1 Introduction

We study social choice in the **metric distortion** framework, where voters and candidates are embedded in an unknown metric space, and voters prefer candidates that are closer to them. A social choice rule sees only the induced rankings and must pick a single winner; its distortion is the worst-case ratio between the social cost, given by the sum of distances to voters, of the chosen winner and that of the optimal candidate. It is well known that any deterministic rule that only uses *ordinal rankings* (e.g., voter v ranks candidate A above candidate B) has distortion at least 3, and deterministic tournament rules face a stronger lower bound of around 3.11.

A natural way to go beyond these limits is to enrich the information available to the rule. Ordinal rankings exhaust all *qualitative* information (for instance, whether voter v prefers A over B) but do not touch upon *quantitative* information (for instance, how strongly v prefers A over B). On the other hand, it is practically impossible to ask voters themselves to reveal precise cardinal utilities: it is unrealistic to expect a voter to answer a question such as "do you prefer candidate A at least $\sqrt{2}$ times as much as candidate B?" A natural workaround, therefore, is to ask questions that act as valid proxies to extract additional cardinal information. In this work, we focus on **two-person deliberation**: for a pair of voters (u, v) and a pair of candidates (A, B), comparing $d(u, A) + d(v, A)$ with $d(u, B) + d(v, B)$ reveals coarse geometric information about the latent metric, essentially indicating on which side of the bisector of (A, B) the pair (u, v)'s barycenter lies. We describe the exact procedure to form such pairwise deliberations later.

This project focuses on the **deliberation-via-matching** protocol and its geometric analysis. On a high level, the write-up is structured as follows.

- We first review the metric and tournament preliminaries in Section 2, and describe the deliberation-via-matching protocol in Section 3. The protocol involves *parsimonious* pairwise deliberations, where each voter participates in at most one deliberation for each pair candidates. In most real-world voting instances, where the number of candidates m is small compared to the number of voters n, this per-voter overhead of $\binom{m}{2}$ is modest.
- In Section 4, we apply the deliberation-via-matching protocol to the two-candidate setting. We show that the deterministic deliberation-via-matching rule achieves distortion 2, which is optimal among *any* deterministic rule that has access to ordinal rankings and outcomes of *any* pairwise deliberations. We then introduce a randomized variant for two candidates and show that it obtains distortion at most 1.53, close to the $3/2$ lower bound, again, for *any* randomized rule that reads ordinal rankings and outcomes of pairwise deliberations.
- In Section 5, we extend this protocol to general instances with m candidates via a weighted *uncovered-set* tournament rule, and show that deliberation-via-matching achieves an overall distortion of 3. This is significant in three ways: (i) it breaks a previously known lower bound of 3.11 for tournament rules that only use ordinal rankings, and (ii) it conceptually shows that tournament rules are just as powerful as general deterministic rules (which are lower bounded by 3) given *minimal* additional cardinal information, and finally, (iii) to our best knowledge, unlike previous literature in this line of work, our proof is the first to be analytically tractable and is *governed by clean geometric intuitions* throughout.

2 Preliminaries

We now fix the geometric setting and notation that will be used throughout the rest of the write-up.

2.1 Metric Model and Social Cost

We first describe the classical, ordinal-ranking only, metric distortion framework. Let $\mathcal{V}$ be a finite set of n voters and C be a finite set of m candidates. We use lowercase letters, most frequently u, v, to denote voters, and uppercase letters to denote candidates. Both $\mathcal{V}$ and C are embedded as points in an unknown metric space with metric d consistent with the voters' ordinal rankings. That is, if voter v prefers candidate X to candidate Y, then $d(v, X) \leqslant d(v, Y)$. The social choice rule never sees the distances $d(v, X)$ themselves; it only observes the induced ordinal rankings.

For two candidates X, Y, let XY denote the set of voters who rank X above Y. We break ties arbitrarily but consistently, so every voter belongs to exactly one of XY or YX. The social cost of X under a metric d is $SC(X) = SC(X, d) = \sum_{v \in V} d(v, X)$. The welfare-optimal candidate is the 1-median of the entire candidate set,

$X^* = \arg\min_{X \in C} SC(X, d)$. Let $\mathcal{S}$ be a social choice rule; it maps each ordinal ranking profile σ to a single winning candidate $\mathcal{S}(\sigma) \in C$. The **distortion** of rule $\mathcal{S}$ is given by its worst case approximation to the 1-median:

$$\text{Distortion}(\mathcal{S}) = \sup_{\sigma} \sup_{d \text{ consistent with } \sigma} \frac{SC(\mathcal{S}(\sigma), d)}{SC(X^*, d)}. \tag{1}$$

A smaller distortion means that the rule is a better geometric approximation to the 1-median using only ordinal information.

