, we give a necessary condition for the edge weights in the weighted tournament graph.

**Lemma 4.4.** *For any  $t$  ( $t \geq 3$ ) candidates  $A_1, \dots, A_t$ , we have*

$$\sum_{i=1}^{t-1} |A_i A_{i+1}| + |A_t A_1| \geq 1.$$

*Proof.* We use induction to prove this claim. First consider three candidates  $A, B, C$ , and we have

$$\begin{aligned}
& |AB| + |BC| + |CA| \\
&= (|ABC| + |ACB| + |CAB|) + (|ABC| + |BAC| + |BCA|) \\
&+ (|CAB| + |CBA| + |BCA|) \\
&= (|ABC| + |ACB| + |CAB| + |BAC| + |BCA| + |CBA|) \\
&+ (|ABC| + |BCA| + |CAB|) \\
&= 1 + (|ABC| + |BCA| + |CAB|) \geq 1.
\end{aligned}$$

Suppose the claim is true when  $t = k$ . Now consider  $t = k + 1$ , and we have

$$\begin{aligned}
& \sum_{i=1}^k |A_i A_{i+1}| + |A_{k+1} A_1| \\
&= \sum_{i=1}^{k-1} (|A_i A_{i+1}| + |A_k A_1|) + \\
&\quad (|A_{k+1} A_1| + |A_1 A_k| + |A_k A_{k+1}|) - 1 \\
&\geq 1 + 1 - 1 = 1,
\end{aligned}$$

which completes the proof.  $\square$

In the remainder of this section, we only consider  $y = 2\sqrt{2} + 1 \approx 3.828$ , and denote  $\lambda = \frac{2}{y+1} = \sqrt{2} - 1$ . Using the above lemma, we prove the distortion for the cases  $n = 4, 5, 6$  in the following.

**Lemma 4.5.** *There is no  $y$ -covered-cycle of length 3 or 4. In addition, when  $n = 4$ , the  $\text{Cov}_{su}(2\sqrt{2} + 1)$ -set is non-empty.*

*Proof.* Suppose there is a  $y$ -covered-cycle  $(A, B, C, A)$  of length 3. We have  $|AB|, |BC|, |CA| > 1 - \lambda = 2 - \sqrt{2}$  and  $|BA| + \sqrt{2}|CA|, |CB| + \sqrt{2}|AB|, |AC| + \sqrt{2}|BC| < 1$ . Summing up the last three inequalities, it derives  $(|BA| + |CB| + |AC|) + \sqrt{2} \cdot (|CA| + |BC| + |AC|) < 3$ . However, by Lemma 4.4, we have

$$\begin{aligned}
& (|BA| + |CB| + |AC|) + \sqrt{2} \cdot (|CA| + |BC| + |AC|) \\
&\geq 1 + \sqrt{2} \cdot (|CA| + |BC| + |AC|) \\
&\geq 1 + 3\sqrt{2}(2 - \sqrt{2}) > 3,
\end{aligned}$$

a contradiction.

Suppose there is a  $y$ -covered-cycle  $(A, B, C, D, A)$  of length 4. Define an edge set  $T := \{(B, A), (C, B), (D, C), (A, D)\}$  in the weighted tournament graph. Clearly all edges in  $T$  have weights less than  $\lambda$ . Let  $k$  be the number of edges in  $E \setminus T$ , whose weights are less than  $\lambda$ . There are 3 cases.

**Case 1:**  $k = 0$ . Then we have  $|BA| + \sqrt{2}|DA|, |CB| + \sqrt{2}|AB|, |DC| + \sqrt{2}|BC|, |AD| + \sqrt{2}|CD| < 1$ . Summing up these 4 inequalities, it derives  $(|BA| + |CB| + |DC| + |AD|) + \sqrt{2} \cdot (|DA| + |CD| + |BC| + |AB|) < 4$ . However, by Lemma 4.4, the left hand side is at least  $1 + 4\sqrt{2} \cdot (2 - \sqrt{2}) > 4$ , a contradiction.

**Case 2:**  $k = 1$ . W.l.o.g. assume  $|CA| < \sqrt{2} - 1$ . We have  $|DC| + \sqrt{2}|BC|, |AD| + \sqrt{2}|CD| < 1$ , which immediately gives  $\sqrt{2} \cdot (|DC| + \sqrt{2}|BC|) + (|AD| + \sqrt{2}|CD|) < \sqrt{2} + 1$ . But the left hand side of this inequality is larger than  $2|BC| + \sqrt{2} > 2 * 0.5 + \sqrt{2} = \sqrt{2} + 1$ , a contradiction.

**Case 3:**  $k = 2$ . The analysis is similar to that in Case 2.

Therefore, there is no  $y$ -covered-cycle of length 3 or 4. When  $n = 4$ , an empty  $Cov_{su}(2\sqrt{2} + 1)$ -set indicates the existence of a  $y$ -covered-cycle of length 3 or 4, since  $A \times B$  and  $B \times A$  cannot hold simultaneously. Thus the  $Cov_{su}(2\sqrt{2} + 1)$ -set is non-empty when  $n = 4$ .  $\square$

We say  $A \rightarrow B$  if  $|BA| < \lambda$ , and  $A \nrightarrow B$  if  $|BA| \geq \lambda$ . In the weighted tournament graph, a  $\lambda$ -path is a vertex sequence  $(A_1, A_2, \dots, A_k)$ , such that  $A_i \rightarrow A_{i+1}$  for  $i = 1, 2, \dots, k - 1$ , in which all vertices are distinct. Further, a  $\lambda$ -cycle is a vertex sequence  $(A_1, A_2, \dots, A_k, A_{k+1} = A_1)$ , such that  $A_i \rightarrow A_{i+1}$  for any  $i = 1, 2, \dots, k$ , in which the only repeated vertices are the first and last vertices. Clearly,  $A \times B$  implies  $A \rightarrow B$ , and any  $y$ -covered-cycle must be a  $\lambda$ -cycle.

**Lemma 4.6.** *There is no  $y$ -covered-cycle of length 5. In addition, when  $n = 5$ , the  $Cov_{su}(2\sqrt{2} + 1)$ -set is non-empty.*

*Proof.* We first prove two useful claims.

**Claim 1.** For any length-3  $\lambda$ -cycle  $(A, B, C, A)$ ,  $A \times B$  does not hold.

For such a  $\lambda$ -cycle, we have  $|BA| + \sqrt{2} \cdot |CA| \geq (1 - |AC| - |CB|) + \sqrt{2} \cdot (2 - \sqrt{2}) \geq (3 - 2\sqrt{2}) + (2\sqrt{2} - 2) \geq 1$  and  $|BC| \geq 1 - \lambda > \lambda$ . Thus,  $A \times B$  does not hold.

