The End Justifies the Mean: A Linear Ranking Rule for Proportional Sequential Decisions

Carmel Baharav
MIT
cbaharav@mit.edu

Niclas Boehmer
Hasso Plattner Institute,
University of Potsdam
niclas.boehmer@hpi.de

Bailey Flanigan
MIT
baileyf@mit.edu

Maximilian T. Wittmann
Hasso Plattner Institute,
University of Potsdam
maximilian.wittmann@hpi.de

Abstract

AI alignment and participatory design motivate a new democratic design problem: how to collectively choose a *decision rule to use repeatedly*. We study this problem for *linear ranking rules*, which repeatedly rank items x_j within batches $X = (x_1, \ldots, x_m) \in (\mathbb{R}^d)^m$, where each item's ranking is dictated by its score $\langle \theta^*, x_j \rangle$ according to a fixed scoring vector θ^*. Given voters' preferred scoring vectors $\theta^{(1)}, \ldots, \theta^{(n)}$ and their population fractions $\alpha^{(1)}, \ldots, \alpha^{(n)}$, we ask how to choose a collective vector θ^* satisfying *individual proportionality (IP)*: the agreement of each voter type i with the resulting rankings should be proportional to $\alpha^{(i)}$, either on average over time (*long-run IP*) or even within each batch (*per batch IP*).

The default rule, the arithmetic mean of the $\theta^{(i)}$, has been shown to be severely majoritarian; more generally, it is not clear that *any* fixed linear rule can balance many voters' disparate opinions. Our main result is that, surprisingly, there *is* a simple rule that does satisfy long-run IP: the *angular mean*, the spherical analog of the arithmetic mean. We then show that exact per-batch IP is impossible for fixed linear rules, but that the gap between per-batch and long-run IP shrinks quickly with batch size. Experiments on three real-world preference datasets show that all rules perform similarly when voters' preferences are homogeneous, while the angular mean substantially improves proportionality in high-disagreement regimes.

1 Introduction

In AI alignment and participatory design, there is growing interest in the democratic design of *decision rules*: mappings that are repeatedly used to map inputs $X \in \mathcal{X}$ to outputs $Y \in \mathcal{Y}$ over the course of repeated decision-making. Examples include collectively choosing objective functions [42], allocation rules for settings such as kidney exchange [16] and food donation [31], and reward or scoring functions used to guide the alignment of language models [19].

Collectively choosing decision rules requires two steps: *eliciting* individuals' preferences, and *aggregating* them into a single decision rule. The literature uses a variety of approaches to this process: classical LLM alignment pipelines pool individual evaluators' raw elicitation data and then fit one collective model [19, 37]; virtual democracy learns a model for every individual agent, has the models vote, and then aggregates these votes with a voting rule [22]; still other work learns an

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

individual model for each agent and then aggregates these models directly into a final model. In order, these approaches utilize increasingly richer information about individuals' beliefs.

We study the aggregation step in the third approach. We focus on the aggregation step because it is comparatively under-studied: there is extensive existing work on eliciting *individuals'* decision rules [16, 31, 3, 24, 23, 35], but even in the few papers that consider how to directly aggregate them, the aggregation method is not the main object of analysis, and they default to taking the arithmetic mean of the parameters [36, 42]. We focus on the third approach because it is a best-case benchmark: if high-quality aggregation is impossible even with known individual rules, then lower-information pipelines cannot be expected to recover it; if it *is* possible, this approach will identify the structure those pipelines should preserve.

We specifically consider aggregating over *linear ranking* decision rules, related to the paradigm of linear social choice [19]. The input to linear ranking rules is a batch of m items $X = (x_1, \ldots, x_m)$, where each item $x \in X$ is described by a feature vector $x \in \mathbb{R}^d$. A rule, parameterized by a score vector $\theta \in \mathbb{R}^d$, assigns each item the score $s_\theta(x) = \langle \theta, x \rangle$ and returns the ranking $\succ_\theta^X$ induced by these scores. Linear ranking rules are a natural first case: they are interpretable, technically tractable, and capture the score-then-rank structure underlying many prioritization systems for, e.g., resource allocation, recipient selection, and preference-based ranking of language model outputs.

For linear ranking rules, the aggregation task is then: given n voter types, their respective fractional shares of the population $\alpha^{(1)}, \ldots, \alpha^{(n)} \in [0, 1]$, and their respective individual scoring vectors $\theta^{(1)}, \ldots, \theta^{(n)} \in \mathbb{R}^d$, we want to choose a scoring vector θ^*. This is a non-standard voting problem because the object being chosen is *reusable*: a decision rule is not evaluated only by whether it reflects the electorate at the moment of aggregation, but by how it treats voters across the sequence of decisions on which it is applied. The central worry is therefore cumulative: a fixed rule may repeatedly agree with some viewpoints and repeatedly override others, producing representational inequality over time. In fact, Feffer et al. [15] document this very problem for the literature's default aggregation approach, the arithmetic mean $\theta_{\text{arith}} = \sum_i \alpha^{(i)} \theta^{(i)}$: with just two voter types and two alternatives, the arithmetic mean can be inherently majoritarian, in some cases producing rankings that never agree with the preferred ranking of the minority group.

This observation motivates the design of *individually proportional* aggregation mechanisms: mechanisms that produce a θ^* such that if voter type i constitutes an $\alpha^{(i)}$ fraction of the electorate, the rankings produced by s_{θ^*} agree with i's preferred ranking on at least an $\alpha^{(i)}$ fraction of item pairs. From the outset, however, it is not at all obvious that *any* fixed scoring rule θ^* should be able to satisfy such a requirement: because a fixed rule must prioritize the same features in the same way across all decision instances, it could easily fail to balance so many competing viewpoints. This leads to our research question below; surprisingly, we find that its answer is *yes*.

Research Question: *Does there exist an aggregation mechanism that, given any $(\theta^{(i)}, \alpha^{(i)})_{i \in [n]}$, produces a θ^* that ranks items proportionally to all voter types i over the long-term?*

1.1 Approach and Contributions

In Section 2, we set up the formal problem and define *long-run* individual proportionality (IP). This notion is measured *in expectation over batches of items* X, where $X = (x_1, \ldots, x_m)$ is repeatedly drawn from some fixed absolutely continuous distribution $\mathcal{D}^m$. For a given X, voter type i's level of agreement with the induced ranking $\succ_{\theta^*}^X$ is measured by the Kendall-Tau (KT) similarity (i.e., the number of pairwise agreements) between $\succ_{\theta^*}^X$ and i's own ranking $\succ_{\theta^{(i)}}^X$. Voter i's proportional agreement entitlement, on average over $X \sim \mathcal{D}^m$, is $\alpha^{(i)} \binom{m}{2}$, so achieving long-run IP means that the following quantity is at least 1, i.e., all voters receive at least their proportional agreement entitlement.

$$\text{Long-Run IP}(\theta^*) = \min_{i \in [n]} \mathbb{E}_{X \in \mathcal{D}^m} \left[\text{ KT similarity of } \succ_{\theta^*}^X \text{ and } \succ_{\theta^{(i)}}^X \right] / \alpha^{(i)} \binom{m}{2}.$$

We then formally re-establish [15]'s computationally derived finding that the arithmetic mean θ_{arith} severely violates Long-Run IP. This motivates our first contribution, a new aggregation mechanism.

Contribution 1: The *angular mean* satisfies long-run proportionality. Our main result is a proof that for generic $(\theta^{(i)}, \alpha^{(i)})_{i \in [n]}$, there is a surprisingly interpretable linear ranking rule that is guaranteed to satisfy long-run IP (Theorem 3.1): the *angular mean* θ_{ang}. It is the angular analog

of the arithmetic mean θ_{arith}, as illustrated by their analogous defining optimization problems (see Appendix B for proof that the arithmetic mean can indeed be formulated this way):

$$\theta_{\text{arith}} := \text{argmin}_\theta \sum_{i\in[n]} \alpha^{(i)} \|\theta - \theta^{(i)}\|^2 \qquad \theta_{\text{ang}} := \text{argmin}_\theta \sum_{i\in[n]} \alpha^{(i)} \measuredangle(\theta, \theta^{(i)})^2$$

We prove θ_{ang} satisfies long-run individual proportionality for *spherically symmetric* distributions $\mathcal{D}$, the distribution class also analyzed by [15].

The main challenge behind proving this result is that the angular mean does not have a known closed-form representation nor is its optimization problem necessarily convex, so reasoning about it is non-trivial. The proof relies on a new key theorem (Theorem 3.3), which establishes a fundamental property of the angular mean in d dimensions. This property may be of independent interest, as it generalizes a similar bound in 2 dimensions from the math literature [21]. To our knowledge, this generalization to higher dimensions was an open question.

As a bonus, we examine when θ_{ang}'s proportionality is *approximable* by other scoring vectors. This leads to a general robustness result: as some vector θ approaches θ_{ang} in angular distance, its long-run proportionality accordingly converges to that of θ_{ang} (Remark 3.4, Proposition C.7).

Contribution 2: The possibility of *per-batch* proportionality. We have so far taken for granted that, given $\theta^{(1)}, \ldots, \theta^{(n)}$, we need to choose a *fixed* and *linear* ranking rule to be applied consistently across inputs X. These two restrictions are well-motivated, serving the interpretability of the decision rule by making it explicit which weights will be applied in every case. However, one may worry that this restriction has major costs to achieving proportionality *per-batch*: instead of making sure that no group is shortchanged on average in the long run, we might want to ensure that within a given batch, we should not expect *any* group to agree too little with the resulting ranking. Technically speaking, we are simply strengthening the definition above by re-ordering the minimum and the expectation in the above definition of Long-Run IP to define Per-Batch IP(θ^*).

In contrast to Long-Run IP, we find that even on spherically symmetric distributions, *no* fixed linear ranking rule can achieve Per-Batch IP of at least 1 (Theorem 4.3). However, we show that for any θ^* and *any* distribution satisfying minimal conditions, Per-Batch IP rapidly approaches Long-Run IP with increasing batch size (Theorem 4.5). Hence, our commitment to a consistent and interpretable rule results in surprisingly little degradation of our ability to achieve per-batch proportionality.

Contribution 3: Empirical analysis. We empirically compare five aggregation mechanisms on three real-world preference datasets [2, 24, 31]: the angular mean, arithmetic mean, geometric median, Borda count, and Proportional Sequential Borda (PSB). We find that on the original datasets, voters' scoring vectors are fairly homogeneous, and all rules easily satisfy individual proportionality. We formalize this finding for θ_{arith}: as voters' scoring vectors converge, the arithmetic mean converges to the angular mean, and thus enjoys near-optimal proportionality in such cases (Proposition E.1).

We then induce greater voter heterogeneity by subsampling and restricting the datasets. Here, the rules separate: the angular mean substantially improves Long-Run IP over the arithmetic mean and, on the most heterogeneous subsamples, over the geometric median and Borda. PSB performs best due to its ability to output rankings not inducible by any scoring vector and to adjust its behavior per batch. This demonstrates the loss associated with restricting to fixed, linear ranking rules. We also demonstrate robustness of the angular mean to non-spherical batch distributions and confirm empirically that each rule's Long-Run IP and Per-Batch IP converge quickly as the batch size grows.

1.2 Related Work

Existing work on incorporating public input into the design of decision rules differs mainly along three dimensions. First, the elicited input may take the form of a scoring vector for each voter, or of ordinal preference information such as pairwise comparisons or rankings. Second, the aggregation procedure may output a single collective scoring vector that is applied repeatedly, or it may aggregate preferences separately for each decision instance. Third, the downstream decision task may be to score, rank, or select candidates. We use these dimensions to structure the following discussion of related work and to highlight how our setting differs from existing approaches.

Virtual democracy and participatory design. The most closely related literature is *virtual democracy*, a paradigm that uses models of each voter's preferences to make collective decisions on their

behalf at decision time [22]. Closest to our work, Noothigattu et al. [36] use the arithmetic mean to aggregate individual scoring vectors, which Feffer et al. [15] critique, noting that the arithmetic mean is majoritarian and suffers under strategic manipulation. Freedman et al. [16] take an alternative approach: they learn a collective scoring vector directly by pooling pairwise comparisons across voters before fitting a single linear model. Kahng et al. [22] and Lee et al. [31] pursue a third paradigm in which, at every decision instance, each voter's preferences over the current alternatives are inferred from a learned model and aggregated via a single-winner voting rule, producing decisions that need not be consistent across instances.

Aggregating scoring vectors in social choice. In divisible participatory budgeting, also known as portioning, each voter proposes a budget allocation across projects [14, 18, 13, 7, 4, 17], and these allocations are subsequently aggregated. This problem is formally equivalent to aggregating scoring vectors. In this literature, the arithmetic mean is the canonical proportional aggregator, and the central concern is its lack of strategyproofness. In contrast to our work, however, the downstream task is a one-shot cardinal task rather than a repeated ordinal one.

Proportionality in repeated decision making. A separate line of work studies individual and group-level proportional fairness in repeated decision making [27, 26, 9, 5, 28, 38]. We draw our conceptual notion of proportionality from these papers (i.e., "an α fraction of the population deserves an α fraction of agreement") [9, 29]. Yet, the models studied in these works fundamentally differ from ours, among others, because they typically allow the use of non-fixed, non-linear voting rules.

Proportional rank aggregation. Individual and group-level proportionality in rank aggregation, i.e., the task of aggregating a given set of rankings into a single collective one, has been the subject of recent work [30, 29]. Lederer [29] introduces the Proportional Sequential Borda (PSB) rule, which guarantees that any α-fraction of the electorate agrees with roughly an α-fraction of the pairwise comparisons in the collective ranking. PSB can be applied to our setting per batch, but the resulting decision rule is neither linear nor consistent across batches. We benchmark against PSB empirically in Section 5. Also from this literature comes the *Squared Kemeny* objective [30], whose optimizers also satisfy certain proportionality guarantees; interestingly, we show in Appendix C.4 that optimizing this objective converges to the same behavior as θ_{ang} as the batch size m grows, but Squared Kemeny is potentially more majoritarian for small m with respect to Long-Run IP.

RLHF and alignment from a social-choice perspective. In reinforcement learning from human feedback, the reward model plays the role of a scoring rule, and standard practice trains the reward model on pooled annotator data [11, 37, 40]. A growing line of work studies alignment as a collective decision-making problem [19, 20, 44, 39, 32, 45, 12]. Closest to us, Ge et al. [19] give an axiomatic analysis of rules that aggregate individual rankings into a collective weight vector under a linear functional form for the reward model; their axioms, however, target majoritarian and efficiency properties rather than proportionality. Shirali et al. [43] argue that the arithmetic mean of per-type reward functions is a principled aggregator when the downstream task uses cardinal scores. Our angular mean offers an alternative in the ordinal regime.

2 Model & Preliminaries

General Notation. Let there be $d \in \mathbb{N}_{\geq 2}$ dimensions, and let S^{d-1} be the unit sphere in $\mathbb{R}^d$. For a vector $z \in \mathbb{R}^d$ and $j \in [d]$, we let z_j be the j-th entry of z. The *angular distance* between two vectors $x, y \in \mathbb{R}^d$ is the (minimal) angle between them, $d_{\measuredangle}(x, y) := \cos^{-1}(\langle x, y\rangle / (\|x\|_2 \|y\|_2))$.

Items and Batches. Let an *item* $x \in \mathbb{R}^d$ be a covariate vector, and let $\mathcal{X} \subseteq \mathbb{R}^d$ be the set of feasible covariate vectors. A *batch* of items consists of $m \in \mathbb{N}_{\geq 2}$ items, represented as a tuple $X = (x_1, \ldots, x_m) \in \mathcal{X}^m$. We assume that batches X are formed by sampling m items i.i.d. from an *item distribution* $\mathcal{D} \in \Delta(\mathcal{X})$ that is absolutely continuous, and has PDF $f_{\mathcal{D}} \colon \mathbb{R}^d \to \mathbb{R}$; we denote the induced *batch distribution* as $\mathcal{D}^m$. Additionally, we assume that $f_{\mathcal{D}}$ is strictly positive and continuous on some open set in $\mathbb{R}^d$.

For some results, we will additionally assume that $\mathcal{D}$ is spherically symmetric, which we will emphasize notationally as $\mathcal{D}_\circ$. The PDF $f_{\mathcal{D}_\circ}$ of a spherically symmetric distribution $\mathcal{D}_\circ$ is rotationally invariant about the origin and hence can be written as a function solely of an item's Euclidean norm,

so $f_{D_\circ}(x) = g(\|x\|)$ for some function $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, i.e., the PDF assigns the same probability to any two items with the same Euclidean norm, regardless of their rotation in space. For instance, spherically symmetric distributions include isotropic Gaussian distributions, standard multivariate Student's t distributions, or uniform distributions on spheres centered at the origin.

Scoring Vectors and Rankings. A *scoring vector* $\theta \in S^{d-1}$ is a vector of linear weights to be placed on the covariates in x. Accordingly, the *score* of x according to θ is $s_\theta(x) := \langle x, \theta \rangle$. These scores naturally induce a ranking over the elements of $\mathcal{X}$: Let $x, y \in \mathcal{X}$, and fix some arbitrary tie-breaking ranking over $\mathcal{X}$ denoted as $\succ_{\text{tiebreak}}$. Then, the ranking over $\mathcal{X}$ induced by scoring vector $\theta \in S^{d-1}$, denoted as $\succ_\theta$, is such that for all $x, y \in \mathcal{X}$ (where $x \succ y$ denotes that x is preferred to y),

$$x \succ_\theta y \iff s_\theta(x) > s_\theta(y) \vee (s_\theta(x) = s_\theta(y) \wedge x \succ_{\text{tiebreak}} y).$$

For any $\theta \in S^{d-1}$ and $X \in \mathcal{X}^m$, $\succ_\theta^X$ denotes the restriction of the ranking $\succ_\theta$ to the items in X.

Voter Types and Aggregation Mechanisms. There are $n \geq 2$ distinct *types* of voters $1, \ldots, n$, where i has an ideal scoring vector $\theta^{(i)} \in S^{d-1}$ and constitutes an $\alpha^{(i)} \in (0, 1]$ fraction of the population (so $\sum_{i \in [n]} \alpha^{(i)} = 1$). We let $\boldsymbol{\theta} = (\theta^{(i)})_{i \in [n]}$ and $\boldsymbol{\alpha} = (\alpha^{(i)})_{i \in [n]}$, and we call $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ a *profile*. Let P be the set of all possible profiles, and $P_{n,d} \subseteq P$ be all those with n voter types in d dimensions.

An *aggregation mechanism* is a mapping that takes in a profile $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ and outputs some $\theta \in S^{d-1}$. A generic aggregation mechanism is denoted as θ_f, so $\theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha})$ is the scoring vector produced by some underlying function f.[1] Given a profile $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$, the two aggregation mechanisms we focus on — the (normalized) *arithmetic mean* θ_{arith} and *angular mean* θ_{ang} — are respectively defined as

$$\theta_{\text{arith}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \operatorname*{arg\,min}_{\theta \in S^{d-1}} \sum_{i=1}^{n} \alpha^{(i)} \|\theta - \theta^{(i)}\|_2^2 \qquad \theta_{\text{ang}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \operatorname*{arg\,min}_{\theta \in S^{d-1}} \sum_{i=1}^{n} \alpha^{(i)} d_\measuredangle(\theta, \theta^{(i)})^2$$

(see Figure 1(a) for a visual comparison). We show that this formulation of θ_{arith} is equivalent to the direct arithmetic mean formulation in Proposition B.1. Because S^{d-1} is compact and the minimized objective functions are continuous, minimizers always exist. In case there are multiple global minimizers for either function, we assume some fixed tie-breaking scheme is applied.

Remark 2.1 (*Computability of Angular Mean*). How to compute the angular mean is a well-studied problem [6, 1, 10, 8]. Because it has no known closed-form representation, computation generally relies on gradient-based approximation schemes. The objective is not necessarily convex so convergence is not guaranteed [6, 1]; however, in practice, this is often addressed by the fact that the arithmetic mean often serves as a good initial point for optimization (cf. Appendix E), in which case there exist strong theoretical guarantees that gradient-based methods will converge [6, 1].

***Long-Run* Individual Proportionality.** In words, an aggregation mechanism θ_f is *long-run* individually proportional if, over many random draws of $X \sim \mathcal{D}^m$, voter type i "agrees with" at least an $\alpha^{(i)}$ fraction of pairwise comparisons in the collective ranking $\succ_{\theta_f}^X$. "Agreement" is formalized as Kendall-Tau (KT) agreement. For two scoring vectors $\theta, \psi \in S^{d-1}$, their induced rankings on a batch $X \in \mathcal{X}^m$ have *KT-agreement*

$$a_{KT}(\theta, \psi, X) := |\{x, y\} \in \tbinom{X}{2} \colon (x \succ_\theta y \wedge x \succ_\psi y) \vee (y \succ_\theta x \wedge y \succ_\psi x)|.$$

An individual's proportional *entitlement* of KT-agreement is then equal to $\alpha^{(i)}\binom{m}{2}$, i.e., an $\alpha^{(i)}$ fraction of all $\binom{m}{2}$ pairwise comparisons in a ranking over m items.

For an aggregation mechanism θ_f, a profile $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$, a batch $X \in \mathcal{X}^m$, and voter type $i \in [n]$, *i's individual proportionality (IP) level* is

$$\text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) := \frac{a_{KT}(\theta^{(i)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)}{\alpha^{(i)}\binom{m}{2}}.$$

Then, the *Long-Run IP Level* (LONG-IP) is the minimum *expected* individual proportionality level received by any voter type.

[1] We remark that requiring the output vector to be normalized (i.e., in S^{d-1}) is without loss of generality, since rescaling a scoring vector does not change the rankings it induces.

