1 Sampling Deliberations

One major concern over the deliberation-via-matching protocol is that in order to form a tournament graph over m candidates, each voter needs to participate in up to m-choose-2 or $\Theta(m^2)$ deliberations.

So, given two candidates A, B, what if instead of requiring a maximum matching on $AB \times BA$, we use significantly fewer pairs of deliberations? For now I revisit the two-candidate warm-up case with $\lambda = 0.5$ and $w = 1$.

1.1 Revisiting the Two-Candidate Instance

Fix a two-candidate instance with candidates A, B and n voters and reuse all of our previous definitions on ordinal preferences and deliberations. With deliberation defined on a maximum matching on $AB \times BA$, we said that A wins against B if $|AB| + W_A \geqslant |BA| + W_B$. We wanted to analyze the supremum of $SC(A)/SC(B)$ subject to A winning (and B being the social optimal).

Let $m_{AB} = \min\{|AB|, |BA|\}$ be the size of any maximum matching on $AB \times BA$. The previous idea of relying on an arbitrary maximum matching can be equivalently (and tautologically) phrased as: pick an arbitrary maximum matching M on $AB \times BA$, then sample m_{AB} disjoint pairs of voters from it. Likewise, to randomly sample k pairs without replacement, an alternate perspective is to pick a random maximum matching, then random pick k disjoint matched pairs in it. This naturally lets us consider the following rule:

> **Sampling deliberation without replacement.** Let $k \in \{1, \ldots, m_{AB}\}$ be given. Fix any maximum matching M on $AB \times BA$. Sample k disjoint edges uniformly without replacement from M. Let W'_A be the number of sampled pairs whose deliberation favors A and $W'_B = k - W'_A$. Define
>
> $$\text{score}'(AB) = \frac{|AB| + (m_{AB}/k)W'_A}{n}, \qquad \text{score}'(BA) = \frac{|BA| + (m_{AB}/k)W'_B}{n}$$
>
> and declare A as winner if $\text{score}'(AB) \geqslant \text{score}'(BA)$.

In other words, we approximate the ground-truth deliberation win-rate of A (represented by $W_A/m_{AB} = W_A/(W_A + W_B)$) with an empirical observation ($W'_A/k = W'_A/(W'_A + W'_B)$), and we properly up-weight it by (m/k) so that the denominator equals $(m/k)(W'_A + W'_B) = m_{AB}$ (i.e. total effect from deliberation remains the same, and this is controlled by w, which we set to $w = 1$ for the ease of analysis). It shouldn't be surprising that with sufficiently many samples, the estimation becomes close, and $f(AB)$ is also approximated well. We formalize this via the following concentration result.

Theorem 1.1. ***Under a two-candidate setting with*** $\lambda = 0.5, w = 1$***, each voter participating in*** $o(1)$ ***deliberations suffice to ensure a distortion of*** $2 + \epsilon$ ***with high probability.***

In more details, for fixed ϵ, δ***, if we sample*** $k = \Theta(\epsilon^{-2} \log(1/\delta))$ ***deliberations, then***

- ***(1)*** $\mathbb{P}($***rule outputs*** A ***and*** $SC(A)/SC(B) > 2 + \epsilon) \leqslant \delta$***, and***
- ***(2) on average, each voter participates in*** $\Theta(n^{-1}\epsilon^{-2} \log(1/\delta)) = o(1)$ ***deliberations.***

To avoid nested subscripts, we use m to denote m_{AB} throughout this proof. This is not to be confused with the number of candidates, denoted also m, in the next section.

Lemma 1.2. *Hoeffding's ineqaulity for sampling without replacement,* $\{0, 1\}$ *case Let* $y_1, \ldots, y_m \in \{0, 1\}$ *have mean* $\mu = m^{-1} \sum_{i=1}^m y_i$*. Sample* k *elements uniformly without replacement and let* $\overline{Y}$ *be the sample mean. Then, for all* $t > 0$*,*

$$\mathbb{P}(|\overline{Y} - \mu| \geqslant t) \leqslant 2\exp(-2kt^2).$$

Proof of Theorem 1.1. First, an easy assumption to make: $m = \Theta(n)$. If $|AB| < n/3$ then regardless of deliberation outcomes, $\text{score}'(AB) < \text{score}'(BA)$. Conversely, if $|AB| > 2n/3$ then $\text{score}'(AB) > \text{score}'(BA)$ automatically, and $SC(A)/SC(B) \leqslant 2/\text{score}(AB) - 1 \leqslant 2/(2/3) - 1 = 2$ (recall this is equation 5 in [MYZ25]). Therefore, we assume $m \in [n/3, n/2]$.

