Contents

This is an informal compilation of current major results. Important sections & paragraphs are highlighted. Please expect typos, repetitions, and other random stuff :-)

1 Copeland-Deliberation Rule on 2 Candidates

1.1 The Copeland-Deliberation Rule

In the notation of ABE18, let us revisit the 2-candidate voting problem. Let A, B be the candidates. We partition the voters into AB and BA where a voter in the former prefers A over B and vice versa in BA. The standard Copeland rule declares A as winner if $a = |AB| \geqslant |BA| = b$. Let $m = \min\{|AB|, |BA|\} = \min\{a, b\}$. Our deliberation step arbitrarily forms a maximal size-m matching $M = \{(s_j, t_j)\}_{j=1}^m \subset AB \times BA$. For each pair (s_j, t_j), we say that A wins the deliberation if

$$d(s_j, A) + d(t_j, A) \leqslant d(s_j, B) + d(t_j, B).$$

Say there are x_a pairs in P that favor A, and x_b pairs that favor B. Then our rule declares A as winner if and only if $a + x_a \geqslant b + x_b$. *Compare this against standard Copeland, which simply compares a against b.*

1.2 Copeland-Deliberation Analysis: 2 Candidates

For convenience, let $\Delta = d(A, B)$, and let M be the matching formed by deliberations. Define

$$M_A = \{(u, v) \in M : A \text{ wins}\} = \{(u, v) \in M : d(u, A) + d(v, A) \leqslant d(u, B) + d(v, B)\}$$

$$M_B = \{(u, v) \in M : B \text{ wins}\} = \{(u, v) \in M : d(u, A) + d(v, A) > d(u, B) + d(v, B)\}.$$

Observe $M_A \cap M_B = \emptyset$ and $M_A \cup M_B = M$. By definition, $|M_A| = x_a$ and $|M_B| = x_b$. Define the "remainders" $R_A = AB \backslash \{u : (u, v) \in M\}$ and $R_B = BA \backslash \{v : (u, v) \in M\}$. Immediately, one (or both) of R_A or R_B must be empty. We WLOG assume that the algorithm declares A as winner. To bound the distortion, we aim to maximize $SC(A)/SC(B)$ and so we upper bound $SC(A)$ and lower bound $SC(B)$.

STEP 1. UPPER-BOUNDING $SC(A)$. Clearly, for every voter v, by triangle inequality

$$d(v,A) \leqslant d(v,B) + \mathbf{1}[v \in BA] \cdot \Delta = \begin{cases} d(v,B) & \text{if } v \in AB \\ d(v,B) + d(B,A) & \text{if } v \in BA. \end{cases} \tag{1}$$

Using M, R to partition the voters, we obtain

$$\begin{aligned} SC(A) &= \sum_{(u,v)\in M} [d(u,A) + d(v,A)] + \sum_{v \in R} d(v,A) \\ &= \sum_{(u,v)\in M_A} [d(u,A) + d(v,A)] + \sum_{(u,v)\in M_B} [d(u,A) + d(v,A)] + \sum_{v\in R_A} d(v,A) + \sum_{v \in R_B} d(v,A). \end{aligned} \tag{2}$$

We analyze the costs term by term.

(A1) On each $(u,v) \in M_A$: As A wins the deliberation, $d(u,A) + d(v,A) \leqslant d(u,B) + d(v,B)$.

(A2) On each $(u,v) \in M_B$: Two applications of (1) give $d(u,A) + d(v,A) \leqslant d(u,B) + d(v,B) + \Delta$.

(A3) (1) is also directly applicable on the last two terms in $R_A \subset AB$ and $R_B \subset BA$.

Summing over everything,

$$SC(A) \leqslant SC(B) + (x_b + |R_b|) \cdot \Delta = SC(B) + (b - x_a) \cdot \Delta \tag{3}$$

as $x_b + |R_b| = (m - x_a) + (b - m) = b - x_a$. [Recall $m = \min(a,b)$ the matching size.]

STEP 2. LOWER-BOUNDING $SC(B)$. For any pair $(u,v) \in M_A$, two applications of the triangle inequality gives

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

With this in mind, we consider certain subsets of voters and analyze their contribution to $SC(B)$:

(B1) Each $(u,v) \in M_A$ collectively contributes Δ to $SC(B)$ by (4). There are x_a such pairs.

(B2) The remaining $a - x_a$ voters in AB each contribute at least $\Delta/2$ to $SC(B)$.

Ignoring everything else, overall

$$SC(B) \geqslant x_a \cdot \Delta + (a - x_a)\Delta/2 = (a + x_a)/2 \cdot \Delta. \tag{5}$$

Combining (3) and (5), we see that

$$\begin{aligned} \frac{SC(A)}{SC(B)} &\leqslant \frac{SC(B) + (b - x_a)\Delta}{SC(B)} \leqslant 1 + \frac{(b - x_a)\cdot\Delta}{(a + x_a)/2 \cdot \Delta} \\ &\leqslant 1 + \frac{2(b - x_a)}{a + x_a} = \frac{a + x_a + 2b - 2x_a}{a + x_a} \\ &= \frac{2a + 2b - a - x_a}{a + x_a} = \frac{2n}{a + x_a} - 1. \end{aligned} \tag{6}$$

Theorem 1.1. *Suppose there are two candidates A, B and n voters. Among the electorate, suppose $a \leqslant n$ voters prefer A. Further, suppose that during the deliberation round, x_a pairs end up favoring A. Define $\text{score}(A) = a + x_a$ and likewise for $\text{score}(B)$. Then,*

$$\frac{SC(A)}{SC(B)} \leqslant 1 + \frac{b - x_a}{(a + x_a)/2} = \frac{2n}{a + x_a} - 1 = \frac{2n}{\text{score}(A)} - 1. \tag{7}$$

Proposition 1.2. *Under this rule, the distortion for any 2-candidate instance is bounded by 2.*

Proof. By the theorem, it suffices to show that if A wins, then $\text{score}(A) \geqslant 2n/3$.
To prove this claim, we WLOG assume $|AB| \leqslant |BA|$. If A is the winner, then $a + x_a \geqslant b + x_b = b + (a - x_a)$ iff $2x_a \geqslant b = n - a$. But we also have $x_a \leqslant a$, so $2x_a \leqslant 2a$ and $n - a \leqslant 2a$ implies $a \geqslant n/3$. Hence

$$\text{score}(A) = a + x_a \geqslant a + \frac{n-a}{2} = \frac{n+a}{2} \geqslant \frac{n + (n/3)}{2} = \frac{2n}{3}.$$

If instead $|AB| \geqslant |BA|$, then the winner condition forces $a + x_a \geqslant b + x_b = b + (b - x_a) = 2b - x_a$, or equivalently $2x_a \geqslant 2b - a$. If $a \geqslant 2n/3$ there is nothing to show, so we assume $n/2 \leqslant a \leqslant 2n/3$. In this case, the requirement becomes $2x_a \geqslant 2(n - a) - a = 2n - 3a$. Then,

$$\text{score}(A) = a + x_a \geqslant a + \frac{2n - 3a}{2} = \frac{2n - a}{2} \geqslant \frac{2n - 2n/3}{2} = \frac{2n}{3}.$$ □

1.3 Comparison Against Standard Copeland

Recall that the standard Copeland rule only compares $a = |AB|$ against $b = |BA|$. It is well known that given two candidates A, B, we always have $SC(A)/SC(B) \leqslant 2n/|AB| - 1$, to which our Theorem 1.1 is analogous. To prove this, we note that

$$\begin{aligned} SC(A) &= \sum_{v \in AB} d(v, A) + \sum_{v \in BA} d(v, A) \leqslant \sum_{v \in AB} d(v, B) + \sum_{v \in BA} [d(v, A) + d(v, B)] \\ &= SC(B) + |BA| \cdot d(A, B), \end{aligned}$$

so $SC(A)/SC(B) \leqslant 1 + |BA| \cdot d(A, B)/SC(B)$. Like before, we note that each AB voter contributes at least $d(A, B)/2$ to $SC(B)$, so

$$\frac{SC(A)}{SC(B)} \leqslant 1 + \frac{|BA| \cdot d(A, B)}{SC(B)} \leqslant 1 + \frac{(n - |AB|) \cdot d(A, B)}{|AB| \cdot d(A, B)/2} = \frac{2n}{|AB|} - 1.$$