Definition 2.2 (Long-Run IP Level). For any batch size m, item distribution $\mathcal{D}$, and profile $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$, an aggregation mechanism θ_f has

$$\textit{Long-Run IP Level} \ \text{LONG-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) := \min_{i \in [n]} \mathbb{E}_{X \sim \mathcal{D}^m} \left[\text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right]$$

$$\& \textit{ Worst-Case Long-Run IP Level} \ \text{LONG-IP}^*(\theta_f, \mathcal{D}^m) := \inf_{(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P} \text{LONG-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})).$$

Given that we aim to bound LONG-IP*, we now characterize what is possible: we show that the most proportional aggregation mechanism can achieve LONG-IP* of no more than 1, and the least can achieve LONG-IP* no less than 0. The lower bound is by nonnegativity of IP$_i$; see Appendix B for proof of the upper bound. Accordingly, we say that θ_f *satisfies* LONG-IP iff LONG-IP$^*(\theta_f, \mathcal{D}^m) = 1$, which means that in every profile, θ_f gives every voter type at least their proportional entitlement.

Proposition 2.3 *For any θ_f, m, and $\mathcal{D}$,* LONG-IP$^*(\theta_f, \mathcal{D}^m) \in [0, 1]$.

2.1 Preliminaries

We now establish that the arithmetic mean has the worst possible worst-case Long-Run IP Level:

Theorem 2.4 (Figure 2, [15]) *For any batch size m, item distribution $\mathcal{D}$,* LONG-IP$^*(\theta_{\text{arith}}, \mathcal{D}^m) = 0$.

The proof (Appendix B) is a simple profile with $n = 2$, where the majority ($\alpha^{(1)} > 1/2$) and minority ($\alpha^{(2)} < 1/2$) types have diametrically opposed scoring vectors, and thus the arithmetic mean overlaps perfectly with $\theta^{(1)}$. We name this profile The Antipodal Construction for later use (Figure 1(b)).

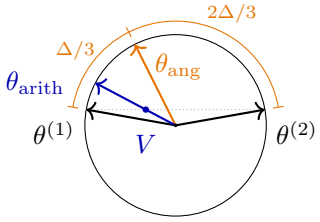


Figure 1(a): Angular and arithmetic means, $\alpha^{(1)} = 2/3$; $V = \sum_i \alpha^{(i)}\theta^{(i)}$, $\Delta = d_\angle(\theta^{(1)}, \theta^{(2)})$

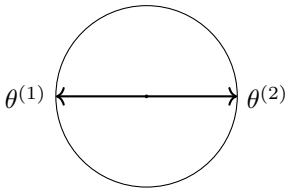


Figure 1(b): The Antipodal Construction

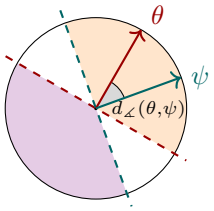


Figure 1(c): Geometric interpretation of agreement

In addition to this critique of the arithmetic mean, Feffer et al. [15] also observe a strong relation between two scoring vectors' angular distance and their expected KT-agreement on batches drawn from spherically symmetric distributions when $m = 2$:

Proposition 2.5 (Proposition 1, [15]) *For any spherically-symmetric item distribution $\mathcal{D}_\circ$ and any $\theta, \psi \in S^{d-1}$, we have $\mathbb{E}_{X \sim \mathcal{D}_\circ^2}[a_{KT}(\theta, \psi, X)] = {}^{\pi - d_\angle(\theta,\psi)}/_{\pi}$.*

The intuition for the argument will be useful, so we summarize it here. Let $X = (x_1, x_2)$ and define the difference vector $\tilde{x} = x_1 - x_2$. Note that $\succ_\theta$ and $\succ_\psi$ agree on the ranking of x_1 and x_2 iff $\text{sign}(\langle\theta, \tilde{x}\rangle) = \text{sign}(\langle\psi, \tilde{x}\rangle)$. Geometrically (see Figure 1(c)), note that the set $H^\theta_> = \{y : \text{sign}(\langle\theta, y\rangle) > 0\}$ forms a halfspace, and so does the similarly defined $H^\psi_>$. The intersection $H^\theta_> \cap H^\psi_>$ then forms a cone representing the set of $\tilde{x}$ for which θ and ψ agree that $x_1 \succ x_2$. Via the symmetric argument for when θ, ψ agree that $x_2 \succ x_1$, it follows that the set of all $\tilde{x}$ on which θ and ψ agree forms a double-wedge. The intersection of that double-wedge with any sphere centered at 0 takes up a ${}^{\pi - d_\angle(\theta,\psi)}/_{\pi}$ fraction of the sphere's total surface area. Finally, observe that the expected agreement between θ and ψ is equal to the probability of $\tilde{x}$ landing in the double-wedge. By spherical symmetry, $\tilde{x}$ itself is drawn from a spherically symmetric distribution $\tilde{\mathcal{D}}_\circ$, and this probability is thus equal to the size of the double-wedge, concluding the proof.

3 The Angular Mean Satisfies Long-Run Individual Proportionality

The dramatic disproportionality of θ_{arith} motivates the question of whether there exists *any* fixed linear decision rule satisfying LONG-IP. We now answer this in the affirmative for the angular mean θ_{ang}, assuming a spherically symmetric item distribution.

Theorem 3.1 (**Main Result**) *For any batch size m and spherically symmetric item distribution $D_\circ$,*

$$\text{LONG-IP}^*(\theta_{\text{ang}}, \mathcal{D}_\circ^m) = 1.$$

The first step of the proof is to extend Proposition 2.5 to the general m-item case, which follows by the intuition above along with linearity of expectation (full proof in Appendix C):

Lemma 3.2 *For any batch size m, spherically symmetric item distribution $D_\circ$, and $\theta, \psi \in S^{d-1}$,*

$$\mathbb{E}_{X\sim\mathcal{D}_\circ^m}\left[a_{KT}(\theta,\psi,X)\right] = \tfrac{\pi - d_\angle(\theta,\psi)}{\pi}\binom{m}{2}.$$

Plugging in $\theta^{(i)}$ and θ_{ang} for θ and ψ and writing our desired bound on LONG-IP* in terms of $\mathbb{E}_{X\sim\mathcal{D}_\circ^m}\left[a_{KT}(\theta,\psi,X)\right]$, we deduce that it suffices to show that $d_\angle(\theta^{(i)}, \theta_{\text{ang}}) \le (1-\alpha^{(i)})\pi$ for all $i \in [n]$. We now show that this fundamental property of the angular mean holds, i.e., the angular mean is guaranteed to simultaneously lie within a $(1-\alpha^{(i)})\pi$ angle from any $\theta^{(i)}$. This property appears natural: For the weighted arithmetic mean $\bar{z} = \sum_j \alpha_j z_j$ in Euclidean space, $\|\bar{z} - z_i\| \le (1-\alpha_i)\max_j \|z_j - z_i\|$. The angular bound is its spherical counterpart in that the maximum possible angular distance, π, takes the place of the Euclidean distance bound. Still, to our knowledge, this was unknown, and generalizes the known version on S^1 [21]. The full proof is located in Appendix C.

Theorem 3.3 (**Fundamental Property of Angular Mean**) *For any profile $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$,*

$$d_\angle(\theta_{\text{ang}}(\boldsymbol{\theta},\boldsymbol{\alpha}), \theta^{(i)}) \le (1-\alpha^{(i)})\pi \quad \textit{for all } i \in [n].$$

Proof Sketch. Let $F(\theta) := \sum_{i=1}^n \alpha^{(i)} d_\angle(\theta,\theta^{(i)})^2$ be the value of the angular mean objective for a generic θ. Suppose for contradiction that there exists $i \in [n]$ such that $d_\angle(\theta_{\text{ang}}, \theta^{(i)}) > (1-\alpha^{(i)})\pi$. We will find another score vector θ' such that $F(\theta') < F(\theta_{\text{ang}})$, contradicting the optimality of θ_{ang}.

We illustrate the intuition with Figure 2. To construct θ', let G_i be the unique great circle passing through $\theta^{(i)}$ and θ_{ang} (left image). Then, let θ' be positioned on the great circle G_i, but swooped around to be within a $(1-\alpha_i)\pi$ distance of $\theta^{(i)}$ but on the other side. We want to show that

$$F(\theta') - F(\theta_{\text{ang}}) = \textstyle\sum_{j\in[n]} \alpha^{(j)}(d_\angle(\theta',\theta^{(j)})^2 - d_\angle(\theta_{\text{ang}},\theta^{(j)})^2) < 0.$$

The $j = i$ term of this sum will be handled at the end. We first bound $d_\angle(\theta',\theta^{(j)})^2$ for each $j \ne i$, corresponding to bounding the angular length of the yellow arc (left image). We do this by applying Toponogov's inequality and the Euclidean law of cosines, which together give a spherical analog of the Pythagorean Theorem showing that the yellow arc is upper bounded by its Euclidean counterpart when the other two sides of the triangle are preserved (middle image). As will be important in the next step, one term of this bound tracks whether the move from θ_{ang} toward θ' initially moves toward or away from $\theta^{(j)}$, as captured by $\cos(\gamma_j)$ (see diagram for γ_j).

With a bound on $d_\angle(\theta',\theta^{(j)})^2$ for each j, we then take the final weighted sum over voters to bound $F(\theta') - F(\theta_{\text{ang}})$ as above. Here we add one final key ingredient: the local optimality of θ_{ang}. Because F has zero derivative at θ_{ang} in the direction of the move from θ_{ang} to θ', the weighted sum of the directional terms involving $\cos(\gamma_j)$, as carried through our bound above, sum to 0 (right image). The remaining terms then reduce to an expression involving only the move length $d_\angle(\theta_{\text{ang}},\theta')$, the original distance $d_\angle(\theta_{\text{ang}},\theta^{(i)})$, and $\alpha^{(i)}$, which is negative by the choice of θ' and the assumed violation $d_\angle(\theta_{\text{ang}},\theta^{(i)}) > (1-\alpha^{(i)})\pi$. We conclude $F(\theta') < F(\theta_{\text{ang}})$, our contradiction. □

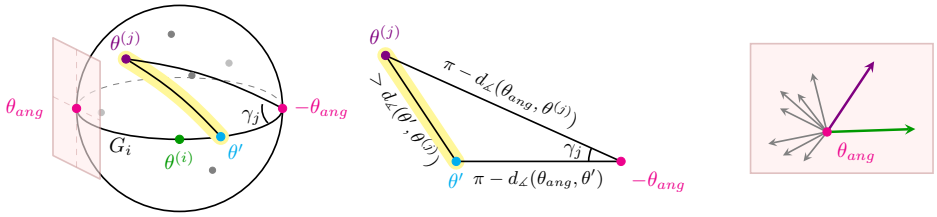


Figure 2: Diagrams for proof sketch of Theorem 3.3.

Remark 3.4 (**Robustness**). A natural question is whether a score vector *close to* θ_{ang} will exhibit similar proportionality. The answer is yes: if a score vector is within angular distance γ of θ_{ang}, each i's IP level can decrease (additively) by at most $\gamma/(\pi\alpha^{(i)})$ (Proposition C.7).

4 The Possibility of Per-Batch Individual Proportionality

So far, we have shown that θ_{ang} ensures that voters will receive their proportional entitlement *in expectation over batches*, and thus on average over sufficient rounds by the Law of Large Numbers. However, one may also want a stronger guarantee: proportionality on *every item batch*. The strongest notion would guarantee that for any $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ with n voter types and any $m \in \mathbb{N}_{\geq 2}$, θ_f is such that

$$\min_{X \in \mathcal{X}^m} \min_{i \in [n]} \text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \geq 1.$$

One can immediately see that this is too strong for simple divisibility reasons: e.g., consider the Antipodal Construction with $m = 2$, where any deterministic rule must output the same ranking each time, thereby giving one group 0 agreement. A similar construction can be used to show that this issue can occur even without divisibility issues (i.e., even when all voters' entitlements $\alpha_i \binom{m}{2}$ are integral): simply let m be large but make the items in the batch collinear, and so again, one group in the Antipodal Construction must end up with 0 agreement (Example D.1).

In light of this impossibility, we pursue a weakened version of per-batch proportionality, which takes the *expectation* over item batches instead of the *minimum*:

Definition 4.1. (**Per-Batch IP Level**) For any batch size m, item distribution $\mathcal{D}$, and profile $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$, an aggregation function θ_f has *per-batch proportionality level* of

$$\text{BATCH-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) := \mathbb{E}_{X \sim \mathcal{D}^m} \left[\min_{i \in [n]} \text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \right].$$

The expectation allows us to avoid over-indexing on the worst-case item batches, while moving the minimum over $i \in [n]$ inside the expectation ensures that, in most batches, no voter type is significantly shortchanged. The following proposition, proved in Appendix D.2, formalizes this intuition that BATCH-IP is a strict strengthening of LONG-IP.

Proposition 4.2 *For any* $\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})$, $\text{BATCH-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \leq \text{LONG-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha}))$.

Also note that this result — and all in this section unless otherwise stated — holds for *all* item distributions (meeting the regularity conditions in Section 2), not just spherically symmetric ones.

Unfortunately, unlike LONG-IP, we show that there exists *no* fixed linear ranking rule that can perfectly satisfy BATCH-IP. This can be shown trivially via divisibility issues, but we prove a stronger impossibility requiring that in the constructed profile, all voter type entitlements are integral:

Theorem 4.3 (**Impossibility of BATCH-IP**) *For any batch size* $m \geq 3$[2]*, there exists a profile* $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$ *with* $\alpha^{(i)} \binom{m}{2} \in \mathbb{Z}\ \forall i \in [n]$ *such that, for any* θ_f *and any* $\mathcal{D}$, $\text{BATCH-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) < 1$.

We prove this result in Appendix D.3 via a similar intuition as used above for Example D.1: we show that on batches X that are *nearly* collinear (and thus occur with nonzero probability), at least one voter type receives $\text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) = 0$. We then upper-bound this type's individual proportionality level in all other batches, demonstrating the bound.

Remark 4.4. This impossibility is universal only over fixed linear ranking rules: the Proportional Sequential Borda voting rule introduced by Lederer [29] is guaranteed to achieve BATCH-IP ≥ 1 on all profiles with integer entitlements.

The bad news is that the above remark establishes a proportionality cost of restricting to fixed linear ranking rules in some instances. The good news, which we show next, is that this cost is surprisingly small and diminishes quickly in the batch size m. Intuitively, the BATCH-IP improves in m because, as m grows, the likelihood that all $x \in X$ are collinear decreases, causing BATCH-IP to approach LONG-IP. The full proof with explicit constants is in Appendix D.4.

Theorem 4.5 (**Main Result**) *For all* $\theta_f, m, \mathcal{D}$, *and* $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$,

$$\text{BATCH-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \geq \text{LONG-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) - \tfrac{1}{\alpha_{min}} \cdot O\left(\sqrt{\tfrac{\min\{\log(n), d \log(m/d)\}}{m}} \right).$$

Proof Sketch. The fundamental goal in this proof is to show that $\text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))$ concentrates around its expectation over $X \sim \mathcal{D}^m$ as m grows, in which case by definition BATCH-IP converges to LONG-IP. We do this by first rewriting i's IP level as

$$\text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) = \tfrac{1}{\alpha^{(i)} \binom{m}{2}} \sum_{1 \leq j < \ell \leq m} \mathbb{I}\left(\text{sign}(\langle \theta^{(i)}, x_j - x_\ell \rangle) = \text{sign}(\langle \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), x_j - x_\ell \rangle) \right),$$

[2]Note that for $m = 2$, there is no profile satisfying the integral proportionality entitlement condition.

the intuition being that agreement between $\theta^{(i)}$ and θ_{ang} on a pairwise comparison of items x_j and x_ℓ just comes down to whether the two scoring vectors have the same signed dot product with the item difference vector. This is now the sum of indicators, which is ripe for a concentration bound except that these indicators have a particular kind of dependence: while the x_j are individually independent, their *differences* are not. To handle this dependence, we reframe individual proportionality level as a U-statistic: we rewrite this sum as the average of $m!$ sums of indicators acting on $\lfloor m/2 \rfloor$ mutually exclusive pairs of items, which lets us show that the sum is sub-Gaussian with variance inversely proportional to $\lfloor m/2 \rfloor$. We use this setup to get both convergence rates in the minimum: The $\log(n)$ rate comes from taking the maximum over the n sub-Gaussian random variables corresponding to IP_i for each $i \in [n]$, and the d-dependent bound is via the Rademacher complexity, the intuition being that in low dimensions, even an arbitrary number of voters must be "close." □

As a byproduct of this proof, we are also able to give tail bounds showing that each i's IP level is close to its expectation with high probability (Corollary D.5). We can additionally apply this bound to show that θ_{ang} converges to the optimal BATCH-IP as m grows, noting that by Proposition 4.2, BATCH-IP of 1 is the best we can hope for.

Corollary 4.6 *For spherically symmetric $\mathcal{D}_\circ$, Theorem 3.1 gives that* $\text{LONG-IP}(\theta_{\text{ang}}, \mathcal{D}_\circ^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \geq 1$*; Theorem 4.5 then immediately implies that*

$$\text{BATCH-IP}(\theta_{\text{ang}}, \mathcal{D}_\circ^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \geq 1 - o(1) \qquad \text{as } m \to \infty$$

5 Empirical Results

Finally, we compare aggregation mechanisms on real and semi-synthetic data. Wherever an experiment requires an expectation $\mathbb{E}_{X \sim \mathcal{D}^m}[\cdot]$, we estimate it by Monte Carlo: we draw 2000 batches independently and average. Unless otherwise stated, we use batch size $m = 10$ and sample each batch's items i.i.d. uniformly from S^{d-1}, so $\mathcal{D}_\circ^m$ is spherically symmetric. We test robustness to non-spherically symmetric batch distributions in Appendix E.4.2: θ_{ang} remains proportional throughout except under strong anisotropy, highlighting its appeal beyond the regime covered in Section 3.

Rules. We consider three fixed linear ranking rules: $\theta_{\text{arith}}, \theta_{\text{ang}}$, and the *geometric median* $\theta_{\text{med}} := \arg\min_{\theta \in S^{d-1}} \sum_i \alpha^{(i)} \|\theta - \theta^{(i)}\|_2$, where the median is a reasonable candidate for its insensitivity to outliers. We also consider two rules that are not linear or fixed, and thus may circumvent the per-batch impossibility of Theorem 4.3. The first is *Borda*, the canonical aggregator of predicted rankings in virtual democracy [22, 31]: given a batch X, each voter type i awards $\alpha^{(i)}(m-j)$ points to the item ranked at position j of $\succ_{\theta^{(i)}}^X$; Borda then returns a ranking over items by total score. The second is *Proportional Sequential Borda* (PSB), introduced by Lederer [29], which iteratively selects the Borda winner and then reduces each voter type's weight in proportion to its contribution to the winner's score. PSB satisfies a strong batch-level guarantee: on any X, every i agrees with the computed ranking on at least $\lfloor \alpha^{(i)} \binom{m}{2} \rfloor$ pairwise comparisons.

Warm-up: Two-Voter Case. In Appendix E.2, we run synthetic experiments in the $n = 2$ case. As expected, we find that as $d_\measuredangle(\theta^{(1)}, \theta^{(2)})$ approaches 180°, θ_{arith} (along with θ_{med} and *Borda*) become extremely majoritarian, while θ_{ang} and PSB maintain expected IP_i of at least 1 (Figure 7).

Real-World Data. We study three main datasets (with two extra in Appendix E.3.2), each including a learned $\theta^{(i)} \in S^{d-1}$ per voter (see Appendix E.3.1 for details). They are *Moral Machine* (MMACHINE) [36, 2] ($n = 137$, $d = 24$); *Kidney Exchange* (KIDNEY) [24] ($n = 404$, $d = 8$); and 412 Food Rescue (FOODRESCUE) [31] ($n = 19$, $d = 7$). We use uniform voter weights $\alpha^{(i)} = 1/n$, and show that this choice is robust to randomly varied weights in Appendix E.4.1.

Results. In this data, we find that voters' score vectors are all highly similar: in degrees, the *mean, standard deviation, and max* of the pairwise angles between any two $\theta^{(i)}$ are $21, 8, 54$ in MMACHINE; $62, 18, 110$ in FOODRESCUE; and $57, 21, 147$ in KIDNEY. Empirically, this translates to all three linear ranking rules $\theta_{\text{ang}}, \theta_{\text{arith}}, \theta_{\text{med}}$ being highly coincident, falling within 1, 5, and 6 degrees of one another (pairwise) across the respective datasets. We formally explain this finding in Proposition E.1, proving that coincident $\theta^{(i)}$'s lead to coincident θ_{arith} and θ_{ang}. Per our robustness results in Remark 3.4, this means that these three rules — and in fact, empirically all five — have almost identical proportionality behavior on the raw datasets (see Appendix E.3.3).

Table 1: Proportionality levels in 2D data. $m = 10$. Cell shading is proportional to $\log_{10}(\text{value})$.

	dataset	arith.	ang.	med.	Borda	PSB
LONG-IP	FOODRESCUE-2D	1.9	2.8	0.7	1.7	5.7
	KIDNEY-2D	0.3	3.3	4.5	11	76
	MMACHINE-2D	0.4	1.9	0.9	4.5	30

	dataset	arith.	ang.	med.	Borda	PSB
BATCH-IP	FOODRESCUE-2D	1.9	2.6	0.7	1.7	5.4
	KIDNEY-2D	0.1	2.1	2.6	6.6	66
	MMACHINE-2D	0.2	1.0	0.5	2.6	26

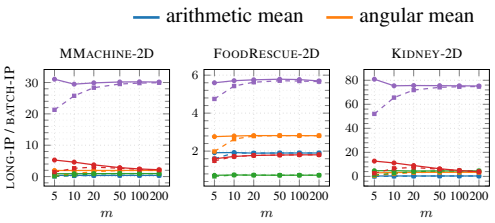


Figure 3: LONG-IP (solid) and BATCH-IP (dashed) on 2D data.

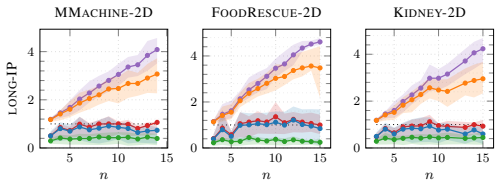


Figure 4: LONG-IP on filtered 2D data. Shading shows inter-quartile range across 100 subsamples.