**Claim 2.** For any  $\lambda$ -path  $(A, B, C, D)$ , if  $A \nrightarrow C$  and  $B \nrightarrow D$ , then  $B \times C$  and  $C \times D$  cannot hold at the same time.

Note that  $\sqrt{2}(|CB| + \sqrt{2}|AB|) + (|DC| + \sqrt{2}|BC|) = \sqrt{2} + 2|AB| + |DC| \geq \sqrt{2} + 2(1 - \lambda) + 0 > \sqrt{2} + 1$ . Then, for  $\lambda$ -path  $(A, B, C, D)$ , it must have either  $|CB| + \sqrt{2} \cdot |AB| \geq 1$  or  $|DC| + \sqrt{2} \cdot |BC| \geq 1$ , implying this claim.

Now we assume for contradiction that there is a  $y$ -covered-cycle  $\Gamma = (A, B, C, D, E, A)$  of length 5, and define edge set  $T = \{(A, C), (B, D), (C, E), (D, A), (E, B)\}$ , as shown in Figure 2. Let  $k$  be the number of edges of  $T$  whose weight is larger than  $1 - \lambda$ . In the following we discuss 4 cases w.r.t. the value of  $k$ .

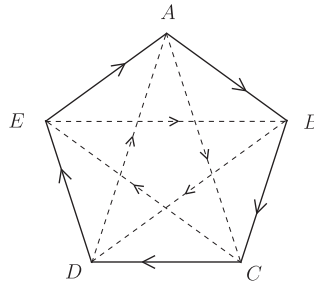


Figure 2: A  $y$ -covered-cycle  $\Gamma = (A, B, C, D, E, A)$ . Edge set  $T$  consists of dashed edges.

**Case 1:**  $k = 0$ . We have  $A \nrightarrow C$  and  $B \nrightarrow D$ . By Claim 2,  $B \times C$  and  $C \times D$  cannot hold at the same time, a contradiction with the assumption of  $y$ -covered-cycle  $\Gamma$ .

**Case 2:**  $k = 1$ . W.l.o.g. assume  $A \rightarrow C$ . We have  $B \nrightarrow D$  and  $C \nrightarrow E$ . By Claim 2,  $C \times D$  and  $D \times E$  cannot hold at the same time, a contradiction with the assumption of  $y$ -covered-cycle  $\Gamma$ .

**Case 3:**  $k = 2$ . There are two different subcases. The first one is typically represented by  $A \rightarrow C, B \rightarrow D$ . Then we have  $C \nrightarrow E$  and  $D \nrightarrow A$ , which derives a contradiction in  $\lambda$ -path  $(C, D, E, A)$  by Claim 2. The second one is typically represented by  $A \rightarrow C \rightarrow E$ . Then we have  $D \nrightarrow A$  and  $E \nrightarrow B$ , which derives a contradiction in  $\lambda$ -path  $(D, E, A, B)$  by Claim 2.

**Case 4:**  $k \geq 3$ . There must be a length-2  $\lambda$ -path using edges in  $T$ , without loss of generality assumed to be  $A \rightarrow C \rightarrow E$ . Then there is a  $\lambda$ -cycle  $(E, A, C, E)$ . However, by Claim 1,  $E \times A$  does not hold, a contradiction with  $\Gamma$ .

When  $n = 5$ , if the  $Cov_{su}(2\sqrt{2} + 1)$ -set is empty, then there must be a  $y$ -covered-cycle of length at most 5. However, such a  $y$ -covered-cycle does not exist by the above discussions. So the  $Cov_{su}(2\sqrt{2} + 1)$ -set is non-empty.  $\square$

**Lemma 4.7.** *There is no  $y$ -covered-cycle of length 6. In addition, when  $n = 6$ , the  $Cov_{su}(2\sqrt{2} + 1)$ -set is non-empty.*

*Proof.* Now we assume for contradiction that there is a  $y$ -covered-cycle  $(A, B, C, D, E, F, A)$  of length 6. Define edge set  $T = \{(A, C), (B, D), (C, E), (D, F), (E, A), (F, B)\}$ . Let  $k$  also be the number of edges of  $T$  whose weight is larger than  $1 - \lambda$ . In the same way as length 5, we divide it into 7 cases. But according to the proof of last lemma, it's easy to know that when  $k = 0, 1$ , the proof is completely the same, and that when  $n = 2$ , the proof is nearly the same except an additional similar situation, thus omitted, leaving other 4 cases i.e., for  $k = 3, 4, 5, 6$ .

**Case 1:**  $k = 3$ . Because there should not be 4-vertices  $\lambda$ -path according to the last lemma (or it will be a contradiction), the three edges can only be  $A \rightarrow C \rightarrow E \rightarrow A$ . If  $A \rightarrow D$ , then there will be a length-3  $\lambda$ -cycle  $(A, D, E, A)$ , causing a contradiction with  $D \times E$ . So by symmetry, we have  $A \nrightarrow D, C \nrightarrow F, E \nrightarrow B$ . Also,  $B \nrightarrow F, F \nrightarrow D, D \nrightarrow B$ , or it will also lead to a length-3  $\lambda$ -cycle.

We have  $|DB|, |FD|, |BD| \geq \sqrt{2} - 1$ , and  $|BA| + \sqrt{2} \cdot |FA|, |DC| + \sqrt{2} \cdot |BC|, |FE| + \sqrt{2} \cdot |DE| < 1$ , or it will be contradict with the  $y$ -covered cycle. So  $|FA|, |BC|, |DE| < \frac{\sqrt{2}}{2}$ , i.e.,  $|AF|, |CB|, |ED| > 1 - \frac{\sqrt{2}}{2}$ . Besides, we have  $|CB| + \sqrt{2}|DB|, |CB| + \sqrt{2}|FB|, |ED| + \sqrt{2}|FD|, |ED| + \sqrt{2}|BD|, |AF| + \sqrt{2}|BF|, |AF| + \sqrt{2}|DF| < 1$ . Add these 6 inequalities together, and we get  $2 \cdot (|CB| + |ED| + |AF|) + 3 \cdot \sqrt{2} < 6$ . Also because  $2 \cdot (|CB| + |ED| + |AF|) > 6 - 3 \cdot \sqrt{2}$ , we get  $6 < 6$ , with contradiction.

**Case 2:**  $k = 4$ . Considering cyclic symmetry, we should consider 2 situations and others can be proved to a contradiction with the same method as 3-vertices cycle and 4-vertices path.

First focus on  $A \rightarrow C \rightarrow E \rightarrow A$  and  $D \rightarrow F$ . This is easier than 3 elements. We should have  $|BA| + \sqrt{2} \cdot |FA| < 1$  and  $|AF| + \sqrt{2}|DF| < 1$ . So add these two to get  $2 > |BA| + \sqrt{2} \cdot |FA| + |AF| + \sqrt{2} \cdot |DF| > 1 + |BA| + (\sqrt{2} - 1) \cdot |FA| + \sqrt{2} \cdot |DF| > 1 + 0 + (2\sqrt{2} - 1)(2 - \sqrt{2}) = 5\sqrt{2} - 5 > 2$ , with contradiction.

