1 Domination Graphs with Matching Certificates

1.1 High-Level Walkthrough

So far, every known deterministic rule with optimal worse-case distortion (of 3) relies, either explicitly or implicitly, on the *domination graph*. The common theme is to show that for some winner A, the corresponding domination graph G_A admits a perfect matching, which then implies that the distortion of A is no more than 3. Intuitively we knew (or hope) that pairwise deliberation may provide additional information beyond what ordinal rankings can provide. So this naturally brings out the following question:

> **Deliberation & domination graphs.** How does pairwise deliberation interact with matching-based, optimal deterministic rules? Specifically, what is a "natural" notion to incorporate deliberation into domination graphs?

To this end I considered the following "upgrade" to the domination graph of a candidate A:

- Original: an ordinal edge $i \to j$ that says voter i ranks candidate A at least as high as $\mathrm{top}(j)$.
- New: retain only those ordinal edges that are additionally certified by a two-person deliberation between voters (i, j) on candidates $(A, \mathrm{top}(j))$.

Formally, fix a profile and an unknown metric d and fix a candidate A. Let G_A be the corresponding (standard, ordinal) **domination graph**, bipartite with left and right copies of voters, where an edge $i \to j$ is present if voter i ranks A at least as high as $\mathrm{top}(j)$. We define the **deliberation-refined subgraph** $H_A \subset G_A$ by keeping an edge $i \to j$ only if the deliberation between (i, j) comparing candidates $(A, \mathrm{top}(j))$ weakly prefers A, i.e.,

$$d(i, A) + d(j, A) \leqslant d(i, \mathrm{top}(j)) + d(j, \mathrm{top}(j)).$$

The hoped-for certificate is: find a candidate A such that H_A admits a perfect matching, then run algebra. If this worked, each certified edge would additionally encode two-voter cardinal inequalities, so one could hope to re-run the matching argument and potentially improve the constant below 3. Pushing through triangle inequality, it is not difficult to show the following result, which shows that deliberation can reduce distortion by 1 (much alike 4.236 vs. 3.236 for $\lambda = 0.618$ uncovered set).

Theorem 1.1. ***If for some candidate*** A***,*** H_A ***has a perfect matching, then the distortion of*** A ***is at most*** 2***.***

However, there are many major setbacks, and I have not found a fix once and for all. As of writing this summary, I tend to believe that this direction (augmenting along domination graph) does not have much potential.

- First and foremost, unlike G-graphs, **there need not exist a candidate** A **whose** H**-graph admits a perfect matching** (Theorem 1.2).
- **Even if** $S = \{A \in V : H_A$ **has a perfect matching**$\}$ **is non-empty, any well-defined deterministic rule selecting a winner from this set inevitably has supremum distortion** 3 (Theorem 1.3). This is proven by designing a class of instances whose S are singletons, so a deterministic selector has only one choice, but this candidate in S can be parameterized to have distortion $> 3 - \epsilon$ for any $\epsilon > 0$.
- A compromise I sought was a weaker notion: **select a candidate** A **whose** H**-graph is *close* to admitting a perfect matching**, if not already.

  We consider the "edit distance" of H_A to a perfectly matchable bipartite graph by considering both ordinal and deliberation information: most costly (2) when we wish to add an edge that is absent from both G_A and H_A, and cheapest (0) when the edge is already present in both. Otherwise the edge must have been present in G_A but not in H_A, and this costs 1. It is more convenient to describe this exact idea using a weighted bipartite graph. For a fixed candidate A, we define a complete bipartite graph on (V_ℓ, V_r) with edge weights

  – (already good) $w_A(i, j) = 0$ if $(i, j) \in H_A$,
  – (need to fix deliberation) $w_A(i, j) = 1$ if $(i, j) \in G_A \backslash H_A$, and

– (need to fix ordinal and deliberation) $w_A(i, j) = 2$ if $(i, j) \notin G_A$.

Then $\rho(A)$, candidate A's distance to the "nearest" perfectly matchable graph, can be expressed as

$$\rho(A) = \min\{\text{total weight of a perfect matching } (V_\ell, V_r) \text{ under } w_A\}.$$

This approach fixes the previous problems and has distortion at most $7/3$ when $|V| = 2$. (Theorems 1.4 to 1.6). In comparison, domination-graph-based, ordinal-information-only algorithms have distortion 3 even when $|V| = 2$. **However, for general instances this rule still has supremum distortion** 3 (Theorem 1.7), and I currently believe the obstacle is structural (inherent to the rule) rather than just a bad choice of ρ.