2D Data. To investigate what happens under greater heterogeneity — the regime where the rule choice becomes nontrivial — we construct *2D* versions of the raw datasets (denoted MMACHINE-2D, etc.) by retaining only the pair of features that maximize the resulting variance of the pairwise angular distances between voters. This produces voter populations that span a wide arc of S^1 (see Figure 10). We report the resulting (larger) angular distances among the $\theta^{(i)}$'s and $\theta_{\text{ang}}, \theta_{\text{arith}}, \theta_{\text{med}}$ in Table 2.

Results. Table 1 reports LONG-IP (left) and BATCH-IP (right) across rules and datasets. We see that θ_{arith} and θ_{med} both violate voters' proportional entitlements substantially; on KIDNEY-2D the LONG-IP of the arithmetic mean drops as low as 0.3, and expectedly, the BATCH-IP is even lower at 0.1. In contrast, θ_{ang} satisfies both BATCH-IP and LONG-IP (i.e., attains levels of at least 1) on all three datasets. As expected, the more powerful non-fixed, non-linear rules perform better, with PSB being 3-10x better than Borda. Looking across the two tables, one may also notice that BATCH-IP and LONG-IP are highly similar, as consistent with Theorem 4.5. Figure 3 shows just how quickly BATCH-IP approaches LONG-IP, empirically: for most rules, convergence happens at extremely small m. As expected, these quantities converge the slowest for the non-fixed, non-linear rules, as these rules are making batch-level adjustments.

Heterogeneity-Filtered 2D Data. For a final stress test, we subsample voters within each 2D dataset to filter for further heterogeneity: for various n, we draw n voters uniformly at random and accept the sample only if the standard deviation of pairwise angular distances in the subsample is $\geq 65°$. This creates instances with few, highly separated voter *clusters* — the hardest regime for proportionality.

Results. Figure 4 shows LONG-IP on these filtered subsamples as n grows (per-batch IP behaves similarly; see Figure 9). Across datasets, the rules separate into two distinct groups: θ_{ang} and PSB scale roughly linearly with n, while θ_{arith}, θ_{med}, and Borda hover near and oftentimes below 1. Note that these distinct groups do not materialize as clearly in Table 1 and Figure 3, where there is less voter heterogeneity. This reflects that when proportionality is more demanding, θ_{ang}'s dedication to proportionality becomes far more important than Borda's ability to be non-fixed / non-linear.

6 Discussion

Zooming out, there are many ways to collectively decide a fixed linear ranking rule: one could pool raw elicitation data to fit a collective model, as in alignment; learn noisy $\hat{\theta}^{(i)}$ that are correct up to rankings, and then simulate a vote, as in virtual democracy; directly aggregate noisy $\hat{\theta}^{(i)}$ with robust methods; or collect enough individual data that more precise $\theta^{(i)}$ can be learned and aggregated, as we study here. Given our finding that there exists a linear ranking rule θ_{ang} that achieves LONG-IP under spherically symmetric item distributions and quickly converges in m to achieving BATCH-IP, a natural next question is how well linear ranking rules learned through each of these less information-rich regimes can approximate these notions of proportionality. Many applications may also merit extending these ideas to nonlinear fixed rules and downstream decision tasks other than ranking.

Acknowledgments

The project was supported by the MIT and HPI AI and Creativity Hub, which is generously funded by the Hasso Plattner Foundation. We also gratefully acknowledge funding from the Chai Family Graduate Fellowship (CB).

References

[1] Bijan Afsari, Roberto Tron, and René Vidal. On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM Journal on Control and Optimization*, 51(3):2230–2260, 2013.

[2] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. The moral machine experiment. *Nature*, 563(7729): 59–64, 2018.

[3] Kyle Boerstler, Vijay Keswani, Lok Chan, Jana Schaich Borg, Vincent Conitzer, Hoda Heidari, and Walter Sinnott-Armstrong. On the stability of moral preferences: A problem with computational elicitation methods. In *Proceedings of the 2024 AAAI/ACM Conference on AI, Ethics, and Society (AIES)*, pages 156–167, 2024.

[4] Felix Brandt, Matthias Greger, Erel Segal-Halevi, and Warut Suksompong. Optimal budget aggregation with single-peaked preferences. In *Proceedings of the 25th ACM Conference on Economics and Computation, (EC)*, page 49, 2024.

[5] Laurent Bulteau, Noam Hazon, Rutvik Page, Ariel Rosenfeld, and Nimrod Talmon. Justified representation for perpetual voting. *IEEE Access*, 9:96598–96612, 2021.

[6] Samuel R. Buss and Jay P. Fillmore. Spherical averages and applications to spherical splines and interpolation. *ACM Transactions on Graphics*, 20(2):95–126, 2001.

[7] Ioannis Caragiannis, George Christodoulou, and Nicos Protopapas. Truthful aggregation of budget proposals with proportionality guarantees. *Artificial Intelligence*, 335:104178, 2024.

[8] Frédéric Cazals, Bernard Delmas, and Timothee O'Donnell. Fréchet mean and p-mean on the unit circle: decidability, algorithm, and applications to clustering on the flat torus. In *19th Symposium on Experimental Algorithms (SEA)*, 2021.

[9] Nikhil Chandak, Shashwat Goel, and Dominik Peters. Proportional aggregation of preferences for sequential decision making. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, pages 9573–9581, 2024.

[10] Benjamin Charlier. Necessary and sufficient condition for the existence of a Fréchet mean on the circle. *ESAIM: Probability and Statistics*, 17:635–649, 2013.

[11] Paul F. Christiano, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, pages 4302–4310, 2017.

[12] Vincent Conitzer, Rachel Freedman, Jobst Heitzig, Wesley H. Holliday, Bob M. Jacobs, Nathan Lambert, Milan Mossé, Eric Pacuit, Stuart Russell, Hailey Schoelkopf, Emanuel Tewolde, and William S. Zwicker. Position: Social choice should guide AI alignment in dealing with diverse human feedback. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, pages 9346–9360, 2024.

[13] Mark de Berg, Rupert Freeman, Ulrike Schmidt-Kraepelin, and Markus Utke. Truthful budget aggregation: Beyond moving-phantom mechanisms. *arXiv preprint arXiv:2405.20303*, 2024.

[14] Edith Elkind, Matthias Greger, Patrick Lederer, Warut Suksompong, and Nicholas Teh. Settling the score: Portioning with cardinal preferences. *Artificial Intelligence*, 352:104487, 2026.

[15] Michael Feffer, Hoda Heidari, and Zachary C Lipton. Moral machine or tyranny of the majority? In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5974–5982, 2023.

[16] Rachel Freedman, Jana Schaich Borg, Walter Sinnott-Armstrong, John P Dickerson, and Vincent Conitzer. Adapting a kidney exchange algorithm to align with human values. *Artificial Intelligence*, 283:103261, 2020.

[17] Rupert Freeman and Ulrike Schmidt-Kraepelin. Project-fair and truthful mechanisms for budget aggregation. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, pages 9704–9712, 2024.

[18] Rupert Freeman, David M. Pennock, Dominik Peters, and Jennifer Wortman Vaughan. Truthful aggregation of budget proposals. *Journal of Economic Theory*, 193:105234, 2021.

[19] Luise Ge, Daniel Halpern, Evi Micha, Ariel D. Procaccia, Itai Shapira, Yevgeniy Vorobeychik, and Junlin Wu. Axioms for AI alignment from human feedback. In *Proceedings of the 38th International Conference on Neural Information Processing Systems (NeurIPS)*, pages 80439–80465, 2024.

[20] Alexandros Hollender and Sonja Kraiczy. Enforcing axioms for AI alignment under loss-based rules. In *The 14th International Conference on Learning Representations (ICLR)*, 2026.

[21] Thomas Hotz and Stephan Huckemann. Intrinsic means on the circle: uniqueness, locus and asymptotics. *Annals of the Institute of Statistical Mathematics*, 67(1):177–193, 2015.

[22] Anson Kahng, Min Kyung Lee, Ritesh Noothigattu, Ariel D. Procaccia, and Christos-Alexandros Psomas. Statistical foundations of virtual democracy. In *Proceedings of the 36th International Conference on Machine Learning, (ICML)*, pages 3173–3182, 2019.

[23] Vijay Keswani, Vincent Conitzer, Hoda Heidari, Jana Schaich Borg, and Walter Sinnott-Armstrong. On the pros and cons of active learning for moral preference elicitation. In *Proceedings of the 2024 AAAI/ACM Conference on AI, Ethics, and Society (AIES)*, pages 711–723, 2024.

[24] Vijay Keswani, Cyrus Cousins, Breanna K. Nguyen, Vincent Conitzer, Hoda Heidari, Jana Schaich Borg, and Walter Sinnott-Armstrong. Moral change or noise? On problems of aligning AI with temporally unstable human feedback. In *Proceedings of the 40th AAAI Conference on Artificial Intelligence (AAAI)*, 2026.

[25] Richard Kim, Max Kleiman-Weiner, Andrés Abeliuk, Edmond Awad, Sohan Dsouza, Joshua B. Tenenbaum, and Iyad Rahwan. A computational model of commonsense moral decision making. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES)*, pages 197–203, 2018.

[26] Alexander Kozachinskiy, Alexander Shen, and Tomasz Steifer. Optimal bounds for dissatisfaction in perpetual voting. In *Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI)*, pages 13977–13984, 2025.

[27] Martin Lackner. Perpetual voting: Fairness in long-term decision making. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 2103–2110, 2020.

[28] Martin Lackner and Jan Maly. Proportional decisions in perpetual voting. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5722–5729, 2023.

[29] Patrick Lederer. Proportional representation in rank aggregation. *arXiv preprint arXiv:2508.16177*, 2025.

[30] Patrick Lederer, Dominik Peters, and Tomasz Was. The squared kemeny rule for averaging rankings. In *Proceedings of the 25th ACM Conference on Economics and Computation (EC)*, page 755. ACM, 2024.

[31] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, and Ariel D. Procaccia. WeBuildAI: Participatory framework for algorithmic governance. *Proceedings of the ACM on human-computer interaction*, 3(CSCW):1–35, 2019.

[32] Roberto-Rafael Maura-Rivero, Marc Lanctot, Francesco Visin, and Kate Larson. Jackpot! alignment as a maximal lottery. *arXiv preprint arXiv:2501.19266*, 2025.

[33] Wolfgang Meyer. Toponogov's theorem and applications. Lecture Notes, Trieste, 1989.

[34] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.

[35] Farhad Mohsin, Inwon Kang, Pin-Yu Chen, Francesca Rossi, and Lirong Xia. Learning individual and collective priorities over moral dilemmas with the life jacket dataset. In *13th Multidisciplinary Workshop on Advances in Preference Handling*, 2022.

[36] Ritesh Noothigattu, Snehalkumar Gaikwad, Edmond Awad, Sohan Dsouza, Iyad Rahwan, Pradeep Ravikumar, and Ariel Procaccia. A voting-based system for ethical decision making. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, 2018.

[37] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. In *Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS)*, 2022.

[38] Bradley Phillips, Edith Elkind, Nicholas Teh, and Tomasz Was. Strengthening proportionality in temporal voting. In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 3823–3832. ACM, 2026.

[39] Ariel D. Procaccia, Benjamin Schiffer, and Shirley Zhang. Clone-robust AI alignment. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, 2025.

[40] Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D. Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS)*, pages 53728–53741, 2023.

[41] Philippe Rigollet and Jan-Christian Hütter. High-dimensional statistics. *arXiv preprint arXiv:2310.19244*, 2023.

[42] Ali Shirali, Jessie Finocchiaro, and Rediet Abebe. Participatory objective design via preference elicitation. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, pages 1637–1662, 2024.

[43] Ali Shirali, Arash Nasr-Esfahany, Abdullah Omar Alomar, Parsa Mirtaheri, Rediet Abebe, and Ariel D. Procaccia. Direct alignment with heterogeneous preferences. In *Proceedings of the 39th International Conference on Neural Information Processing Systems (NeurIPS)*, 2025.

[44] Anand Siththaranjan, Cassidy Laidlaw, and Dylan Hadfield-Menell. Distributional preference learning: Understanding and accounting for hidden context in RLHF. In *The 12th International Conference on Learning Representations, (ICLR)*, 2024.

[45] Taylor Sorensen, Jared Moore, Jillian Fisher, Mitchell L. Gordon, Niloofar Mireshghallah, Christopher Michael Rytting, Andre Ye, Liwei Jiang, Ximing Lu, Nouha Dziri, Tim Althoff, and Yejin Choi. Position: A roadmap to pluralistic alignment. In *Proceedings of the 41st International Conference on Machine Learning, (ICML)*, volume 235, pages 46280–46302, 2024.

[46] Victor Andreevich Toponogov. Riemannian spaces having their curvature bounded below by a positive number. *Twenty-Two Papers on Algebra, Number Theory and Differential Geometry*, pages 291–336, 1964.

[47] David E. Tyler. Statistical analysis for the angular central Gaussian distribution on the sphere. *Biometrika*, 74(3):579–589, 1987.

[48] Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*. Cambridge University Press, 2019.

A Appendix

In this portion of the Appendix, we put examples and small technical results that we reference throughout the remainder of the Appendix.

A.1 Tied scores

Proposition A.1 *(Score ties occur with probability 0) For any item distribution $\mathcal{D}$, finite $d, m \in \mathbb{N}$, and finite set of scoring vectors V*

$$\Pr_{X \sim \mathcal{D}^m}(\exists i, j \in [m], i \neq j \text{ and } \theta \in V \colon s_\theta(x_i) = s_\theta(x_j)) = 0$$

Proof. We can union bound the expression and get the following:

$$\Pr_{X \sim \mathcal{D}^m}(\exists i, j \in [m], i \neq j \text{ and } \theta \in V \colon s_\theta(x_i) = s_\theta(x_j)) \leq \sum_{1 \leq i < j \leq m} \sum_{\theta \in V} \Pr_{x_i, x_j \overset{\text{iid}}{\sim} \mathcal{D}}(s_\theta(x_i) = s_\theta(x_j))$$

As the total number of terms, $\binom{m}{2} \cdot |V|$, is finite by our specification of finite $m, |V|$, it suffices to show that for any $\theta \in V$, $\Pr_{x, y \overset{\text{iid}}{\sim} \mathcal{D}}(s_\theta(x) = s_\theta(y)) = 0$. Equivalently, $\Pr_{x_i, x_j \overset{\text{iid}}{\sim} \mathcal{D}}(\langle \theta, x_i \rangle = \langle \theta, x_j \rangle) = 0$. Fix some $\theta \in V$. By conditioning on a specific realization of x_j, we can rewrite our expression as

$$\Pr_{x_i, x_j \overset{\text{iid}}{\sim} \mathcal{D}}(\langle \theta, x_i \rangle = \langle \theta, x_j \rangle) = \int_{\mathbb{R}^d} \Pr_{x_i \sim \mathcal{D}}(\langle \theta, x_i \rangle = \langle \theta, c \rangle | x_j = c) f_{\mathcal{D}}(c) dc$$

Now observe that for $\langle \theta, x_i \rangle$ to be equal to $\langle \theta, c \rangle$ for some fixed $c \in \mathbb{R}^d$, it must be that x_i is in a specific hyperplane $H_{\theta,c} = \{z \in \mathbb{R}^d \colon \langle \theta, z \rangle = \langle \theta, c \rangle\}$. $H_{\theta,c}$ an affine subspace of dimension $d - 1$, and hence has d-dimensional Lebesgue measure 0: $\lambda^d(H_{\theta,c}) = 0$. The item distribution $\mathcal{D}$ is absolutely continuous with respect to the Lebesgue measure, which tells us that sets with Lebesgue measure 0 have probability measure 0. Putting it all together

$$\Pr_{x_i \sim \mathcal{D}}(\langle \theta, x_i \rangle = \langle \theta, c \rangle | y = c) = \Pr_{x_i \sim \mathcal{D}}(x_i \in H_{\theta,c}) = 0$$

Plugging back into our integral, we get that

$$\int_{\mathbb{R}^d} \Pr_{x_i \sim \mathcal{D}}(\langle \theta, x_i \rangle = \langle \theta, c \rangle | x_j = c) f_{\mathcal{D}}(c) dc = \int_{\mathbb{R}^d} 0 \cdot f_{\mathcal{D}}(c) dc = 0$$

As $\Pr_{x_i, x_j \overset{\text{iid}}{\sim} \mathcal{D}}(s_\theta(x_i) = s_\theta(x_j)) = 0$ holds for all $\theta \in V$, we have that the expression derived from the union bound sums to 0. Hence, for a finite set of vectors, and for a finite item batch size[3] and dimension, ties between object scores occur with probability 0. □

Corollary A.2 *(Expectation conditioned on no ties) Given a profile $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$ (with finite $n, d \in \mathbb{N}$), aggregation mechanism θ_f, item distribution $\mathcal{D}$, and finite item batch size $m \in \mathbb{N}$, we say that for an item batch $X \in \mathcal{X}^m$, $\mathcal{E}_{\text{no-ties}}$ is the event where none of the scoring rules are tied on any items in $X = (x_1, \ldots, x_m)$:*

$$\mathcal{E}_{\text{no-ties}} = \bigcap_{\theta \in (\boldsymbol{\theta} \cup \{\theta_f\})} \bigcap_{1 \leq i < j \leq m} \{s_\theta(x_i) \neq s_\theta(x_j)\}$$

Then, for any function h that can be written as a function of these variables, the expectation over random draws of $X \sim \mathcal{D}^m$ is equivalent to the expectation conditioned on $\mathcal{E}_{\text{no-ties}}$:

$$\mathbb{E}_{X \sim \mathcal{D}^m}[h(X, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \theta_f)] = \mathbb{E}_{X \sim \mathcal{D}^m}[h(X, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \theta_f) | \mathcal{E}_{\text{no-ties}}]$$

Proof. Observe that by Proposition A.1, we have that $\Pr(\overline{\mathcal{E}_{\text{no-ties}}}) = 0$. Then, by law of total expectation:

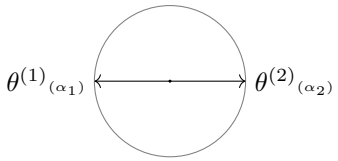
$$\begin{aligned}
&\mathbb{E}_{X \sim \mathcal{D}^m}[h(X, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \theta_f)] \\
&= \mathbb{E}_{X \sim \mathcal{D}^m}[h(X, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \theta_f) | \mathcal{E}_{\text{no-ties}}] \Pr(\mathcal{E}_{\text{no-ties}}) + \mathbb{E}_{X \sim \mathcal{D}^m}\left[h(X, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \theta_f) | \overline{\mathcal{E}_{\text{no-ties}}}\right] \underbrace{\Pr(\overline{\mathcal{E}_{\text{no-ties}}})}_{0} \\
&= \mathbb{E}_{X \sim \mathcal{D}^m}[h(X, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \theta_f) | \mathcal{E}_{\text{no-ties}}] \Pr(\mathcal{E}_{\text{no-ties}})
\end{aligned}$$

□

[3] The proof can be extended to countably infinite item batch sizes if desired.

A.2 The Antipodal Construction

This profile, called the Antipodal Construction, is introduced in Figure 1(b). This profile can be realized in any number of dimensions for $d \geq 2$, where $\theta^{(1)} = (-1, \underbrace{0, \ldots, 0}_{d-1})$ and $\theta^{(2)} = (1, \underbrace{0, \ldots, 0}_{d-1})$. By default we will assume $d = 2$ in its instantiations, and we leave the $\boldsymbol{\alpha}$ up to specification. We use it in numerous proofs, so we place it here again for reference and show a simple fact about the profile.



Proposition A.3 *For $\theta^{(1)}$ and $\theta^{(2)}$ in the Antipodal Construction, and any additional scoring vector $\theta^* \in S^{d-1}$, for any $X \in \mathcal{X}^m$ such that there are no scoring vector ties,*

$$a_{KT}(\theta^{(2)}, \theta^*, X) = \binom{m}{2} - a_{KT}(\theta^{(1)}, \theta^*, X)$$

Proof. Observe that $\theta^{(1)} = -\theta^{(2)}$, so for any $\{x, y\} \in \binom{X}{2}$, $s_{\theta^{(1)}}(x) > s_{\theta^{(1)}}(y) \iff s_{\theta^{(2)}}(x) < s_{\theta^{(2)}}(y)$. With no tied scores, this means that $\succ^X_{\theta^{(1)}}$ and $\succ^X_{\theta^{(2)}}$ disagree on all $\binom{m}{2}$ pairwise comparisons. Therefore, for any pair $\{x, y\} \in \binom{X}{2}$, θ^* must agree with exactly one of $\theta^{(1)}$ or $\theta^{(2)}$ on the ranking. □

B Appendix for Section 2

Proposition B.1 (Arithmetic Mean Definition Equivalence) *In the event that $\| \sum_{i=1}^n \alpha^{(i)} \theta^{(i)} \|_2 > 0$, we have that*

$$\theta_{\text{arith}} = \operatorname*{arg\,min}_{\theta \in S^{d-1}} \sum_{i \in [n]} \alpha^{(i)} \|\theta - \theta^{(i)}\|_2^2 = \frac{\sum_{i \in [n]} \alpha^{(i)} \theta^{(i)}}{\| \sum_{i \in [n]} \alpha^{(i)} \theta^{(i)} \|_2}$$

Proof. Call $\arg\min_{\theta \in S^{d-1}} \sum_{i \in [n]} \alpha^{(i)} \|\theta - \theta^{(i)}\|_2^2$ the minimizing form and $\frac{\sum_{i \in [n]} \alpha^{(i)} \theta^{(i)}}{\| \sum_{i \in [n]} \alpha^{(i)} \theta^{(i)} \|_2}$ the closed form. Equality when the minimizing form minimizes over all $\theta \in \mathbb{R}^d$ and when the closed form is unnormalized is commonly known, but this variation is less common. We start by expanding the expression we're minimizing:

$$\sum_{i \in [n]} \alpha^{(i)} \|\theta - \theta^{(i)}\|_2^2 = \sum_{i \in [n]} \alpha^{(i)} (\underbrace{\|\theta\|_2^2}_{1} + \underbrace{\|\theta^{(i)}\|_2^2}_{1} - 2\langle \theta, \theta^{(i)} \rangle) = 2 \left(\sum_{i \in [n]} \alpha^{(i)} \right) - 2 \langle \theta, \sum_{i \in [n]} \alpha^{(i)} \theta^{(i)} \rangle$$

Then, as $\sum_{i \in [n]} \alpha^{(i)} = 1$, we have that the minimizing form simplifies to $\arg\min_{\theta \in S^{d-1}} 2 - 2\langle \theta, \sum_{i \in [n]} \alpha^{(i)} \theta^{(i)} \rangle$. This is then equivalent to $\arg\max_{\theta \in S^{d-1}} \langle \theta, \sum_{i \in [n]} \alpha^{(i)} \theta^{(i)} \rangle$. Let $V = \sum_{i \in [n]} \alpha^{(i)} \theta^{(i)}$ and ϕ be the angle between θ and V. We can rewrite

$$\langle \theta, V \rangle = \|\theta\|_2 \|V\|_2 \cos(\phi)$$

Observe that because we require $\theta \in S^{d-1}$, we have that $\|\theta\|_2 = 1$. So $\langle \theta, V \rangle = \|V\|_2 \cos(\phi)$. Under our assumption that $\|V\|_2 > 0$, this expression is uniquely maximized when $\cos(\phi) = 1$, which implies that $\phi = 0$. Hence it must be the case that θ_{arith} is in the same direction as V. In combination with the knowledge that θ_{arith} is a unit vector, this gives us everything we need to write it down in closed form:

$$\theta_{\text{arith}} = \frac{V}{\|V\|_2} = \frac{\sum_{i \in [n]} \alpha^{(i)} \theta^{(i)}}{\| \sum_{i \in [n]} \alpha^{(i)} \theta^{(i)} \|_2}$$

□

Proposition 2.3 *For any θ_f, m, and $\mathcal{D}$,* LONG-IP$^*(\theta_f, \mathcal{D}^m) \in [0, 1]$.