An interesting connection is as follows. In standard Copeland, we naturally define a normalized score $g(AB) = |AB|/n$ so that $g(AB) + g(BA) = 1$. Likewise, as seen in the following section, we can define a normalized Copeland-deliberation score $f(AB) = \text{score}(AB)/[\text{score}(AB) + \text{score}(BA)]$. Below we show that the distortion as functions of $f(AB)$ and $g(AB)$, differ by precisely 1 on $[0, 1/2]$.
From above, we already know that if $g(AB) = |AB|/n \geqslant r$, then the distortion is bounded by

$$G(\lambda) = 2/r - 1. \tag{8}$$

On the other hand, if given $f(AB) \geqslant r$, to use Theorem 1.1 we need to minimize $\text{score}(AB)$. Thus, we solve the following program

$$\begin{aligned} \text{minimize} \quad & a + x_a \\ \text{subject to} \quad & a + b = 1 \\ & x_a + x_b = \min(a, b) \\ & (a + x_a) \geqslant \lambda(a + b + x_a + x_b) \\ & a, b, x_a, x_b \geqslant 0. \end{aligned} \tag{9}$$

which equals $2r/(2-r)$ if $r \in [0, 1/2]$ and $2r/(1+r)$ if $r \in [1/2, 1]$ (note the value w.r.t. r is continuous). Applying the Theorem, we see that if $f(AB) \geqslant r$, then the distortion is bounded by

$$F(\lambda) = \begin{cases} (2-r)/r - 1 = 2/r - 2 & \text{if} \quad 0 \leqslant r \leqslant 1/2 \\ (1+r)/r - 1 = 1/r & \text{if} \quad 1/2 \leqslant r \leqslant 1. \end{cases} \tag{10}$$

Observe that G and F differ by precisely 1 on $[0, 1/2]$, and as $r \to 1$, their difference converges to 0.

2 Extending Copeland-Deliberation to General Instance

Now we consider instances with more candidates. Notation-wise, we still let $\text{score}(AB) = |AB| + x_{AB}$ to be the deliberation-augmented score of A when compared against B. We also introduce the notion of normalized score, $f(AB) := \text{score}(AB)/[\text{score}(AB) + \text{score}(BA)]$, the benefits being that $f(AB) + f(BA) = 1$ for all distinct pairs of candidates A, B.

To construct a tournament graph, we will pick two candidates, X, Y, define a maximal matching between XY and YX. Below we show that it suffices to choose the maximal matching that is X-optimal (cf. Theorem 2.2). Then, using the rules of Copeland-deliberation as outlined in the previous section, we compute $f(XY), f(YX)$ accordingly. Now, we repeat this process for each pair of candidates X, Y until the entire tournament graph is defined.

2.1 λ-Weighted Uncovered Sets

Standard literature defines a candidate A to be in the λ-weighted uncovered set (WUS) if for all other B, either

- (easy case) $f(AB) \geqslant 1 - \lambda$, or
- (hard case) there exists another C such that $f(AC) \geq 1 - \lambda$ and $f(CB) \geq \lambda$.

It is known that when $\lambda \in [0.5, 1)$, the λ-WUS is nonempty. As an λ-weighted uncovered candidate can reach any other candidate in at most 2 hops via one of the above cases, to perform distortion analysis on any element in the λ-WUS, it suffices to look at a 3-candidate instance.

By substituting $1 - \lambda \in (0, 1/2)$ into Equation (10), the easy case immediately yields a distortion upper bound of $2\lambda/(1 - \lambda)$. From now on, we will focus exclusively on the hard case: We aim to calculate the supremum of $SC(A)/SC(B)$ subject to $f(AC) \geqslant 1 - \lambda$ and $f(CB) \geqslant 1 - \lambda$.

The proof is rather long, but here is a high-level roadmap. We argue that in order to find the supremum distortion, it suffices to conduct a series of reductions. In the end, we only consider only two parametric families, which are sufficiently simple in structure that we can directly use LP to model and solve. More specifically, we show the following reductions in the order they are listed:

(Current) We want to find $\sup SC(A)/SC(B)$ such that the instance admits an $A-C$ deliberation with $f(AC) \geqslant 1-\lambda$ and a $C - B$ deliberation with $f(CB) \geqslant \lambda$.

(Theorem 2.2) It suffices to consider the A-optimal deliberation for $f(AC)$ and the C-optimal deliberation for $f(CB)$, i.e., given a fixed instance, consider the largest $f(AC), f(CB)$ attainable.

(Theorem 2.4) It suffices to assume $f(AC) = 1 - \lambda$ and $f(CB) = \lambda$ instead of $\geqslant$.

(Theorem 2.8) It suffices to additionally assume that the sets $AC = BC$ and $CB = CA$.

(Theorem 2.10) Given λ, it suffices to consider the following two parameterized families of instances (feasibility is discussed in detail later):

- ($AC_{\min}, CB_{\max}$, i.e., C-loses-all) Let $|AC|$ be the smallest feasible quantity to meet $f(AC) = 1 - \lambda$ (by letting A win all $A - C$ deliberations) and $|CB|$ be the largest feasible quantity to meet $f(CB) = \lambda$ (by letting C lose all $C - B$ deliberations).

- ($AC_{\max}, CB_{\min}$, i.e., C-wins-all) Let $|AC|$ be the largest feasible quantity to meet $f(AC) = 1 - \lambda$ (A loses all against C) and $|CB|$ the smallest feasible quantity to meet $f(CB) = \lambda$ (C wins all against B).

Once established, Theorem 2.10 gives a very strong structural characterization of our instances of interest, from which we conclude with Theorem 2.13 that

$$\sup_{\substack{f(AC)\geqslant 1-\lambda \\ f(CB)\geqslant\lambda}} \frac{SC(A)}{SC(B)} = \max\left\{\frac{2}{\lambda}, \frac{2\lambda}{1-\lambda}\right\}.$$

2.2 Measure-Space Model and Other Technical Preliminaries

Section 1 worked with finite electorates, where sums, orderings, and matchings are purely combinatorial. Section 2, however, pursues a rather more analytical approach where we constantly use continuous approximations. To use these tools freely, we therefore pass to an equivalent measure-space model: every finite instance embeds as an atomic probability space, and all statements below specialize back to the finite case by interpreting integrals as discrete sums.

To enhance readability, in the following sections, we will use simple terms as long as it doesn't hurt rigor, and only resort to technical statements when necessary. We first describe our generalized setup.

Voters & candidates. The **electorate** is a probability space (V, μ) with $\mu(V) = 1$. For a measurable set $S \subset V$, write $|S| = \mu(S)$. An **instance** consists of an electorate V, three **candidates** A, B, C, and a metric defined on $V \cup \{A, B, C\}$. We define the **social cost** of a candidate X as $SC(X) = \int_V d(v, X)\, \mathrm{d}\mu(v)$.

For candidates $X, Y \in \{A, B, C\}$ and $v \in V$, define the **intensity** $\Delta_{XY}(v) = d(v, Y) - d(v, X)$ which quantifies how much v prefers X over Y. The preference set partitions remain unchanged: $XY = \{v : \Delta_{XY}(v) \geqslant 0\}$ and $YX = \{v : \Delta_{XY}(v) < 0\}$, and let $m_{XY} = \min\{|XY|, |YX|\}$.

Deliberation & maximal matchings. The first real generalization appears in defining a maximal matching in the deliberation step. A **maximal matching** between XY and YX is a finite measure γ on $XY \times YX$ such that there exist measurable $S \subset XY, T \subset YX$ with $|S| = |T| = m_{XY}$ and

$$\gamma(A \times YX) = |A \cap S|, \qquad \gamma(XY \times B) = |B \cap T|, \qquad \text{for all measurable } A, B.$$

Given a maximal matching γ, the **win mass** for X is

$$W_{XY}(\gamma) = \gamma(\{(u, v) : \Delta_{XY}(u) + \Delta_{XY}(v) \geqslant 0\}).$$

An ***X*-optimal** matching is a maximizer $\gamma^* = \operatorname{argmax}_\gamma W_{XY}(\gamma)$ (existence & characterization given by Theorem 2.2).