Second focus on  $A \rightarrow C \rightarrow E$  and  $D \rightarrow F \rightarrow B$ . Similar to last situation, it is easy to know  $|FE| + \sqrt{2}|CE| < 1, |AF| + \sqrt{2}|EF| < 1$ . So add these together to get  $2 > |FE| + \sqrt{2} \cdot |CE| + |AF| + \sqrt{2} \cdot |EF| > 1 + |AF| + (\sqrt{2} - 1) \cdot |EF| + \sqrt{2} \cdot |CE| > 1 + 0 + (2\sqrt{2} - 1)(2 - \sqrt{2}) = 5\sqrt{2} - 5 > 2$ , with contradiction.

**Case 3:**  $k = 5$ . W.l.o.g we assume  $E \rightarrow A$  is not in the graph. We can get contradiction with the same proof as the second situation in case 2.

**Case 4:**  $k = 6$ . If  $A \rightarrow D$ , then we can get length-3  $\lambda$ -cycle  $(A, D, E, A)$ , leading to a contradiction. Therefore, we have  $A \nrightarrow D, D \nrightarrow A, B \nrightarrow E, E \nrightarrow B, C \nrightarrow F, F \nrightarrow C$ . Also we have  $|BA| + \sqrt{2}|FA|, |CB| + \sqrt{2}|AB|, |DC| + \sqrt{2}|BC|, |ED| + \sqrt{2}|CD|, |FE| + \sqrt{2}|DE|, |AF| + \sqrt{2}|EF| < 1$ .



Add them together, and we can get  $6 > (|BA| + |CB| + |DC| + |ED| + |FE| + |AF|) + \sqrt{2} \cdot (6 - |CA| - |DB| - |EC| - |FD| - |AE| - |BF|)$ . Because according to Lemma 4.4,  $|CA| < |BA| + |CB|$ , we have  $6 > 6\sqrt{2} - (\sqrt{2} - 0.5) \cdot (|CA| + |DB| + |EC| + |FD| + |AE| + |BF|) > 6\sqrt{2} - 6(\sqrt{2} - 0.5) \cdot (\sqrt{2} - 1) = 15 \cdot (\sqrt{2} - 1) > 6$ , with contradiction.

When  $n = 6$ , if the  $Cov_{su}(2\sqrt{2} + 1)$ -set is empty, then there must be a  $y$ -covered-cycle of length 3 or 4 or 5 or 6, since  $A \times B$  and  $B \times A$  cannot hold simultaneously. So the  $Cov_{su}(2\sqrt{2} + 1)$ -set is nonempty.  $\square$

**Theorem 4.8.** *If all  $\lambda$ -cycles in the weighted tournament graph have length at most 6, then the  $Cov_{su}(2\sqrt{2} + 1)$ -set is non-empty, and  $Cov_{su}(2\sqrt{2} + 1)$  rule has distortion at most  $2\sqrt{2} + 1 \approx 3.828$  for the social utility objective.*

*Proof.* Note that any  $y$ -covered-cycle must be a  $\lambda$ -cycle, indicating that the length of any  $y$ -covered-cycle is at most 6. However, by Lemma 4.5, 4.6 and 4.7, there is no  $y$ -covered-cycle of length 3, 4, 5 or 6. Therefore, there exists no  $y$ -covered-cycle in the weighted tournament graph. We can easily find a candidate  $A$  such that for any  $B \neq A$ ,  $A \times B$  does not hold. Then  $A$  is in the  $Cov_{su}(y)$ -set. By Theorem 4.1, the distortion is at most  $2\sqrt{2} + 1$ .  $\square$

We believe that the  $Cov_{su}(2\sqrt{2} + 1)$ -set is non-empty for more general situations when there is some  $\lambda$ -cycle of length larger than 6. We are not able to provide a proof or a counterexample to show whether there is  $y$ -covered-cycle of length at least 7, since there are more and more cases to discuss and it is hard to generalize. We leave this as an open question.

The following example shows the tightness of the distortion  $2\sqrt{2} + 1$  for  $Cov_{su}(2\sqrt{2} + 1)$  rule.

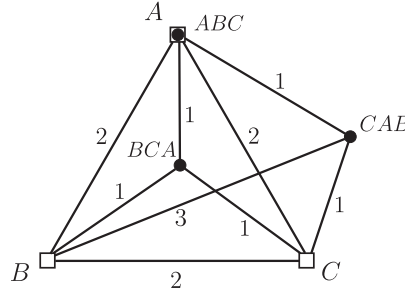


Figure 3: Tight example for  $Cov_{su}(2\sqrt{2} + 1)$  rule.

*Example 4.9.* Let  $\mathcal{C} = \{A, B, C\}$ . Each voter belongs to one set of  $BCA, ABC, CAB$ , and define  $|BCA| = 3 - 2\sqrt{2}$ ,  $|ABC| = |CAB| = \sqrt{2} - 1$ . It is easy to see that candidate  $A$  is in the  $Cov_{su}(2\sqrt{2} + 1)$  set, and can be selected by  $Cov_{su}(2\sqrt{2} + 1)$  rule. However, the underlying metric space may be as shown in Figure 3. At this time, we have  $S(A, d) = 2 - \sqrt{2}$  and  $S(B, d) = 3\sqrt{2} - 2$ , indicating a distortion at least  $\frac{S(B, d)}{S(A, d)} = 2\sqrt{2} + 1$ .

This example can be easily extended to that with more than three candidates, by setting extra candidates very far away.

## 5 Extension: Minimizing Social Cost

In this section, we extend the results in Section 4 to the problem of minimizing social cost. The cost of each voter is the distance to the winner selected by the rule, and the social cost is the total cost of all voters. Under the social cost objective, the lower bound on distortion of voting rules is 3 [2], and the best known distortion is  $2 + \sqrt{5} \approx 4.236$ , achieved by the Weighted Uncovered rule [17].

Some special cases have improved distortion. Anshelevich *et al.* [2] proved that the Ranked Pairs rule achieves the best possible distortion bound 3 when the tournament graph has maximum cycle size no more than 4. It indicates that this rule is optimal when  $n = 3$  or 4. However, its distortion is at least 5 in general. We note that when  $n = 3$ , Ranked Pairs is equivalent to the following rule: select the candidate with maximum min-weight, where the min-weight of a candidate is defined as the minimum weight of all its outgoing edges in the weighted tournament graph.