Proof. The lower bound follows because for all $i \in [n]$, i's individual proportionality level is always nonnegative.

As LONG-IP$^*(\theta_f, \mathcal{D})$ minimizes over all profiles, we can upper bound it by constructing a specific profile, $(\boldsymbol{\theta}, \boldsymbol{\alpha})$, and evaluating LONG-IP$(\theta_f, \mathcal{D}^m(\boldsymbol{\theta}, \boldsymbol{\alpha}))$. Let $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ be the Antipodal Construction shown in Figure 1(b) with $\alpha^{(1)} = \alpha^{(2)} = 1/2$.

$$\begin{aligned}
&\text{LONG-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \\
&= \min\left(\mathbb{E}_{X\sim\mathcal{D}^m}\left[\text{IP}_1(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right], \mathbb{E}_{X\sim\mathcal{D}^m}\left[\text{IP}_2(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right]\right) \\
&= \min\left(\mathbb{E}_{X\sim\mathcal{D}^m}\left[\frac{a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)}{\alpha^{(1)}\binom{m}{2}}\right], \mathbb{E}_{X\sim\mathcal{D}^m}\left[\frac{a_{KT}(\theta^{(2)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)}{\alpha^{(2)}\binom{m}{2}}\right]\right)
\end{aligned}$$

Now we can make two simplifications. The first is that $\alpha^{(1)} = \alpha^{(2)}$, so we can just take the denominator out of the expectations and outside of the min.

$$= \frac{1}{\alpha^{(1)}\binom{m}{2}} \min\left(\mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right], \mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(2)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right]\right)$$

The other simplification is that via $Corollary\ A.2$ we can assume no scoring vector ties between $\theta_f, \theta^{(1)}, \theta^{(2)}$ on X in both of our expectations, and by Proposition A.3, this means that $a_{KT}(\theta^{(2)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X) = \binom{m}{2} - a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)$. Then we can rewrite the expression as:

$$\begin{aligned}
&= \frac{1}{\alpha^{(1)}\binom{m}{2}} \min\left(\mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right], \mathbb{E}_{X\sim\mathcal{D}^m}\left[\binom{m}{2} - a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right]\right) \\
&= \frac{1}{\alpha^{(1)}\binom{m}{2}} \min\left(\mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right], \binom{m}{2} - \mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right]\right)
\end{aligned}$$

Observe that this minimum is now just a function of $\mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right]$, and is maximized when the two expressions are equal. $\mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right] = \binom{m}{2} - \mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right] \implies \mathbb{E}_{X\sim\mathcal{D}^m}\left[a_{KT}(\theta^{(1)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right] = \frac{1}{2}\cdot\binom{m}{2}$. Plugging this all in along with our value of $\alpha^{(1)}$ gives that

$$\text{LONG-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \leq \frac{1}{\alpha^{(1)}\binom{m}{2}} \cdot \frac{\binom{m}{2}}{2} = 1$$

And hence we can conclude that

$$\text{LONG-IP}^*(\theta_f, \mathcal{D}^m) \leq 1$$

□

Theorem 2.4 (Figure 2, [15]) *For any batch size m, item distribution $\mathcal{D}$,* LONG-IP$^*(\theta_{\text{arith}}, \mathcal{D}^m) = 0$.

Proof. As with the proof of $Proposition\ 2.3$, we can upper bound LONG-IP$^*(\theta_f, \mathcal{D})$ by evaluating LONG-IP$(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha}))$ on a given profile $(\boldsymbol{\theta}, \boldsymbol{\alpha})$. Once again, let $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ be the Antipodal Construction in Figure 1(b), this time with $\alpha^{(1)} = 0.3$ and $\alpha^{(2)} = 0.7$. Then we have that:

$$\theta_{\text{arith}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \frac{\alpha^{(1)}\theta^{(1)} + \alpha^{(2)}\theta^{(2)}}{\|\alpha^{(1)}\theta^{(1)} + \alpha^{(2)}\theta^{(2)}\|} = \frac{(-0.3, 0) + (0.7, 0)}{\|(-0.3, 0) + (0.7, 0)\|} = (1, 0) = \theta^{(2)}$$

Hence, once we normalize, we have that $\theta_{\text{arith}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \theta^{(2)}$. As discussed in Proposition 2.3, we have that for $X \sim \mathcal{D}^m$, with no tied scores, $\succ^X_{\theta^{(1)}}$ and $\succ^X_{\theta^{(2)}}$ will disagree on all pairwise comparisons. As $\theta_{\text{arith}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \theta^{(2)}$, we know that $\succ^X_{\theta^{(1)}}$ and $\succ^X_{\theta_{\text{arith}}(\boldsymbol{\theta}, \boldsymbol{\alpha})}$ will disagree on all pairwise comparisons. Hence, by Corollary A.2 we can assume no ties, and we have that

$$\mathbb{E}_{X\sim\mathcal{D}^m}[a_{KT}(\theta^{(1)}, \theta_{\text{arith}}(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)] = 0$$

Therefore,

$$\begin{aligned}\textsc{long-ip}^*(\theta_{\text{arith}}, \mathcal{D}^m) &\le \textsc{long-ip}(\theta_{\text{arith}}, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \\ &= \min_{i\in\{1,2\}} \mathbb{E}_{X\sim\mathcal{D}^m}[a_{KT}(\theta^{(i)}, \theta_{\text{arith}}(\boldsymbol{\theta}, \boldsymbol{\alpha}), X)] \\ &= 0\end{aligned}$$

□

C Appendix for Section 3

C.1 Proof of Lemma 3.2.

Lemma 3.2 *For any batch size m, spherically symmetric item distribution $D_\circ$, and $\theta, \psi \in S^{d-1}$,*

$$\mathbb{E}_{X\sim\mathcal{D}_\circ^m}\big[a_{KT}(\theta, \psi, X)\big] = \tfrac{\pi - d_\measuredangle(\theta,\psi)}{\pi}\tbinom{m}{2}.$$

Proof. Recall the relation between angles and expected agreement from Proposition 2.5.

$$\mathbb{E}_{x\sim\mathcal{D}, y\sim\mathcal{D}}\big[\mathbf{1}(x \succ_\theta y \wedge x \succ_\psi y)\big] = \mathbb{E}_{X\sim\mathcal{D}_\circ^2}[a_{KT}(\theta, \psi, X)] = \frac{\pi - d_\measuredangle(\theta,\psi)}{\pi}.$$

For $X = (x_1, \ldots, x_m)$ we then apply linearity of expectation to $\mathbb{E}_{X\sim\mathcal{D}^m}[a_{KT}(\theta, \psi, X)]$ and obtain

$$\begin{aligned}\mathbb{E}_{X\sim\mathcal{D}^m}\big[a_{KT}(\theta, \psi, X)\big] &= \mathbb{E}_{X\sim\mathcal{D}_\circ^m}\left[\left|\{x_i, x_j\} \in \tbinom{X}{2} : (x_i \succ_\theta x_j \wedge x_i \succ_\psi x_j) \vee (x_j \succ_\theta x_i \wedge x_j \succ_\psi x_j)\right|\right] \\ &= \mathbb{E}_{X\sim\mathcal{D}_\circ^m}\left[\sum_{\{x_i,x_j\}\in\binom{X}{2}} \mathbf{1}\big((x_i \succ_\theta x_j \wedge x_i \succ_\psi x_j) \vee (x_j \succ_\theta x_i \wedge x_j \succ_\psi x_j)\big)\right] \\ &= \sum_{\{i,j\}\in\binom{[m]}{2}} \mathbb{E}_{X\sim\mathcal{D}_\circ^m}\left[\mathbf{1}\big((x_i \succ_\theta x_j \wedge x_i \succ_\psi x_j) \vee (x_j \succ_\theta x_i \wedge x_j \succ_\psi x_j)\big)\right] \\ &= \sum_{\{i,j\}\in\binom{[m]}{2}} \mathbb{E}_{x_i\sim\mathcal{D}_\circ, x_j\sim\mathcal{D}_\circ}\left[\mathbf{1}\big((x_i \succ_\theta x_j \wedge x_i \succ_\psi x_j) \vee (x_j \succ_\theta x_i \wedge x_j \succ_\psi x_j)\big)\right] \\ &= \sum_{\{i,j\}\in\binom{[m]}{2}} \frac{\pi - d_\measuredangle(\theta,\psi)}{\pi} \\ &= \frac{\pi - d_\measuredangle(\theta,\psi)}{\pi}\binom{m}{2}.\end{aligned}$$

□

C.2 Proof of Theorem 3.1

Theorem 3.1 (Main Result) *For any batch size m and spherically symmetric item distribution $D_\circ$,*

$$\textsc{long-ip}^*(\theta_{\text{ang}}, \mathcal{D}_\circ^m) = 1.$$

Proof. We can apply the definition of LONG-IP*, as well as Lemma 3.2 and Theorem 3.3 to obtain.

$$\begin{aligned}
\text{LONG-IP}^*(\theta_{\text{ang}}, \mathcal{D}_\circ^m) &= \inf_{(\boldsymbol{\theta},\boldsymbol{\alpha})\in P} \text{LONG-IP}(\theta_{\text{ang}}, \mathcal{D}^m, (\boldsymbol{\theta},\boldsymbol{\alpha})) \\
&= \inf_{(\boldsymbol{\theta},\boldsymbol{\alpha})\in P} \min_{i\in[n]} \mathbb{E}_{X\sim\mathcal{D}^m}[\text{IP}_i(\theta_{\text{ang}}, X, (\boldsymbol{\theta},\boldsymbol{\alpha}))] \\
&= \inf_{(\boldsymbol{\theta},\boldsymbol{\alpha})\in P} \min_{i\in[n]} \mathbb{E}_{X\sim\mathcal{D}^m}\left[\frac{a_{KT}(\theta^{(i)}, \theta_{\text{ang}}((\boldsymbol{\theta},\boldsymbol{\alpha})), X)}{\alpha^{(i)}\binom{m}{2}}\right] \\
&= \inf_{(\boldsymbol{\theta},\boldsymbol{\alpha})\in P} \min_{i\in[n]} \mathbb{E}_{X\sim\mathcal{D}^m}\left[\frac{(\pi - d_\measuredangle(\theta^{(i)}, \theta_{\text{ang}}))\binom{m}{2}}{\pi\alpha^{(i)}\binom{m}{2}}\right] \\
&\geq \inf_{(\boldsymbol{\theta},\boldsymbol{\alpha})\in P} \min_{i\in[n]} \mathbb{E}_{X\sim\mathcal{D}^m}\left[\frac{\pi - (1-\alpha^{(i)})\pi\binom{m}{2}}{\pi\alpha^{(i)}\binom{m}{2}}\right] \\
&= \inf_{(\boldsymbol{\theta},\boldsymbol{\alpha})\in P} \min_{i\in[n]} \mathbb{E}_{X\sim\mathcal{D}^m}\left[\frac{\pi\alpha^{(i)}\binom{m}{2}}{\pi\alpha^{(i)}\binom{m}{2}}\right] = 1
\end{aligned}$$

Due to Proposition 2.3, we also get that LONG-IP$^*(\theta_{\text{ang}}, \mathcal{D}_\circ^m) \leq 1$, so LONG-IP$^*(\theta_{\text{ang}}, \mathcal{D}_\circ^m) = 1$. □

C.3 Proof of Theorem 3.3

In this section we provide the proof details for Theorem 3.3.

Theorem 3.3 (Fundamental Property of Angular Mean) *For any profile $(\boldsymbol{\theta},\boldsymbol{\alpha}) \in P_{n,d}$,*

$$d_\measuredangle(\theta_{\text{ang}}(\boldsymbol{\theta},\boldsymbol{\alpha}), \theta^{(i)}) \leq (1-\alpha^{(i)})\pi \quad \textit{for all } i \in [n].$$

The primary idea behind the proof is to show our main auxiliary Lemma, Lemma C.1. Essentially, it states that the distance bound from Theorem 3.3 remains valid for $\theta^{(i)}$, even if we may only select a minimizer θ^* of $F(\theta) = \sum_{i=1}^n \alpha^{(i)} d_\measuredangle(\theta, \theta^{(i)})^2$ from some great circle G_i that contains $\theta^{(i)}$.

Lemma C.1 *For a given profile $(\boldsymbol{\theta},\boldsymbol{\alpha})$, let F be the objective function of the angular mean,*

$$F(\theta) = \sum_{i\in[n]} \alpha^{(i)} d_\measuredangle(\theta, \theta^{(i)})^2.$$

Let $G_i \subseteq S^d$ be any great circle containing $\theta^{(i)}$, and write $F_{|G_i} : G_i \to \mathbb{R}$ for the restriction of F to G_i.

Every minimizer $\theta^* \in G_i$ of $F_{|G_i}$ fulfills $d_\measuredangle(\theta^{(i)}, \theta^*) \leq (1-\alpha^{(i)})\pi$.

Let us first formally establish why Theorem 3.3 follows from Lemma C.1.

Proof of Theorem 3.3. Let $(\boldsymbol{\theta},\boldsymbol{\alpha})$ be a profile in S^d. Consider any voter $i \in [n]$ and let G_i be a great circle passing through both $\theta^{(i)}$ and θ_{ang}. Because θ_{ang} is a global minimizer of F, it also minimizes F on G_i. Thus, by Lemma C.1,

$$d_\measuredangle(\theta_{\text{ang}}, \theta^{(i)}) \leq (1-\alpha^{(i)})\pi.$$

□

The full proof of Lemma C.1 is more technical. The high-level idea is a proof by contradiction. We will show that, for a global minimizer θ^* of $F_{|G_i}$ and a vector $\theta^{(i)} \in G_i$, that if $d_\measuredangle(\theta^*, \theta^{(i)}) > (1-\alpha^{(i)})\pi$ then we can construct another vector θ' with a lower objective value with $d_\measuredangle(\theta', \theta^{(i)}) < d_\measuredangle(\theta^*, \theta^{(i)})$. In the proof we will leverage a distance bound inspired by the Euclidean law of cosines and Toponogov's Inequality (Lemma C.5) and a first order local optimality condition on θ^* Employing these two tools allows us to greatly simplify the the expression $F(\theta') - F(\theta)$, which then makes it easy to verify that θ' does indeed achieve an improved objective value.

To use the first order optimality condition on the derivative of $F_{|G_i}$ at θ^*, we must first establish that F is differentiable at θ^* in the first place. Because each $d_i(\theta) = d_\measuredangle(\theta, \theta^{(i)})^2$ is differentiable everywhere other than $-\theta^{(i)}$, we just need to show that θ^* cannot be antipodal to any of the $\theta^{(i)}$.

Lemma C.2 *Let G_i and $F_{|G_i}$ be as in Lemma C.1. No minimizer $\theta^* \in G_i$ of $F_{|G_i}$ is antipodal to any of the $\theta^{(i)}$.*

Proof. First observe that the map $d_i(\theta) = d_\measuredangle(\theta, \theta^{(i)})$ is differentiable everywhere other than $-\theta^{(i)}$, where only the directional derivatives exist. So suppose that θ^* is antipodal to some of the $\theta^{(i)}$. Let $I = \{i \in [n] : \theta^{(i)} = -\theta^*\}$ be the set of indices of these antipodal $\theta^{(i)}$.

Let $x : \mathbb{R}/2\pi\mathbb{Z} \to G_i$ be a unit-speed parametrization of G_i such that $x(0) = \theta^*$. We will decompose F into a singular and non-singular part as $F(\theta) = F_{\text{sin}}(\theta) + F_{\text{diff}}(\theta)$, with

$$F_{\text{sin}} = \sum_{i \in I} \alpha^{(i)} d_\measuredangle(\theta, \theta^{(i)}) \quad \text{and} \quad F_{\text{diff}} = \sum_{i \notin I} \alpha^{(i)} d_\measuredangle(\theta, \theta^{(i)})$$

The right-sided derivative of $F(x(\varphi))$ at 0 is then given by

$$\frac{d}{d\varphi} F(x(\varphi))\bigg|_{\varphi=0^+} = \left(\frac{d}{d\varphi} F_{\text{sin}}(x(\varphi))\bigg|_{\varphi=0^+} \right) + F_{\text{diff}}(x(\varphi))' = F_{\text{sin}}^+ + F_{\text{diff}}(x(\varphi))',$$

while the left-sided derivative is given by

$$\frac{d}{d\varphi} F(x(\varphi))\bigg|_{\varphi=0^-} = \left(\frac{d}{d\varphi} F_{\text{sin}}(x(\varphi))\bigg|_{\varphi=0^-} \right) + F_{\text{diff}}(x(\varphi))' = F_{\text{sin}}^- + F_{\text{diff}}(x(\varphi))'.$$

Now, because $\theta^{(i)}$ is antipodal to $\theta^* = x(0)$ for all $i \in I$, we get that $F_{\text{sin}}(\varphi) = F_{\text{sin}}(-\varphi)$ and hence the left and right-sided derivative of $F_{\text{sin}}(x(\varphi))$ at 0 are identical up to their sign, i.e. $F_{\text{sin}}^+ = -F_{\text{sin}}^-$. In particular $F_{\text{sin}}^+ < 0$ and $F_{\text{sin}}^- > 0$, as F_{sin} only decreases when moving away from $x(0)$. $x(0)$ can only be a minimum when the right-side derivative non-negative and the left-sided derivative is non-positive, so

$$\begin{aligned} 0 \leq F_{\text{sin}}^+ + F_{\text{diff}}(x(0))' &\implies -F_{\text{diff}}(x(0))' \leq F_{\text{sin}}^+ \\ 0 \geq F_{\text{sin}}^- + F_{\text{diff}}(x(0))' &\implies F_{\text{diff}}(x(0))' \leq -F_{\text{sin}}^- \end{aligned}$$

Combining these two inequalities then yields

$$0 < F_{\text{sin}}^- \leq F_{\text{diff}}(x(0))' \leq -F_{\text{sin}}^- < 0,$$

a contradiction. Therefore θ^* cannot be antipodal to $\theta^{(i)}$.

□

Let us give a quick introduction to the aforementioned Toponogov's Inequality. Intuitively it states that spherical triangles are "tighter" than flat euclidean triangles, because the edges can curve towards each other. More formally, we compare a spherical triangle and a euclidean triangle that are both constructed using the same two side lengths and one angle. Toponogov's Inequality then states that the last side of the spherical triangle is shorter than that of the euclidean triangle.

Theorem C.3 *[[46, 33]] Let M be an m-dimensional Riemannian manifold with sectional curvature K satisfying $K \geq \delta$. Let pqr be a geodesic triangle, i.e. a triangle whose sides are geodesics, in M, such that the geodesic pq is minimal and if $\delta > 0$, the length of the side pr is less than $\pi/\sqrt{\delta}$. Let $p'q'r'$ be a geodesic triangle in the model space M_δ, i.e. the simply connected space of constant curvature δ, such that the lengths of sides $p'q'$ and $p'r'$ are equal to that of pq and pr respectively and the angle at p' is equal to that at p. Then*

$$d(q, r) \leq d(q', r').$$

We combine this theorem with the euclidean law of cosines to bound the side lengths of spherical triangles.

Corollary C.4 *Consider three points $p, q, r \in S^n$ forming a geodesic spherical triangle. Then*

$$d_\measuredangle(q, r)^2 \leq d_\measuredangle(p, q)^2 + d_\measuredangle(p, r)^2 - 2 d_\measuredangle(p, q) d_\measuredangle(p, r) \cdot \cos(\gamma),$$

where γ is the angle between $\overline{pq}$, and $\overline{pr}$ at p.

Proof. We use the law of cosines in conjunction with Toponogov's theorem. First, note that the sectional curvature K of the unit sphere is constantly equal to 1. Applying Toponogov's theorem with lower curvature bound $\delta = 1$, we obtain $d(q,r) \leq d(q',r')$, where $p'q'r'$ is the comparison triangle in the model space $M_1 = S^n$ with the same two side lengths and included angle γ at p'. Applying the Euclidean law of cosines to this Euclidean triangle then yields

$$d(q,r)^2 \leq d(q',r')^2 = d(p,q)^2 + d(p,r)^2 - 2d(p,q)d(p,r)\cos(\gamma).$$

□

The specific bound we use in our main proof is related, but slightly stronger. The distance bound should also apply in case we do not consider strictly geodesic triangles. The precise formulation is as follows.