The $f(\cdot)$ (Copeland-deliberation) function and the objective. Given X, Y and a maximal matching γ on X, Y, we defined $f(XY) = (|XY| + W_{XY}(\gamma))/(1 + m_{XY})$. Our goal is to compute $\sup SC(A)/SC(B)$ subject to $f(AC) \geqslant 1 - \lambda, f(CB) \geqslant \lambda$. Observe that it is WLOG to assume that we take the X-optimal matching γ when computing $f(XY)$: indeed, tautologically V admits an $XY - YX$ matching with $f(XY) \geqslant c$ if and only if its X-optimal matching satisfies $f(XY) \geqslant c$. Therefore, to require $f(AC) \geqslant 1 - \lambda$, it suffices to require that the A-optimal $AC - CA$ matching satisfies this, and likewise for $f(CB)$.

Therefore, from now on, we define

$$f(XY) = \frac{|XY| + W_{XY}(\gamma^*)}{1 + m_{XY}} \qquad \text{where } \gamma^* \text{ is the } X\text{-optimal } XY - YX \text{ maximal matching,}$$

and we seek to upper bound, with the updated $f(\cdot)$-constraints, the same objective:

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

X-optimal matching and greedy pairing. We now prove the existence and characterize the structure of an X-optimal matching γ^* of on (XY, YX). The intuition is that we need to "sort" XY from the most pro-X to the most indifferent voters (start with $u \in XY$ who *really* prefers X over Y, or equivalently, large $\Delta_{XY}(u)$), and "sort" YX from the most indifferent to the most pro-Y voters, and pair them up "greedily." The discrete version of this claim is easy to establish and prove via a standard exchange argument.

We first formalize the notion of "sorting" as well as "rank" in a continuum.

Definition 2.1 (Tail Counts and Ascending Quantiles)**.** *Let X, Y be given, and Δ_{XY} be defined as above. For $t \geqslant 0$, we define the **tail counts** as*

$$\varphi^+(t) = \{u \in XY : \Delta_{XY}(u) \geqslant t\}, \qquad \varphi^-(t) = |\{v \in YX : -\Delta_{XY}(t) \geqslant t\}$$

*to be the subsets of XY, YX whose Δ intensity is at least as strong as t in magnitude. We also define the **ascending quantiles** of Δ_{XY} on XY and of $-\Delta_{XY}$ on YX as follows:*

$$Q^+(q) = \inf\{t : |\{u \in XY : \Delta_{XY}(u) \leqslant t\}| \geqslant q\}, \qquad 0 \leqslant q \leqslant |XY|$$

$$Q^-(q) = \inf\{t : |\{v \in YX : -\Delta_{XY}(v) \leqslant t\}| \geqslant q\}, \qquad 0 \leqslant q \leqslant |YX|.$$

Intuitively, this describes the continuous "rank" of a voter whose intensity's magnitude $= t$.

Theorem 2.2 (Characterization of $\boldsymbol{X}$-Optimal Matchings)**.** *Fix X, Y, and let $m = m_{XY} = \min\{|XY|, |YX|\}$. Define*

$$\alpha = \sup_{t \geqslant 0} \max\left\{0, \underbrace{\max\{|\varphi^-(t)| - (|YX| - m), 0\}}_{\text{hard } YX \text{ that must be included}} - \underbrace{\min\{|\varphi^+(t)|, m\}}_{\text{strong enough } XY}\right\} \in [0, m]. \tag{11}$$

Let γ be any matching on (XY, YX). Then the X-win mass $W_{XY}(\gamma) \leqslant m - \alpha$. Moreover, there exists a γ^* attaining this equality, so in particular an X-optimal matching γ^* exists ($\sup_\gamma W_{XY}(\gamma)$ is attained). Furthermore, in plain terms, such γ^* greedily matches the most pro-X voter in XY with the least pro-Y voter in YX until one side of XY, YX is exhausted. We note that the monstrosity of Equation (11) is not important; its sole purpose is to prove the existence of γ^*.

Proof. We first prove that $W_{XY}(\gamma) \leqslant m - \alpha$. Let γ be any maximal matching with marginals $S \subset XY, T \subset YX$. Since

$$|\varphi^-(t) \cap T| \geqslant |\varphi^-(t)| - |YX \backslash T| = |\varphi^-(t)| - (|YX| - m),$$

we have in particular $|\varphi^-(t) \cap T| \geqslant \max\{0, |\varphi^-(t)| - (|YX| - m)\}$. Intuitively, voters in $\varphi^-(t)$ are the sufficiently pro-Y (with intensity $\leqslant -t$) so they are "hard" to be convinced into W-wins. The inequality above lower bounds the total number of "hard" YX voters that must be present in a maximal matching. On the other hand, $|\varphi^+(t) \cap S| \leqslant \min\{|\varphi^+(t)|, m\}$ upper bounds the number of XY voters that are sufficiently pro-X (with intensity $\geqslant t$) to possibly beat the "hard" voters. Every voter in $\varphi^-(t) \cap T$ must be matched, and an A-loss is guaranteed if matched to a voter outside $\varphi^+(t)$. Hence, the number of A-*losses* is at least $(|\varphi^-(t) \cap T| - |\varphi^+(t) \cap S|)$ which upper bounds Equation (11) after taking supremum over $t \geqslant 0$. Hence $W_{XY}(\gamma) \leqslant m - \alpha$.

We now exhibit a maximal matching γ^* that incurs precisely α X-losses. First assume $|XY| \leqslant |YX|$ so $m = |XY|$. Let $S = XY$ and choose $T \subset YX$ to be the "easiest" m-mass of YX, i.e., the subset with measure m and the smallest values of Δ_{XY}. Intuitively, we greedily pair the most pro-X voters in S with the least pro-Y voters in T until we exhaust both sets and show this matching has precisely α X-losses. The rest of the proof is to formalize this in measure theory. Observe for all t, $|\varphi^-(t) \cap T| = \max\{\varphi^-(t) - (|YX| - m), 0\}$. This allows us to define CDFs on S and T via

$$F^+(t) = |\{u \in S : \Delta_{XY}(u) \leqslant t\}| = |XY| - S^+(t), \qquad F^-(t) = |\{v \in T : -\Delta_{XY}(v) \leqslant t\}| = m - |\varphi^-(t) \cap T|. \tag{12}$$

By the definition of α, we have $F^-(t) \leqslant F^+(t) + \alpha$ for all t. Now let $Q^+(q), Q^-(q)$ be the ascending quantiles of Δ_{XY} on S and of $-\Delta_{XY}$ on T. Define measure-preserving parameterizations $\pi^+ : [0, m] \to S, \pi^- : [0, m] \to T$ with

$$\Delta_{XY}(\pi^+(q)) = Q^+(q), \qquad -\Delta_{XY}(\pi^-(r)) = Q^-(r) \qquad \text{a.e.}$$

and a shift map $f : [0, m] \to [0, m]$ by $f(q) = q - \alpha \mod m$. Finally, set γ^* to be the pushforward of Lebesgue measure on $[0, m]$ by $q \mapsto (\pi^+(q), \pi^-(f(q)))$. By construction, γ^* has marginals S and T.

We claim that for almost every $q \in [\alpha, m]$, $Q^+(q) \geqslant Q^-(q - \alpha)$. Set $t = Q^-(q - \alpha)$. By definition of Q^-, for every $\epsilon > 0$, we have $F^-(t - \epsilon) < q - \alpha$ and $F^-(t) \geqslant q - \alpha$. Recall that $F^-(t) \leqslant F^+(t) + \alpha$, so

$$F^+(t - \epsilon) \leqslant F^-(t - \epsilon) + \alpha < q.$$

Letting $\epsilon \searrow 0$ and using right-continuity of F^+ yields $F^+(t-) \leqslant q$, hence $Q^+(q) \geqslant t$ by the definition of the quantile as a generalized inverse. Hence, for a.e. $q \in [\alpha, m]$ we have

$$\Delta_{XY}(\pi^+(q)) + \Delta_{XY}(\pi^-(f(q))) = Q^+(q) - Q^-(q - \alpha) \geqslant 0,$$

so those pairs are X-wins. Therefore $W_{XY}(\gamma^*) \geqslant m - \alpha$. Combining this with the upper bound $W_{XY} \leqslant m - \alpha$ completes the proof.