1.2 Detailed Theorem Statements

Theorem 1.2. ***There exists a family of instances in which H_A has no perfect matching for any candidate A. Furthermore, these instances can parameterized to have distortion $\Omega(n)$.***

This claim is proven by the following construction. Let L be large. Consider a star metric with a center and n leaves labeled $1, \ldots, n$, such that for the first $n-1$ branches, center-leaf distances are 1, and for the last one, L. For each $k \in [n]$, place one voter v_k and one candidate c_k right at leaf k.

Theorem 1.3. ***For all $\epsilon > 0$, there exists an instance with two candidates A, B satisfying the following:***

- ***G_A admits a perfect matching, while H_A, G_B, and H_B don't, and***
- ***$3 > SC(A)/SC(B) > 3 - \epsilon$.***

This claim is proven by the following construction. Consider an instance on $\mathbb{R}$ where candidate A is located at 0 and B at 1. Place $(k+1)$ voters at $1/2 - \delta$, and k voters at $B = 1$. Then let $k \to \infty$.

Theorem 1.4. ***By selecting the minimizer of $\arg\min_{v \in V} \rho(v)$ as winner, all instances considered in the proofs of Theorems 1.1 to 1.3 result in distortion $\leqslant 2$.***

Theorem 1.5. ***The counterexample considered in the proof of Theorem 1.3 can be parameterized to have distortion up to $7/3$ under the rule of selecting $\arg\min_{v \in V} \rho(v)$.***

Theorem 1.6. ***The rule selecting $\arg\min_{v \in V} \rho(v)$ has distortion at most $7/3$ when there are only two candidates.***

Theorem 1.7. ***For all $\epsilon > 0$, there exists an instance in which the ρ-minimizer has distortion $> 3 - \epsilon$.***

This claim is proven by another variant of the star metric. Let L, k be given. Let the metric be a star with center o, with short leaves $s_1, \ldots, s_{3k}$ at distance $d(o, s_i) = 1$, and one far leaf f with $d(o, f) = L$. For each short leaf s_i, place one candidate and one voter. For the far leaf, place one candidate but $k+1$ voters. It can then be shown that short candidates have ρ-value $4k - 1$ whereas the far candidates have ρ-value $4k - 2$.

2 Matching Uncovered Set + Deliberation

2.1 A More Relaxed Notion

Another notation to consider is the **Matching Uncovered Set**, a weaker notion of domination graphs. Specifically, given candidates A, B, (the classic definition is that) we define a voter-voter bipartite graph $G(A, B)$, in which an edge (u, v) exists if there exists a candidate C such that:

(1) (left) Voter u prefers C over B, and

(2) (right) Voter v prefers A over C.

Then we select a candidate A such that for all B, the graph $G(A, B)$ has a perfect matching. With deliberation, we define $H(A, B)$ by replacing (2) and keeping (1) unchanged, and still select a candidate A such that $H(A, \cdot)$ always has a perfect matching: an edge (u, v) exists if

(1) (left) Voter u prefers C over B, and

(2') (right deliberation) The voter pair (u, v) favors A in the (A, C) deliberation.

It is easy to prove the following results:

Theorem 2.1. ***If there exists a candidate A such that $H(A, B)$ has a perfect matching for each $B \neq A$, then the distortion of A is at most 2.***

Theorem 2.2. ***If for some candidate A, the deliberation-augmented domination graph H_A admits a perfect matching as in Theorem 1.1, then $H(A, B)$ admits a perfect matching for each B.***

However, even with this weaker notation of matching uncovered set, there need not be a candidate A satisfying the conditions of Theorem 2.1.