Now we apply a dual idea of Mechanism 2 to the social cost objective, and provide improved distortion for some special cases. In particular, we show that there is a mechanism with distortion at most 3.828 when  $n = 5$  or  $n = 6$ , improving the best known distortion 4.236 before.

**Mechanism 3.** ( $Cov_{sc}(y)$  RULE.) Let  $y \geq 1$  be a constant. The  $Cov_{sc}(y)$ -set is the set of candidates  $A \in \mathcal{C}$  such that for any candidate  $B \neq A$ , we have

1. either  $|AB| \geq \frac{2}{y+1}$ ,
2. or there exists an intermediary candidate  $C \neq A, B$ , such that  $|AC| \geq \frac{2}{y+1}$  and  $|AB| + \frac{y-1}{2} \cdot |CB| \geq 1$ .

The rule selects an arbitrary candidate in the  $Cov_{sc}(y)$ -set if possible.

**Theorem 5.1.** *If the  $Cov_{sc}(y)$ -set is non-empty, then  $Cov_{sc}(y)$  rule has distortion of at most  $y$ .*

*Proof.* Let  $A \in \mathcal{C}$  be a candidate in the  $Cov_{sc}(y)$ -set. Consider an arbitrary candidate  $B \neq A$ . If  $|AB| \geq \frac{2}{y+1}$ , by Lemma 3.1, we have  $\frac{S(A,d)}{S(B,d)} \leq y$ . If there is an intermediary  $C$  such that  $|AC| \geq \frac{2}{y+1}$  and  $|AB| + \frac{y-1}{2} \cdot |CB| \geq 1$ , by Lemma 3.2, we also have  $\frac{S(A,d)}{S(B,d)} \leq y$ , which completes the proof.  $\square$

Next, we fix  $y = 2\sqrt{2} + 1$ , and  $\lambda = \frac{2}{y+1} = \sqrt{2} - 1$ . In the weighted tournament graph, a  $\lambda'$ -cycle is a vertex sequence  $(A_1, A_2, \dots, A_k, A_{k+1} = A_1)$ , such that  $|A_i A_{i+1}| \leq \lambda$  for any  $i = 1, 2, \dots, k$ , in which the only repeated vertices are the first and last vertices. And a  $\lambda'$ -path is a path  $(A_1, A_2, \dots, A_k, A_{k+1})$  such that any  $|A_i A_{i+1}| < \lambda$  holds. We call  $A \times' B$  if  $|AB| < \lambda$  and there is no candidate  $C$  such that  $|AC| \geq \lambda$  and  $|AC| + \frac{y-1}{2} |CB| \geq 1$ . Then a  $y'$ -covered-cycle is defined as a  $\lambda'$ -cycle such that  $A_i \times' A_{i+1}$  for any  $i = 1, 2, \dots, k$ .

**Lemma 5.2.** *There is no  $y'$ -covered-cycle of length 3, 4, 5 or 6. In addition, when  $n = 3, 4, 5$  or 6, the  $Cov_{sc}(2\sqrt{2} + 1)$ -set is non-empty.*

The proof of Lemma 5.2 is completely symmetric to that of Lemma 4.3, 4.5, 4.6 and 4.7. Then we have the following theorem.

**Theorem 5.3.** *If all  $\lambda'$ -cycles in the weighted tournament graph have length at most 6, then the  $Cov_{sc}(2\sqrt{2} + 1)$ -set is non-empty, and  $Cov_{sc}(2\sqrt{2} + 1)$  rule has distortion at most  $2\sqrt{2} + 1 \approx 3.828$  for the social cost objective.*

*Proof.* Since any  $y'$ -covered-cycle must be a  $\lambda'$ -cycle, the length of any  $y'$ -covered-cycle is at most 6. However, by Lemma 5.2, there is no  $y'$ -covered-cycle of length 3, 4, 5 or 6. Therefore, there exists no  $y'$ -covered-cycle in the weighted tournament graph. It is easy to find a candidate  $A$  such that for any  $B \neq A$ ,  $A \times' B$  does not hold. Then  $A$  is in the  $Cov_{sc}(y)$ -set. By Theorem 5.1, the distortion is at most  $2\sqrt{2} + 1$ .  $\square$

Proof of Lemma 5.2 can be divided into the following lemmas. Notice that if the  $Cov_{sc}(y)$ -set is empty, then there must be a  $y'$ -covered-cycle. And we only need to get contradiction when the length of this cycle is no more than 6.

**Lemma 5.4.** *When  $n = 3$ , the  $Cov_{sc}(3)$ -set is non-empty, and the distortion of  $Cov_{sc}(3)$  rule is at most 3.*

*Proof.* Assume it's empty and let  $A, B, C$  be the three candidates in this election. If there exists a candidate  $A$  such that  $|AB|, |AC| \geq 0.5$ , then it will be a contradiction. Therefore, w.l.o.g, we can assume  $|AB|, |BC|, |CA| < 0.5$ . Because  $|AC|, |CB|, |BA| > 0.5$ , we should let  $|AB| + |CB|, |BC| + |AC|, |CA| + |BA| < 1$ . But add these 3 together, we will get  $3 = |AB| + |CB| + |BC| + |AC| + |CA| + |BA| < 3$ , with a contradiction.  $\square$

**Lemma 5.5.** *There is no  $y'$ -covered-cycle of length 3 or 4.*

*Proof.* We call  $A \prec_\lambda B$  if  $|AB| < \lambda$ . If there is a  $y'$ -covered-cycle of length 3 such as  $(A, B, C, A)$ , then  $A \prec_\lambda B \prec_\lambda C \prec_\lambda A$  and  $|AC|, |CB|, |BA| > 1 - \lambda$ . This time  $|AB| + \sqrt{2} \cdot |CB| < 1, |BC| + \sqrt{2} \cdot |AC| < 1, |CA| + \sqrt{2} \cdot |BA| < 1$  cannot hold together, or we can add these 3 to get  $3 < 3$ , with a contradiction.

If there is a  $y'$ -covered-cycle of length 3 such as  $(A, B, C, D, A)$ . Define set  $T = \{A \prec_\lambda C, C \prec_\lambda A, B \prec_\lambda D, D \prec_\lambda B, \}$ . Let  $k$  denote the number of the elements that hold in this situation and obviously  $0 \leq k \leq 2$ . There are 3 cases.

**Case 1:**  $k = 0$ . Similar to the last paragraph,  $|AB| + \sqrt{2} \cdot |CB| < 1, |BC| + \sqrt{2} \cdot |DC| < 1, |CD| + \sqrt{2} \cdot |AD| < 1, |DA| + \sqrt{2} \cdot |BA| < 1$  cannot hold together. Assume  $|AB| + \sqrt{2} \cdot |CB| \geq 1$ , and because  $|AC| \geq \lambda$ , we know  $A \times' B$  does not hold.