Next, to show $SC(A)/SC(B) \leqslant 2+\epsilon$, it suffices that $\text{score}(AB) \geqslant 2/(3+\epsilon)$. Set $\gamma = \epsilon/(9+3\epsilon)$ for reasons that will become clear later. Define the "good" event

$$\mathcal{E} = \{|\text{score}'(AB) - \text{score}(AB)| \leqslant \gamma\} = \{|\text{score}'(BA) - \text{score}(BA)| \leqslant \gamma\}.$$

If the sampled rule outputs A, by definition $\text{score}'(AB) \geqslant \text{score}'(BA)$. Then,

$$\max\{\text{score}(AB), \text{score}(BA)\} \leqslant \max\{\text{score}'(AB), \text{score}'(BA)\} + \gamma = \text{score}'(AB) + \gamma \leqslant \text{score}(AB) + 2\gamma.$$

Hence $\text{score}(AB) \geqslant \max\{\text{score}(AB), \text{score}(BA)\} - 2\gamma \geqslant 2/3 - 2\gamma = 2/(3+\epsilon)$, where the second $\geqslant$ is due to theorem 4.1 of [MYZ25]. As

$$\mathbb{P}(\text{rule outputs } A \text{ and } SC(A)/SC(B) > 2+\epsilon) \leqslant \mathbb{P}(\mathcal{E}^c),$$

it remains to bound the RHS. By definition,

$$\text{score}'(AB) - \text{score}(AB) = \frac{|AB| + (m/k)W'_A}{n} - \frac{|AB| + W_A}{n} = \frac{m}{n}\left(\frac{W'_A}{k} - \frac{W_A}{m}\right).$$

The m matched-pair outcomes in M form a fixed $\{0,1\}$ population with mean W_A/m, and W'_A/k is the sample mean from k draws without replacement. Applying Lemma 1.2 with $t = (n/m)\gamma$ gives

$$\mathbb{P}\left(\left|\frac{W'_A}{k} - \frac{W_A}{m}\right| > \frac{n}{m} \cdot \gamma\right) \leqslant 2\exp\left(-2k\left(\frac{n}{m} \cdot \gamma\right)^2\right).$$

Because $m = \Theta(n)$, the ratio $n/m = \Theta(1)$, and $(n/m \cdot \gamma)^2 = \Theta(\epsilon^2)$. Choosing $k = \Theta(\epsilon^{-2}\log(1/\delta))$ makes the RHS at most δ. This concludes the proof of (1). As this asymptotic expression of k is independent of n, (2) immediately follows (and note that no voter can appear in more than one matched edge). □

1.2 Sampling Deliberations on General Instances

Now we generalize the idea of sampling deliberations to an arbitrary instance, where there are n voters and m candidates:

> **Sampling-based Deliberation via Matching.** Let k be given and keep λ^*, w^* unchanged. For each pair of candidates X, Y, run the sampling protocol in Section 1.1 for a total of k pairs and correspondingly compute a score $\tilde{f}(XY)$. Repeat for each candidate pair (X, Y). Build the tournament graph based on the aggregated $\tilde{f}$-values and output any λ^*-WUS as winner.