Theorem 2.3. ***There exists an instance in which every candidate fails to satisfy the conditions of Theorem 2.1.***

Theorem 2.4. *For every $\epsilon > 0$, there exists a metric instance with two candidates A, B such that:*

- *The (standard) matching-uncovered-set graph $G(A, B)$ admits a perfect matching, but $G(B, A)$ does not.*
- *Neither deliberation graph admits a perfect matching: $H(A, B)$ and $H(B, A)$ both fail to admit a perfect matching.*
- *B is socially optimal and $3 > SC(A)/SC(B) > 3 - \epsilon$.*

2.2 What Went Wrong?

Conceptually, I think this is due to **discontinuity** of the objective (existence of a perfect matching) and the unerlying structure (bipartite graphs). **In our previous project, the rules were continuous w.r.t. perturbations.** Recall in our paper, we were able to continuously adjust quantities such as $f(AC)$ by shifting the marginal variables like $X_t(v) = X_0(v) - t$. When there were ties/jump discontinuities, we overcame this problem by arbitrating the ties. **The perfect matching based methods are not continuous in this regard,** as they are integral ("does there exist a matching or no?"). Small purturbations can essentially ruin a perfect matching and delete edges in a way that creates major imbalances on Hall conditions (i.e., the in/out neighbor of a set is far smaller, so no perfect matching can exist), see below. I think this is the key cause of why it is so hard to optimize in the worst case.

Additionally, changing (2) to (2') also doesn't sound canonical. An edge in the standard MUS $(u, v) \in G(A, B)$ is inherently stating a witness condition on a three-candidate chain $A - C - B$, that u is responsible for certifying "there exists a C that beats B for u," and v is responsible for certifying A beats that C for v." The whole point is that these two pieces of evidence come from different voters without interfering what each other thinks.

In the definition of $H(A, B)$, however, we lose this independence. By replacing (2) with (2'), we drag the left voter u (who was only supposed to witness $C \succ B$) into the A vs. C certificate itself. In some sense, u is allowed to "contaminate" the A vs. C comparison, and because of this, **an adversary can select instances where the voters C**

who can witness (1) are exactly the ones who force (2') to fail for almost every partner v. This structural weakness was consistently exploited by Theorem 2.3 and in many examples in Section 1. The construction in Theorem 2.3 contains "die-hard" fans who are co-located with candidate B. For any such left-side voter u, condition (1) forces $C = B$, but then for any other v, condition (2') can never hold. **Combinatorially, this creates vertices with small neighborhoods in $H(A, B)$ which prevents perfect matching from happening,** tying back to my first argument on discontinuity (as a perturbation by δ in Theorem 2.3 instantly recovers all edges needed for perfect matchings).

(That said, I believe the following notion also doesn't work. Define $G(A, B)$ via (1) and (2), and define $H(A, B)$ via (1), (2), and additionally edge (i, j) needs to satisfy that the deliberation between (i, j) over (A, B) favors A. Pick the candidate A who is "closest" to having perfect matching in $H(A, \cdot)$ in the sense of ρ defined in Section 1. More likely than not, I think need a new perspective.)

A Proofs for Section 1

Proof of Theorem 1.1. Let the perfect matching M on H_A be a bisecting $\pi : V \to V$, so edges are $(i, \pi(i)) \in M$. By assumption we know $d(i, A) + d(\pi(i), A) \leqslant d(i, \mathrm{top}(\pi(i))) + d(\pi(i), \mathrm{top}(\pi(i)))$. Summing over $i \in V$, the LHS becomes

$$\sum_i d(i, A) + \sum_i d(\pi(i), A) = 2SC(A)$$

since π is a permutation. For the RHS, manipulating triangle inequalities gives

$$\begin{aligned}
\mathrm{RHS} &= \sum_i d(i, \mathrm{top}(\pi(i))) + \sum_i d(\pi(i), \mathrm{top}(\pi(i))) \\
&\leqslant \sum_i [d(i, \mathrm{OPT}) + d(\mathrm{OPT}, \mathrm{top}(\pi(i)))] + \sum_j d(j, \mathrm{top}(j)) \\
&\leqslant SC(\mathrm{OPT}) + \sum_i d(\mathrm{OPT}, \mathrm{top}(\pi(i))) + \sum_j d(j, \mathrm{OPT}) \\
&= SC(\mathrm{OPT}) + \sum_j d(\mathrm{OPT}, \mathrm{top}(j)) + SC(\mathrm{OPT}) \\
&= SC(\mathrm{OPT}) + \sum_j [d(\mathrm{OPT}, j) + d(j, \mathrm{top}(j))] + SC(\mathrm{OPT}) \\
&\leqslant SC(\mathrm{OPT}) + \sum_j [d(\mathrm{OPT}, j) + d(j, \mathrm{OPT})] + SC(\mathrm{OPT}) \\
&= SC(\mathrm{OPT}) + 2SC(\mathrm{OPT}) + SC(\mathrm{OPT}) = 4SC(\mathrm{OPT}).
\end{aligned}$$

Consequently $2SC(A) \leqslant 4SC(\mathrm{OPT})$, and this finishes the proof. □

Proof of Theorem 1.2. First consider the star metric with a center and n leaves labeled $1, \ldots, n$, such that center-leaf distances are 1, and leaf-leaf distances are 2 (for distinct leaves). For each $k \in [n]$, place one voter v_k and one candidate c_k right at leaf k. Then $\mathrm{top}(v_k) = c_k$ and $d(v_k, \mathrm{top}(v_k)) = 0$.