**Case 2:**  $k = 1$ . Assume  $A \prec_\lambda C$ . Observe the cycle  $A \prec_\lambda C \prec_\lambda D \prec_\lambda A$ , we can have  $|CD| \geq 1 - |AC| - |DA| \geq 3 - 2\sqrt{2}$ . So  $|CD| + \sqrt{2} \cdot |AD| \geq 3 - 2\sqrt{2} + 2\sqrt{2} - 2 = 1$  and  $C \times' D$  does not hold.

**Case 3:**  $k = 2$ . The method is the same as case 2.  $\square$

**Lemma 5.6.** *There is no  $y'$ -covered-cycle of length 5.*

*Proof.* First we give two claims.

**Claim 1.** For any length-3  $\lambda'$ -cycle  $(A, B, C, A)$ ,  $A \times' B$  does not hold.

For such a  $\lambda$ -cycle, we have  $|AB| + \sqrt{2} \cdot |CB| \geq (1 - |CA| - |BC|) + \sqrt{2} \cdot (2 - \sqrt{2}) \geq (3 - 2\sqrt{2}) + (2\sqrt{2} - 2) \geq 1$  and  $|AC| \geq 1 - \lambda > \lambda$ . Thus,  $A \times B$  does not hold.

**Claim 2.** For any  $\lambda'$ -path  $(A, B, C, D)$ , if neither  $A \prec_\lambda C$  nor  $B \prec_\lambda D$ , then  $A \times' B$  and  $B \times' C$  cannot hold at the same time.

Note that  $\sqrt{2}(|BC| + \sqrt{2}|DC|) + (|AB| + \sqrt{2}|CB|) = \sqrt{2} + 2|DC| + |AB| \geq \sqrt{2} + 2(1 - \lambda) + 0 > \sqrt{2} + 1$ . Then, for  $\lambda$ -path  $(A, B, C, D)$ , it must have either  $|AB| + \sqrt{2} \cdot |CB| \geq 1$  or  $|BC| + \sqrt{2} \cdot |DC| \geq 1$ , implying this claim.

Back to the 5 vertices loop, we call  $\{A \prec_\lambda C, B \prec_\lambda D, C \prec_\lambda E, D \prec_\lambda A, E \prec_\lambda B\}$  a set  $T$ . Let  $k$  denotes the number of the elements in  $T$  that appear in the graph and we divide this into 4 cases.

**Case 1:**  $k = 0$ . Observe a 4-vertices  $\lambda'$ -path  $(A, B, C, D)$  and obviously at least one of  $A \times' B$  and  $B \times' C$  does not hold, with a contradiction.

**Case 2:**  $k = 1$ . Without loss of generality, we can assume only  $E \prec_\lambda B$  holds. At this time, we can also observe the 4-vertices  $\lambda'$ -path  $(A, B, C, D)$  and use the same method to get a contradiction.

**Case 3:**  $k = 2$ . According to the symmetry, w.l.o.g, we should discuss 2 different situations. The former one is  $A \prec_\lambda C$  and  $B \prec_\lambda D$ . At this time we only need to consider a 4-vertices  $\lambda'$ -path  $(C, D, E, A)$  to get a contradiction. The latter one is  $A \prec_\lambda C \prec_\lambda E$  and this time we have a 3-vertices  $\lambda'$ -cycle  $(A, C, E, A)$  and therefore  $E \times' A$  cannot hold, leading to a contradiction.

**Case 4:**  $k \geq 3$ . Considering the symmetry, there must be a 3-vertices  $\lambda'$ -path such as  $(A, C, E)$ . Because  $E \prec_\lambda A$ , we have a 3-vertices  $\lambda'$ -cycle  $(A, C, E, A)$  and therefore  $E \times' A$  cannot hold, leading to a contradiction.  $\square$

**Lemma 5.7.** *There is no  $y'$ -covered-cycle of length 6.*

*Proof.* Assume  $A \times' B \times' C \times' D \times' E \times' F \times' A$ , and let set  $T$  be  $\{A \prec_\lambda C, B \prec_\lambda D, C \prec_\lambda E, D \prec_\lambda F, E \prec_\lambda A, F \prec_\lambda B\}$ . In the same way as above, at least 3 elements in the set should appear in the graph, or we will easily find a 4-vertices  $\lambda'$ -path to get a contradiction. Also let  $k$  denotes the number of the elements in  $T$  that appear in the graph. We divide this into 4 cases.

**Case 1:**  $k = 3$ . If there are 3 elements, then it can only be  $A \prec_\lambda C \prec_\lambda E \prec_\lambda A$ . If  $A \prec_\lambda D$ , then we can have a 3-vertices  $\lambda'$ -cycle  $(A, D, E, A)$ . So there should not be  $A \prec_\lambda D, C \prec_\lambda F, E \prec_\lambda B$ . Also, there should not be  $B \prec_\lambda F, F \prec_\lambda D, D \prec_\lambda B$ .

Because  $|BD|, |DF|, |FB| > \sqrt{2} - 1$ , we have  $|BC| + \sqrt{2} \cdot |DC|, |DE| + \sqrt{2} \cdot |FE|, |FA| + \sqrt{2} \cdot |BA| < 1$ . So  $|BA|, |DC|, |FE| < \frac{\sqrt{2}}{2}$ , i.e.,  $|AB|, |CD|, |EF| > 1 - \frac{\sqrt{2}}{2}$ . Besides, we have  $|AB| + \sqrt{2}|DB|, |AB| + \sqrt{2}|FB|, |CD| + \sqrt{2}|FD|, |CD| + \sqrt{2}|BD|, |EF| + \sqrt{2}|BF|, |EF| + \sqrt{2}|DF| < 1$ .

Add these 6 inequalities together, we get  $2 \cdot (|AB| + |CD| + |EF|) + 3 \cdot \sqrt{2} < 6$ . Also because  $2 \cdot (|AB| + |CD| + |EF|) > 6 - 3 \cdot \sqrt{2}$ , we get  $6 < 6$ , with contradiction.

**Case 2:**  $k = 4$ . If there are 4 elements, we should consider 2 situations.

First focus on  $A \prec_\lambda C \prec_\lambda E \prec_\lambda A$  and  $D \prec_\lambda F$ . The method is easier than 3 elements. We should have  $|BC| + \sqrt{2} \cdot |DC| < 1$ . Besides,  $|CD| + \sqrt{2}|FD| < 1$ , so we have  $|CD| < 1 - \sqrt{2}|FD| < 1 - \sqrt{2} \cdot (2 - \sqrt{2}) < 3 - 2\sqrt{2}$ . Thus  $|BC| + \sqrt{2} \cdot |DC| > \sqrt{2} \cdot |DC| > \sqrt{2} \cdot (1 - (3 - 2\sqrt{2}))4 - 2\sqrt{2} > 1$ , with contradiction.