For most parts of the analysis, it is convenient to normalize the total voter mass to 1 and view voters as a probability distribution over the latent metric space. In that view, social costs are viewed as expectations $SC(X) = \mathbb{E}_{v \in \mathcal{V}}[d(v, X)]$.

2.2 Weighted Tournaments and the Weighted Uncovered Set

A tournament rule aggregates pairwise comparison statistics into a weighted directed graph, called a *tournament graph*, where the weight $f(XY) \in [0, 1]$ satisfies the strength of X against Y, satisfying $f(XY) + f(YX) = 1$. In our framework, we build a tournament graph over the set of candidates C, detailed in Section 5. We utilize the **λ-weighted uncovered set (λ-WUS)** for aggregation. A candidate X belongs to the λ-WUS if, for every opponent Y, one of the following holds:

- (Direct weak dominance.) $f(XY) \geqslant 1 - \lambda$, or
- (Two-step weak-strong dominance.) There exists a pivot candidate Z such that $f(XZ) \geqslant 1 - \lambda$ and $f(ZY) \geqslant \lambda$.

It is established that for $\lambda \in [0.5, 1]$, the λ-WUS is nonempty cite.

3 The *Deliberation via Matching* Protocol

We propose the deliberation-via-matching protocol. This is a tournament-based social choice mechanism that refines ordinal preferences using selective pairwise deliberations. The protocol is governed by two scalar parameters:

- *Deliberation weight* $w \geqslant 0$: controls the strength of deliberation outcomes relative to ordinal information;
- *Uncovering parameter* $\lambda \in [0.5, 1]$: determines the final λ-WUS from which a winner is chosen.

The protocol proceeds in three steps.

Step 1. Matching and Deliberation. For every (unordered) pair of distinct candidates (X, Y), we identify the sets of voters who disagree on their relative ranking of X, Y: this partitions $\mathcal{V}$ into XY (those preferring X) and $YX = \mathcal{V} \backslash XY$ (those preferring Y). We construct an *arbitrary* maximum cardinality matching M_{XY} between those two sets. This results in $M_{XY} = \min\{|XY|, |YX|\}$ disjoint pairs of voters with opposing preferences.

Each matched pair $(u, v) \in M_{XY}$ deliberates. The pair favors X if $d(u, X) + d(v, X) \leqslant d(u, Y) + d(v, Y)$, Y otherwise, with ties handled arbitrarily, similar to the membership status of XY and YX. Let W_{XY} denote the number of matched pairs that favor candidate X.

Step 2. Tournament Construction. We aggregate the results into a direct, weighted tournament graph. As mentioned earlier, we assume that $|\mathcal{V}|$ has been normalized to a total mass of 1. The saw score for X against Y is defined as a weighted sum between the ordinal count and deliberation outcomes: $\text{score}(XY) = \text{score}(XY; w) = |XY| + w \cdot W_{XY}$. Note that unmatched voters contribute only to their ordinal vote (with weight 1, while matched voters contribute to their ordinal vote and a weighted share of deliberation outcome). It naturally follows that we can define the normalized edge weight $f(XY) = f(XY; w)$ for the tournament graph as

$$f(XY) = f(XY; w) = \frac{\text{score}(XY; w)}{\text{score}(XY; w) + \text{score}(YX; w)} = \frac{|XY| + w \cdot W_{XY}}{1 + w \cdot M_{XY}}. \tag{2}$$

Step 3. Winner Selection. Once the weighted tournament graph is constructed, we select any candidate in the λ-WUS as the winner. Recall λ-WUS is nonempty given $\lambda \in [0.5, 1]$.

4 *Deliberation via Matching* on 2 Candidates

We first consider a simple scenario in which there are only two candidates A, B, so the tournament graph consists of only two vertices. The λ-WUG winner, WLOG assumed to be A, must satisfy $f(AB) \geqslant 1 - \lambda$. Subject to this constraint, we aim to find the supremum of $SC(A)/SC(B)$.

In this section, we establish tight bounds for this setting. We first show in Section 4.1 that the deterministic deliberation-via-matching rule achieves a distortion of exactly 2, matching the theoretical lower bound for deterministic rules. Then, in Section 4.2, we introduce a randomized variant that achieves a distortion of approximately 1.523, nearly matching the theoretical randomized lower bound of 1.5.

Clearly, among two-candidate instances and subject to $\lambda \in [0.5, 1]$, the set $\{\text{instance } I : f(AB) \geqslant 1 - \lambda\}$ is minimized when $\lambda = 0.5$. Therefore, to maximize $SC(A)/SC(B)$ it suffices to assume $\lambda = 0.5$.

4.1 An Optimal Deterministic Algorithm

We consider the simplest setup with $w = 1$, equal weight for information obtained from ordinal rankings and deliberations. With $\lambda = 0.5$ we essentially consider the classic Copeland/majority tournament rule applied to the weighted graph. Assume A is the winner. Then to upper bound the distortion, we aim to adversarially maximize $SC(A)/SC(B)$ (for clearly $SC(A)/SC(A) = 1$). This is done by upper bounding $SC(A)$ and lower bounding $SC(B)$.