The intuition is to hope that $f(XY)$ (maximum matching version) is sufficiently approximated by $\tilde{f}(XY)$ (sampling-based version) for each edge. Suppose that $|f(XY) - \tilde{f}(XY)| < \epsilon$ for all candidates X, Y. Then, to analyze

$$\sup \frac{SC(A)}{SC(B)} \qquad \text{subject to} \qquad \begin{cases} \tilde{f}(AC) \geqslant 1 - \lambda^* \\ \tilde{f}(CB) \geqslant \lambda^*, \end{cases} \tag{1}$$

it suffices to analyze

$$\sup \frac{SC(A)}{SC(B)} \qquad \text{subject to} \qquad \begin{cases} f(AC) \geqslant 1 - \lambda^* - \epsilon \\ f(CB) \geqslant \lambda^* - \epsilon. \end{cases} \tag{2}$$

We claim the following theorem.

Theorem 1.3. *Fix λ^*, w^*. There exists an absolute constant $C > 0$ such that for all sufficiently small $\epsilon, \delta > 0$, if $k \geqslant C\epsilon^{-2}\log(m/\delta)$, then for every metric space,*

$$\mathbb{P}\left(\frac{SC(\textit{output})}{SC(\text{OPT})} > 3 + \epsilon\right) \leqslant \delta.$$

In particular, to ensure distortion $\leqslant 3+\epsilon$ with probability at least $1-\delta$, under the sampling-based deliberation via matching, it suffices that each voter participates in $\Theta(m^2/n \cdot \epsilon^{-2}\log(m/\delta))$ deliberations in expectation.

Proof. STEP 1. CONCENTRATION ON ONE EDGE. Suppose we sample k deliberations for each pair of candidates (X, Y). We first analyze the concentration of a single edge $f(X, Y)$. Let W'_{XY} be the number of sampled pairs that prefer X and W_{XY} the total number of pairs in the underlying maximum matching that favor X. To approximate W_{XY}/m_{XY} with W'_{XY}/k, we upweight the observed number of wins W'_{XY} into $\widehat{W}_{XY} = W'_{XY} \cdot m_{XY}/k$ and compare it against W_{XY}. This estimate satisfies

$$\widehat{W}_{XY} - W_{XY} = m_{XY}\left(\frac{W'_{XY}}{k} - \frac{W_{XY}}{m_{XY}}\right),$$

hence

$$|\tilde{f}(XY) - f(XY)| = \frac{w^*}{n + w^* m_{XY}} \cdot |\widehat{W}_{XY} - W_{XY}| = \frac{w^* m_{XY}}{n + w^* m_{XY}} \cdot \left|\frac{W'_{XY}}{k} - \frac{W_{XY}}{m_{XY}}\right|.$$

Since $m_{XY} \leqslant n/2$ the pre-factor is at most $w^*/2$. We therefore obtain

$$\mathbb{P}(|\tilde{f}(XY) - f(XY)| \geqslant \eta) \leqslant 2\exp\left(-2k\left(\frac{2\eta}{w^*}\right)^2\right) = 2\exp\left(-\frac{8k\eta^2}{(w^*)^2}\right). \tag{3}$$

STEP 2. UNION BOUND. The current analysis uses the crude bound to ensure that $\tilde{f}$ approximates f sufficiently well on *all* edges: by union bound,

$$\mathbb{P}(\text{there exists } (X, Y) \text{ such that } |\tilde{f}(XY) - f(XY)| > \eta) \leqslant 2\binom{m}{2}\exp\left(-\frac{8k\eta^2}{(w^*)^2}\right)$$

so choosing $k = \Theta(\eta^{-2}\log(m/\delta))$ makes the LHS at most δ. This explains the choice of k.

STEP 3 (MAIN PART). THE DISTORTION IS LIPSCHITZ TO PERTURBATIONS. Now suppose $|\tilde{f}(XY) - f(XY)| \leqslant \eta$ for each (X, Y). To upper bound the distortion achieved by the tournament graph defined on $\tilde{f}$, it suffices to analyze

$$\sup \frac{SC(A)}{SC(B)} \qquad \text{subject to} \qquad \begin{cases} f(AC) \geqslant 1 - \lambda^* - \eta \\ f(CB) \geqslant \lambda^* - \eta \end{cases} \tag{4}$$

in the fully deterministic, maximum matching case. We observe that [MYZ25] already provides all machinery needed for this analysis. The overall theme, as described in Section 5.4 and 5.5, was to run a series of reduction (counter-monotone coupling, prefix-suffix and optimal matching, tightening the f-constraints, and compacting voters within each interval in the bilinear program), and then solve a relatively simple bilinear program.