Lemma C.5 *Let $\theta^*, \theta^{(j)} \in S^d$ with spherical distance $l_j = d_\measuredangle(\theta^*, \theta^{(j)})$. Let $x : \mathbb{R}/2\pi\mathbb{Z} \to S^d$ be a unit-speed geodesic with $x(0) = \theta^*$, and let γ_j be the angle between the geodesic $\overline{\theta^*\theta^{(j)}}$ and the tangent $x'(0)$ in the positive direction of x.*

For any $\varphi \in [-2\pi, 0]$, $x_j(\varphi)^2 = d_\measuredangle(x(\varphi), \theta^{(j)})^2$ satisfies

$$x_j(\varphi)^2 \leq l_j^2 + \varphi^2 - 2l_j\varphi\cos(\gamma_j).$$

Proof. Since $\varphi \leq 0$, φ can be seen to travel backward when compared to the forward facing tangent $x'(0)$. We consider two cases for φ:

Case 1: $\varphi \in (-\pi, 0]$**.** In this case the geodesic $\overline{\theta^* x(\varphi)}$ runs directly opposite to $x'(0)$, so the angle between the geodesics $\overline{x(\varphi)\theta^*}$ and $\overline{\theta^*\theta^{(j)}}$ is equal to $\pi - \gamma_j$. The lengths of these geodesics are $d_\measuredangle(x(\varphi), \theta^*) = |\varphi| = -\varphi$ and $d_\measuredangle(\theta^*, \theta^{(j)}) = l_j$, so by applying Corollary C.4 to these two geodesics with the angle $\pi - \gamma_j$ we obtain

$$\begin{aligned} x_j(\varphi^2) &\leq l_j^2 + (-\varphi)^2 - 2l_j(-\varphi)\cos(\pi - \gamma_j) \\ &= l_j^2 + \varphi^2 - 2l_j(-\varphi)(-\cos(\gamma_j)) \\ &= l_j^2 + \varphi^2 - 2l_j\varphi\cos(\gamma_j), \end{aligned}$$

which is exactly the desired bound for $\varphi \in [-\pi, 0)$.

Case 2: $\varphi \in (-\pi, -2\pi]$**.** Here, we employ Corollary C.4 for a different triangle. We replace the vertex θ^* by $-\theta^*$ when compared to the spherical triangle in case 1. So we will consider the two geodesics $\overline{-\theta^* x(\varphi)}$ and $\overline{-\theta^*\theta^{(j)}}$. Note that the angle between them is also γ_j, because it sits at the unique other intersection between the great circle parametrized by x and the geodesic $\overline{\theta^*, \theta^{(j)}}$. The lengths are then $d_\measuredangle(-\theta^*, x(\varphi)) = -\varphi - \pi$ and $d_\measuredangle(-\theta^*, \theta^{(j)}) = \pi - l_j$. Applying Corollary C.4

$$\begin{aligned} x_j(\varphi^2) &\leq (\pi - l_j)^2 + (-\varphi - \pi)^2 - 2(\pi - l_j)(-\varphi - \pi)\cos(\gamma_j) \\ &\leq (\pi - l_j)^2 + (-\varphi - \pi)^2 + 2(\pi - l_j)(\varphi + \pi)\cos(\gamma_j) \\ &= (\underbrace{\pi^2 - 2\pi l_j}_{(B)} + \underbrace{l_j^2}_{(A)}) + (\underbrace{\varphi^2}_{(A)} + \underbrace{2\pi\varphi + \pi^2}_{(B)}) + 2(\underbrace{\pi^2 - l_j\pi + \varphi\pi}_{(C)} \underbrace{-l_j\varphi}_{(A)})\cos(\gamma_j) \\ &= \underbrace{l_j^2 + \varphi^2 - 2l_j\varphi\cos(\gamma_j)}_{(A)} + \underbrace{2\pi^2 - 2\pi l_j + 2\pi\varphi}_{(B)} + 2(\underbrace{pi^2 - l_j\pi + \varphi\pi}_{(C)})\cos(\gamma_j) \\ &= (l_j^2 + \varphi^2 - 2l_j\varphi\cos(\gamma_j)) + 2\pi(\pi - l_j + \varphi + \cos(\gamma_j)(\pi - l_j + \varphi)) \end{aligned}$$

The expression in the first parentheses is what we want to obtain as an upper bound, so the statement is proven, if we can show the remainder is at most 0. We can rewrite this remainder via

$$2\pi(\pi - l_j + \varphi + \cos(\gamma_j)(\pi - l_j + \varphi)) = 2\pi(\pi - l_j + \varphi)(\cos(\gamma_j) + 1).$$

Note that $\cos(\gamma_j) + 1 \geq 0$, so we only need to certify that $\pi - l_j + \varphi \leq 0$. This is true, since $\varphi < -\pi$ and $l_j \geq 0$ by definition. $\square$

Having collected all the ingredients we can now get into the full proof of Lemma C.1.

Proof of Lemma C.1. Let θ^* be a global minimizer of $F_{|G_i}$. Because $\theta^{(i)} \in G_i$, we can define a unit-speed parametrization $x : \mathbb{R}/2\pi\mathbb{Z} \to G_i$ of G_i such that $x(0) = \theta^*$ and there is a unique $l_i \in [0, \pi]$ such that $x(l_i) = \theta^{(i)}$. Since G_i is a geodesic, for φ in a neighborhood of 0 we have $d_{\measuredangle}(\theta^*, x(\varphi)) = |\varphi|$, and in particular $d_{\measuredangle}(\theta^*, \theta^{(i)}) = l_i$.

To reason about the minima of $F(x(\varphi))$, we first want to analyze the derivative of $F(x(\varphi))$ at 0. By Lemma C.2, there is no point antipodal to θ^*, hence the derivative of $F(x(\varphi))$ at 0 is well defined. If we write $l_j \coloneqq d_{\measuredangle}(\theta^*, \theta^{(j)})$ and define $x_j(\varphi) = d_{\measuredangle}(x(\varphi), \theta^{(j)})$, then the spherical law of cosines states that

$$\cos(x_j(\varphi)) = \cos(l_j)\cos(\varphi) + \sin(l_j)\sin(\varphi)\cos(\gamma_j),$$

where γ_j is the angle between the two geodesics $\overline{\theta^*\theta^{(j)}}$ and $\overline{\theta^* x(\varphi)}$. Taking the derivative on both sides gives

$$\begin{aligned} -\sin(x_j(\varphi)) \cdot x_j(\varphi)' &= -\cos(l_j)\sin(\varphi) + \sin(l_j)\cos(\varphi)\cos(\gamma_j) \\ \implies x_j(\varphi)' &= \frac{\cos(l_j)\sin(\varphi) - \sin(l_j)\cos(\varphi)\cos(\gamma_j)}{\sin(x_j(\varphi))}. \end{aligned}$$

Since $l_j = x_j(0)$ by definition, plugging in $\varphi = 0$ yields $x_j'(0) = -\cos(\gamma_j)$. Furthermore, because $\theta^* = x(0)$ is a global optimum of F along G_i, first-order optimality implies

$$0 = \frac{1}{2}F'(x(0)) = \sum_{j=1}^{n} \alpha^{(j)} x_j(0) x_j'(0) = -\sum_{j=1}^{n} \alpha^{(j)} l_j \cos(\gamma_j),$$

and hence

$$0 = \sum_{j=1}^{n} \alpha^{(j)} l_j \cos(\gamma_j) = \sum_{j \neq i} \alpha^{(j)} l_j \cos(\gamma_j) + \alpha^{(i)} l_i,$$

where the last equality follows from $\gamma_i = 0$ so $\cos(\gamma_i) = 1$.

In addition, if we apply Lemma C.5 to θ^*, $x(\varphi)$, and $\theta^{(j)}$ then we obtain

$$x_j(\varphi)^2 \leq l_j^2 + \varphi^2 - 2 l_j \varphi \cos(\gamma_j).$$

For $x_i(\varphi)$ specifically, we know that both $\theta^{(i)}$ and $x(\varphi)$ are on the circle G_i, thus the distance is given by $x_i(\varphi) = d_{\measuredangle}(\theta^{(i)}, x(\varphi)) = \min_{k \in \mathbb{Z}} |(l_i - \varphi - 2k\pi)|$. Thus the squared distance $x_i(\varphi)^2$ can be bounded by

$$x_i(\varphi)^2 = \min_{k \in \mathbb{Z}} (l_i - \varphi - 2k\pi)^2 \leq (l_i - \varphi - 2\pi)^2.$$

With all these bounds established, we will now derive a contradiction. Comparing the objective value of $F(x(0))$ to that of another proposed solution $F(x(\varphi))$ results in

$$
\begin{aligned}
&F(x(\varphi)) - F(x(0)) \\
&= \sum_{j\neq i} \alpha^{(j)} x_j(\varphi)^2 + \alpha^{(i)} x_i(\varphi)^2 - \sum_{j=1}^{n} \alpha^{(j)} l_j^2 \\
&\leq \sum_{j\neq i} \alpha^{(j)} \left(l_j^2 + \varphi^2 - 2 l_j \varphi \cos(\gamma_j)\right) + \alpha^{(i)} (l_i - \varphi - 2\pi)^2 - \sum_{j=1}^{n} \alpha^{(j)} l_j^2 \\
&= \sum_{j\neq i} \alpha^{(j)} \Big(\underbrace{l_j^2}_{(A)} + \underbrace{\varphi^2}_{(B)} \underbrace{- 2 l_j \varphi \cos(\gamma_j)}_{(C)} \Big) + \alpha^{(i)} \Big(\underbrace{l_i^2}_{(A)} + \underbrace{\varphi^2}_{(B)} \underbrace{+ 4\pi^2}_{(D)} \underbrace{- 2 l_i \varphi}_{(C)} \underbrace{- 4\pi l_i + 4\pi\varphi}_{(D)} \Big) - \underbrace{\sum_{j=1}^{n} \alpha^{(j)} l_j^2}_{(A)} \\
&= \underbrace{\sum_{j=1}^{n} \alpha^{(j)} \left(l_j^2 - l_j^2\right)}_{(A)} + \underbrace{\varphi^2 \sum_{j=1}^{n} \alpha^{(j)}}_{(B)} \underbrace{- 2\varphi \left(\sum_{j\neq i}^{n} \alpha^{(j)} l_j \cos(\gamma_j) + \alpha^{(i)} l_i \right)}_{(C)} + \underbrace{\alpha^{(i)} \left(4\pi^2 + 4\pi\varphi - 4\pi l_i\right)}_{(D)}
\end{aligned}
$$

Using $l_j^2 - l_j^2 = 0$, $\sum_{j=1}^{n} \alpha^{(j)} = 1$, and the local optimality condition $\sum_{j\neq i} \alpha^{(j)} l_j \cos(\gamma_j) + \alpha^{(i)} l_i = 0$, this reduces to $F(x(\varphi)) - F(x(0)) = \varphi^2 + \alpha^{(i)} 4\pi(\pi + \varphi - l_i)$.

By using the assumption that $l_i > (1 - \alpha^{(i)})\pi$ and setting $\varphi = -2\pi\alpha^{(i)}$, we then get

$$
\begin{aligned}
F(x(\varphi)) - F(x(0)) &= \varphi^2 + \alpha^{(i)} 4\pi(\pi + \varphi - l_i) \\
&< 4\pi^2(\alpha^{(i)})^2 + \alpha^{(i)} 4\pi(\pi - 2\pi\alpha^{(i)} - (1 - \alpha^{(i)})\pi) \\
&= 4\pi^2(\alpha^{(i)})^2 + 4\pi\alpha^{(i)}(-\alpha^{(i)}\pi) \\
&= 4\pi^2(\alpha^{(i)})^2 - 4\pi^2(\alpha^{(i)})^2 = 0.
\end{aligned}
$$

But this means that $F(x(\varphi)) < F(x(0)) = F(\theta^*)$, thus contradicting our initial assumption that θ^* minimizes F on G_i. □

Approximation

C.4 Squared Kemeny Objective

Lederer et al. [30] present the Squared Kemeny rule as a means to obtaining a proportional ranking. Their setting is different — they receive voter rankings on every batch, and directly aggregate the rankings themselves. That is, given a profile $(\boldsymbol{\theta}, \boldsymbol{\alpha})$, and batch $X \in \mathcal{X}^m$, their rule would output $\succ_{SqK}$ as follows

$$
\succ^X_{\text{sqK}} = \operatorname*{arg\,min}_{\succ^X} \sum_{i=1}^{n} \alpha^{(i)} \cdot \text{KT-dist}(\succ^X, \succ^X_{\theta^{(i)}})^2
$$

where $\text{KT-dist}(\succ^X, \succ^X_{\theta^{(i)}}) = \sum_{1\leq j<k\leq m} \mathbb{I}((x_j \succ x_k \wedge x_k \succ^X_{\theta^{(i)}} x_j) \vee (x_k \succ x_j \wedge x_j \succ^X_{\theta^{(i)}} x_k))$ is the Kendall-Tau distance between $\succ^X$ and $\succ^X_{\theta^{(i)}}$.

As we are looking to optimize for a similar notion of proportionality, what if we select our linear scoring rule based off of minimizing the expected Squared Kemeny objective?

If we optimize for the Squared Kemeny objective on a known item distribution $\mathcal{D}$, we have

$$
\theta_{\text{sqK}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) \coloneqq \operatorname*{arg\,min}_{\theta \in S^{d-1}} \mathbb{E}_{X\sim\mathcal{D}^m} \left[\sum_{i=1}^{n} \alpha^{(i)} \left(\binom{m}{2} - a_{KT}(\theta^{(i)}, \theta, X) \right)^2 \right]
$$

Interestingly, we find that on spherically symmetric item distributions, θ_{sqK} approaches θ_{ang} for sufficiently large batch sizes.

Theorem C.6 *For spherically symmetric $\mathcal{D}_\circ^m$, we have that*

$$\theta_{\text{sqK}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \operatorname*{arg\,min}_{\theta \in S^{d-1}} \left(\Theta(m^2) \sum_{i \in [n]} d_{\measuredangle}(\theta^{(i)}, \theta) + \Theta(m^3) c(\theta) + \Theta(m^4) \sum_{i \in [n]} d_{\measuredangle}(\theta^{(i)}, \theta)^2 \right)$$

where $c(\theta)$ is a value that does not depend on m and is upper bounded by $\sum_{i \in [n]} d_{\measuredangle}(\theta^{(i)}, \theta)$. Importantly this means that as m increases, $\theta_{\text{sqK}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) \to \theta_{\text{ang}}$.

Proof. We simply rewrite the objective that θ_{sqK} minimizes. All expectations are with respect to $X \sim \mathcal{D}_\circ^m$. First for any $i \in [n]$, and $j, k \in [m]$ such that $j < k$, let D_{jk} be the event that $\theta^{(i)}$ and θ disagree on the ranking of x_j and x_k, i.e. $D_{jk} = \mathbb{I}\{(x_j \succ^X_{\theta^{(i)}} x_k \wedge x_k \succ^X_\theta x_j) \vee (x_k \succ^X_{\theta^{(i)}} x_j \wedge x_j \succ^X_\theta x_x)\}$. Then observe that for any $X \in \mathcal{X}^m$,

$$\binom{m}{2} - a_{KT}(\theta^{(i)}, \theta, X) = \sum_{1 \le j < k \le m} D_{jk}$$

Substituting this into our objective, we have that $\mathbb{E}\left[\sum_{i=1}^n \alpha^{(i)} \left(\binom{m}{2} - a_{KT}(\theta^{(i)}, \theta, X)\right)^2\right]$:

$$= \mathbb{E}\left[\sum_{i=1}^n \alpha^{(i)} \left(\sum_{1 \le j < k \le m} D_{jk}\right)^2\right] = \sum_{i=1}^n \alpha^{(i)} \mathbb{E}\left[\alpha^{(i)} \left(\sum_{1 \le j < k \le m} D_{jk}\right)^2\right]$$

When we expand this squared term, there are 3 different types of terms: those in which there is an indicator variable squared, those in which the two indicator variables being multiplied together share exactly one index, and those in which the indicator variables refer to disjoint sets of items. Informally we can write this in the following way:

$$= \sum_{i=1}^n \alpha^{(i)} \mathbb{E}\left[\sum_{j,k} D_{jk}^2 + \sum_{j,k,\ell} D_{jk} D_{k\ell} + \sum_{j,k,\ell,p} D_{jk} D_{\ell p}\right]$$

How many of each term is there? There are $\binom{m}{2}$ of the first type, and then $\binom{m}{3} \cdot 3 \cdot 2$ of the second type (picking three distinct indices, picking which is repeated, and picking which set of two corresponds to the first indicator), and $\binom{m}{4} \cdot \binom{4}{2}$ of the last type (picking four distinct indices, and picking which correspond to the first indicator). As the x_i are drawn iid, the indicators do not depend on the specific choice of indices (e.g. D_{jk} is the same for any choice of $j \neq k$). So by linearity of expectation, for some choice of distinct j, k, ℓ, p:

$$= \sum_{i=1}^n \alpha^{(i)} \left(\binom{m}{2} \mathbb{E}[D_{jk}^2] + 6\binom{m}{3} \mathbb{E}[D_{jk} D_{k\ell}] + 6\binom{m}{4} \mathbb{E}[D_{jk} D_{\ell p}]\right)$$

Observe that as D_{jk} is an indicator variable, $\mathbb{E}[D_{jk}^2] = \mathbb{E}[D_{jk}]$. This is just $1 -$ { the probability that $\theta^{(i)}$ and θ agree on a pairwise ranking}, so $1 - \frac{\pi - d_{\measuredangle}(\theta^{(i)}, \theta)}{\pi}$ by Proposition 2.5. Hence $\mathbb{E}[D_{jk}] = \frac{d_{\measuredangle}(\theta^{(i)}, \theta)}{\pi}$. Additionally, D_{jk} is independent of $A_{\ell p}$ when $\{j, k\}$ and $\{\ell, p\}$ are disjoint, so $\mathbb{E}[A_{jk} A_{\ell p}] = \mathbb{E}[D_{jk}] \mathbb{E}[D_{\ell p}] = \frac{d_{\measuredangle}(\theta^{(i)}, \theta)^2}{\pi^2}$. So we have that

$$= \left(\sum_{i=1}^n \alpha^{(i)} \binom{m}{2} \frac{d_{\measuredangle}(\theta^{(i)}, \theta)}{\pi}\right) + \left(\sum_{i=1}^n \alpha^{(i)} 6\binom{m}{3} \mathbb{E}[D_{jk} D_{k\ell}]\right) + \left(\sum_{i=1}^n \alpha^{(i)} 6\binom{m}{4} \frac{d_{\measuredangle}(\theta^{(i)}, \theta)^2}{\pi^2}\right)$$

Let $c(\theta) = \sum_{i=1}^n \alpha^{(i)} \mathbb{E}[D_{jk} D_{k\ell}]$ and substitute in to get the final expression:

$$= \binom{m}{2} \frac{1}{\pi} \left(\sum_{i=1}^n \alpha^{(i)} d_{\measuredangle}(\theta^{(i)}, \theta)\right) + 6\binom{m}{3} c(\theta) + \frac{6}{\pi^2} \binom{m}{4} \left(\sum_{i=1}^n \alpha^{(i)} d_{\measuredangle}(\theta^{(i)}, \theta)^2\right)$$

Note that for $m = 2$, we have that $\theta_{\text{sqK}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \arg\min_{\theta \in S^{d-1}} \sum_{i=1}^n \alpha^{(i)} d_{\measuredangle}(\theta^{(i)}, \theta)$. Let $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ be the Antipodal Construction instantiated with $\alpha^{(1)} > 0.5$ and $\alpha^{(2)} = 1 - \alpha^{(1)} < 0.5$. Minimizing

the weighted sum of *unsquared* angular distances recovers exactly the majority vector (in this case $\theta^{(1)}$). Hence, $\theta_{\mathrm{sqK}}(\boldsymbol{\theta}, \boldsymbol{\alpha})$ gives voter type 2 individual proportionality level of 0.

However, for large m, the final term dominates, and minimizing this objective is equivalent to finding θ_{ang}. We know then that in the regime of large m, $\theta_{\mathrm{sqK}}(\boldsymbol{\theta}, \boldsymbol{\alpha}) \approx \theta_{\mathrm{ang}}(\boldsymbol{\theta}, \boldsymbol{\alpha})$. By Proposition C.7, this tells us that $\theta_{\mathrm{sqK}}(\boldsymbol{\theta}, \boldsymbol{\alpha})$ inherits the good proportionality guarantees of θ_{ang} for spherically symmetrical item distributions.

Interestingly, this points to a potential "phase shift" where minimizing the expected Squared Kemeny for small batch sizes yields an overly majoritarian scoring vector, but doing so for large batch sizes yields a vector that gives strong proportionality guarantees.