Let M be the maximal cardinality of $AB \times BA$ that we formed, and decompose the pairs into $M_A = \{(u, v) \in M : (u, v) \text{ favors } A\}$ and $M_B = \{(u, v) \in M : (u, v) \text{ favors } B\}$. Then $\mathcal{V}$ is partitioned into M_A and M_B, both expressed in pairs of voters, as well as the unmatched voters (which must either all belong to AB or to BA).

Upper-bounding $SC(A)$. For every voter v, by triangle inequality,

$$\begin{aligned} SC(A) &= \sum_{(u,v)\in M_A} [d(u, A) + d(v, A)] + \sum_{(u,v)\in M_B} [d(u, A) + d(v, A)] + \sum_{v \text{ unmatched}} d(v, A) \\ &\leqslant \sum_{(u,v)\in M_A} [d(u, B) + d(v, B)] + \sum_{(u,v)\in M_B} [d(u, B) + d(v, B) + d(A, B)] + \sum_{v \text{ unmatched}} [d(v, B) + d(A, B)] \\ &= SC(B) + (|BA| - W_{AB}) \cdot d(A, B). \end{aligned} \tag{3}$$

The first term uses the definition of M_A; the second term uses $u \in AB$ so $d(u, A) \leqslant d(u, B)$, and triangle inequality on (v, B, A); and the third term also uses triangle inequality on (v, B, A).

Lower-bounding $SC(B)$. For any $(u, v) \in M_B$, two applications of triangle inequality imply

$$\begin{cases} d(u, B) + d(v, B) \geqslant d(u, A) + d(v, A) \\ d(u, A) + d(u, B) \geqslant d(A, B) \\ d(v, A) + d(v, B) \geqslant d(A, B) \end{cases} \implies d(u, B) + d(v, B) \geqslant d(A, B). \tag{4}$$

Therefore,

$$\begin{aligned} SC(B) &= \sum_{(u,v)\in M_A} [d(u, B) + d(v, B)] + \sum_{\text{other } v} d(v, B) \\ &\geqslant \sum_{(u,v)\in M_A} d(A, B) + \sum_{(\text{other } v)\cap AB} d(A, B)/2 = (|AB| + W_{AB})/2 \cdot d(A, B). \end{aligned} \tag{5}$$

Combining the two bounds, we see that

$$\begin{aligned} \frac{SC(A)}{SC(B)} &\leqslant \frac{SC(B) + (|BA| - W_{AB}) \cdot d(A, B)}{SC(B)} \leqslant 1 + \frac{(|BA| - W_{AB}) \cdot d(A, B)}{(|AB| + W_{AB})/2 \cdot d(A, B)} \\ &= 1 + \frac{2(|BA| - W_{AB})}{|AB| + W_{AB}} = \frac{2}{|AB| + W_{AB}} = \frac{2}{\text{score}(AB)} - 1 \end{aligned} \tag{6}$$

as we assumed $w = 1$. This implies the following claim.

Theorem 4.1. *The metric distortion of deliberation-via-matching with $w = 1, \lambda = 0.5$ for any 2-candidate instance is bounded by 2.*

Proof. By Equation (6), it suffices to show that if A wins, then $\text{score}(AB) \geqslant 2/3$. This immediately follows from solving tiny program,

$$\begin{aligned} \text{minimize} \quad & a + w_a \\ \text{subject to} \quad & a + b = 1 \\ & w_a \leqslant \min(a, b) \\ & a + w_a \geqslant 0.5 \cdot (a + b + \min(a, b)) \\ & a, b, w_a \geqslant 0, \end{aligned}$$

whose objective is minimized by $(a, b, w_a) = (1/3, 2/3, 1/3)$ or $(2/3, 1/3, 0)$, both of which takes value $2/3$. □

Corresponding Lower Bound. We prove a corresponding lower bound: *any* social choice that only uses voters' ordinal preference and the outcome of pairwise deliberation cannot do better. This proves that Theorem 4.1 is tight for 2 candidates and cannot be improved by running the deliberations differently or using a different social choice rule.

Theorem 4.2. *Any deterministic social choice rule that considers only voters' ordinal rankings and the outcomes of pairwise deliberations have distortion at least 2, even with 2 candidates.*

Proof. Consider two instances on $\mathbb{R}$, each with two candidates, A at -1, and B at 1. To simplify the description we temporarily drop the unit mass assumption on $\mathcal{V}$.

For the first instance, place two voters at $A = -1$ and one voter at $B = 1$, and set the deliberation (between one voter at -1 and one at 1) to prefer B. For the second instance, place two voters at 0 favoring A and one voter at B.