If $|XY| \geqslant |YX|$ then we set $T = YX$ and $S \subset XY$ to be the subset of measure m and largest values of Δ_{XY}. The rest of the proof follows analogously. □

Metric realizability and omission of candidate-candidate distances. So far — and throughout this section — we will only perform analysis, modify instances, or create new ones by changing or defining the candidate-candidate distances $d(X, Y)$ and the voter-candidate distances $d(v, X)$. Voter-voter distance $d(u, v)$ never enters the definitions or analyses we use. We claim and justify that it is harmless to specify just these values and ignore $d(u, v)$ entirely: When a full metric on $V \cup \{A, B, C\}$ is desired, any partial specification that satisfies triangle inequalities on each voter-candidate-candidate triple (v, X, Y) and on $\{A, B, C\}$ extends to a *bona fide* metric via

$$d(u, v) = \min_{X, Y \in \{A, B, C\}} (d(u, X) + d(X, Y) + d(Y, v))$$

which preserves all variables that social costs and $f(\cdot)$-values depend on.

2.3 Reduction I: Making the f-Constraints Tight

In this section, we argue that in order to find $\sup SC(A)/SC(B)$ subject to $f(AC) \geqslant 1 - \lambda$ and $f(CB) \geqslant \lambda$, it suffices to restrict our attention to instance where $f(AC) = 1 - \lambda$ and $f(CB) = \lambda$. Overall, we first show an easy claim (Lemma 2.3): it suffices to assume (at least) one of them is tight. The heavy lifting (Theorem 2.4) is to show that conditioned on one $f(\cdot)$ being tight, it is WLOG to assume that the other one also is.

Lemma 2.3. *For the purpose of maximizing distortion, it suffices to assume that at least one $f(\cdot)$ constraint is tight, i.e.:*

$$\sup_{\substack{f(AC)\geqslant 1-\lambda \\ f(CB)\geqslant\lambda}} \frac{SC(A)}{SC(B)} = \max\left\{ \sup_{\substack{f(AC)=1-\lambda \\ f(CB)\geqslant\lambda}} \frac{SC(A)}{SC(B)}, \sup_{\substack{f(AC)\geqslant 1-\lambda \\ f(CB)=\lambda}} \frac{SC(A)}{SC(B)} \right\}$$

Proof. Pick any instance with $f(AC) > 1-\lambda$ and $f(CB) > \lambda$. Continuously pad voters at B (and re-normalize) so that $SC(A)/SC(B)$ strictly increases. Doing so will drive $f(CB)$ down. Stop when either $f(AC)$ hits $1-\lambda$ or $f(CB)$ hits λ. ◻

The next two claims together will indicate that it suffices to assume that *both* constraints are tight. For convenience, we will normalize the total voter mass to 1 and use $\mu(v)$ to denote the measure/weight of a voter cluster v.

Theorem 2.4 (From one tight constraint to two). $\sup_{\substack{f(AC)=1-\lambda \\ f(CB)\geqslant\lambda}} \frac{SC(A)}{SC(B)} = \sup_{\substack{f(AC)=1-\lambda \\ f(CB)=\lambda}} \frac{SC(A)}{SC(B)}$.

Proof. Throughout, write $\Delta(v) = d(v,B) - d(v,C)$ for the C vs. B intensity.

STEP 1. TIE-BREAKING VIA PRODUCT EXTENSION. We first refine the instance so that $C-B$ related ties occur on sets with measure 0, all while keeping $f(AC)$ fixed and either increasing or keeping $f(CB)$ fixed too.

Fix $\epsilon > 0$, and pass V and μ to the product space. We define

$$V' = V \times (0,\epsilon), \qquad \mathrm{d}\mu'(v,r) = \mathrm{d}\mu(v)\frac{\mathrm{d}r}{\epsilon}.$$

In this product space, define a metric d' on V' by

$$d'((v,r),A) = d(v,A) + r, \qquad d'((v,r),C) = d(v,C) + r, \qquad d'((v,r),B) = d(v,B) + \epsilon$$
$$d'(A,B) = d(A,B) + \epsilon, \qquad d'(C,B) = d(C,B) + \epsilon, \qquad d'(A,C) = d(A,C).$$

Intuitively, for each voter v, we clone it uniformly onto a continuum of $(0,\epsilon)$ and preserve the total mass. Observe that the changes made to $d'(\cdot,A)$ and $d'(\cdot,C)$ are identical, so the sets AC, CA, the $A-C$ maximal matching outcomes, as well as $f(AC)$, all remain unchanged. On the other hand,

$$\Delta'(v,r) = d'((v,r),B) - d'((v,r),C) = \Delta(v) + (\epsilon - r) > \Delta(v),$$

so every voter's CB intensity strictly increases. Under the C-optimal matching, $f(CB)$ cannot decrease.

Crucially, we observe that (CB, BC)-related level sets have measure zero:

- For a voter v and any $x \in \mathbb{R}$,

$$\mu'\left(\{(v,r) : \Delta'(v,r) = x\}\right) = \int_V \underbrace{\text{Lebesgue}\left(\{r \mid \epsilon - r = x - \Delta(v)\}\right)}_{=0} \mathrm{d}\mu(v) = 0,$$

  so the distribution of Δ' is non-atomic and $\mu'\left(\{\Delta' = 0\}\right) = 0$.

- For two voters $u \in CB, v \in BC$, and any $c \in \mathbb{R}$,

$$(\mu'{\times}\mu')\left(\{(u,r_1),(v,r_2) : \Delta'(u,r_1) + \Delta'(v,r_2) = c\}\right) = \int_{CB\times BC} \underbrace{\mu'\left(\{(v,r_2) : \Delta'(v,r_2) = c - \Delta'(u,r_1)\}\right)}_{=0} \mathrm{d}\mu'(u,r_1) = 0$$

  by Fubini and non-atomicity of Δ'.

Therefore, both single ties (with respect to candidate preferences) and pair-sum ties have measure 0 (w.r.t. appropriate measure). Furthermore, because $\mathbb{E}[r] = \epsilon/2$, we have $SC'(A) = SC(A) + \mathbb{E}[r] = \epsilon/2$ and $SC'(B) = SC(B) + \epsilon$. Hence in this step the distortion changes by $SC'(A)/SC'(B) - SC(A)/SC(B) = O(\epsilon)$.

Step 2. Monotonically pushing $f(CB)$ down only. We now describe an operation that preserves $f(AC) = 1 - \lambda$ but continuously pushes down $f(CB)$. Based on d' defined in Step 1, for $t \geqslant 0$, we define

$$\begin{aligned} d_t((v,r),A) &= d'((v,r),A) + t & d_t((v,r),C) &= d'((v,r),C) + t & d_t((v,r),B) &= d'((v,r),B) \\ d_t(A,B) &= d'(A,B) + t & d_t(C,B) &= d'(C,B) + t & d_t(A,C) &= d'(A,C) + 2t \end{aligned} \tag{13}$$

Once again, the changes to $d_t(\cdot, A)$ and $d_t(\cdot, C)$ are identical, so the $f(AC)$ remains unchanged. On the other hand, for every (v,r) we have

$$\Delta_T(v,r) = d_t((v,r),B) - d_t((v,r),C) = \Delta'(v,r) - t.$$

Consider the partition of V by $CB_t = \{(v,r) : \Delta'(v,r) \geqslant t\}$ and $BC_t = \{(v,r) : \Delta'(v,r) < t\}$. Let $m_t = \min\{|CB_t|, |BC_t|\}$. Consider constructing the C-optimal (CB, BC) matching using the quantile argument (Definition 2.1 and Theorem 2.2). Let Q_t^+ be the ascending quantile of Δ' on CB_t and Q_t^- the ascending quantile of $-\Delta'$ on BC_t, both parameterized on $[0, m_t]$. By definition, a matched pair $((u^+, r^+), (u^-, r^-)) \in CB \times BC$ at rank q is a C-win if and only if

$$(\Delta'(u^+, r^+) - t) + (\Delta'(u^-, r^-) - t) \geqslant 0 \iff Q_t^+(q) - Q_t^-(q) \geqslant 2t.$$