Second focus on  $B \prec_\lambda D \prec_\lambda F$  and  $C \prec_\lambda E \prec_\lambda A$ . Similar to last situation, it is easy to know  $|AB| + \sqrt{2}|DB| < 1, |FA| + \sqrt{2}|BA| < 1$ . So  $|AB| < 1 - \sqrt{2}|DB| < 3 - 2\sqrt{2}$ .  $|BA| > 2\sqrt{2} - 2$ . Therefore  $|FA| + \sqrt{2}|BA| > 4 - 2\sqrt{2} > 1$ , with contradiction.

**Case 3:**  $k = 5$ . If there are 5 elements, we assume  $F \prec_\lambda B$  is not in the graph. We can get contradiction with the same proof as the second situation in 4 elements.

**Case 4:**  $k = 6$ . If there are 6 elements, all the edges int the set appear. In the same way as case 1, there is not  $A \prec_\lambda D$ . Therefore, none of  $A \prec_\lambda D, D \prec_\lambda A, B \prec_\lambda E, E \prec_\lambda B, C \prec_\lambda F, F \prec_\lambda C$  appears. We have  $|AB| + \sqrt{2}|DB|, |BC| + \sqrt{2}|EB|, |CD| + \sqrt{2}|FD|, |DE| + \sqrt{2}|AE|, |EF| + \sqrt{2}|BF|, |FA| + \sqrt{2}|CA| < 1$ .

Add them together, and we can get  $6 > (|AB| + |BC| + |CD| + |DE| + |EF| + |FA|) + \sqrt{2} \cdot (6 - |AC| - |BD| - |CE| - |DF| - |EA| - |FB|)$ . Because according to Lemma 4.4p,  $|AC| < |AB| + |BC|$ , we have  $6 > 6\sqrt{2} - (\sqrt{2} - 0.5) \cdot (|AC| + |BD| + |CE| + |DF| + |EA| + |FB|) > 6\sqrt{2} - 6(\sqrt{2} - 0.5) \cdot (\sqrt{2} - 1) = 15 \cdot (\sqrt{2} - 1) > 6$ , with contradiction.  $\square$

Even though we are not able to prove whether there is  $y'$ -covered-cycle of length at least 7, we believe that the  $Cov_{sc}(2\sqrt{2} + 1)$ -set is non-empty for some more general cases.

**Theorem 5.8.** *If all  $\lambda'$ -cycles in the weighted tournament graph have length at most 6, then the  $Cov_{sc}(2\sqrt{2} + 1)$ -set is non-empty, and  $Cov_{sc}(2\sqrt{2} + 1)$  rule has distortion at most  $2\sqrt{2} + 1 \approx 3.828$  for the social cost objective.*

*Proof.* Note that any  $y$ -covered-cycle must be a  $\lambda'$ -cycle, indicating that the length of any  $y'$ -covered-cycle is at most 6. However, we have known that there is no  $y'$ -covered-cycle of length 3, 4, 5 or 6. Therefore, there exists no  $y'$ -covered-cycle in the weighted tournament graph. We can easily find a candidate  $A$  such that for any  $B \neq A$ ,  $A \times' B$  does not hold. Then  $A$  is in the  $Cov_{sc}(y)$ -set. By Theorem 5.1, the distortion is at most  $2\sqrt{2} + 1$ .  $\square$

## 6 Discussion

We discuss the social choice problem for two egalitarian objectives: maximizing the minimum utility and minimizing the maximum cost, and conclude this paper.

### 6.1 Egalitarian Objectives

Define  $M(A, d) = \max_{v \in V} d(v, A)$  to be the maximum distance to candidate  $A$  for metric  $d$ , among all voters. Then the distortion of a voting rule  $f$  under the objective of minimizing maximum cost is

$$\Delta_m(f) = \sup_{\sigma} \sup_{d \in \rho(\sigma)} \frac{M(f(\sigma), d)}{\min_{A \in \mathcal{C}} M(A, d)}.$$

We show that the lower bound and upper bound on the distortion are matching.

**Proposition 6.1.** *For the objective of minimizing maximum cost,*

1. *no rule has distortion better than 3, and*
2. *the rule that selects a candidate  $A$  such that  $|AB| > 0$  for any  $B \neq A$ , has distortion at most 3.*

*Proof.* (1) Consider the election when  $m = n = 2$ , and each candidate is the top choice of a voter. W.l.o.g. assume a rule selects candidate  $A$ . Then the worst case is when two candidates  $A$  and  $B$  are located at point 0 and 2 on the real line, and two voters are located at  $1 - \epsilon$  and 3 for small  $\epsilon > 0$ . The optimal maximum cost is  $M(B, d) = 1$ , while the maximum cost induced by the rule is  $M(A, d) = 3$ .

(2) The existence of such a candidate  $A$  is guaranteed by Lemma 4.4. For any  $B \neq A$ , suppose  $v \in AB$ . Then by triangle inequality, we have  $M(A, d) \leq d(A, B) + M(B, d) \leq 2d(v, B) + M(B, d) \leq 3M(B, d)$ .  $\square$

For the objective of maximizing the minimum utility of all voters, there is no voting rule with bounded distortion: Suppose there are two candidates  $X$  and  $Y$ ; one voter shares the same location with  $Y$ , and the other has the same distance to  $X$  and  $Y$ . Assume w.l.o.g. that the given rule selects  $Y$ . Then the induced minimum utility is 0, and the optimal minimum utility has a certain amount larger than 0.

## 6.2 A New Model

**Lemma 6.2.** *Let  $3 \leq U \leq 2 + \sqrt{5}$ . If  $|AB| \leq \frac{2}{U+1}, [(U+1) - \frac{4}{5-U}] \cdot |AB| + \frac{4}{5-U} \cdot |AD| \geq 2, |AB| + \frac{U-1}{2} \cdot |DC| \geq 1, |AB| + \frac{U-1}{2} \cdot |CB| \geq 1$ , then for any metric  $d$ , we have*

$$\frac{S(A, d)}{S(B, d)} \leq U.$$

*Proof Sketch.* We divide the proof into several cases. First let  $a = d(A, B), b = d(B, C), c = d(B, D)$ .

**Case 1:**  $a \geq b + c$ . If we want  $\frac{S(A)}{S(B)} \leq U$ , then by the properties of linear expression, we only need to have  $[(U+1) - \frac{4}{5-U}] \cdot |AB| + \frac{4}{5-U} \cdot \max\{|AC|, |AD|\} \geq 2, \max\{|AC|, |AD|\} \geq \frac{2}{U+1}, \max\{|AB|, |AD|\} + \frac{U-1}{2} \cdot \max\{|CB|, |CD|\} \geq 1$ , and  $\max\{|AB|, |AC|\} + \frac{U-1}{2} \cdot \max\{|DC|, |DB|\} \geq 1$ .