Observe that with η introduced, the only part that no longer works is the identity $|AC| + W_{AC} = 0.5$. After counter-monotone coupling, optimal matching, and tightened f-constraints, we know that X and Y individually partition $[0, 1]$ into finitely many chunks. We then need to keep the full list of block endpoints that arise from either X or Y, and use them to partition the voters into blocks indiscernible to the rule. With $|AC| + W_{AC} = 0.5$ the structure is further simplified, but without it, the argument carries through.

At this stage, the reduction depends on how these finitely many endpoints are interleaved on $[0, 1]$. We define a *combinatorial type* t to be the relative order (allowing ties) of those finitely many endpoints. Since there are only constantly many endpoints, there are finitely many combinatorial types, and every compacted instance satisfying Equation (4) belongs to one of them.

We also recall the certificate functional Φ_R given by Program (14):

$$\frac{SC(A)}{SC(B)} > R + 1 \qquad \Longrightarrow \qquad \Phi_R(X, Y) = \mathbb{E}X + (R + 1)\cdot\mathbb{E}Y + R\cdot\mathbb{E}Z < 0.$$

Fixed a type t. For this fixed type, the worst-case value of Φ_R can be written as a finite-dimensional bilinear program of form

$$\mathrm{OPT}_{R,t}(\eta) = \mathrm{OPT}(\eta) = \inf\{\langle p, g(q)\rangle : p \in P(\eta), q \in Q\},$$

where we separate the variables into two independent sets:

- p represents a vector of block masses, constrained to a polytope $P(\eta)$;

- q represents finitely many "metric variables" (the compacted X_i, Y_i, Z_i plus the linearization variables for the sup-norms), constrained to lie in a fixed polyhedron Q, and $g(q)$ maps q to the quantity suitable for describing the functional Φ_R.

(We drop the subscripts R, t to ease notation.)

Importantly, the dependence of η enters only through the two relaxed f-constraints in Equation (4), hence only through the linear constraints defining $P(\eta)$, and this dependence is affine in η locally. (Concretely, one can write $P(\eta) = \{p : Ap \leqslant b + \eta d\}$ for some fixed matrices/vectors A, b, d depending on type t, as well as λ^* and w^*, but not on η.)

Recall that an optimal solution for the bilinear program occurs at a pair of extreme points, where we may take p to be a vertex of the polytope $P(\eta)$ and q to be an extreme point of Q. In particular, the optimal value of inf Φ_2 under Equation (4) can be expressed as the minimum of finitely many linear programs in p with feasible region $Ap \leqslant b + \eta d$ and coefficients determined by the finitely many relevant extreme points of Q. Each of these LPs has fixed coefficient matrix and an additive perturbation ηd in the constants on the RHS, so its optimal value is piecewise linear in η. In particular, there exists $K_t > 0$ such that for all sufficiently small η, $\mathrm{OPT}_2(\eta) \geqslant \mathrm{OPT}_2(0) - K_t\eta = -K_t\eta$. Since there are finitely many types, taking $K = \max_t K_t$ gives a uniform lower bound of $-K\eta$, i.e.,

$$\inf\{\Phi_2 : \text{feasible under Equation (4)}\} \geqslant -K\eta.$$

Finally, Φ_R is affine in R with $\phi_{2+\Delta} = \Phi_2 + \Delta \cdot SC(B)$. By scale invariance we may set $SC(B)$ and recover inf $\Phi_{2+K\eta} \geqslant 0$, certifying distortion at most $(2 + K\eta) + 1 = 3 + K\eta$ whenever Equation (4) holds. This completes Step 3.

Step 4. Everything else. Let $\eta = \epsilon/K$. Then Step 3 becomes $SC(\mathit{output})/SC(\mathrm{OPT}) \leqslant 3 + \epsilon$, and by Step 2 this happens with probability $\geqslant 1 - \delta$ provided $k = \Theta(\eta^{-2}\log(m/\delta)) = \Theta(\epsilon^{-2}\log(m/\delta))$, as claimed. □