□

C.5 Robustness

Proposition C.7 *For any $m \in \mathbb{N}$, any spherically symmetric batch distribution $D_\circ^m$, any profile $(\boldsymbol{\theta}, \boldsymbol{\alpha})$, $\theta, \psi \in S^d$ such that $d_\measuredangle(\theta, \psi) \leq \gamma$,*

$$\mathbb{E}_{X \sim \mathcal{D}_\circ^m}\left[\mathrm{IP}_i(\theta, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right] \geq \mathbb{E}_{X \sim D_\circ^m}[\mathrm{IP}_i(\psi, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))] - \gamma/(\pi\alpha^{(i)}) \qquad \forall i \in [n].$$

Proof. By using the identity established in Lemma 3.2 and the reverse triangle inequality we get

$$\begin{aligned}
\mathbb{E}_{X \sim \mathcal{D}^m}\left[\mathrm{IP}_i(\theta, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right] &= \mathbb{E}_{X \sim \mathcal{D}^m}\left[\frac{a_{KT}(\theta_i, \theta, X)}{\alpha^{(i)}\binom{m}{2}}\right] \\
&= \frac{1}{\alpha^{(i)}\binom{m}{2}}\mathbb{E}_{X \sim \mathcal{D}^m}\left[a_{KT}(\theta_i, \theta, X)\right] \\
&= \frac{1}{\alpha^{(i)}\binom{m}{2}}\mathbb{E}_{X \sim \mathcal{D}^m}\left[\frac{\pi - d_\measuredangle(\theta^{(i)}, \theta)}{\pi} \cdot \binom{m}{2}\right] \\
&\geq \frac{1}{\alpha^{(i)}\binom{m}{2}}\mathbb{E}_{X \sim \mathcal{D}^m}\left[\frac{\pi - d_\measuredangle(\theta^{(i)}, \psi) - \gamma}{\pi} \cdot \binom{m}{2}\right] \\
&= \frac{1}{\alpha^{(i)}\binom{m}{2}}\mathbb{E}_{X \sim \mathcal{D}^m}\left[\frac{\pi - d_\measuredangle(\theta^{(i)}, \psi)}{\pi} \cdot \binom{m}{2}\right] - \frac{\gamma}{\pi\alpha^{(i)}} \\
&= \mathbb{E}_{X \sim \mathcal{D}^m}[\mathrm{IP}_i(\psi, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))] - \frac{\gamma}{\pi\alpha^{(i)}}.
\end{aligned}$$

□

D Appendix for Section 4

D.1 Example showing failure of stronger per-batch proportionality

Example D.1 (Impossibility of IP Level ≥ 1 on some X). Fix any aggregation mechanism θ_f, and let $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ be the Antipodal Construction (Figure 1(b)) instantiated with any weights. Let $X \in \mathcal{X}^m$ be a batch where all items are co-linear on some vector v that is not orthogonal to any of $\theta^{(1)}, \theta^{(2)}, \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha})$. Then, a ranking over X by a linear scoring rule is entirely determined by the sign of its inner product with v. As $\theta^{(1)} = -\theta^{(2)}$, they have oppositely signed inner products with v. Without loss of generality, assume that $\mathrm{sign}(\langle\theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), v\rangle) = -\mathrm{sign}(\langle\theta^{(1)}, v\rangle)$, then $\mathrm{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) = 0$.

D.2 Proof of Proposition 4.2

Proposition 4.2 *For any $\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})$,* $\textsc{batch-ip}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \leq \textsc{long-ip}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha}))$.

Proof. Let $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$, and observe that for any $X \in \mathcal{X}^m$

$$\min_{i \in [n]} \mathrm{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \leq \mathrm{IP}_j(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \quad \forall j \in [n]$$

Because this is true for all $X \in \mathcal{X}^m$, this inequality is preserved when we take the expectation on both sides over $X \sim \mathcal{D}^m$:

$$\mathbb{E}_{X\sim\mathcal{D}^m}\left[\min_{i\in[n]} \text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right] \leq \mathbb{E}_{X\sim\mathcal{D}^m}\left[\text{IP}_j(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right] \quad \forall j \in [n]$$

As this inequality is true for all choices of $j \in [n]$ on the RHS, it is also true for the minimizing choice of j. Then we have that

$$\mathbb{E}_{X\sim\mathcal{D}^m}\left[\min_{i\in[n]} \text{IP}_i(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right] \leq \min_{j\in[n]} \mathbb{E}_{X\sim\mathcal{D}^m}\left[\text{IP}_j(\theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))\right]$$

Hence,

$$\text{BATCH-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \leq \text{LONG-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha}))$$

□

D.3 Proof of Theorem 4.3

Theorem 4.3 (Impossibility of BATCH-IP) *For any batch size $m \geq 3$[4], there exists a profile $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$ with $\alpha^{(i)}\binom{m}{2} \in \mathbb{Z}\ \forall i \in [n]$ such that, for any θ_f and any $\mathcal{D}$,* $\text{BATCH-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) < 1$.

Proof. As with the other upper bound proofs, we let $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ be the Antipodal Construction ($Figure\ 1(b)$) and let $\alpha^{(1)} = \frac{1}{\binom{m}{2}}, \alpha^{(2)} = \frac{\binom{m}{2}-1}{\binom{m}{2}}$. Denote $\theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha})$ as θ_f for the remainder of this proof. For notational concision, we define $Z(X)$ to be the minimum of voter 1 and 2's individual proportionality level

$$\begin{aligned} Z(X) &\coloneqq \min\left(\frac{a_{KT}(\theta^{(1)}, \theta_f, X)}{\alpha^{(1)}\binom{m}{2}}, \frac{a_{KT}(\theta^{(2)}, \theta_f, X)}{\alpha^{(2)}\binom{m}{2}}\right) \\ &= \min\left(a_{KT}(\theta^{(1)}, \theta_f, X), \frac{a_{KT}(\theta^{(2)}, \theta_f, X)}{\binom{m}{2}-1}\right) \end{aligned}$$

Then,

$$\text{BATCH-IP}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) = \mathbb{E}_{X\sim\mathcal{D}^m}[Z(X)]$$

Recall from $Corollary\ A.2$ that we can evaluate this expectation assuming that there are no scoring ties for any $\theta \in \boldsymbol{\theta}$ or θ_f over the item batch sampled.

Deterministic upper bound. First, we will show that for any $X \in \mathcal{X}^m$, with no scoring ties, we have that $Z(X) \leq 1$. Recalling via Proposition A.3 that $a_{KT}(\theta^{(2)}, \theta_f, X) = \binom{m}{2} - a_{KT}(\theta^{(1)}, \theta_f, X)$

$$\min\left(a_{KT}(\theta^{(1)}, \theta_f, X), \frac{\binom{m}{2} - a_{KT}(\theta^{(1)}, \theta_f, X)}{\binom{m}{2}-1}\right) \leq 1$$

Observe that $a_{KT}(\theta^{(1)}, \theta_f, X)$ is a nonnegative integer. If $a_{KT}(\theta^{(1)}, \theta_f, X) = 0$, then the minimum of the two terms is 0. If $a_{KT}(\theta^{(1)}, \theta_f, X) \geq 1$, then $\frac{\binom{m}{2}-a_{KT}(\theta^{(1)},\theta_f,X)}{\binom{m}{2}-1} \leq 1$, so the minimum of the two terms is at most 1.

Now that we know that $Z(X) \leq 1$ *deterministically* for all X with no ties, we will show that with with strictly positive probability, $Z(X) = 0$.

Showing positive probability on a near-collinear batch. First observe that if $\theta_f = \theta^{(2)}$, then $a_{KT}(\theta^{(1)}, \theta_f, X) = 0$ for all X with no ties, and we immediately have that $Z(X) = 0$ with positive probability (in fact, with probability 1).

Assume then that $\theta_f \neq \theta^{(2)}$. This means that θ_f and $\theta^{(1)}$ are not antipodal – and as they are both unit vectors, their sum is nonzero. So, if we let v be the unit vector bisecting them, we have that

$$v = \frac{\theta^{(1)} + \theta_f}{\|\theta^{(1)} + \theta_f\|}$$

[4]Note that for $m = 2$, there is no profile satisfying the integral proportionality entitlement condition.

and it follows that

$$\langle \theta^{(1)}, v \rangle = \frac{1 + \langle \theta^{(1)}, \theta_f \rangle}{\|\theta^{(1)} + \theta_f\|} = \langle \theta_f, v \rangle.$$

We remark that here and throughout this proof, whenever there is a norm being used (also to define a ball), it is the 2-norm. Let $\gamma := \langle \theta^{(1)}, v \rangle = \langle \theta_f, v \rangle$ and observe that $\gamma > 0$ because $\langle \theta^{(1)}, \theta_f \rangle > -1$.

The high-level idea is diagrammed in Figure 5, and is as follows: if we sample $X = (x_1, \ldots, x_m)$ such that they are all along v (or some translation/scaling thereof), then θ_f and $\theta^{(1)}$ will agree on every pairwise comparison in X. Consequently, $\theta^{(2)}$ will have KT-agreement of 0 with θ_f on X, and $Z(X) = 0$. However, this outcome has probability measure 0. Therefore, we have to allow for some wiggle room around all of the points lying exactly on v so that such a batch is sampled with positive probability. We construct this wiggle room in such a way that for any items in it, their pairwise difference vectors are still within the region of agreement between $\theta^{(1)}$ and θ_f.

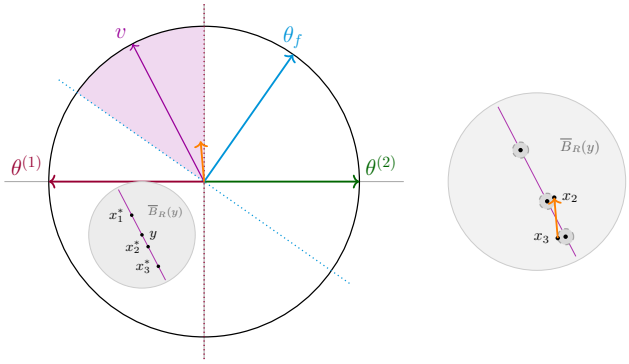


Figure 5: Proof visualization for Theorem 4.3.

By our assumption on $\mathcal{D}$, there exists an open set $U \subseteq \mathbb{R}^d$ in which $f_{\mathcal{D}}$ is strictly positive and continuous. Fix some point $y \in U$ and a radius $R > 0$ such that $\overline{B}_R(y) \subset U$. Let $\eta = \min_{x \in \overline{B}_R(y)} f_{\mathcal{D}}(x) > 0$ be the minimum probability density in this ball. Now we define an ideal series of points for the m items that are perfectly aligned along v. For $k \in \{1, \ldots, m\}$, define x_k^*:

$$x_k^* = y + \left(\frac{m - 2k}{2m} \right) R \cdot v$$

Now let $\varepsilon = \frac{R\gamma}{3m}$. We will show two claims: (1) for any batch $X = (x_1, \ldots, x_m)$ such that $x_i \in B_\varepsilon(x_k^*)$ for all $k \in \{1, \ldots, m\}$, $a_{KT}(\theta^{(2)}, \theta_f, X) = 0$ and consequently $Z(X) = 0$ and(2) the probability of drawing such a batch is strictly positive.

To show claim (1), fix some $X = (x_1, \ldots, x_m)$ such that $x_i \in B_\varepsilon(x_k^*)$ for all $k \in \{1, \ldots, m\}$ and consider $i, j \in [m]$ with $i < j$. We have that $x_i = x_i^* + \delta_i$ and $x_j = x_j^* + \delta_j$ for some $\delta_i, \delta_j \in \mathbb{R}^d$ with $\|\delta_i\|, \|\delta_j\| \leq \varepsilon$. Then:

$$\langle \theta^{(1)}, x_i - x_j \rangle = \langle \theta^{(1)}, x_i^* - x_j^* \rangle + \langle \theta^{(1)}, \delta_i - \delta_j \rangle$$

By observing that for any $i < j$, $x_i^* - x_j^* = \frac{(j-i)R}{m} \cdot v$ and applying Cauchy-Schwarz to the second term, we can rewrite as follows.

$$\begin{aligned}
&\geq \frac{(j-i)R}{m} \langle \theta^{(1)}, v \rangle - \underbrace{\|\theta^{(1)}\|}_{1} \underbrace{\|\delta_i - \delta_j\|}_{\leq 2\varepsilon} \\
&\geq \frac{(j-i)R\gamma}{m} - 2\varepsilon \\
&= \frac{(j-i)R\gamma}{m} - \frac{2R\gamma}{3m} \geq \frac{R\gamma}{3m} > 0
\end{aligned}$$

By the exact same computation, we get that $\langle \theta_f, x_i - x_j \rangle > 0$ — all we needed is that $\|\theta^{(1)}\| = 1$ and $\langle \theta^{(1)}, v \rangle = \gamma$, and both hold when we substitute in θ_f instead. Hence, we have that for any $i < j$,

both $x_i \succ^X_{\theta^{(1)}} x_j$ and $x_i \succ^X_{\theta_f} x_j$. Therefore, $a_{KT}(\theta^{(1)}, \theta_f, X) = \binom{m}{2}$, so $a_{KT}(\theta^{(2)}, \theta_f, X) = 0$ and $Z(X) = 0$.

To show claim (2), first we'll argue that for all $k \in [m]$, $B_\varepsilon(x^*_k) \subset \overline{B}_R(y)$. Fix some $k \in [m]$, and some $x_k \in B_\varepsilon(x^*_k)$. By the triangle inequality, we have that

$$\|x_k - y\| \leq \|x_k - x^*_k\| + \|x^*_k - y\| \leq \frac{R\gamma}{3m} + \left\| \left(\frac{m-2k}{2m} \right) R \cdot v \right\| \leq \frac{R}{3m} + R\left(\frac{|m-2k|}{2m} \right)$$

where we used the fact that $\|v\| = 1$ and $\gamma \leq 1$. Now observe that $|m - 2k| \leq m$ (achieved when $k = m$), and so we have that $\|x_k - y\| \leq \frac{R}{3m} + \frac{R}{2} < R$ as $m \geq 3$. Therefore, $x_k \in \overline{B}_R(y)$. As this was true for arbitrary k and arbitrary $x_k \in B_\varepsilon(x^*_k)$, we can say that for all $k \in \{1, \ldots, m\}$, $B_\varepsilon(x^*_k) \subset \overline{B}_R(y)$. Now recall that $f_{\mathcal{D}}(x) \geq \eta > 0$ for all $x \in \overline{B}_R(y)$ and hence all $x \in B_\varepsilon(x^*_k)$ for some $k \in [m]$. Then we have that

$$\Pr_{X \sim \mathcal{D}^m}(x_k \in B_\varepsilon(x^*_k) \quad \forall k \in [m]) = \prod_{k \in [m]} \Pr_{x_k \sim \mathcal{D}}(x_k \in B_\varepsilon(x^*_k)) \geq \eta^m (\text{Vol}(B_\varepsilon))^m > 0$$

Putting it all together. By conditioning on whether or not we witness a "clustered" item batch, we can derive the desired upper bound on the expectation. Let C be the event that $x_k \in B_\varepsilon(x^*_k) \quad \forall k \in [m]$. Then,

$$\begin{aligned} \mathbb{E}_{X \sim \mathcal{D}^m}[Z(X)] &= \mathbb{E}_{X \sim \mathcal{D}^m}[Z(X)|C] \Pr(C) + \mathbb{E}_{X \sim \mathcal{D}^m}[Z(X)|\overline{C}] \Pr(\overline{C}) \\ &= 0 \cdot \Pr(C) + 1 \cdot (1 - \underbrace{\Pr(C)}_{>0}) < 1 \end{aligned}$$

□

D.4 Proof of Theorem 4.5

Theorem 4.5 (Main Result) *For all $\theta_f, m, \mathcal{D}$, and $(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in P_{n,d}$,*

$$\textsc{batch-ip}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) \geq \textsc{long-ip}(\theta_f, \mathcal{D}^m, (\boldsymbol{\theta}, \boldsymbol{\alpha})) - \tfrac{1}{\alpha_{min}} \cdot O\left(\sqrt{\tfrac{\min\{\log(n), d\log(m/d)\}}{m}} \right).$$

First, define a "deviation quantity" for voter type i:

$$W_i \coloneqq \mathbb{E}_{X \sim \mathcal{D}^m}\left[\text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X)\right] - \text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X).$$

The W_i captures how much the normalized individual agreement for voter type i falls short of the expected normalized individual agreement. Observe that we can then rewrite $\min_{i \in [n]} \text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \mathcal{D})$ in terms of expected normalized individual agreement minus this deviation term as follows:

$$\begin{aligned} \min_{i \in [n]} \text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X) &= \min_{i \in [n]} \left(\mathbb{E}_{X \sim \mathcal{D}^m}[\text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X)] - W_i \right) \\ &\geq \min_{i \in [n]} \mathbb{E}_{X \sim \mathcal{D}^m}[\text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X)] - \max_{i \in [n]} W_i \end{aligned}$$

Taking the expectation of both sides, applying linearity of expectation, and observing that $\mathbb{E}[\min_{i \in [n]} \mathbb{E}[\text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X)]] = \min_{i \in [n]} \mathbb{E}[\text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X)]$ as the inner expectation is a constant yields:

$$\textsc{batch-ip}(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \mathcal{D}^m) \geq \textsc{long-ip}(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), \mathcal{D}^m) - \mathbb{E}[\max_{i \in [n]} W_i]$$

Hence, the remaining task is upper bounding $\mathbb{E}_{X \sim \mathcal{D}^m}[\max_{i \in [n]} W_i]$. We show two different upper bounds, one that has a dependence on n and the other that has a dependence on d, which results in the minimum term in the final expression.

Both upper bounds rely on reformulating normalized individual agreement as a U-statistic (with an addition $\frac{1}{\alpha_i}$ factor). Recall that

$$\text{IP}_i(\theta_f, (\boldsymbol{\theta}, \boldsymbol{\alpha}), X) = \frac{a_{KT}(\theta^{(i)}, \theta_f, X, (\boldsymbol{\theta}, \boldsymbol{\alpha}))}{\alpha_i \binom{m}{2}}$$

As θ_f and $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ are fixed, we omit the function inputs in the rest of this proof for concision, and just write IP_i to refer to the quantity above. Define a kernel function $h_{\theta^{(i)}} : \mathbb{R}^d \to \{0, 1\}$ that takes in a difference vector (between two items) and indicates whether $\theta^{(i)}$ and $\theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha})$ have the same sign of dot product with this difference vector

$$h_{\theta^{(i)}}(z) \coloneqq \mathbb{I}\{\text{sign}(\langle \theta^{(i)}, z \rangle) = \text{sign}(\langle \theta_f(\boldsymbol{\theta}, \boldsymbol{\alpha}), z \rangle)\}$$

Recalling that agreement between scoring vectors on a pairwise comparison between items just comes down to whether the two scoring vectors have the same signed dot product with the item difference vector, we can rewrite the IP_i as

$$\text{IP}_i = \frac{1}{\alpha_i \binom{m}{2}} \sum_{1 \le j < \ell \le m} h_{\theta^{(i)}}(x_j - x_\ell)$$

For all permutations $\sigma \in \Pi_m$, define

$$V_{i,\sigma} = \frac{1}{\lfloor m/2 \rfloor} \left(h_{\theta^{(i)}}(x_{\sigma(1)} - x_{\sigma(2)}) + \cdots + h_{\theta^{(i)}}(x_{\sigma(2\lfloor m/2 \rfloor - 1)} - x_{\sigma(2\lfloor m/2 \rfloor)}) \right)$$

Then we can rewrite IP_i again, this time as a function of the $V_{i,\sigma}$.

$$\text{IP}_i = \frac{1}{\alpha_i \cdot m!} \sum_{\sigma \in \Pi_m} V_{i,\sigma}$$

The key observation is that every $V_{i,\sigma}$ is the sum of $\lfloor m/2 \rfloor$ *independent* variables, because each term is the application of the kernel function to a disjoint pair of items. Therefore, we can more easily apply standard results to show that the all of the $V_{i,\sigma}$ concentrate around their expectation quickly, and translate that into bounds on the deviation of IP_i from its expectation (W_i).

Now the two upper bound methods diverge. For the first bound, resulting in a dependence on n, we use the reformulation of IP_i to show that the W_i are all sub-Gaussian in Lemma D.2. Then, we apply a known bound on the maximum of sub-Gaussian random variables in Lemma D.3.

The second method, found in Lemma D.4, bounds the maximum W_i by taking the supremum over the deviation quantity for *any* scoring vector in S^{d-1} and bounding this using Rademacher complexity. This bound supports the intuition that even if there are many different voter types, the diversity among scoring vectors is limited by the dimension. Scoring vectors that are close together in space have positively correlated deviation quantities — on any given batch, they are likely experience similar deviations from their expected individual proportionality.