In both instances, $|AB| = 2$, $|BA| = 2$, and the deliberation outcomes are identical. Thus no deterministic social choice rule looking only at preference and deliberation profiles can distinguish them. Yet, in the first one $SC(B)/SC(A) = 2$ and in the second, $SC(A)/SC(B) = 2$. □

4.2 A Near-Optimal Randomized Variant

While the deterministic deliberation-via-matching rule achieves an optimal distortion of 2 for 2 candidates, we can further improve this bound by allowing the social choice rule to be randomized. A deterministic rule is vulnerable because it effectively applies a sharp threshold at $s = 0.5$; consequently, an adversary can exploit this by constructing instances where the scores are symmetric, but the underlying metric costs are significantly skewed, as seen in the proof of Theorem 4.2.

To mitigate this, we introduce a randomized rule that selects candidates with probabilities proportional to a function of their scores. We consider a natural choice: the power of k. This "smooths" the decision boundary, ensuring that slight advantages in score translate to marginal increases in winning probability, rather than all-or-nothing outcomes.

The Power-of-k Algorithm. Let $f(AB)$ and $f(BA)$ be the normalized edge weights derived in the previous section, and let $k \geq 0$ be a parameter. Our randomized rule elects A with probability

$$p_A(k) = \frac{f(AB)^k}{f(AB)^k + f(BA)^k},$$

and elects B otherwise. When $k = 1$, this reduces to choosing A with probability proportional to its edge weight. As $k \to \infty$, this rule converges to the deterministic version. For intermediate k, the rule behaves like a "smoothed" majority based on the geometric evidence. Numerical optimization[1] of this parameter yields a worst-case distortion of approximately 1.523, attained around $k \approx 2.84$, significantly improving upon the deterministic limit. We now show that this result is nearly tight.

[1] The optimization objective for the worst-case distortion does not admit a simple closed-form expression.

Theorem 4.3. *Any randomized social choice rule that considers only voters' ordinal rankings and the outcomes of pairwise deliberations has distortion at least* $3/2$*.*

Proof. We use Yao's minimax principle. Let I_1 and I_2 be the two 2-candidate instances constructed in the proof of Theorem 4.2. By construction, the ordinal preferences and all deliberation outcomes are identical in both instances, yet the underlying costs differ: $SC(B)/SC(A) = 2$ in I_1, whereas $SC(A)/SC(B) = 2$ in I_2.

Consider the distribution $\mathcal{D}$ that assigns probability $1/2$ to I_1 and $1/2$ to I_2. Let R be any deterministic rule that uses rankings and deliberation outcomes. Since the observable information for I_1 and I_2 is identical, R must output the same winner for both instances.

- If R always outputs A, the distortion is 1 on I_1 (where A is optimal) and 2 on I_2 (where B is optimal). Thus, the expected distortion of R on $I \sim \mathcal{D}$ is $(1+2)/2 = 3/2$.
- If R always outputs B, the roles of I_1 and I_2 are swapped, and the calculation remains identical.

Therefore, for every deterministic rule R, $\mathbb{E}_{I\sim\mathcal{D}}[\text{distortion of } R \text{ on } I] \geqslant 3/2$. By Yao's minimax principle, this implies that for any randomized rule, there exists some instance on which the rule has expected distortion at least $3/2$. In other words, no randomized rule using only rankings and deliberation outcomes can achieve a distortion below $3/2$. □

5 Distortion 3 for Many Candidates

We now move on from the two-candidate toy setting of Section 4 to general instances with $m \geqslant 3$ candidates. The protocol is exactly deliberation-via-matching from Section 3: we build a weighted tournament graph using ordinal information and pairwise deliberations, then pick any winner from the λ-WUS. Our goal in this section is to show that for a suitable choice of parameters (λ, w), this protocol has metric distortion at most 3 on every instance. Correspondingly, we show that this result is tight, both with respect to the parameter choice (λ, w), and to the choice of matching which justifies the somewhat vague phrase of "forming a *arbitrary* maximal cardinality matchings" described in the protocol.

The proof will reduce the analysis to a carefully structured three-candidate subproblem and then exploit the geometry of that subproblem (via a reparametrization of distances, triangle inequalities, and convexity) to bound distortion. Until the last section, we treat (λ, w) symbolically and use them implicitly to ease notation. We optimally choose these parameters at the end.

5.1 Reduction to a Three-Candidate Witness

Fix parameters (λ, w) and a metric instance. Let B be an optimal candidate (the 1-median among candidates), and let A be the winner returned by the deliberation-via-matching protocol with these parameters. Our goal is to bound $SC(A)/SC(B)$. Because A lies in the λ-WUS of the weighted tournament built in Section 3, applied to candidate B, there must exist a witness candidate C such that either

- A directly dominates B: $f(AB) \geqslant 1 - \lambda$, or
- A dominates B in two steps through C: $f(AC) \geqslant 1 - \lambda$ and $f(CB) \geqslant \lambda$.