The C-win mass and the Copeland-deliberation score can be respectively written as

$$W_C(t) = \int_0^{m_t} \mathbf{1}\{Q_t^+(q) - Q_t^-(q) \geqslant 2t\}\, \mathrm{d}q, \qquad f(CB;t) = \frac{|CB_t| + W_C(t)}{1 + m_t}. \tag{14}$$

Step 3. The monotonic push is continuous. We now argue that $t \mapsto W_C(t)$ is continuous. This, along with $t \mapsto m_t$ being continuous, establishes that $t \mapsto f(CB;t)$ is also continuous. It is clear that when $t = 0$, we recover the results from Step 1, so $f(CB; t = 0) > \lambda$. As t gets sufficiently large, every voter will rank B over C according to Equation (13), at which point $f(CB;t)$ is guaranteed to be 0. Then continuity result implies there exists t^* such that $f(CB;t^*) = \lambda$, and furthermore,

$$\frac{SC_t(A)}{SC_t(B)} \geqslant \frac{SC'(A)}{SC'(B)} \geqslant \frac{SC(A)}{SC(B)} - O(\epsilon),$$

where the first $\leqslant$ is by Equation (13) and the second by the last line of Step 1. As ϵ is arbitrary, the proof is then complete.

Coming back, now we prove the continuity of $t \mapsto W_C(t)$. To this end let $t_n \to t$; we show $W_C(t_n) \to W_C(t)$. Because Δ' is made non-atomic by Step 1, its CDF $F(x) = |\{\Delta' \leqslant x\}|$ is continuous. Thus $|CB_t| = 1 - F(t)$, $|BC_t| = F(t)$, and m_t vary continuously in t. The quantiles $Q_t^{\pm}(\cdot)$ are also continuous a.e. Crucially, from Step 1, the pair-sum tie set $\{q \in [0, m_t] : Q_t^+(q) - Q_t^-(q) = 2t\}$ has Lebesgue measure 0, so the indicators converge pointwise a.e.:

$$1\{Q_{t_n}^+(q) - Q_{t_n}^-(q) \geqslant 2t_n\} \to \mathbf{1}\{Q_t^+(q) - Q_t^-(q) \geqslant 2t\}.$$

The integrands of W_C are bounded by 1, so by dominated convergence, $W_C(t_n) \to W_C(t)$, as desired. □

Corollary 2.5 (From one tight constraint to two, mirrored). $\displaystyle \sup_{\substack{f(AC) \geqslant 1-\lambda \\ f(CB)=\lambda}} \frac{SC(A)}{SC(B)} = \sup_{\substack{f(AC)=1-\lambda \\ f(CB)=\lambda}} \frac{SC(A)}{SC(B)}.$

Proof. The idea of the other case also works here. We only outline the changes made to d' and d_t. To define d' we let

$$d'((v,r),A) = d(v,A), \qquad d'((v,r),C) = d(v,C)+r, \qquad d'((v,r),B) = d(v,B)+r.$$

We realize this by defining $d'(A,B) = d(A,B)+\epsilon$, $d'(C,B) = d(C,B)$, and $d'(A,C) = d(A,C)+\epsilon$. For the second step, for $t \geqslant 0$ we let

$$d_t((v,r),A) = d'((v,r),A)+t, \qquad d_t((v,r),C) = d'((v,r),C), \qquad d_t((v,r),B) = d'((v,r),B),$$

with $d_t(A,B) = d'(A,B)+t$, $d_t(C,B) = d'(C,B)$, and $d_t(A,C) = d'(A,C)+t$. □

Theorem 2.6. *Combining the previous three claims, we achieve the following reduction, and from now on we will assume the two $f(\cdot)$ constraints are tight:*

$$\sup_{\substack{f(AC)\geqslant 1-\lambda \\ f(CB)\geqslant \lambda}} \frac{SC(A)}{SC(B)} = \sup_{\substack{f(AC)=1-\lambda \\ f(CB)=\lambda}} \frac{SC(A)}{SC(B)}.$$

2.4 Reduction II: Reusing the Partition

In this section, we show that we may further restrict our attention to instances free of voters who rank C as their *second* choice; in other words, those who prefer $A > C$ must also rank $B > C$, and those prefer $C > B$ must also rank $C > A$. Hence, the (AC, CA) partition is identical to the (BC, CB) partition.

This claim, once proved, along with the tight constraints as assumed in Theorem 2.6, gives us a very powerful characterization: the mass of A-wins (in the $A-C$ deliberation) plus the mass of C-wins (in the $C-B$ deliberation) must add up to $\min\{|AC|,|CA|\} = \min\{|CB|,|BC|\}$, which we crucially exploit in the next section (Theorem 2.10).

We first present a simple yet powerful Lemma whose result we will repeatedly apply throughout this section.

Lemma 2.7 (Deliberation tuning & scaling)**.** *Fix an ordered pair of candidates $X, Y \in \{A,B,C\}$. Let $s = |XY|$ and write $m = \min\{|XY|, 1-|XY|\}$ as the deliberation matching size. Define the intensity function Δ on the electorate by $\Delta(v) = d(v,Y) - d(v,X)$. Then for any instance with $|XY| = s$, necessarily $f(XY) \in [s/(1+m), (s+m)/(1+m)]$. We claim the following.*

(i) (Full feasible range is attainable by three-level profiles). For any target value $\tau \in [s/(1+m), (s+m)/(1+m)]$, there exists an intensity function Δ taking (at most) three distinct values, such that under the X-optimal matching, $f(XY) = \tau$.

(ii) (Scaling). If we multiply (i)'s Δ by any scalar $c > 0$, the matching outcomes and therefore $f(XY)$ remain unchanged. In particular, for any $\epsilon > 0$ we can realize $f(XY) = \tau$ with $\|\Delta\|_\infty < \epsilon$.

Proof. (ii) is trivial (i) is established. For (i), fix τ and set $\alpha = \tau(1+m) - s \in [0,m]$ to be the total mass of X-wins from the deliberation. Pick constants $x^- < 0 < x_0 < x^+$ such that $x^+ + x^- > 0$ but $x_0 + x^- < 0$. Define Δ as follows: on XY, $\Delta = x^+$ on a set of measure α and x_0 otherwise, and on YX, set $\Delta \equiv x^-$ uniformly. Then the X-optimal matching will prioritize matching $\{X = x^+\} \cap XY$ with YX, resulting in a total win mass of α. This proves the claim. □

Theorem 2.8 (One partition, two uses)**.** *Fix $\lambda \in (0.5, 1)$. In maximizing $SC(A)/SC(B)$ subject to $f(AC) = 1-\lambda$ and $f(CB) = \lambda$, it is WLOG restrict to instances in which $AC = BC$ and $CA = CB$ (up to null sets), i.e., no voter has strict ranking ACB or BCA.*

Proof. STEP 1. $f(AC), f(CB)$ AS ONE-DIMENSIONAL MARGINALS. Define three variables X, Y, Z by

$$X(v) = d(v,C) - d(v,A), \qquad Y(v) = d(v,B) - d(v,C), \qquad Z(v) = d(v,C).$$

Then $AC = \{X \geqslant 0\}, CA = \{X < 0\}, CB = \{Y \geqslant 0\}$, and $BC = \{Y < 0\}$. Also define $m_{XY} = \min\{|XY|, |YX|\}$.
For the $A - C$ comparison, let $Q_X^+(q)$ be the decreasing quantile of X restricted to AC, and $Q_X^-(q)$ the increasing quantile of X on CA, so $0 \leqslant q \leqslant |AC|$ or $0 \leqslant q \leqslant |CA|$ as appropriate. Then the A-win mass in $A - C$ deliberation is

$$W_A(X) = \int_0^{m_{AC}} \mathbf{1}\{Q_X^+(q) + Q_X^-(q) \geqslant 0\}\, \mathrm{d}q$$

and $f(AC) = (|AC| + W_A(X))/(1 + m_{AC})$. Likewise, for the $C - B$, define $Q_Y^+(q)$ as the decreasing quantile of Y on CB and $Q_Y^-(q)$ the increasing quantile of Y on BC, so that

$$W_C(Y) = \int_0^{m_{CB}} \mathbf{1}\{Q_Y^+(q) + Q_Y^-(q) \geqslant 0\}\, \mathrm{d}q, \qquad f(CB) = \frac{|CB| + W_C(Y)}{1 + m_{CB}}.$$

In particular, we see that $f(AC)$ (resp. $f(CB)$) is a functional of the one-dimensional laws of X (resp. Y) only.