Fix any candidate $A = c_t$. We show H_A has no perfect matching. Enumerate the left vertices by $v_\ell(1), \ldots, v_\ell(n)$ and the right ones by v_r, respectively. An edge $(i, j) = (v_\ell(i), v_r(j))$ exists in H_A if and only if the voter pair (v_i, v_j) favors c_t over $\mathrm{top}(v_j) = c_j$. With $d(v_j, \mathrm{top}(v_j)) = 0$, this reduces to

$$(i, j) \in H_A \qquad \text{if and only if} \qquad d(v_i, c_t) + d(v_j, c_t) \leqslant d(v_i, c_j). \tag{1}$$

Now pick any right vertex $v_r(j)$ with $j \neq t$. We claim its neighborhood satisfies $N(v_r(j)) \subset \{v_\ell(t)\}$:

- If $i = t$, then $d(v_i, c_t) = d(v_t, c_t) = 0, d(v_j, c_t) = 2$, and $d(v_t, c_j) = 2$. Then Equation (1) holds and $(v_\ell(t), v_r(j))$ is an edge.
- If $i \neq t$, then both $i \neq t$ and $j \neq t$, so $d(v_i, c_t) = d(v_j, c_t) = 2$. On the other hand, $d(v_i, c_j)$ is other 0 (if $i = j$) or 2 (if $i \neq j$). In either case, Equation (1) cannot hold, so no $i \neq t$ can be adjacent to $v_r(j)$.

Therefore, for every $j \neq t$, the only possible in-neighbor of $v_r(j)$ in H_A is $v_\ell(t)$. Let $R = \{v_r(j) : j \neq t\}$. Then the only possible in-neighbors of R is $v_\ell(t)$, which renders a perfect matching impossible for $n \geqslant 3$.

To get $\Omega(n)$ distortion, we now extend the construction to weighted stars. Keep all center-leaf distances 1 except the last one, which we set to some large L. The proof carries over, except the last candidate now has social cost $(n-1)(L+1)$ whereas all others have $L + 2n - 3$. The ratio $(n-1)(L+1)/(L+2n-3) \to n-1$ as $L \to \infty$. □

Proof of Theorem 1.3. Let k be large and $\delta > 0$ be small. Consider an instance on $\mathbb{R}$ where candidate A is located at 0 and B at 1. Place $(k+1)$ voters at $1/2 - \delta$ (call them $\mathcal{L}$), and k voters at $B = 1$ (call them $\mathcal{R}$). Clearly, as $\delta \downarrow 0$ and $k \uparrow \infty$, the distortion $SC(A)/SC(B)$ converges to 3.

We first show that G_A has a perfect matching. Observe $\text{top}(\mathcal{L}) = A$ and $\text{top}(\mathcal{R}) = B$. Vacuously, everyone ranks A no lower than $\text{top}(\mathcal{L}) = A$, so every vertex in the left copy of $\mathcal{L} \cup \mathcal{R}$ has an edge to any vertex in the right copy of $\mathcal{L}$. On the other hand, a voter ranks A no lower than B if and only if the voter belongs to $\mathcal{L}$, so edges also flow from $\mathcal{L}$ to $\mathcal{R}$. Because $|\mathcal{L}| \geqslant |\mathcal{R}|$ and the right copies of $\mathcal{R}$ can be matched by any left vertex, G_A is perfectly matchable.

To show that H_A does not admit perfect matchings, we show that the right copies of $\mathcal{R}$ are isolated. As shown previously, the in-neighbors of them are the left copies of $\mathcal{L}$. However, after deliberation, the pair cannot favor A as $d(\mathcal{L}, A) + d(\mathcal{R}, A) = 1/2 - \delta + 1 > 1/2 + \delta + 0 = d(\mathcal{L}, B) + d(\mathcal{R}, B)$ for small δ.