Observe that the first case can be viewed as a degenerate instance of the second: if A directly dominates B, we may simply take $C = B$, and the second inequality is then trivially satisfied. Thus it suffices to analyze the two-step witness case, and we henceforth solve the following optimization problem:

$$\text{find } \sup \frac{SC(A)}{SC(B)} \qquad \text{subject to} \qquad f(AC) \geqslant 1-\lambda, f(CB) \geqslant \lambda. \tag{7}$$

In this situation, the fact that A is allowed to beat B in this tournament is fully certified by the triangle $\{A, B, C\}$; the remaining candidates are irrelevant for this particular relation. Therefore, we may WLOG narrow our focus to them, deleting all other candidates while leaving the metric on $\mathcal{V}$ and on $\{A, B, C\}$ unchanged. This preserves both social costs $SC(A), SC(B)$ and the relevant constraints, so the distortion on this reduced 3-candidate instance is the same as the original one.

Consequently, for fixed (λ, w), the worst-case distortion of the protocol is attained on some 3-candidate instance with candidates $\{A, B, C\}$ satisfying one of the two witnesses conditions above.

5.2 Geometric Reparameterization and the Bilinear Objecitive

We now specialize to the three candidates A, B, C as in Equation (7) and reparameterize all voter-candidate distances in a geometric way that makes the distortion objective linear in a small set of expectations.

Change of Variables. For each voter v, define three real-valued functions on the electorate:

$$X(v) = d(v, C) - d(v, A), \qquad Y(v) = d(v, B) - d(v, C), \qquad Z(v) = d(v, C) \tag{8}$$

so equivalently $d(\cdot, A) = Z - X$, $d(\cdot, B) = Z + Y$, and $d(\cdot, C) = Z$. Then X captures voters' relative preference between A and C, positive when voters prefer A, and Y captures the same for C versus B (positive when C is preferred). This reparameterization is tailored to our protocol: the (A, C) matching and its deliberation outcomes depend only on comparisons of form

$$d(v, A) \leqslant d(v, C) \qquad \text{if and only if} \qquad X(v) \geqslant 0$$

and

$$d(u, A) + d(v, A) \leqslant d(u, C) + d(v, C) \qquad \text{if and only if} \qquad X(u) + X(v) \geqslant 0,$$

so the values $\{X(v)\}$ *completely determine* $|AC|, |CA|$, and the result of every (A, C) deliberation. Likewise for Y. Consequently, X, Y encode *all* information of interest to our protocol.

Viewing Voters as a Distribution. As in Section 2, we normalize total voter mass to 1 and view voters as drawn from a probability distribution over the metric space. In this view, X, Y, Z become random variables and social costs expectations:

$$SC(A) = \mathbb{E}_{v \in \mathcal{V}}[d(v, A)] = \mathbb{E}Z - \mathbb{E}X, \qquad SC(B) = \mathbb{E}_{v \in \mathcal{V}}[d(v, B)] = \mathbb{E}Z + \mathbb{E}Y, \tag{9}$$

so that $SC(A)/SC(B) = [\mathbb{E}Z - \mathbb{E}X]/[\mathbb{E}Z + \mathbb{E}Y]$. We are interested in upper bounding this ratio under the constraints $f(AC) \geqslant 1 - \lambda$, $f(CB) \geqslant \lambda$, which are now purely expressed in terms of the distributions of X and Y.

Linearizing the Distortion Objective. Fix a threshold $R > 0$. Observe that if $SC(A)/SC(B) = [\mathbb{E}Z - \mathbb{E}X]/[\mathbb{E}Z + \mathbb{E}Y]$ exceeds $R + 1$, then the functional

$$\Phi_R(X, Y, Z) = \mathbb{E}X + (R + 1) \cdot \mathbb{E}Y + \mathbb{R} \cdot \mathbb{E}Z < 0. \tag{10}$$

Later, we will choose R appropriately and show that the global minimum of the LHS of Equation (10) is at least zero, and this will imply a distortion of at most $R + 1$. Notice that Φ_R is bilinear in the natural variables:

- For *fixed* (X, Y, Z), it is linear in the voter distribution, and
- For *fixed* masses on each voter type, it is linear in the coordinates X, Y, Z themselves.

Our overall optimization problem is therefore

$$\begin{aligned} \text{Minimize} \quad & \Phi_R(X, Y, Z) = \mathbb{E}X + (R + 1) \cdot \mathbb{E}Y + R \cdot \mathbb{E}Z \\ \text{over} \quad & X, Y, Z \text{ and a distribution over voters} \\ \text{Subject to} \quad & \text{(i) } f(AC) \text{ is induced by } \textit{some} \text{ matching determined by } X; \\ & \text{(ii) } f(CB) \text{ is induced by } \textit{some} \text{ matching determined by } Y; \\ & \text{(iii) } f(AC) \geqslant 1 - \lambda, \;\; f(CB) \geqslant \lambda, \end{aligned} \tag{11}$$

and our goal is to plug in $R = 2$ and argue that the objective is nonnegative.