Lemma D.2 *For all $i \in [n]$, $W_i = \mathbb{E}[\text{IP}_i] - \text{IP}_i$ is sub-Gaussian with parameter $\sigma^2 = \frac{1}{4\lfloor m/2 \rfloor \alpha_i^2}$.*

Proof. In order to show sub-Gaussianity, we want to show that $\mathbb{E}[e^{\lambda(W_i - \mathbb{E}[W_i])}] \le e^{\sigma^2 \lambda^2 / 2}$ for all $\lambda \in \mathbb{R}$. Observe that $\mathbb{E}[W_i] = 0$. To bound the MGF,

$$\begin{aligned}
\mathbb{E}[e^{\lambda(W_i - \mathbb{E}[W_i])}] &= \mathbb{E}[e^{\lambda(\mathbb{E}[\text{IP}_i] - \text{IP}_i)}] \\
&= \mathbb{E}\left[\exp\left(\frac{\lambda}{\alpha_i} \sum_{\sigma \in \Pi_m} \frac{1}{m!} (\alpha_i \mathbb{E}[\text{IP}_i] - V_{i,\sigma}) \right) \right]
\end{aligned}$$

By Jensen's inequality:

$$\le \mathbb{E}\left[\sum_{\sigma \in \Pi_m} \frac{1}{m!} \exp\left(\frac{\lambda}{\alpha_i} (\alpha_i \mathbb{E}[\text{IP}_i] - V_{i,\sigma}) \right) \right]$$

Then by linearity of expectation and the observation that all $V_{i,\sigma}$ have the same distribution (so the statement holds for any permutation σ):

$$\begin{aligned}
&= \mathbb{E}\left[\exp\left(\frac{\lambda}{\alpha_i} (\alpha_i \mathbb{E}[\text{IP}_i] - V_{i,\sigma}) \right) \right] \\
&= \mathbb{E}\left[\exp\left(\frac{-\lambda}{\alpha_i} (V_{i,\sigma} - \alpha_i \mathbb{E}[\text{IP}_i]) \right) \right]
\end{aligned}$$

Note that $\alpha_i \mathbb{E}_{x_1,\ldots,x_m}[\text{IP}_i] = \mathbb{E}_{x_1,x_2}[h_{\theta^{(i)}}(x_1 - x_2)] = \mathbb{E}_{x_1,\ldots,x_m}[V_{i,\sigma}]$. So the expression $V_{i,\sigma} - \alpha_i \mathbb{E}[\text{IP}_i]$ can be rewritten as $V_{i,\sigma} - \mathbb{E}[V_{i,\sigma}]$. Now we can see that the task has come down to bounding the MGF of $V_{i,\sigma}$. Recalling the definition of $V_{i,\sigma}$ and plugging in yields:

$$= \mathbb{E}\left[\exp\left(\frac{-\lambda}{\alpha_i}\left(\sum_{j=1}^{k}\frac{h_{\theta^{(i)}}(x_{\sigma(2j-1)} - x_{\sigma(2j)})}{k} - \frac{\mathbb{E}[V_{i,\sigma}]}{k}\right)\right)\right]$$
$$= \mathbb{E}\left[\prod_{j=1}^{k}\exp\left(\frac{-\lambda}{\alpha_i}\left(\frac{h_{\theta^{(i)}}(x_{\sigma(2j-1)} - x_{\sigma(2j)})}{k} - \frac{\mathbb{E}[V_{i,\sigma}]}{k}\right)\right)\right]$$

Now we can utilize the independence of $h_{\theta^{(i)}}(x_{\sigma(2j-1)}, x_{\sigma(2j)})$ for different values of j. Recall that these are independent because each of the $\lfloor m/2 \rfloor$ terms are functions of disjoint subsets of the randomly drawn $x_1, \ldots, x_m$.

$$= \prod_{j=1}^{k}\mathbb{E}\left[\exp\left(\frac{-\lambda}{\alpha_i}\left(\frac{h_{\theta^{(i)}}(x_{\sigma(2j-1)} - x_{\sigma(2j)})}{k} - \frac{\mathbb{E}[V_{i,\sigma}]}{k}\right)\right)\right]$$

Moving the factor of $1/k$ outside and recalling that $\mathbb{E}[V_{i,\sigma}] = \mathbb{E}\left[h_{\theta^{(i)}}(x_{\sigma(2j-1)}, x_{\sigma(2j)})\right]$ gives us:

$$= \prod_{j=1}^{k}\mathbb{E}\left[\exp\left(\frac{-\lambda}{\alpha_i k}\left(h_{\theta^{(i)}}(x_{\sigma(2j-1)} - x_{\sigma(2j)}) - \mathbb{E}[V_{i,\sigma}]\right)\right)\right]$$

Then, we can apply Hoeffding's lemma to each of these k terms, which states that for a random variable X that is bounded in $[a, b]$ w.p. 1, we have that $\mathbb{E}[\exp(c(X - \mathbb{E}[X]))] \leq e^{c^2(b-a)^2/8}$ for any $c \in \mathbb{R}$. Observe that $h_{\theta^{(i)}}(x_{\sigma(2j-1)} - x_{\sigma(2j)}) \in [0, 1]$, hence:

$$\leq \prod_{j=1}^{k} e^{\lambda^2/(8\alpha_i^2 k^2)} = e^{\lambda^2/(8\alpha_i^2 k)}$$

Putting it altogether and plugging in for k, we have that:

$$\mathbb{E}[e^{\lambda(W_i - \mathbb{E}[W_i])}] \leq e^{\frac{\lambda^2}{8\alpha_i^2 \lfloor m/2 \rfloor}}$$

So, W_i is sub-Gaussian with parameter $\sigma^2 = \frac{1}{4\alpha_i^2 \lfloor m/2 \rfloor}$. □

Lemma D.3 *We can bound the expected max fluctuation in individual proportionality as follows:*

$$\mathbb{E}_X[\max_{i\in[n]} W_i] = \mathbb{E}_X\left[\max_{i\in[n]}\left(\mathbb{E}[\text{IP}_i] - \text{IP}_i\right)\right] \leq \frac{1}{\alpha_{min}}\sqrt{\frac{\log(n)}{2\lfloor m/2 \rfloor}}$$

Proof. In combination with our result showing that the individual W_i are sub-Gaussian in Lemma D.2, we can directly apply a known bound on the maximum of (not necessarily independent) sub-Gaussian RVs (e.g. Rigollet and Hütter [41] Thm 1.14). We can upper bound the variance parameter of all of these individual variables by $\max_{i\in[n]} \sigma_i^2 \leq \max_{i\in[n]} \frac{1}{4\lfloor m/2 \rfloor \alpha_i^2} = \frac{1}{4\lfloor m/2 \rfloor \alpha_{min}^2}$. So we have that

$$\mathbb{E}_X[\max_{i\in[n]} W_i] \leq \sqrt{2\sigma^2 \log(n)} \leq \frac{1}{\alpha_{min}}\sqrt{\frac{\log(n)}{2\lfloor m/2 \rfloor}}$$

□

Lemma D.4 *We can bound the expected maximum deviation in individual proportionality in terms of d as follows:*

$$\mathbb{E}_{X\sim\mathcal{D}^m}\left[\max_{i\in[n]} W_i\right] \leq \frac{1}{\alpha_{min}}\sqrt{\frac{d \cdot 8\log(e\lfloor m/2 \rfloor / d)}{\lfloor m/2 \rfloor}}$$

Proof. Observe that we can upper bound the maximum deviation experienced by any voter type, by considering the maximum deviation that would be experienced for *any* scoring vector on the sphere, if it were associated with the minimum weight $\alpha_{min} = \min_{i\in[n]} \alpha_i$.

$$\mathbb{E}\left[\max_{i\in[n]} W_i\right] = \mathbb{E}\left[\max_{i\in[n]}\left(\mathbb{E}\left[\frac{a_{KT}(\theta^{(i)},\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))}{\alpha_i\binom{m}{2}}\right] - \frac{a_{KT}(\theta^{(i)},\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))}{\alpha_i\binom{m}{2}}\right)\right]$$
$$\leq \mathbb{E}\left[\sup_{\theta\in S^{d-1}}\left(\mathbb{E}\left[\frac{a_{KT}(\theta,\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))}{\alpha_{min}\binom{m}{2}}\right] - \frac{a_{KT}(\theta,\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))}{\alpha_{min}\binom{m}{2}}\right)\right]$$

We can use our method of rewriting individual proportionality for an arbitrary vector $\theta \in S^{d-1}$:

$$\frac{a_{KT}(\theta^{(i)},\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))}{\binom{m}{2}} = \frac{1}{m!}\sum_{\sigma\in\Pi_m}\frac{1}{\lfloor m/2\rfloor}\sum_{j=1}^{\lfloor m/2\rfloor} h_\theta(x_{\sigma(2j-1)} - x_{\sigma(2j)})$$

recalling that $h_\theta(x-y)$ is the kernel function, and is an indicator function for whether θ and $\theta_f(\boldsymbol{\theta},\boldsymbol{\alpha})$ agree on the ranking of x and y. Observe that all of the x_i are iid, so for any permutation σ, all of the difference vectors $x_{\sigma(2j-1)} - x_{\sigma(2j)}$ are also identically distributed. Then, by linearity of expectation $\mathbb{E}\left[\frac{1}{\lfloor m/2\rfloor}\sum_{j=1}^{\lfloor m/2\rfloor} h_\theta(x_{\sigma(2j-1)} - x_{\sigma(2j)})\right] = \mathbb{E}[h_\theta(x_1 - x_2)]$. We can continue in our bound on $\mathbb{E}\left[\max_{i\in[n]} W_i\right]$ follows:

$$\leq \frac{1}{\alpha_{min}}\mathbb{E}\left[\sup_{\theta\in S^{d-1}}\left(\mathbb{E}\left[\frac{a_{KT}(\theta,\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))}{\binom{m}{2}}\right] - \frac{a_{KT}(\theta,\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))}{\binom{m}{2}}\right)\right]$$
$$= \frac{1}{\alpha_{min}}\mathbb{E}\left[\sup_{\theta\in S^{d-1}}\frac{1}{m!}\sum_{\sigma\in\Pi_m}\left(\mathbb{E}[h_\theta(x_1-x_2)] - \frac{1}{\lfloor m/2\rfloor}\sum_{j=1}^{\lfloor m/2\rfloor} h_\theta(x_{\sigma(2j-1)} - x_{\sigma(2j)})\right)\right]$$

The permutation average can be thought of as the expectation over a permutation drawn uniformly at random. Then, by the convexity of supremum and Jensen's inequality, we can upper bound this expression by pulling the supremum inside the permutation average:

$$\leq \frac{1}{\alpha_{min}}\mathbb{E}\left[\frac{1}{m!}\sum_{\sigma\in\Pi_m}\sup_{\theta\in S^{d-1}}\left(\mathbb{E}[h_\theta(x_1-x_2)] - \frac{1}{\lfloor m/2\rfloor}\sum_{j=1}^{\lfloor m/2\rfloor} h_\theta(x_{\sigma(2j-1)} - x_{\sigma(2j)})\right)\right]$$

Note that this supremum is the same for all permutations, as all of them share the same distribution. So, we can fix any permutation σ, and write our expression in terms of it:

$$= \frac{1}{\alpha_{min}}\mathbb{E}\left[\sup_{\theta\in S^{d-1}}\left(\mathbb{E}[h_\theta(x_1-x_2)] - \frac{1}{\lfloor m/2\rfloor}\sum_{j=1}^{\lfloor m/2\rfloor} h_\theta(x_{\sigma(2j-1)} - x_{\sigma(2j)})\right)\right]$$

If we let $\mathcal{H} = \{h_\theta \colon \theta \in S^{d-1}\} = \{z \mapsto \mathbb{I}(\text{sign}(\langle z,\theta\rangle) = \text{sign}(\langle z,\theta_f(\boldsymbol{\theta},\boldsymbol{\alpha})\rangle) \colon \theta \in S^{d-1}\}$ be our function class, then it is a known result that we can upper bound this expected supremum by twice the Rademacher complexity of $\mathcal{H}$ (e.g. [48] Prop 4.11).

$$\leq \frac{2}{\alpha_{min}}\mathcal{R}_{\lfloor m/2\rfloor}(\mathcal{H})$$

In order to bound $\mathcal{R}_{\lfloor m/2\rfloor}(\mathcal{H})$, we first bound the VC-dimension of $\mathcal{H}$ and then use a known connection between the two.

Consider the class of linear halfspaces that pass through the origin, $\mathcal{F} = \{z \mapsto \text{sign}(\langle\theta, z\rangle) \mid \theta \in S^{d-1}\}$. Because our function class $\mathcal{H}$ is formed by checking equality against a fixed boolean label pre-assigned by $\theta_f(\boldsymbol{\theta},\boldsymbol{\alpha})$, $\mathcal{H}$ cannot shatter any set of points that $\mathcal{F}$ cannot shatter. While the VC dimension of general affine halfspaces in $\mathbb{R}^d$ is $d+1$ ([34], Example 3.12 and Thm 3.13), our class $\mathcal{F}$ has the additional restriction that the halfspaces must pass through the origin, making it geometrically isomorphic to the class of general affine halfspaces in $\mathbb{R}^{d-1}$. Therefore, $\text{VCdim}(\mathcal{F}) = d$. Thus, $\text{VCdim}(\mathcal{H}) \leq \text{VCdim}(\mathcal{F}) = d$.

Finally, combining Massart's lemma and Sauer's lemma (and their implications in [34] Corollaries 3.8 and 3.18), we get that

$$\frac{2}{\alpha_{min}}\mathcal{R}_{\lfloor m/2\rfloor}(\mathcal{H}) \leq \frac{2}{\alpha_{min}}\sqrt{\frac{2\log((e\lfloor m/2\rfloor/d)^d)}{\lfloor m/2\rfloor}} = \frac{1}{\alpha_{min}}\sqrt{\frac{d\cdot 8\log(e\lfloor m/2\rfloor/d)}{\lfloor m/2\rfloor}}$$

□

Returning to the proof of Theorem 4.5, we can now plug in our upper bounds on $\mathbb{E}[\max_{i\in[n]} W_i]$ from Lemma D.2 and Lemma D.4, and recover the desired bound:

$$\text{BATCH-IP}(\theta_f,(\boldsymbol{\theta},\boldsymbol{\alpha}),\mathcal{D}^m) \geq \text{LONG-IP}(\theta_f,(\boldsymbol{\theta},\boldsymbol{\alpha}),\mathcal{D}^m) - \frac{1}{\alpha_{min}}\sqrt{\frac{\min(\log(n), d\cdot 16\log(e\lfloor m/2\rfloor/d))}{2\lfloor m/2\rfloor}}$$

D.5 Tail bound on i's individual proportionality level

Corollary D.5 *For any $\theta_f, \mathcal{D}^m, (\boldsymbol{\theta},\boldsymbol{\alpha}), i \in (\boldsymbol{\theta},\boldsymbol{\alpha})$, and any $t>0$ we have that*

$$\Pr_{X\sim\mathcal{D}^m}\left(\text{IP}_i(\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha})) < \mathbb{E}[\text{IP}_i(\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))] - t\right) \leq e^{-2\lfloor m/2\rfloor\alpha_i^2 t^2}$$

Proof. We can reformulate this as a tail bound on W_i:

$$\begin{aligned}
&\Pr_{X\sim\mathcal{D}^m}\left(\text{IP}_i(\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha})) < \mathbb{E}[\text{IP}_i(\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))] - t\right)\\
&= \Pr_{X\sim\mathcal{D}^m}\left(\mathbb{E}[\text{IP}_i(\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha}))] - \text{IP}_i(\theta_f,X,(\boldsymbol{\theta},\boldsymbol{\alpha})) > t\right)\\
&= \Pr_{X\sim\mathcal{D}^m}(W_i > t)
\end{aligned}$$

We know from Lemma D.2 that W_i is sub-Gaussian with variance parameter $\sigma^2 = \frac{1}{4\lfloor m/2\rfloor\alpha_i^2}$. This immediately gives us the following tail bound (e.g. [48] Equation 2.9):

$$\begin{aligned}
&\leq e^{\frac{-t^2}{2\sigma^2}}\\
&= e^{-1\lfloor m/2\rfloor\alpha_i^2 t^2}
\end{aligned}$$

□

E Appendix for Section 5

Note that, on a 12-core machine, running all experiments takes approximately one hour. No special computing resources are needed.

E.1 Approximation of θ_{arith} to θ_{ang} under high agreement

Proposition E.1 *Fix any profile $(\boldsymbol{\theta},\boldsymbol{\alpha})$ and $R\in[0,\pi/2)$ for which there exists $\theta\in S^{d-1}$ such that for all $i\in[n]$, $d_\measuredangle(\theta,\theta^{(i)})\leq R$. Then, $d_\measuredangle(\theta_{\text{ang}},\theta_{\text{arith}}) \leq \arctan\left(\frac{2R-\sin(2R)}{\cos(R)}\right)$.*

Proof. Without loss of generality, suppose $\theta_{\text{ang}} = e_1$ where $e_1 = (1,0,\ldots,0)\in S^{d-1}$. If we write $d_i := d_\measuredangle(\theta_{\text{ang}},\theta^{(i)})$ we get that $\cos(d_i) = \langle\theta_{\text{ang}},\theta^{(i)}\rangle = \theta_1^{(i)}$ for every $i\in\mathbb{N}$. Next we set $w_i := \theta^{(i)} - e_1\cos(d_i)$, and $v_i := w_i/||w_i||$. Then $w_i = v_i\sin(d_i)$ only consists of the last $d-1$ components of $\theta^{(i)}$, and v_i is w_i normalized. We use this to rewrite

$$\theta^{(i)} = \begin{pmatrix}\cos(d_i)\\ v_i\sin(d_i)\end{pmatrix}.$$

The unnormalized arithmetic mean can then be written as

$$\theta^*_{\text{arith}} = \sum_{i=1}^n \alpha^{(i)}\theta^{(i)} = \sum_{i=1}^n \alpha^{(i)}\begin{pmatrix}\cos(d_i)\\ v_i\sin(d_i)\end{pmatrix} = \begin{pmatrix}\sum_{i=1}^n\alpha^{(i)}\cos(d_i)\\ \sum_{i=1}^n\alpha^{(i)}v_i\sin(d_i)\end{pmatrix}.$$

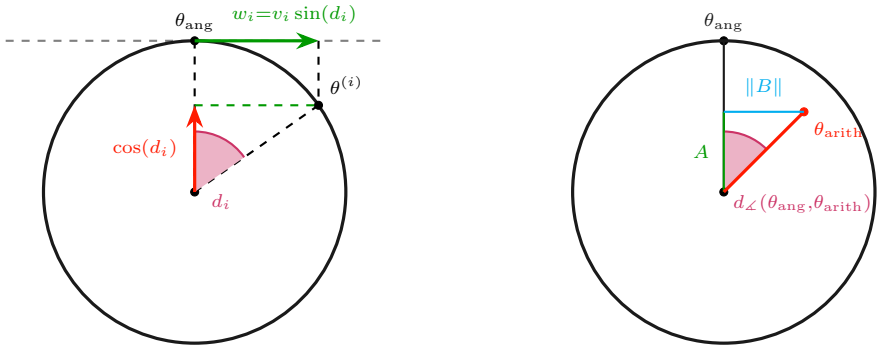


Figure 6: *Left:* $\theta^{(i)}$ decomposes into $(\cos(d_i), v_i \sin(d_i))$, *Right:* Through considering the right angled triangle formed through the origin, θ_{arith} and the projection of θ_{arith} to θ_{ang}, we can rewrite $d_\measuredangle(\theta_{\text{ang}}, \theta_{\text{arith}}) = \arctan(\|B\|/A)$.

Because θ_{arith} is now just the normalized vector of θ^*_{arith}, the distance $d_\measuredangle(\theta_{\text{arith}}, \theta_{\text{ang}})$ is the angle between θ_{ang} and θ^*_{arith}.

We write $A = \sum_{i=1}^n \alpha^{(i)} \cos(d_i)$, and $B = \sum_{i=1}^n \alpha^{(i)} v_i \sin(d_i)$. From Figure 6, we can then see that we can therefore determine the angle $d_\measuredangle(\theta_{\text{ang}}, \theta_{\text{arith}})$ as $\arctan\left(\frac{||B||}{A}\right)$. Because $\arctan$ is monotonic, we will separately bound A from below and $||B||$ from above.

Bounding $||B||$. We will prove that $||B|| \leq (2R) - \sin(2R)$. To see this observe that

$$||B|| = \left|\left|\sum_{i=1}^n \alpha^{(i)} v_i \sin(d_i)\right|\right| = \left|\left|\sum_{i=1}^n \alpha^{(i)} v_i (\sin(d_i) - d_i)\right|\right|$$
$$\leq \sum_{i=1}^n \alpha^{(i)} ||v_i|| (d_i - \sin(d_i)) \leq \sum_{i=1}^n \alpha^{(i)} (2R - \sin(2R)) = 2R - \sin(2R).$$

Where we use the local optimality condition in the second equality, as well as the fact that the function $f(x) = x - \sin(x)$ and $d_i = d_\measuredangle(\theta_{\text{ang}}, \theta^{(i)}) \leq 2R$, since θ_{ang} must be in the spherically convex hull of the $\theta^{(i)}$ [6].

Bounding A. We will show that $A = \sum_{i=1}^n \alpha^{(i)} \cos(d_i) \geq \cos(R)$. If we define $r : [0, \pi] \to \mathbb{R}, x \mapsto \cos(x) - \cos(R)$ and $r_i := \cos(d_i) \cos(R)$, then this is equivalent to proving

$$\sum_{i=1}^n \alpha^{(i)} r_i > 0, \tag{1}$$

Now observe that r is a decreasing function as $\cos(x)$ is decreasing on $[0, \pi]$. Due to Chebyshev's Sum Inequality, it suffices to prove $\sum_{i=1}^n \alpha^{(i)} w(d_i) r_i = \sum_{i=1}^n \alpha^{(i)} w(d_i) r(d_i) \geq 0$ for some increasing function $w : [0, \pi] \to \mathbb{R}$ to show Equation (1). We will prove that this holds for the function $w(x) = x/\sin(x)$ (defining $w(0) = 1$) which is increasing on $[0, \pi]$. We write $w_i := w(d_i)$ as a shorthand.

It holds that

$$\sum_{i=1}^n \alpha^{(i)} w_i \theta^{(i)} = \left(\sum_{i=1}^n \alpha^{(i)} w_i \cos(d_i)\right) \theta_{\text{ang}} + \sum_{i=1}^n \alpha^{(i)} \sin(d_i) w_i v_i$$
$$= \left(\sum_{i=1}^n \alpha^{(i)} w_i \cos(d_i)\right) \theta_{\text{ang}} + \sum_{i=1}^n \alpha^{(i)} d_i v_i = \theta_{\text{ang}} \sum_{i=1}^n \alpha^{(i)} w_i \cos(d_i).$$

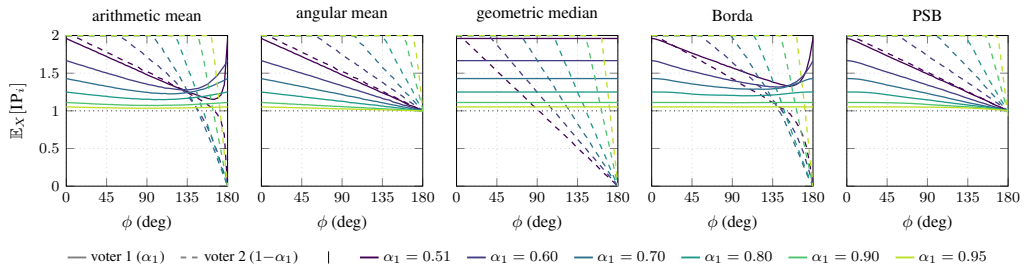


Figure 7: Each voter type's expected individual proportionality level $\mathbb{E}_X[\mathsf{IP}_i]$ as a function of the angle ϕ between $\theta^{(1)}$ and $\theta^{(2)}$, on the two-voter profile for different weight splits.