Similarly, in G_B, the in-neighbors of the right copies of $\mathcal{L}$ are the left copies of $\mathcal{R}$, but $|\mathcal{L}| > |\mathcal{R}|$, so some right vertex is guaranteed to be isolated. Therefore, neither G_B nor its subgraph H_B admits perfect matchings. □

Proof of Theorem 1.4. We make the following claims on the examples considered in Theorems 1.2 and 1.3. For an n-leaf star instance used in Theorem 1.2, $\rho(A) = n - 2$ for candidates on short leaves and $\rho(A) \geqslant 2(n-2)$ for the far candidate. For the two-candidate instance considered in Theorem 1.3, the "bad", chosen candidate A has $\rho(A) = k \approx n/2$ while the "good" candidate has $\rho(B) = 2$.

We first consider the examples used in the proof of Theorem 1.2, starting with the unweighted one. Recall for $A = c_t$, we showed that for every right vertex $v_r(j)$ with $j \neq t$, its in-neighborhood in H_A is contained in the singleton $\{v_\ell(t)\}$. Consequently, without modification, the size of H_A's maximum matching is at most 2 (at best, match one non-t right vertex using left t, and match right t using some other left). Hence the matching deficiency is $n-2$, meaning $\rho(A) \geqslant n - 2$. We claim this bound is attainable. First, recall H_A consists exactly of edges of form (t, j) for every j, and (i, t) for every i, i.e., every edge is incident to left or right t. Pick any ordering of the $n - 1$ indices other than t, say $p_1, p_2, \ldots, p_{n-1} \in [n] \backslash \{t\}$. Define a matching M by the directed n-cycle

$$t \to p_1 \to p_2 \to \ldots \to p_{n-1} \to t$$

or in left-to-right edge notation, $M = \{(t, p_1), (p_1, p_2), \ldots, (p_{n-2}, p_{n-1}), (p_{n-1}, t)\}$. We now check costs:

- (t, p_1), (p_{n-1}, t) clearly cost 0 as they are in H_A.
- An intermediate edge (p_r, p_{r+1}) belongs to G_A because $d(v_{p_r}, A) = 2$ and $d(v_{p_r}, c_{p_{r+1}}) = 2$ (distinct leaves), so p_r ranks A at least as high as $c_{p_{r+1}} = \text{top}(v_{p_{r+1}})$, hence the ordinal edge exists. Yet, it is not in H_A because neither endpoint is t. Therefore this edge costs 1.

Summing up, M has exactly $(n-2)$ cost-1 edges and 2 cost-0 edges, and this shows $\rho(A) = n - 2$.

For the weighted variant, for a non-far candidate $A = c_t$ with $t \neq n$, the analysis is the same as in the unweighted case. In the directed n-cycle $t \to p_1 \to p_2 \to \ldots \to p_{n-1} \to t$, the two edges incident to t have cost 0 as before, and each intermediate edge (p_r, p_{r+1}) is not in H_A. If neither endpoint is n then the edge is in G_A as before; if $p_r = n$, then voter n ties c_t with any non-far $c_{p_{r+1}}$, and the edge lies in G_A again. Therefore every intermediate edge costs 1 and $\rho(A) = n - 2$ for all $t \neq n$. For the far candidate c_n, however, every non-far voter strictly prefers *any* non-far candidate to c_n. Consequently, for $j \neq n$ and $i \neq n$, the ordinal edge $i \to j$, interpreted as "i ranks c_n at least as high as c_j," is *not* present in G_{c_n}. In this weighted star, one has $H_{c_n} = G_{c_n}$. Recall the matching deficiency of $n - 2$: this means every edge added incurs cost 2, so $\rho(c_n) \geqslant 2(n-2)$. This bound is attained by an arbitrary n-cycle $n \to p_1 \to \ldots \to p_{n-1} \to n$ where $p_1, \ldots, p_{n-1}$ is any ordering of $[n-1]$ (two cost-0 endpoint edges and $(n-2)$ intermediate cost-2 edges).

For Theorem 1.3, let $n = 2k + 1$. We first analyze candidate A. Recall that in H_A, everything on the left connects to right-$\mathcal{L}$, but every vertex in right-$\mathcal{R}$ is isolated, so $\rho(A) \geqslant |\mathcal{R}| = k$. On the other hand, because G_A admitted a perfect matching, we can simply add back the k cost-1 edges used to match to right-$\mathcal{R}$ in G_A, so $\rho(A) = k$.