5.3 Optimal Coupling of the Reparameterized Variables

We now simplify the metric side of Program (11). The two steps are:

(1) Use triangle inequalities to pin down, for each voter v, the smallest possible $d(v, C)$, compatible with X and Y;

(2) Show that, for fixed one-dimensional *distributions* of X and Y (recall we can view them as distributions over $\mathcal{V}$), the worst (most adversarial) case occurs when X and Y are coupled in an *anti-monotone* way: large $X(v)$ paired with small $Y(v)$.

Triangle Inequalities and the Function $Z_{\min}$. Fix functions X, Y on the electorate. From Equation (10), we should pointw-ise minimize Z such that $\{(X(v), Y(v), Z(v))\}_{v\in\mathcal{V}}$ is still metric feasible in the sense that Equation (8) can still be realized in some latent metric space. This observation leads to the following key lemma. By $\|X\|_\infty$ we mean $\max_{v\in\mathcal{V}}|X(v)|$.

Lemma 5.1. *Fix real-valued functions X, Y on the electorate V. For any real-valued function Z on V, in order for (X, Y, Z) to be realized by some metric d under Equation (8), it is necessary and sufficient that*

$$Z(v) \geqslant Z_{\min}(v) = \max\left\{\frac{\|X\|_\infty + X(v)}{2}, \frac{\|Y\|_\infty - Y(v)}{2}, \frac{\|X+Y\|_\infty + X(v) - Y(v)}{2}\right\} \qquad \textit{for all } v. \tag{12}$$

Intuitively, each term comes from one family of triangle inequalities, (v, A, C), (v, B, C), and (v, A, B), respectively. When written in terms of X, Y, Z, these inequalities impose the following constraints:

- For each voter v, $Z(v)$ must be at least the three simple linear expressions in $X(v), Y(v)$, reflecting the three triangles we can draw with v and any two candidates;
- Globally, the ranges of X and Y are bounded by the side lengths of the triangle (A, B, C), and this explains the presence of the $\|\cdot\|_\infty$ terms.

Geometrically $Z_{\min}(x, y)$ can be viewed as the upper envelope of three planes in the (x, y, z)-space, such that any metric-feasible triple must lie above it. Since our objective Equation (10) is increasing in Z, the worst-case (smallest value) of Φ_R for given X, Y is always obtained by pushing each voter down onto this envelope. Consequently, for our task, we can therefore treat Z as a deterministic function of (X, Y).

Couplings and Counter-Monotone Structure. Once we fix the functions X and Y on the electorate, the metric constraints imply it suffices to assign Z pointwise by $Z_{\min}(X, Y)$. Recall that the value of $f(AC)$ depends only on the multiset $\{X(v)\}$, and $f(CB)$ depends only on $\{Y(v)\}$. Under the perspective of viewing X, Y as distributions over $\mathcal{V}$, the tournament constraint fix the marginal distributions of X and Y, but they do not constrain how the values of X and Y are paired across voters. We now analyze this.

For worst-case analysis, we are free to change the **coupling** between X and Y: we may reorder the X- and Y-values arbitrarily among voters. Under such rearrangements, $\mathbb{E}X, \mathbb{E}Y$ remain fixed; the only term in Equation (10) subject to change is $\mathbb{E}[Z_{\min}(X, Y)]$. We are interested in the smallest possible value of Φ_R, so we want the coupling that minimizes $\mathbb{E}[Z_{\min}(X, Y)]$. A key result from this write-up is the following structural lemma.

Lemma 5.2. *Fix one-dimensional distributions of X, Y on $\mathcal{V}$ as well as a parameter $R > 0$. Among all joint distributions (couplings) with these marginals, the functional $\Phi_R(X, Y)$ is minimized when X, Y are **counter-monotone**: there is an indexing of the electorate $\mathcal{V}$ by $t \in [0, 1]$ such that $X(t)$ is non-increasing in t, and $Y(t)$ is non-decreasing in t. Equivalently, whenever two voters v_1, v_2 satisfy $X(v_1) < X(v_2)$, we also have $Y(v_1) \geqslant Y(v_2)$.*

Proof sketch. The key geometric fact is that $Z_{\min}(x, y)$ is the maximum of three affine functions in (x, y) in a specific way, so that the mapping $(x, y) \mapsto Z_{\min}(x, y)$ is supermodular. Consequently, "aligning" large X with large Y *tends* to increase $\mathbb{E}[Z_{\min}(X, Y)]$. This gives rise to a standard exchange argument that iteratively swaps "out-of-order" pairs that is not yet coupled counter-monotonically.