STEP 2. CHOOSING MASSES AND TARGETS. We now aim to enforce $|AC| = |BC| = m \in (0, 1/2]$ (so $m_{AC} = m_{CB} = m$, and $|CA| = |CB| = 1 - m$) while also preserving $f(AC) = 1 - \lambda, f(CB) = \lambda$. Hence we need to enforce win masses $0 \leqslant W_A, W_C \leqslant m$ such that

$$\frac{m + W_A}{1 + m} = 1 - \lambda, \qquad \frac{(1 - m) + W_C}{1 + m} = \lambda.$$

These along with $m \leqslant 1/2$ require $m \in [(1-\lambda)/(1+\lambda), \min\{1/2, (1-\lambda)/\lambda\}]$. This interval is nonempty for $\lambda \in (0.5, 1)$. Pick any feasible m and compute W_A, W_C correpsondingly.
STEP 3. THE OBJECTIVE DEPENDS ONLY ON MARGINALS. In this part, we show that we can in fact make $AC = BC$ and $CA = CB$ up to null sets by decoupling X and Y as we wish. Indeed,

$$SC(A) = \int_V (Z - X)\, \mathrm{d}\mu = \mathbb{E}Z - \mathbb{E}X, \qquad SC(B) = \int_V (Z + Y)\, \mathrm{d}\mu = \mathbb{E}Z + \mathbb{E}Y,$$

so $SC(A)/SC(B) = [\mathbb{E}Z - \mathbb{E}X]/[\mathbb{E}Z + \mathbb{E}Y]$ which depends only on the marginal expectations. Consequently, altering the coupling of X, Y across voters will not alter the constraint or the objective. With $|AC| = |BC| = m$ as in STEP 3, let $B \sim \text{Bernoulli}(m)$. Conditioned on B, independently draw X, Y by

$$\begin{cases} B = 1 : X \text{ from its law on } AC;\ Y \text{ from its law on } BC \\ B = 0 : X \text{ from its law on } CA;\ Y \text{ from its law on } CB. \end{cases}$$

This preserves the marginals of X and Y and hence preserves $f(AC), f(CB), \mathbb{E}X, \mathbb{E}Y$, as well as the overall distortion $SC(A)/SC(B)$. Further, it enforces that $\mathbf{1}\{X \geqslant 0\} = B = \mathbf{1}\{Y < 0\}$ almost surely, so $AC = BC$ and $CA = CB$ up to null sets, as claimed. This establishes STEP 3.

STEP 4. REALIZATION IN A METRIC. To complete the proof, we show that the post-STEP 3 triple (X, Y, Z) (where $AC = BC, CA = CB$ up to null sets) can be realized by some metric d' while preserving $f(\cdot)$ and arbitrarily approximates $SC(A)/SC(B)$. Pick any constant $T \geqslant \sup|X| + \sup|Y|$ and define

$$\begin{cases} d'(v,A) = Z(v) + T - X(v) \\ d'(v,C) = Z(v) + T \\ d'(v,B) = Z(v) + T + Y(v) \end{cases} \quad \text{and} \quad \begin{cases} d'(A,C) = \sup|X| \\ d'(C,B) = \sup|Y| \\ d'(A,B) = \sup|X + Y|. \end{cases}$$

It is easy to verify that d' obeys triangle inequality among $\{A, B, C\}$, as well as among any triplets of form $\{v, X, Y\}$. On the other hand, Lemma 2.7 implies that $\sup|X|$ and $\sup|Y|$ can be made arbitrarily small. Therefore we may push $T \searrow 0$ and see that

$$\frac{SC'(A)}{SC'(B)} = \frac{\mathbb{E}Z + T - \mathbb{E}X}{\mathbb{E}Z + T + \mathbb{E}Y} = \frac{SC(A) + T}{SC(B) + T} \to \frac{SC(A)}{SC(B)}.$$

Taking the supremum completes the proof. □

2.5 Reduction III: Two Types of Instances Suffice

In this section, we argue that two types of instances suffice (in addition to previous assumptions: $f(AC) = 1 - \lambda$, $f(CB) = \lambda$, and $AC = BC, CA = CB$): we need to check either instances where C wins all deliberations from *both* $A - C$ and $C - B$ contests, or ones where C loses all.

The key benefit of this result is that both deliberations are made unilateral, which greatly simplifies the structures of the instances we wish to analyze (Theorem 2.11 from next section).

Remark 2.9. Fix a $\lambda \in (1/2, 1)$. With $f(AC) = 1 - \lambda$, one must have $AC_{\min} \leqslant |AC| \leqslant AC_{\max}$ where

$$AC_{\min} = \frac{1-\lambda}{1+\lambda}, \qquad AC_{\max} = \begin{cases} (2-2\lambda)/(2-\lambda) & \lambda \leqslant 2/3 \\ (1-\lambda)/\lambda & \text{otherwise.} \end{cases}$$

Likewise, $|CB|$ is bounded between

$$CB_{\min} = \begin{cases} \lambda/(2-\lambda) & \lambda \leqslant 2/3 \\ 2 - 1/\lambda & \text{otherwise,} \end{cases} \qquad CB_{\max} = \frac{2\lambda}{1+\lambda}.$$

These are the quantities that allow the $f(\cdot)$ constraints to be satisfied by winning all deliberations (min) or winning zero deliberation (max).

Theorem 2.10 (Extremal instances suffice). *Fix $\lambda \in (1/2, 1)$. For $a \in [0, 1]$, let $I_\lambda(a)$ be the class of instances satisfying the following:*

- *Set size $|AC| = a$,*
- *Tight deliberation constraints $f(AC) = 1 - \lambda, f(CB) = \lambda$, and*
- *Overlap: $AC = BC$ and $CA = CB$ up to null sets.*

(Note tight constraints immediately imply that $AC_{\min}(\lambda) \leqslant a \leqslant AC_{\max}(\lambda)$.) Then,

$$\sup_{a \text{ feasible}} \sup_{I_\lambda(a)} \frac{SC(A)}{SC(B)} = \max\left\{ \sup_{I_\lambda(AC_{\min})} \frac{SC(A)}{SC(B)}, \sup_{I_\lambda(AC_{\max})} \frac{SC(A)}{SC(B)} \right\}.$$

In other words, for any λ, to find the supremum distortion, it suffices to analyze two cases: one where $|AC| = |BC| = AC_{\min}$ and $|CA| = |CB| = CB_{\max}$, the other where $|AC| = |BC| = AC_{\max}$ and $|CA| = |CB| = CB_{\min}$.

Proof. With tight constraints $f(AC) = 1 - \lambda, f(CB) = \lambda$, as well as $AC = BC, CA = CB$, we are in fact able to completely characterize the deliberations: Indeed, the A-win mass in the $A - C$ deliberation and the C-win mass in the $C - B$ deliberation, denoted $W_A(a, \lambda)$ and $W_C(a, \lambda)$, are fixed by

$$\frac{a + W_A}{1 + \min\{a, 1-a\}} = 1 - \lambda, \qquad \frac{(1-a) + W_C}{1 + \min\{a, 1-a\}} = \lambda, \tag{*}$$

which implies the identity $W_A(a,\lambda) + W_C(a,\lambda) = \min\{a, 1-a\}$; note this is independent of λ.