We now claim $\rho(B) = 2$. First observe $H_B = G_B$, as the deliberation inequality for the edges from $\mathcal{R}$ to $\mathcal{L}$ favors B, so no edges get deleted relative to G_B. The only obstruction is that right-$\mathcal{L}$ has size $k + 1$, yet its in-neighborhood in G_B is left-$\mathcal{R}$, whose size is k, so we are short by 1. As there are no cost-1 edges available (since $H_B = G_B$), $\rho(B) \geqslant 2$. Conversely, adding one edge from *some* left-$\mathcal{L}$ to *some* right-$\mathcal{L}$ (an edge not in G_B) suffices: match that chosen right-$\mathcal{L}$ to the chosen left-$\mathcal{L}$, match the other k right-$\mathcal{L}$ vertices using the left-$\mathcal{R}$ vertices (these edges already exist), and match all right-$\mathcal{R}$ vertices using the remaining k left-$\mathcal{L}$ vertices (recall they already connect to all of right-$\mathcal{R}$). So $\rho(B) = 2$. □

Proof of Theorem 1.4. Consider a metric instance on the real line with two candidates A at 0 and B at 1. Voters are partitioned into $\mathcal{L}$ at some $x \in (0, 1/2]$ and $\mathcal{R}$ located at 1, so $\text{top}(\mathcal{R}) = B$. Assume $|\mathcal{L}| \geqslant |\mathcal{R}|$. This generalizes the construction used in Theorem 1.3.

As in the proof of Theorem 1.3, $\rho(A) = |\mathcal{R}|$. (Every right-$\mathcal{R}$ is isolated in H_A, yet they are matched in G_A. They can be repaired by adding $|\mathcal{R}|$ cost-1 edges.) On the other hand, deliberation comparing (b, A) always favors B in this geometry, so $H_B = G_B$ as in Theorem 1.3. In G_B, the only in-neighbors of right-$\mathcal{L}$ are left-$\mathcal{R}$, so to match all $|\mathcal{L}|$ vertices of right-$\mathcal{L}$ with only $|\mathcal{R}|$ vertices in left-$\mathcal{R}$, we must add $(|\mathcal{L}| - |\mathcal{R}|)$ edges not in G_B, each costing 2. Thus $\rho(B) = 2(|\mathcal{L}| - |\mathcal{R}|)$, generalizing the $\rho(B) = 2$ computation in Theorem 1.3.

The rule picks A if and only if $\rho(A) \leqslant \rho(B)$, i.e., $|\mathcal{R}|/|\mathcal{L}| \leqslant 2/3$, and social costs can be expressed by

$$SC(A) = |\mathcal{L}| \cdot x + |\mathcal{R}|, \qquad SC(B) = |\mathcal{L}| \cdot (1 - x).$$

If the rule picks A, then from $|\mathcal{R}|/|\mathcal{L}| \leqslant 2/3$ and $x \leqslant 1/2$ we have

$$\frac{SC(A)}{SC(B)} = \frac{x + |\mathcal{R}|/|\mathcal{L}|}{1 - x} \leqslant \frac{1/2 + 2/3}{1/2} = 7/3,$$

and if the rule picks B, then using $|\mathcal{R}|/|\mathcal{L}| \geqslant 2/3$ and $x \geqslant 0$, we have

$$\frac{SC(B)}{SC(A)} = \frac{1 - x}{x + |\mathcal{R}|/|\mathcal{L}|} \leqslant \frac{1}{2/3} < 7/3. \qquad \square$$

Proof of Theorem 1.6. Assume the rule outputs A while B is socially optimal. Scale $d(A, B) = 1$. For each voter, define $\delta_v = d(v, A) - d(v, B) \in [-1, 1]$. For ease of notation, let $\Delta = SC(A) - SC(B)$ and $S = SC(A) + SC(B) = \sum_{v \in V} \delta_v$. Finally, because $d(v, A) + d(v, B) \geqslant d(A, B)$ we have $S \geqslant n$. We may rewrite social costs and distortion as

$$SC(A) = \frac{S + \Delta}{2}, \qquad SC(B) = \frac{S - \Delta}{2}, \qquad \frac{SC(A)}{SC(B)} = \frac{S + \Delta}{S - \Delta} \leqslant \frac{n + \Delta}{n - \Delta}, \tag{2}$$

so it suffices to show $\Delta \leqslant 2n/5$.