In slightly more details, we note that this explanation is a slight oversimplification, for swapping the coupling may lead to different $\|X+Y\|_\infty$. The full proof proceeds in two steps: first, it shows that against a *frozen* c, the mapping

$$(x, y) \mapsto h_c(x, -y) = \max\{\|X\|_\infty + x, \|Y\|_\infty + (-y), c + x + (-y)\} \tag{13}$$

is supermodular; second, it shows that a local swap does not increase $\|X+Y\|_\infty$. These together complete the proof. □

Lemma 5.2 allows us to impose a very clean structure on worst-case instances. We can now index voters by a parameter $t \in [0, 1]$ such that $X(t)$ is nonincreasing in t, $Y(t)$ is nondecreasing, and $Z(t) = Z_{\min}(X(t), Y(t))$. In other words, the geometric worst case is completely captured by a *one-dimensional* parameterized curve $t \mapsto (X(t), Y(t), Z_{\min}(t))$ running along the upper envelop surface in the (x, y, z)-space, with X decreasing and Y increasing. The remaining work will take place entirely in this 1D picture.

5.4 Optimal Matchings and Tight Constraints

At this point, we have made voter-candidate distances deterministic with respect to X and Y. The remaining degree of freedom is the choice of *maximum matchings* we use between (A, C) and between (C, B). Different matchings can give different numbers of deliberation wins and hence different values of $f(AC)$ and $f(CB)$, even for the voters and candidates. First observe that when everything is fixed, the constraints for $f(AC), f(CB)$ in Equation (11) are made most slack by choosing the matchings with the most number of A-wins for X (resp. C-wins for Y). Call them the **A-optimal (A, C) matching** and the **C-optimal (C, B) matching**, respectively. We will WLOG assume that these are the matchings we seek and use.

This section proves two additional structural reductions, that it is WLOG to assume:

(1) The optimal matchings additionally admit a simple **prefix-suffix** structure along the $[0, 1]$ line, and

(2) The constraints $f(AC) \geqslant 1 - \lambda$ and $f(CB) \geqslant \lambda$ attain equality.

We focus on establishing both for (A, C); the results, once established, directly apply to the (C, B) side. Notation wise, we say a voter pair (u, v) is an **$\mathbf{A \succ C}$ pair** if the pair deliberates in the (A, C) matching and favors A. Similarly, a $C \succ A$ pair is one that deliberates in the (A, C) matching but favors C. The $C \succ B, B \succ C$ pairs are defined analogously.

Prefix Property of Matchings. Even among all A-optimal matchings for (A, C), the pattern of who is matched with whom need not be unique. Our next lemma says that we can choose a particularly nice one: it pairs "most extreme" voters on each side.

Lemma 5.3. *Let W_{AC} be the mass of $A \succ C$ pairs in some A-optimal (A, C) matching. By definition, all A-optimal (A, C) matchings have the same $A \succ C$ mass. In particular, there exists another A-optimal (A, C) matching in which:*

(i) All $A \succ C$ pairs are formed by the leftmost W_{AC} mass (voters with the largest $X(t) \geqslant 0$) to the rightmost W_{AC} mass of the CA side (voters with the most negative $X(t) < 0$), and

(ii) These two blocks are matched counter-monotonically: the largest X on AC pairs with the smallest X on CA, and so on.

It might be helpful to simply refer to Figure 1, where the blue blocks are involved in the $A \succ C$ pairs, whereas the white ones are unmatched or contribute to $C \succ A$.

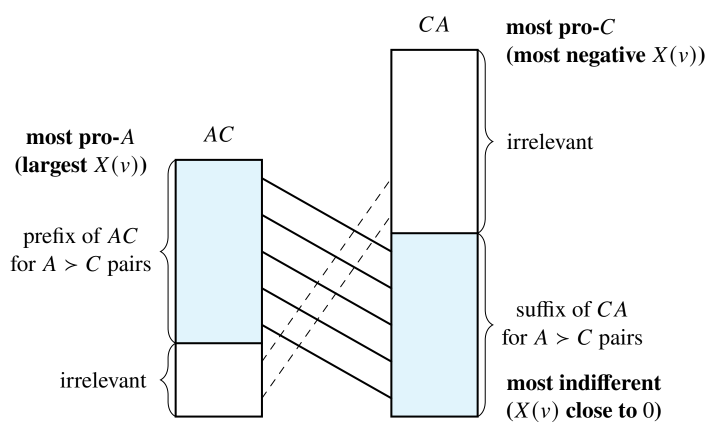

Figure 1: The prefix-suffix structure of $A \succ C$ pairs.