**Case 2:**  $\max\{b, c\} \leq a < b + c$ . In the same way, if we want  $\frac{S(A)}{S(B)} \leq U$ , we only need to have  $[(U+1) - \frac{4}{5-U}] \cdot |AB| + \frac{4}{5-U} \cdot \max\{|AC|, |AD|\} \geq 2, \max\{|AB|, |AD|\} + \frac{U-1}{2} \cdot \max\{|CB|, |CD|\} \geq 1, \max\{|AB|, |AC|\} + \frac{U-1}{2} \cdot \max\{|DC|, |DB|\} \geq 1$ , and  $|AB| + \frac{U-1}{2} \cdot \max\{|CB|, |DB|\} \geq 1$ .

**Case 3:**  $\min\{b, c\} \leq a < \max\{b, c\}$ . If we want  $\frac{S(A)}{S(B)} \leq U$ , we only need to have  $[(U+1) - \frac{4}{5-U}] \cdot |AB| + \frac{4}{5-U} \cdot \max\{|AC|, |AD|\} \geq 2, \max\{|AB|, |AD|\} + \frac{U-1}{2} \cdot \max\{|CB|, |CD|\} \geq 1, \max\{|AB|, |AC|\} + \frac{U-1}{2} \cdot \max\{|DC|, |DB|\} \geq 1$ , and  $|AB| + \frac{U-1}{2} \cdot \max\{|CB|, |DB|\} \geq 1$ . In fact, the condition is the same as case 2.

**Case 4:**  $0 \leq a < \min\{b, c\}$ . If we want  $\frac{S(A)}{S(B)} \leq U$ , we only need to have  $|AB| + \frac{U-1}{2} \cdot \max\{|CB|, |DB|\} \geq 1$ .

Considering all the 4 cases, we can merge and simplify some expressions. If  $|AB| \leq \frac{2}{U+1}, [(U+1) - \frac{4}{5-U}] \cdot |AB| + \frac{4}{5-U} \cdot |AD| \geq 2$ , then  $\max\{|AC|, |AD|\} \geq \frac{2}{U+1}$ . If  $|AB| + \frac{U-1}{2} \cdot |DC| \geq 1$ , then  $\max\{|AB|, |AC|\} + \frac{U-1}{2} \cdot \max\{|DC|, |DB|\} \geq 1$ . If  $|AB| + \frac{U-1}{2} \cdot |CB| \geq 1$ , then  $|AB| + \frac{U-1}{2} \cdot \max\{|CB|, |DB|\} \geq 1$ .

Therefore, we can conclude that if  $|AB| \leq \frac{2}{U+1}, [(U+1) - \frac{4}{5-U}] \cdot |AB| + \frac{4}{5-U} \cdot |AD| \geq 2, |AB| + \frac{U-1}{2} \cdot |DC| \geq 1, |AB| + \frac{U-1}{2} \cdot |CB| \geq 1$ , then for any metric  $d$ , we have  $\frac{S(A, d)}{S(B, d)} \leq U$ .  $\square$

Then we introduce two mechanisms based on Lemma 6.2.

**Mechanism 4.** (FOUR-VERTEX UTILITY MODEL.) Select a candidate  $A \in \mathcal{C}$  such that, for any candidate  $B \neq A$ , we have

1. either  $|BA| \geq \sqrt{2} - 1 \approx 0.414$ ,
2. or there exists an intermediary candidate  $C \neq A, B$  such that  $|BA| + \sqrt{2} \cdot |CA| \geq 1$  and  $|BC| \geq \sqrt{2} - 1$ .
3. or there exists two intermediary candidates  $C, D \neq A, B$  such that  $|BA| + (\sqrt{2} + 1) \cdot |BC| \geq 2$ ,  $|BA| + \sqrt{2} \cdot |CD| \geq 1$  and  $|BA| + \sqrt{2} \cdot |DA| \geq 1$ .

**Mechanism 5.** (FOUR-VERTEX COST MODEL.) Select a candidate  $A \in \mathcal{C}$  such that, for any candidate  $B \neq A$ , we have

1. either  $|AB| \geq \sqrt{2} - 1 \approx 0.414$ ,
2. or there exists an intermediary candidate  $C \neq A, B$  such that  $|AB| + \sqrt{2} \cdot |CB| \geq 1$  and  $|AC| \geq \sqrt{2} - 1$ .
3. or there exists two intermediary candidates  $C, D \neq A, B$  such that  $|AB| + (\sqrt{2} + 1) \cdot |AD| \geq 2$ ,  $|AB| + \sqrt{2} \cdot |DC| \geq 1$  and  $|AB| + \sqrt{2} \cdot |CB| \geq 1$ .

Obviously, if there is a candidate that satisfies the two mechanisms above, then according to Lemma 6.2, the upper bound of distortion in both utility and cost views will be  $2\sqrt{2} + 1 \approx 3.828$ .

Consider the following cyclically symmetric weighted tournament graph. There are  $n$  vertex ( $n$  is big enough to make  $1/n \rightarrow 0$ )  $A_1, \dots, A_n$ . Let  $A_{i+n} = A_i$ , then obviously  $|A_i A_j| = |A_{i+k} A_{j+k}|$  for  $\forall i, j, k \in [1, n]$ . Here we give another setting that  $\sum_i^n |A_i A_{i+1}| = 1$ . We call this graph  $G1$ . Then every candidate in  $G1$  fits both Mechanism 4 and 5. Because in  $G1$ , Mechanism 4 and 5 are symmetric, we only need to give the proof of Mechanism 5.

Firstly it is easy to see that  $|A_i A_{i+1}| \leq \frac{1}{n} < \sqrt{2} - 1$ , or it will contradict with  $\sum_i^n |A_i A_{i+1}| = 1$ . Secondly considering Lemma 4.4, we can find that for  $\forall i, k \in [1, n]$ , we have  $|A_i A_{i+k}| = \sum_{j=i}^{i+k-1} |A_j A_{j+1}|$ . W.l.o.g, we focus on candidate  $A_1$ . Then we can get  $\exists m_0, \forall 2 \leq m \leq m_0, |A_1 A_m| < \sqrt{2} - 1$  and  $\forall m_0 < m \leq n, |A_1 A_m| \geq \sqrt{2} - 1$ . Because  $\forall m > m_0$ , the first item of Mechanism 5 is satisfied, let  $m \in [2, m_0]$ . And  $|A_1 A_m| = \frac{m-1}{n}$ .

We use contradiction to prove, and divide the situation into two cases.