We first reduce Δ to a "matching deficiency." Partition voters into $AB = \{v : \delta_v \leqslant 0\}$ and $BA = \{v : \delta_v > 0\}$. Consider the bipartite graph with left vertices AB and right vertices BA, containing an edge (i, j) iff the deliberation between (i, j) comparing (A, B) weakly prefers A, i.e., $d(i, A) + d(j, A) \leqslant d(i, B) + d(j, B)$, i.e., $\delta_i + \delta_j \leqslant 0$. Let μ_A be the maximum matching size in this graph and fix a maximum matching M. Summing $\delta_i + \delta_j \leqslant 0$ over $(i, j) \in M$ gives $\sum_{v \in V(M)} \delta_v \leqslant 0$. Therefore,

$$\Delta = \sum_{v \in V} \delta_v \leqslant \sum_{v \notin V(M)} \delta_v.$$

All unmatched voters in AB have $\delta_v \leqslant 0$, so only unmatched voters in BA can increase the sum, and each satisfies $\delta_v \leqslant 1$. Hence

$$\Delta \leqslant \sum_{j \in BA \setminus V(M)} \delta_j \leqslant |BA \setminus V(M)| = |BA| - \mu_A.$$

Therefore, it suffices to show $|BA| - \mu_A \leqslant 2n/5$.

We define μ_B analogously to be the maximum matching size in the bipartite graph with left BA, right AB, that contains edge (i, j) iff deliberation between (i, j) comparing (B, A) weakly prefers B, i.e., $\delta_i + \delta_j \geqslant 0$.

Let $m = \min\{|AB|, |BA|\}$. Trivially, $\mu_A, \mu_B \leqslant m$. We first claim that $\mu_A + \mu_B \geqslant m$. To see this, recall the maximum matching M in the graph $\delta_i + \delta_j \leqslant 0$ from AB to BA. Let $U_{AB} \subset AB$ and $U_{BA} \subset BA$ be the unmaatched vertices. By definition, for $i \in U_{AB}, j \in U_{BA}$, we must have $\delta_i + \delta_j > 0$, so (j, i) is an edge of the reverse graph $\delta_j + \delta_i \geqslant 0$ from BA to AB. Therefore, the reverse graph contains all edges between U_{BA} and U_{AB}, so it has a matching of size (at least) $\min\{|U_{BA}|, |U_{AB}|\} = m - \mu_A$. Therefore $\mu_A + \mu_B \geqslant m$.

In the two candidate setting, $\rho(\cdot)$ admits closed forms

$$\rho(A) = 2(|BA| - |AB|)^+ + m - \mu_A, \qquad \rho(B) = 2(|AB| - |BA|)^+ + m - \mu_B,$$

because only right vertices that ever incur positive cost for A are those in BA (and symmetrically for B), and within those, μ_A (resp. μ_B) is exactly the maximum number that can be matched using cost-0 edges. Any remaining deficit is paid using cost-1 edges when possible and otherwise forced to cost-2 due to the imbalance of $|AB|$ versus $|BA|$.

Since the rule outputs A, $\rho(A) \leqslant \rho(B)$. As $\mu_A + \mu_B \geqslant m$, the closed forms reduce to

$$2(|BA| - |AB|)^+ + m - \mu_A \leqslant 2(|AB| - |BA|)^+ + \mu_A. \tag{3}$$

We now bound $|BA| - \mu_A$ depending on whether $|AB| \geqslant |BA|$.

- If $|AB| \geqslant |BA|$, then Equation (3) becomes $2\mu_A \geqslant 5|BA| - 2n$. If $|BA| \leqslant 2n/5$ then $|BA| - \mu_A \leqslant 2n/5$. Otherwise,
$$|BA| - \mu_A \leqslant |BA| - \frac{5|BA| - 2n}{2} = \frac{2n - 3|BA|}{2} \leqslant \frac{2n - 3 \cdot 2n/5}{2} = \frac{2n}{5}.$$
- If $|AB| < |BA|$, then Equation (3) becomes $2\mu_A \geqslant 2|BA| - |AB| = 3|BA| - n$. Therefore
$$|BA| - \mu_A \leqslant |BA| - \frac{3|BA| - n}{2} = \frac{n - |BA|}{2} = \frac{|AB|}{2} < \frac{n}{4} < \frac{2n}{5}.$$