Proof. This proof is an exchange argument at heart, similar to that of Lemma 5.2,

Suppose $X(u_1) \geqslant X(u_2) \geqslant 0 \geqslant X(v_1) \geqslant X(v_2)$. Then we have $X(u_1) + X(v_1) \geqslant 0$ and $X(u_2) + X(v_2) \geqslant 0$. It is easy to check that $X(u_1) + X(v_2) \geqslant 0$ and $X(u_2) + X(v_1) \geqslant 0$. This means we can replace the matchings with (u_1, v_2) and (u_2, v_1). This means the matchings can be made *counter-monotone*. Further, suppose $X(u_1) \geqslant X(u_2) \geqslant 0$ and u_1 does not participate in an $A \succ C$ pair, while u_2 is matched to v_2 in an $A \succ C$ pair. Then we can replace (u_2, v_2) with (u_1, v_2).

Analogously, if $0 \geqslant X(v_1) \geqslant X(v_2)$ and v_1 does not participate in an $A \succ C$ pair while v_2 is matched to u_2 in an $A \succ C$ pair, we can replace (u_2, v_2) with (u_2, v_1). This is feasible for the f constraint since W_{CA} cannot increase in this process, and W_{AC} is preserved.

Iterating this process, we obtain a new (A, C) matching with W_{AC} pairs satisfying $A \succ C$ that additionally meets the criteria described in the lemma statement. This concludes the proof. □

Tightening the f-constraints. The second structural simplification is that we replace $f(AC) \geqslant 1-\lambda, f(CB) \geqslant \lambda$ with $f(AC) = 1-\lambda$ and $f(CB) = \lambda$. This intuitively makes sense, as larger values of $f(AC)$ and $f(CB)$ mean that the electorate likes A over C and C over B more strongly, which overall helps A and hurts B. On the other hand, the distortion is large exactly when A is *unpopular* and B is *popular* (large $SC(A)$, small $SC(B)$). Therefore, it makes sense that the worst case is obtained when these f-constraints are as weak as possible.

Lemma 5.4. *Given any instance with (X,Y) and couplings as above satisfying $f(AC) \geqslant 1-\lambda, f(CB) \geqslant \lambda$, there exists another instance (with the same three candidates and voter mass) such that*

(i) $f(AC) = 1-\lambda$ and $f(CB)$ is unchanged,

(ii) The new (X',Y') still comes from some latent metric, and

(iii) distortion of the new instance is at least that of the old instance.

A symmetric statement holds for tightening $f(CB) \geqslant \lambda$.

Proof sketch. The idea is purely geometric: we "slide" all X-values uniformly away from A by a parameter $t \geqslant 0$, so that over time, $f(AC)$ decreases, and argue that there exists a moment along this trajectory at which $f(AC)$ hits exactly $1-\lambda$. Formally, we replace $X(v)$ by $X(v)-t$, leave Y unchanged, and make other adjustments to the underlying metric as necessary (e.g. candidate-candidate distances). Given $t \geqslant 0$, the adjustments can be realized by

$$\begin{aligned} d_t(v,A) &= d(v,A)+t \qquad d_t(v,C) = d(v,C), \qquad d_t(v,B) = d(v,B) \\ d_t(A,C) &= d(A,C)+t \qquad d_t(C,B) = d(C,B), \qquad d_t(A,B) = d(A,B)+t, \end{aligned} \tag{14}$$

which proves (ii). The comparisons between C and B is unaffected, so under C-optimal matchings, the value of $f(CB)$ stays fixed. For (A,C), as we increase t, more voters and matched pairs move from "A strictly better to C" eventually to "C strictly better than A," so $f(AC)$, as a function of t, is nonincreasing. When t is sufficiently large, everyone prefers C over A, at which point $f(AC)$ becomes 0. The complication is that $f(AC)$, defined by Equation (2), need not be continuous when $|AC|$ or W_{AC} admits a jump. We omit the technical proof here, but the high-level idea is that we can smooth out these jumps by apportioning ordinal preference and/or deliberation ties carefully. This proves (i). Finally, (iii) trivially follows from Equation (14) as $SC(A)$ increases in t but $SC(B)$ stays unchanged. □

We now restate Program (11) in its updated form.

$$\begin{array}{rl} \text{Minimize} & \Phi_R(X,Y) = \mathbb{E}X + (R+1)\cdot\mathbb{E}Y + R\cdot\mathbb{E}Z \\ \text{over} & X, Y \text{ on } \mathcal{V}, \quad Z = Z_{\min}(X,Y) \ \text{ from Equation (12)} \\ \text{Subject to} & \text{[Lemma 5.3]} \ \ f(AC) \text{ is induced by an } A\text{-optimal } (AC, CA) \text{ matching;} \\ & \text{[Lemma 5.3]} \ \ f(CB) \text{ is induced by a } C\text{-optimal } (CB, BC) \text{ matching;} \\ & \text{[Lemma 5.4]} \ \ f(AC) = 1-\lambda, \quad f(CB) = \lambda. \end{array} \tag{15}$$