In this proof, we resort to the following well-known result. (With abuse of notation) if $\varphi : [a,b] \to \mathbb{R}$ is continuous, and if for all $x \in (a,b)$ there exists $\epsilon > 0$ such that either $\varphi(x+\epsilon) > \varphi(x)$ or $\varphi(x-\epsilon) > \varphi(x)$, then φ attains its maximum on an endpoint. Likewise, in this proof, given $a \in (AC_{\min}, AC_{\max})$ and any instance in $I_\lambda(a)$ and a small $\epsilon > 0$, we consider the following operations:

- Move an ϵ-mass of voters from AC to CA, so $|AC|$ decreases by ϵ, or
- Move an ϵ-mass of voters from CA to AC, so $|AC|$ increases by ϵ.

We show that (at least) one of them strictly increases the distortion $D = SC(A)/SC(B)$. This will help us conclude that the supremum is attained by an endpoint where $|AC| \in \{AC_{\min}, AC_{\max}\}$, as claimed. We continue using our previous definitions of X, Y, Z and the observation that $SC(A)/SC(B) = [\mathbb{E}Z - \mathbb{E}X]/[\mathbb{E}Z + \mathbb{E}Y]$.

Step 1. What happens when a changes? From (*), we can obtain closed-form expression for $W_A(a,\lambda)$ and $W_C(a,\lambda)$ as functions of a:

$$\begin{cases} W_A(a,\lambda) = 1-\lambda-\lambda a & a \leqslant 1/2 \\ W_A(a,\lambda) = (1-\lambda)(2-a) - a & a > 1/2, \end{cases} \quad \text{and} \quad \begin{cases} W_C(a,\lambda) = -1+\lambda+(1+\lambda)a & a \leqslant 1/2 \\ W_C(a,\lambda) = 2\lambda - 1 + (1-\lambda)a & a > 1/2. \end{cases}$$

Differentiating with respect to a gives $W_A'(a,\lambda) = -\lambda$ if $a \leqslant 1/2$ and $-(2-\lambda)$ otherwise, as well as $W_C'(a,\lambda) = 1+\lambda$ if $a \leqslant 1/2$ and $1-\lambda$ otherwise. Thus, when $m \mapsto m \pm \epsilon$, we much change W_A, W_C by $O(\epsilon)$.

Step 2. Implementing the update. Concretely, we carry out the following procedure so the constraints (set sizes and deliberation outcomes) remain exactly tight while the change in overall distortion is $o(\epsilon)$:

- Prepare low-intensity buffers. Inside AC and CB, reserve buffers $\mathcal{B}_{AC} \subset AC$ and $\mathcal{B}_{CB} \subset CB$ of measure $\gg \epsilon$. Using Lemma 2.7 we may rescale the magnitudes of $|X|$ and $|Y| \leqslant c$ for arbitrarily small $c > 0$ without changing $|AC|, |CB|$, or the $f(\cdot)$ constraints.
- Compute $\Delta W_A = W_A(a \pm \epsilon, \lambda) - W_A(a,\lambda)$ and ΔW_C analogously. They are $O(\epsilon)$.
- Flip winners on $\Theta(\epsilon)$-mass only. Again recall Lemma 2.7, in which we parameterize X by three values: a "high" x^+ and a "low" x_0 positive value for AC, and a negative value x^- for CA, such that "high" corresponds to A-wins ($x^+ + x^- > 0$) and "low" corresponds to C-wins ($x_0 + x^- < 0$). If $\Delta W_A > 0$, select a "low" subset of $\mathcal{B}_{AC}$ with measure ΔW_A and raise its X levels to the "high," so these pairs now become A-wins. If $\Delta W_A < 0$, demote a set of measure $|\Delta W_A|$ from "high" to "low" so they become C-wins. Perform actions symmetrically on $\mathcal{B}_{CB}$ to adjust ΔW_C. Doing so preserves $f(AC) = 1-\lambda$ and $f(CB) = \lambda$.

Because $|X|, |Y| \leqslant c$, we have $|\Delta\mathbb{E}X| + |\Delta\mathbb{E}Y| \leqslant O(c\epsilon) = o(\epsilon)$ as $c \searrow 0$. Since the ratio depends only on $(\mathbb{E}X, \mathbb{E}Y, \mathbb{E}Z)$, using Theorem 2.8 Step 4, we may recouple (X,Y) with Z to keep $\mathbb{E}Z$ unchanged. We quantify these effects below.

Step 3. First-order effects of the update. We define four new quantities

$$\mu_X^+ = \mathbb{E}[X \mid AC], \qquad \mu_X^- = \mathbb{E}[X \mid CA], \qquad \mu_Y^- = \mathbb{E}[Y \mid AC], \qquad \mu_Y^+ = \mathbb{E}[Y \mid CA]$$

so that $\mu_X^+ \geqslant 0 \geqslant \mu_X^-$ and $\mu_Y^+ \geqslant 0 \geqslant \mu_Y^-$. We consider two actions: $+\epsilon$ or $-\epsilon$ mass to AC.

(i) (Move from AC to CA) Take $S \subset AC$ of measure ϵ with means $x_S = \mathbb{E}[X \mid S]$, $y_S = \mathbb{E}[Y \mid S]$. After STEP 2,

$$\Delta\mathbb{E}X = \epsilon(\mu_X^- - x_S) + o(\epsilon), \qquad \Delta\mathbb{E}Y = \epsilon(\mu_Y^+ - y_S) + o(\epsilon).$$

(i) (Move from CA to AC) Analogously, take $T \subset CA$ of measure ϵ with means x_T, y_T. Then

$$\Delta\mathbb{E}X = \epsilon(\mu_X^+ - x_T) + o(\epsilon), \qquad \Delta\mathbb{E}Y = \epsilon(\mu_Y^- - y_T) + o(\epsilon).$$

Let $D = SC(A)/SC(B)$ (old distortion) and $D' = [SC(A) - \Delta\mathbb{E}X]/[SC(B) + \Delta\mathbb{E}Y]$ (new). First-order expansion gives

$$D' - D = \frac{SC(A) - \Delta\mathbb{E}X}{SC(B) + \Delta\mathbb{E}Y} - \frac{SC(A)}{SC(B)} = \frac{[\mathbb{E}Z + \mathbb{E}Y](-\Delta\mathbb{E}X) - [\mathbb{E}Z - \mathbb{E}X](\Delta\mathbb{E}Y)}{[\mathbb{E}Z + \mathbb{E}Y]^2} + o(\epsilon). \qquad (**)$$

To determine the sign of $D' - D$, we introduce

$$g(v) = (\mathbb{E}Z + \mathbb{E}Y)X(v) + (\mathbb{E}Z - \mathbb{E}X)Y(v),$$

and for any set A write $\overline{g}_A = [\mathbb{E}Z + \mathbb{E}Y]\mathbb{E}[X \mid A] + (\mathbb{E}Z - \mathbb{E}X)\mathbb{E}[Y \mid A]$ as the weighted average of g over A. This gives us a clean way to characterize the sign of (**):

- (Move from AC to CA) $\operatorname{sgn}(D' - D) = \operatorname{sgn}\left([\mathbb{E}Z + \mathbb{E}Y](x_S - \mu_X^-) + [\mathbb{E}Z - \mathbb{E}X](y_S - \mu_Y^+)\right) = \operatorname{sgn}(\overline{g}_S - \overline{g}_{CA})$;
- (Move from CA to AC) $\operatorname{sgn}(D' - D) = \operatorname{sgn}\left([\mathbb{E}Z + \mathbb{E}Y](\mu_X^+ - x_T) + [\mathbb{E}Z - \mathbb{E}X](\mu_Y^- - y_T)\right) = \operatorname{sgn}(\overline{g}_{AC} - \overline{g}_T)$.

Intuitively, these sign equations indicate that a tiny perturbation increases distortion iff the perturbed set's g-average beats the destination/source g-average.

STEP 4. EXISTENCE OF A STRICTLY IMPROVING DIRECTION. Since the electorate is atomless, we may take small sets with g-means as close as possible to the essential suprema/infima.