Case 1:  $m \geq 3$ .  $\exists p, q, (\sqrt{2} + 1) \cdot |A_1 A_q| \geq \sqrt{2}, (\sqrt{2} + 1) \cdot |A_1 A_{q-1}| < \sqrt{2}, \sqrt{2} \cdot |A_p A_m| \geq 1, \sqrt{2} \cdot |A_p A_{m-1}| < 1$ . Therefore,  $(\sqrt{2} + 1) \cdot \frac{q-1}{n} \geq \sqrt{2}, (\sqrt{2} + 1) \cdot \frac{q-2}{n} < \sqrt{2}, \sqrt{2} \cdot \frac{m+n-p}{n} \geq 1, \sqrt{2} \cdot \frac{m+n-p}{n} < 1$ . Notice that  $|A_1 A_m| + |A_m A_p| + |A_q A_1| \leq 3 - 1.5\sqrt{2} < 1$ . Add the second inequality  $\times (\sqrt{2} - 1)$  and forth inequality  $\times \frac{\sqrt{2}}{2}$  together, and we will find that if  $|A_q A_p| = \frac{p+n-q}{n} < \frac{\sqrt{2}}{2}$ , adding this inequality  $\times \frac{\sqrt{2}}{2}$  can lead to  $m < 3$ , with a contradiction.

Case 2:  $m = 2$ . In the same way,  $\exists p, q, |A_1 A_m| + (\sqrt{2} + 1) \cdot |A_1 A_q| \geq \sqrt{2}, |A_1 A_m| + (\sqrt{2} + 1) \cdot |A_1 A_{q-1}| < \sqrt{2}, |A_1 A_m| + \sqrt{2} \cdot |A_p A_m| \geq 1, |A_1 A_m| + \sqrt{2} \cdot |A_p A_{m-1}| < 1$ . Therefore,  $\frac{m-1}{n} + (\sqrt{2} + 1) \cdot \frac{q-1}{n} \geq \sqrt{2}, \frac{m-1}{n} + (\sqrt{2} + 1) \cdot \frac{q-2}{n} < \sqrt{2}, \frac{m-1}{n} + \sqrt{2} \cdot \frac{m+n-p}{n} \geq 1, \frac{m-1}{n} + \sqrt{2} \cdot \frac{m+n-p}{n} < 1$ . Notice that  $|A_1 A_m| + |A_m A_p| + |A_q A_1| \leq 3 - 1.5\sqrt{2} + \frac{1.5\sqrt{2}}{n} \approx 3 - 1.5\sqrt{2} < 1$ . Add the second inequality  $\times (\sqrt{2} - 1)$  and forth inequality  $\times \frac{\sqrt{2}}{2}$  together, and we will find that if  $|A_q A_p| = \frac{p+n-q}{n} < \frac{\sqrt{2}}{2}$ , adding this inequality  $\times \frac{\sqrt{2}}{2}$  can lead to  $m < \frac{1+3\sqrt{2}}{4} < 2$ , with a contradiction.  $\square$

Therefore we come up with a conjecture.

**Conjecture 6.3.** For  $n \geq 4$ , in every weighted tournament graph, there exists a candidate  $A$  which satisfies Mechanism 4 and Mechanism 5. Therefore, the upper bound of distortion in both utility and cost views is  $2\sqrt{2} + 1 \approx 3.828$ .

### 6.3 Conclusions

In this paper, we study the social choice problem under metric preferences of agents, and focus on the social utility objective that aims at maximizing the total distance of agents to the winner. We propose a C2 rule (a rule making decisions only based on the weighted tournament graph) that has

distortion  $2 + \sqrt{5}$ , and there is still a gap between this distortion and the lower bound 3. To narrow this gap, we propose another C2 rule (Mechanism 2) that improves the distortion to  $2\sqrt{2} + 1$  when every  $\lambda$ -cycle has length no more than 6, and it may work well for more general cases or even for all instances. This remains an open question:

*How well can a C2 rule (or any deterministic voting rule) do for the social choice problem?*

This work also contributes to the social cost objectives, which are widely studied in literature. We adapt Mechanism 2 and provide improved distortion of 3.828 when there are 5 or 6 candidates.

## References

- [1] Ben Abramowitz, Elliot Anshelevich, and Wennan Zhu. Awareness of voter passion greatly improves the distortion of metric social choice. In *International Conference on Web and Internet Economics*. Springer, 2019.
- [2] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. Approximating optimal social choice under metric preferences. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 777–783, 2015.
- [3] Elliot Anshelevich and John Postl. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research*, 58:797–827, 2017.
- [4] Elliot Anshelevich and Wennan Zhu. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. In *International Conference on Web and Internet Economics*, pages 3–20. Springer, 2018.
- [5] Elliot Anshelevich and Wennan Zhu. Tradeoffs between information and ordinal approximation for bipartite matching. *Theory of Computing Systems*, 63(7):1499–1530, 2019.
- [6] Umang Bhaskar and Abheek Ghosh. On the welfare of cardinal voting mechanisms. In *38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018)*, volume 122 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:22, 2018.
- [7] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213, 2015.
- [8] Ioannis Caragiannis and Ariel D Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175(9-10):1655–1671, 2011.
- [9] Yu Cheng, Shaddin Dughmi, and David Kempe. Of the people: voting is more effective with representative candidates. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 305–322. ACM, 2017.
- [10] Yu Cheng, Shaddin Dughmi, and David Kempe. On the distortion of voting with multiple representative candidates. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- [11] Edith Elkind, Piotr Faliszewski, Jean-François Laslier, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. What do multiwinner voting rules do? an experiment over the two-dimensional euclidean domain. In *Thirty-First AAAI Conference on Artificial Intelligence*, pages 494–501, 2017.



- [12] Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 269–286. ACM, 2016.
- [13] Ashish Goel, Reyna Hulett, and Anilesh K Krishnaswamy. Relating metric distortion and fairness of social choice rules. In *Proceedings of the 13th Workshop on Economics of Networks, Systems and Computation*, page 4. ACM, 2018.
- [14] Ashish Goel, Anilesh K Krishnaswamy, and Kamesh Munagala. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 287–304. ACM, 2017.
- [15] Stephen Gross, Elliot Anshelevich, and Lirong Xia. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [16] Hervé Moulin. Choosing from a tournament. *Social Choice and Welfare*, 3(4):271–291, 1986.
- [17] Kamesh Munagala and Kangning Wang. Improved metric distortion for deterministic social choice rules. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, EC ’19, pages 245–262. ACM, 2019.
- [18] Joel Oren and Brendan Lucier. Online (budgeted) social choice. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 1456–1462, 2014.
- [19] Ariel D Procaccia and Jeffrey S Rosenschein. The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents*, pages 317–331. Springer, 2006.
- [20] Markus Schulze. A new monotonic, clone-independent, reversal symmetric, and condorcet-consistent single-winner election method. *Social Choice and Welfare*, 36(2):267–303, 2011.
- [21] Piotr Krzysztof Skowron and Edith Elkind. Social choice under metric preferences: scoring rules and stv. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [22] T Nicolaus Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.