In both cases $|BA| - \mu_A \leqslant 2n/5$, so by Equation (2) this implies $SC(A)/SC(B) \leqslant 7/3$. Tightness follows from Theorem 1.5. □

Proof of Theorem 1.7. To be completed. □

B Proofs for Section 2

Proof of Theorem 2.3. Work with $\mathbb{R}$ and $d(x, y) = |x - y|$. Consider an instance with 2 candidates and 3 voters, as follows: $A = 0, B = 3$, and $v_1 = 0, v_2 = 2, v_3 = 2$. Then v_1 ranks A over B, whereas v_2, v_3 rank B over A. We show that neither candidate satisfies the said requirement. In particular, neither $H(A, B)$ nor $H(B, A)$ admits a perfect matching.

First consider $H(A, B)$. For a left voter $u \in \{v_2, v_3\}$, their top choice is already B, so to satisfy (1), the witness C must be $C = B$. Thus, for left vertex $u \in \{v_2, v_3\}$, an edge (u, v) exists iff the deliberation between (u, v) comparing (A, B) weakly prefers A. It is easy to see that the only viable choice is $v = v_1$ on the right. In other words, each of the left v_2, v_3 can only be matched against the right v_1, so no perfect matching is possible.

Similarly, in $H(B, A)$, for the left voter v_1, A is top, so (1) implies that the witness C must be $C = A$. Then, an out-edge from left v_1 exists iff the deliberation between (v_1, v) comparing (B, A) weakly prefers B. This is violated by every voter, so left v_1 has no neighbors, and $H(B, A)$ has no perfect matching.

□

Proof of Theorem 2.4. The construction and argument is essentially identical to that of Theorem 1.3. We provide a thorough statement for the sake of completeness. We work on $\mathbb{R}$. Let k be large and $\delta > 0$ be small. Consider the instance where candidate A is placed at 0, B at 1, $(k + 1)$ voters at $1/2 - \delta$ (denoted $\mathcal{L}$), and k voters at 1 (denoted $\mathcal{R}$). Then as $\delta \downarrow 0$ and $k \uparrow \infty$, the distortion $SC(A)/SC(B)$ converges to 3.

First, we show that $G(A, B)$ has a perfect matching. If $u \in \mathcal{L}$ then u ranks A over B. Choose the witness $C = A$. Condition (ii) becomes "v ranks A weakly above A," which always holds. Therefore every left vertex in $\mathcal{L}$ is connected to every right vertex. On the other hand, if $u \in \mathcal{R}$, then u strictly ranks B over A, so the only possible witness is $C = B$. Condition (ii) becomes "v ranks A over B," which holds iff $v \in \mathcal{L}$. So every left vertex in $\mathcal{R}$ connects to the right vertices in $\mathcal{L}$. Since $|\mathcal{L}| \geqslant |\mathcal{R}|$, $G(A, B)$ has a perfect matching.

We now claim $G(B, A)$ has no perfect matching. By symmetry, left $\mathcal{R}$ connects to all rights (taking $C = B$), whereas left L connects to only right $\mathcal{R}$ (take $C = A$, require "v ranks B over A," which happens iff $v \in \mathcal{R}$). Since $|\mathcal{R}| < |\mathcal{L}|$ there is no way to ensure all left-$\mathcal{L}$ vertices are matched. So $G(B, A)$ has no perfect matching.

Now we look at the H graphs. First, $H(A, B)$. As before, left $\mathcal{L}$ has edges to every right vertex via witness $C = A$, as deliberation between (A, A) is a tie so it weakly favors A. For a left voter $u \in \mathcal{R}$, the only possible witness is $C = B$. Then an edge (u, v) exists iff the pair (u, v) weakly favors A over B in deliberation. It is easy to see this is impossible, as u is a "die-hard" fan of B ($d(u, B) = 0$), which can only be countered by a "die-hard" fan of A, but this type of voter doesn't exist in this instance. So $H(A, B)$ has no perfect matching.

For $H(B, A)$, left $\mathcal{R}$ still connects to every right vertex via witness $C = B$ as deliberation between (B, B) is a tie. For a left voter $u \in \mathcal{L}$, the only possible witness is $C = A$. A pair (u, v) exists if the pair weakly prefers B over A in deliberation, and this happens iff v belongs to the right $\mathcal{R}$. But $|\mathcal{L}| > |\mathcal{R}|$, so no perfect matching exists in $H(B, A)$. □