If g is essentially constant, apply Lemma 2.7 to perturb the magnitudes of X and Y inside AC, CA. This changes μ_X, μ_Y, and hence $\overline{g}_{AC}, \overline{g}_{CA}$, by $o(1)$, while keeping all $|XY|$ and $f(XY)$ unchanged. After that, only two possibilities exist. If $\operatorname{ess\,sup}_{AC} g > \overline{g}_{CA}$, choose $S \subset AC$ with sufficiently large $\overline{g}_S$ so $\overline{g}_S - \overline{g}_{CA} > 0$; this guarantees that the $AC \to CA$ move strictly increases distortion. Otherwise, we must have $\operatorname{ess\,inf}_{CA} g < \overline{g}_{AC}$, so pick $T \subset CA$ sufficiently near the infimum, which guarantees that the $CA \to AC$ move strictly improves distortion.

Wrapping up, we showed that for any interior $|AC| \in (AC_{\min}, AC_{\max})$, the distortion can always be strictly increased along at least one side locally. This, along with continuity of distortion as a function of a continuously changing AC, conclude the proof. □

2.6 Proving Bound $\max\{2/\lambda, 2\lambda/(1-\lambda)\}$

Theorem 2.11 (Distortion of C-loses-all instances). *Let $\lambda \in (0.5, 1)$. Consider a tight C-loses-all instance where:*

- *$AC = BC, CA = CB$,*
- *$f(AC) = 1 - \lambda$ and A wins every deliberation in the $A - C$ matching, so $|AC| = AC_{\min} = (1-\lambda)/(1+\lambda)$, and*
- *$f(CB) = \lambda$ and B wins every deliberation in the $C - B$ matching, so $|CB| = CB_{\max} = 1 - AC_{\min} = 2\lambda/(1+\lambda)$.*

Among all such instances, the supremum distortion is given by

$$\sup_{C\text{-loses-all}} \frac{SC(A)}{SC(B)} = \begin{cases} 3 & \text{if } \lambda \leqslant 3/5 \\ 2\lambda/(1-\lambda) & \text{if } \lambda > 3/5. \end{cases}$$

Proof. We prove this by writing a linear program. First observe that $AC_{\min}$ never exceeds $1/2$ for all $\lambda \in (0.5, 1)$, so the AC/BC side is always exhausted in each matching. This allows us to partition the electorate into three parts:

- $S = AC = BC$ of mass $(1-\lambda)/(1+\lambda)$,
- $T_1 \subset CA = CB$ also of mass $(1-\lambda)/(1+\lambda)$ that is matched against S in the $A - C$ deliberation, and
- $T_2 = CA \backslash T_1$ (this set has mass $(3\lambda - 1)/(1+\lambda)$).

Note that S, T_1, T_2 cleanly partition the electorate by all events we use (individual preference, as well as matching role). For instance, everyone in S ranks both A, B over C, and is paired with someone of type ranks C over both $\{A, B\}$; further, everyone in S is matched in both deliberations and achieves the same outcome (wins $AC - CA$, loses $CB - BC$). In particular, no finer subdivision of S, T_1, T_2 changes which inequalities apply.
Thus, we may compress each part $R \in \{S, T_1, T_2\}$ to a point r by $d(r, X) = |R|^{-1} \int_R d(v, X)\, \mathrm{d}v$ for $X \in \{A, B, C\}$. Doing so does not break any constraints and enables us to instead analyze instances with just three atoms s, t_1, t_2 representing S, T_1, T_2, respectively.

Now we describe the linear program, with base variables A, B, C (candidates) and s, t_1, t_2 (voter locations). The weights are fixed: $w(s) = w(t_1) = (1-\lambda)/(1+\lambda)$, and $w(t_2) = (3\lambda - 1)/(1+\lambda)$. Consider the following program. Throughout, v denotes arbitrary elements in $\{s, t_1, t_2\}$ and X, Y denote candidates $\in \{A, B, C\}$.

$$\begin{aligned}
\text{maximize} \quad & SC(A) = \sum_{v \in \{s, t_1, t_2\}} w(v) d(v, A) \\
\text{subject to} \quad & d(A, B) \leqslant d(A, C) + d(C, B) \text{ and cyclic permutations} \\
& |d(v, X) - d(v, Y)| \leqslant d(X, Y) \leqslant d(v, X) + d(v, Y) && (\Delta\text{-ineq}) \\
& d(s, A) \leqslant d(s, C), \qquad d(s, B) \leqslant d(s, C) && (s \in AC \cap BC) \\
& d(t_i, C) \leqslant d(t_i, A), \qquad d(t_i, C) \leqslant d(t_i, B) \qquad i \in \{1, 2\} && (t_i \in CA \cap CB) \\
& d(s, A) + d(t_1, A) \leqslant d(s, C) + d(t_1, C) && (A - C \text{ deliberation \& } A\text{-win}) \\
& d(s, B) + d(t_i, B) \leqslant d(s, C) + d(t_i, C) \qquad i \in \{1, 2\} && (C - B \text{ deliberation \& } B\text{-win}\dagger) \\
& SC(B) = \sum_{v \in \{s, t_1, t_2\}} w(v) d(v, B) = 1. && (\text{normalization})
\end{aligned}$$

We put $i \in \{1, 2\}$ in (†) because we assumed zero C-win in the $C - B$ matching even if the matching is C-optimal (cf. **??**), so if we pair any BC voter with any CB voter, the result must be a B-win. (Metric realizability is given by **??**.) Solving this program gives the claimed upper bounds. TODO: I think this problem is now much, much easier and can be proven analytically quite easily. Will replace once I have that. (Observe the transition at $\lambda = 3/5$ is precisely when $|T_1| = |T_2|$.) □

Theorem 2.12 (Distortion of ***C***-wins-all instances)**.** *Let $\lambda \in (0.5, 1)$. Symmetrically, consider a tight C-wins-all instance where:*

- *$AC = BC, CA = CB,$*
- *$f(AC) = 1 - \lambda$ is realized with zero A-wins, so*

$$|AC| = AC_{\max} = \begin{cases} (2 - 2\lambda)/(2 - \lambda) & \lambda \leqslant 2/3 \\ (1 - \lambda)/\lambda & \textit{otherwise}, \end{cases}$$

- *and $f(CB) = \lambda$ is realized with all C-wins, so*

$$|CB| = CB_{\min} = 1 - AC_{\max} = \begin{cases} \lambda/(2-\lambda) & \lambda \leqslant 2/3 \\ 2 - 1/\lambda & \textit{otherwise}. \end{cases}$$

Among all feasible instances, the supremum distortion is given by

$$\sup_{\textit{C-wins-all}} \frac{SC(A)}{SC(B)} = \begin{cases} 2/\lambda & \textit{if } \lambda \leqslant 2/3 \\ (3\lambda - 1)/(1-\lambda) & \textit{if } \lambda > 2/3. \end{cases}$$

Proof. Analogous to the previous proof, with the caveat that as λ exceeds 2/3, the CB/CA side now becomes smaller and exhausted in matchings, so two separate LPs may be needed. □

We are now ready to restate the main theorem by combining Theorem 2.11 and Theorem 2.12:

Theorem 2.13 (Distortion of λ-WUS Copeland-Deliberation). *Fix $\lambda \in (1/2, 1)$. Consider any metric instance with candidates A, B, C, an electorate V, and the deliberation-Copeland scores $f(AC), f(CB)$ as specified by* **??**. *If $f(AC) \geqslant 1 - \lambda$ and $f(CB) \geqslant \lambda$, then the distortion satisfies*

$$\sup \frac{SC(A)}{SC(B)} = \max\left\{\frac{2}{\lambda}, \frac{2\lambda}{1-\lambda}\right\} = \begin{cases} 2/\lambda & \lambda \in (1/2, \lambda^*] \\ 2\lambda/(1-\lambda) & \lambda \in (\lambda^*, 1) \end{cases}$$

where $\lambda^* = (\sqrt{5} - 1)/2 \approx 0.618$. In particular, at λ^*, the distortion $\max\{2/\lambda, 2\lambda/(1-\lambda)\}$ is minimized with value $1 + \sqrt{5} \approx 3.236$.