Taking the scalar product with u on both sides and using that $d_{\measuredangle}(\theta^{(i)}, u) < R$, so $\langle \theta^{(i)}, u\rangle > \cos(R)$, as well as $\langle \theta_{\text{ang}}, u\rangle \leq 1$ we get

$$\sum_{i=1}^{n} \alpha^{(i)} w_i \cos(R) < \sum_{i=1}^{n} \alpha^{(i)} w_i \langle \theta^{(i)}, u\rangle = \langle \theta_{\text{ang}}, u\rangle \sum_{i=1}^{n} \alpha^{(i)} w_i \cos(d_i) \leq \sum_{i=1}^{n} \alpha^{(i)} w_i \cos(d_i),$$

and henceforth

$$\sum_{i=1}^{n} \alpha^{(i)} w_i < \alpha^{(i)} w_i \cos(d_i) \iff 0 < \sum_{i=1}^{n} \alpha^{(i)} w_i (\cos(d_i) - \cos(R)) = \sum_{i=1}^{n} \alpha^{(i)} w_i r_i.$$

□

E.2 Warm-up: Expected Proportionality in the Two-Voter Case

We begin in the simple setting that originally exposed the disproportionality of the arithmetic mean: $n = 2$ voter types in $d = 2$ dimensions, with a majority group and a minority group [15]. We let voter type 1 be the majority and sweep its weight over $\alpha_1 \in \{0.51, 0.60, 0.70, 0.80, 0.90, 0.95\}$; we let $\phi := d_{\measuredangle}(\theta^{(1)}, \theta^{(2)}) \in [0°, 180°]$ denote the angle between the two scoring vectors. Together (α_1, ϕ) parametrize every instance of this 2-voter, 2-dimensional problem up to rotation. Figure 7 shows the expected individual proportionality level $\mathbb{E}_X[\mathsf{IP}_i]$ for the majority (solid) and minority (dashed) voter type under each of the five rules. We observe that for $\phi \leq 90°$, every rule serves both voter types above their entitlement and the choice of rule has little impact; an observation in line with our approximation results from Proposition E.1. As $\phi \to 180°$, the picture changes sharply: the arithmetic mean, Borda, and the geometric median all move towards the majority voter type's preferred ranking, and the minority's individual proportionality level falls toward 0. In contrast, as guaranteed, the angular mean and PSB provide a individual proportionality level of at least 1 for both voter types across the entire range of ϕ.

Note that Figure 7 reports voter type's individual proportionality level $\mathsf{IP}_i = a_{KT}/\alpha^{(i)}\binom{m}{2}$, which counts every pairwise comparison the rule agrees with voter type i on. This includes pairs on which the two voter types already agree with each other, a free "win" for any reasonable aggregation method that inflates individual proportionality levels. We now define a disagreement-restricted variant $\widetilde{\mathsf{IP}}_i$ that ignores these free wins and credits the rule only for its agreement on *contested* pairs. For a batch $X = (x_1, \ldots, x_m) \in \mathcal{X}^m$, a pair of items $\{x_j, x_k\} \subseteq X$ is *contested* iff the two voter types rank it differently: $(x_j \succ^X_{\theta^{(1)}} x_k) \Leftrightarrow (x_k \succ^X_{\theta^{(2)}} x_j)$. Let $s_i(\theta_f, X)$ be the number of contested pairs on which θ_f's ranking sides with voter type i. Voter type i's disagreement-restricted individual proportionality level $\widetilde{\mathsf{IP}}_i$ is the ratio

$$\widetilde{\mathsf{IP}}_i(\theta_f, X) \;=\; \frac{s_i(\theta_f, X)}{\alpha^{(i)} \cdot \big|\{\{x_j, x_k\} \subseteq X : \{x_j, x_k\} \text{ is contested}\}\big|}.$$

The denominator is voter type i's proportional share of the contested pairs in this batch. The numerator is the number of contested pairs the rule actually awards to voter type i. The expected value of $\widetilde{\mathsf{IP}}_i$ is plotted in Figure 8.

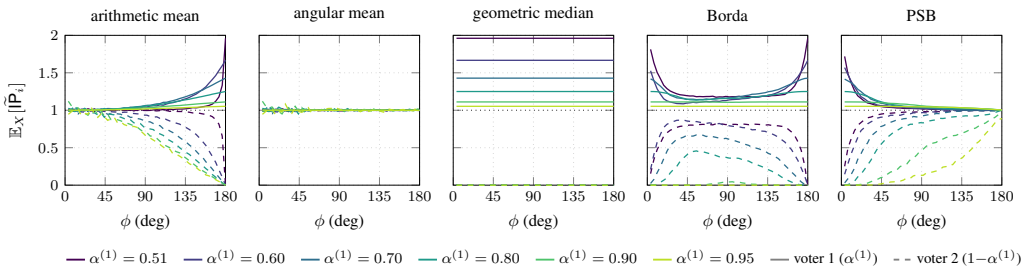


Figure 8: Experiment on two voter types. Setup as Figure 7, but with an adapted definition of the individual proportionality level: each voter type demands an α fraction of the pairs on which the two voter types disagree.

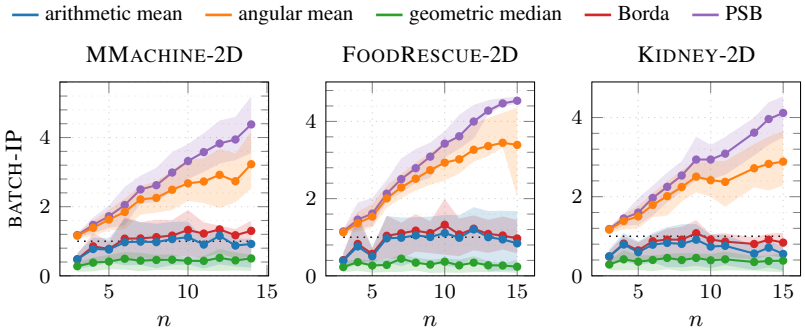


Figure 9: BATCH-IP on heterogeneity-filtered voter subsamples (std of pairwise angular distances $\geq 65°$), as a function of subsample size n, on the 2D variants of the three main-body datasets. Counterpart of Figure 4. Shaded bands show the inter-quartile range across 100 subsamples per n.

E.3 Additional Information on Datasets

E.3.1 Dataset Description

Across all five datasets, each voter's scoring vector is learned from binary pairwise comparisons using the same procedure. Given a voter i who answered T_i comparisons between alternatives $a^{(t)}$ and $b^{(t)}$, we fit a per-voter logistic regression $P\big(\text{voter } i \text{ chooses } a^{(t)}\big) = \sigma\big(\beta_i^\top (x_{a^{(t)}} - x_{b^{(t)}})\big)$, where $x_a \in \mathbb{R}^d$ is the feature vector of alternative a, and σ is the logistic function. This corresponds to a Thurstone–Mosteller random utility model with linear utilities $u_i(a) = \beta_i^\top x_a$. Feature differences are standardized per voter (zero mean, unit variance) before fitting, and we use ℓ_2 regularization with default strength $C = 1$ (scikit-learn's `LogisticRegression`). The learned coefficient vector $\beta_i \in \mathbb{R}^d$ is the voter's scoring vector.

The five datasets differ in the alternatives that are being compared, and the feature set used.

MMACHINE (Moral Machine) [2, 36]. $n = 137$ voters, $d = 24$ features. Users compared autonomous-vehicle moral dilemmas in which a self-driving car must choose which of two groups of pedestrians (or passengers) to prioritize when an accident is unavoidable. Each alternative is encoded as a vector of demographic and contextual attributes of the corresponding group. Following Kim et al. [25], we fit one scoring vector per country by pooling all responses from that country; we only include countries for which at least 1000 responses have been recorded. The 24 features are: 4 scenario-level attributes (Intervention, Barrier, CrossingSignal, NumberOfCharacters) and 20 pedestrian-group composition attributes (Man, Woman, Pregnant, Stroller, OldMan, OldWoman, Boy, Girl, Homeless, LargeWoman, LargeMan, Criminal, MaleExecutive, FemaleExecutive, FemaleAthlete, MaleAthlete, FemaleDoctor, MaleDoctor, Dog, Cat). The 2D variant MMACHINE-2D uses the feature pair LargeWoman and MaleExecutive.

KIDNEY (Kidney) [24]. $n = 404$ voters, $d = 8$ features. Voters compared hypothetical kidney-allocation candidates described by clinical and behavioral attributes, choosing which of two patients

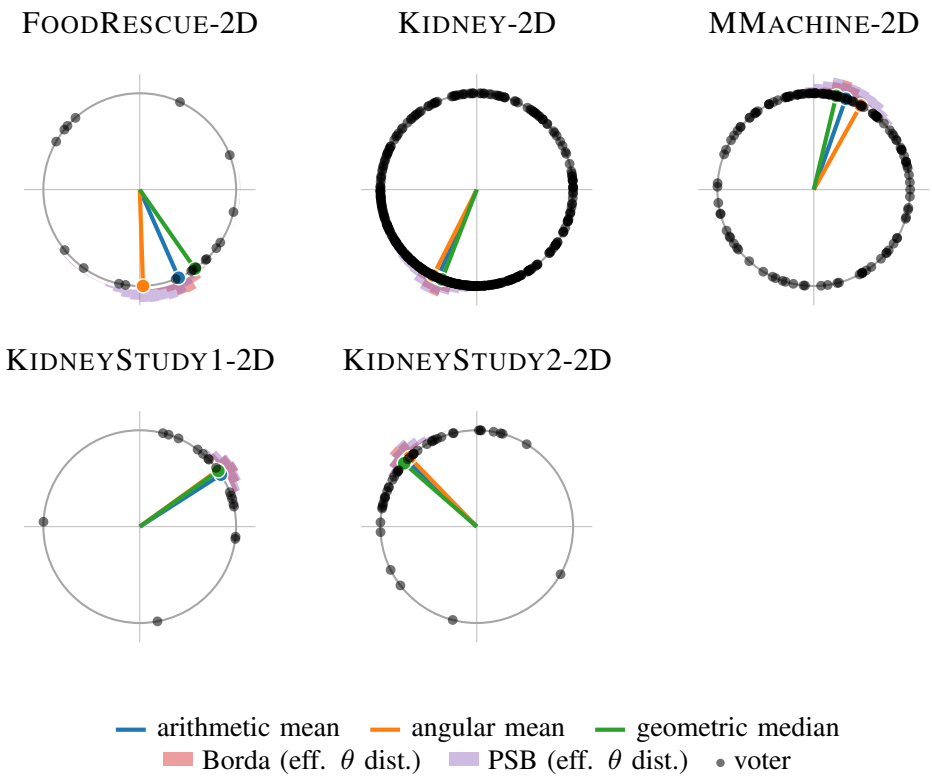


Figure 10: Voter and mechanism positions on S^1 for the 2D variants of all five datasets. Black dots: voters. Radial lines: arithmetic mean (blue), angular mean (orange), geometric median (green). Shades outside the unit circle: histograms of the effective direction $\theta_{\text{eff}}(X)$ recovered from the rankings produced by Borda (red) and PSB (purple) over $R = 2000$ batches per dataset.

should receive an available organ. The 8 features are: dep (number of dependents), life (life expectancy with the transplant), years_waiting (years on the waiting list), obesity (BMI status), alco (prior alcoholism), work_hours (weekly work hours), crim (prior criminal record), and reject_chance (estimated probability of organ rejection). The 2D variant KIDNEY-2D uses the feature pair alco and crim.

FOODRESCUE (412 Food Rescue) [31]. $n = 19$ voters, $d = 7$ features. Voters compared candidate recipient organizations for food donations, deciding which of two organizations should be prioritized for an incoming rescue. The 7 features are: size (organization size), access (alternative-food access score), income (median income of the served population), poverty (poverty rate of the served population), last_donation (time since last received donation), total (total donations received historically), and dist (delivery distance). The 2D variant FOODRESCUE-2D uses the feature pair last_donation and total.

KIDNEYSTUDY1 [3]. $n = 17$ voters, $d = 4$ features. An earlier pilot study on kidney-allocation preferences with a smaller voter population and a narrower feature set than the main KIDNEY dataset. The 4 features are: alco, dep, life, and crim (definitions as in KIDNEY). The 2D variant KIDNEYSTUDY1-2D uses the feature pair dep and life.

KIDNEYSTUDY2 [3]. $n = 43$ voters, $d = 5$ features. A follow-up kidney-allocation preference study with a partially overlapping but distinct feature set. The 5 features are: elderlyDep (number of elderly dependents), lifeYearsGained (expected life-years gained from the transplant), obesity, weeklyWorkhours, and yearsWaiting. The 2D variant KIDNEYSTUDY2-2D uses the feature pair obesity and weeklyWorkhours.

Table 2: Heterogeneity of voter types' scoring vectors and pairwise angular distances between aggregation rules. All values are in degrees.

	MMACHINE		FOODRESCUE		KIDNEY		KIDNEYSTUDY1		KIDNEYSTUDY2	
	Orig.	2D	Orig.	2D	Orig.	2D	Orig.	2D	Orig.	2D
Heterogeneity of $\theta^{(i)}$										
max pairwise angle between $\theta^{(i)}$	53.5	180.0	109.6	178.8	147.3	180.0	85.26	176.91	88.00	179.86
avg of pairwise angular distances between $\theta^{(i)}$	20.7	80.4	61.7	85.7	57.1	77.5	48.72	53.29	39.87	49.58
stdev of pairwise angular distances between $\theta^{(i)}$	8.3	52.7	17.9	56.7	20.6	52.5	20.07	46.13	17.73	40.51
Rule coincidence										
$d_{\angle}(\theta_{\text{ang}}, \theta_{\text{arith}})$	0.03	9.70	1.10	21.86	1.18	2.85	0.33	3.23	0.66	2.30
$d_{\angle}(\theta_{\text{ang}}, \theta_{\text{med}})$	0.93	15.57	4.23	32.93	5.69	5.62	3.38	0.55	6.41	4.73
$d_{\angle}(\theta_{\text{arith}}, \theta_{\text{med}})$	0.91	5.87	3.16	11.07	4.54	2.76	3.06	2.68	5.77	2.43

Dataset Properties Table 2 depicts the pairwise similarity between scoring vectors and the similarity between $\theta_{\text{ang}}, \theta_{\text{arith}}, \theta_{\text{med}}$ on our dataset.

Figure 10 visualizes the voter and mechanism positions on S^1 for the 2D variants of all five datasets. Visualizing Borda and PSB on S^1 requires care. Both rules produce an arbitrary ranking per batch, not a single direction, and the ranking they produce on a given batch need not be induced by any scoring vector. To place them on the same circle as the fixed linear rules, we recover for each batch X an *effective direction* $\theta_{\text{eff}}(X) \in S^1$, defined as

$$\theta_{\text{eff}}(X) \;\in\; \operatorname*{arg\,min}_{\theta \in S^1} \; d_{KT}\big(\succ_{\theta}^{X},\; \succ_{f}^{X}\big),$$

where $\succ_f^X$ is the ranking produced by the rule $f \in \{\text{Borda}, \text{PSB}\}$ on X and d_{KT} is Kendall-tau distance.

E.3.2 Results on Kidney Study 1 and 2

We replicate the main-body results (Table 1, Figure 4, Figure 3) on the 2D variants of KIDNEYSTUDY1 and KIDNEYSTUDY2. Table 3 is the analogue of Table 1, Figure 11 the analogue of Figure 3, and Figure 12 the analogue of Figure 4.

Table 3: Long-run IP (LONG-IP) and per-batch IP (BATCH-IP) for each rule on KIDNEYSTUDY1 and KIDNEYSTUDY2 and its 2D variant. Cell shading is proportional to $\log_{10}(\text{value})$: darker means more proportional.

	arithmetic		angular		geometric		Borda		PSB	
dataset	long	batch	long	batch	long	batch	long	batch	long	batch
KIDNEYSTUDY1	12	11	12	11	12	11	12	11	12	11
KIDNEYSTUDY2	29	26	29	26	27	25	28	26	29	26
KIDNEYSTUDY1-2D	3.4	3.3	3.7	3.6	3.6	3.5	3.5	3.4	3.8	3.8
KIDNEYSTUDY2-2D	3.3	3.3	3.9	3.8	2.8	2.8	3.2	3.2	4.3	4.3

E.3.3 Results on Full Datasets

Table 4: Long-run IP (LONG-IP) and per-batch IP (BATCH-IP) for each rule on each dataset. Cell shading is proportional to $\log_{10}(\text{value})$: darker means more proportional.

	long					batch				
dataset	arith.	ang.	geom.	Borda	PSB	arith.	ang.	geom.	Borda	PSB
FOODRESCUE	11	11	11	11	11	9.7	9.8	9.6	9.9	10
KIDNEY	134	135	132	133	135	112	112	110	112	115
MMACHINE	111	111	111	111	111	98	98	98	98	98

Table 4 shows that on the original (high-dimensional) feature sets, all five rules attain comparable IP. Figure 13 confirms this in the same convergence-in-m plot used in Figure 3. We omit a figure

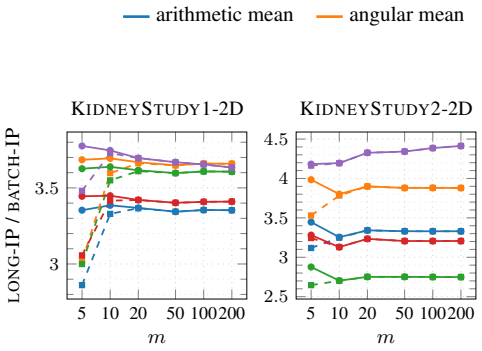


Figure 11: LONG-IP (solid) and BATCH-IP (dashed) on the full electorates, as a function of batch size m. Counterpart of Figure 3 for the 2D variants of the two kidney study datasets.

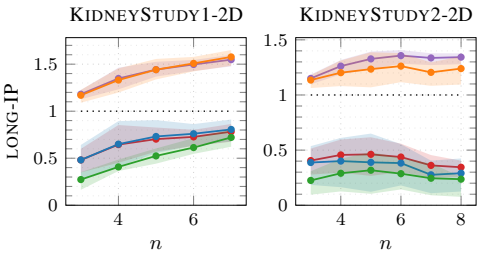


Figure 12: Counterpart of Figure 4 for the two kidney study datasets: LONG-IP on heterogeneity-filtered voter subsamples (std of pairwise angular distances $\geq 65°$), as a function of subsample size n, with shaded bands showing the inter-quartile range across 100 subsamples per n. The maximum n is smaller than in Figure 4 because the kidney-study electorates are less diverse: at large n, no subsample of voters spans a wide enough arc on S^1 to satisfy the filter.

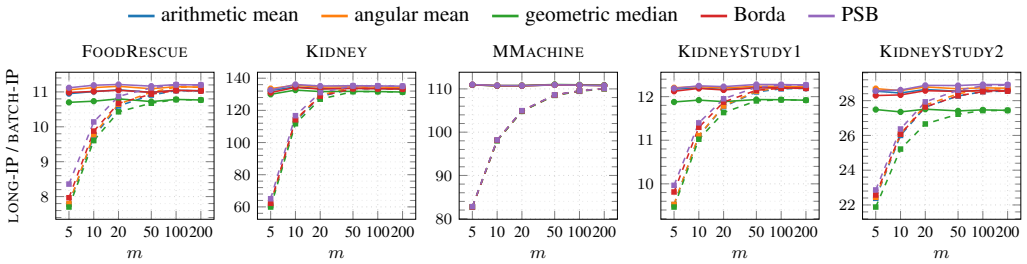


Figure 13: LONG-IP (solid) and BATCH-IP (dashed) on the full feature sets of all five datasets, as a function of batch size m. As in Figure 3, the per-batch gap shrinks rapidly as m grows. The five rules are nearly indistinguishable: the full-feature electorates are not heterogeneous enough to separate them, consistent with the homogeneity statistics in Table 2.

parallel to Figure 4 concerning heterogeneity-filtered subsamples: at the original feature dimensions, voters' scoring vectors are too clustered for the $65°$ pairwise-distance filter to admit any subsample, so the filtered analysis cannot be run on the original feature sets.

E.4 Ablations

Our empirical results presented in the main body use uniform voter weights $\alpha^{(i)} = 1/n$ and a uniform distribution on $\mathbb{S}^{d-1}$. We now proceed to relax each in turn.

E.4.1 Non-Uniform Weights

In Section 5 we used uniform voter weights. Here we draw weights from a $\mathrm{Dirichlet}(1)$ distribution—that is, uniformly at random from the simplex of weight vectors—once per dataset. Table 5 reports LONG-IP and BATCH-IP for each rule on each of the five datasets and their 2D variants under random weights, in the same format as Table 1 in the main body.

Figure 14 repeats the convergence-in-m analysis of Figure 3 on the full feature sets of all five datasets under random weights.

The qualitative conclusions of Section 5 carry over. In Table 5, the relative ordering of rules within each dataset and the relative ordering of datasets stay largely unchanged from Table 1, the main exception being MMACHINE-2D, where the arithmetic mean now attains a higher LONG-IP than the angular mean. On the full feature sets, all rules satisfy IP comfortably and are nearly indistinguishable;

Table 5: Long-run IP (LONG-IP) and per-batch IP (BATCH-IP) for each rule on each dataset and its 2D variant. Cell shading is proportional to $\log_{10}$(value): darker means more proportional. Voter weights: Dirichlet(1) weights.

	arithmetic		angular		geometric		Borda		PSB	
dataset	long	batch	long	batch	long	batch	long	batch	long	batch
FOODRESCUE	3.0	3.0	3.0	3.0	2.8	2.8	3.0	3.0	3.0	3.0
KIDNEY	40	40	40	40	40	40	40	40	40	40
MMACHINE	22	22	22	22	22	22	22	22	22	22
KIDNEYSTUDY1	3.5	3.5	3.5	3.5	3.4	3.4	3.5	3.5	3.5	3.5
KIDNEYSTUDY2	7.1	7.1	7.1	7.1	7.3	7.3	7.2	7.2	7.2	7.2
FOODRESCUE-2D	3.3	2.9	3.9	3.7	2.6	2.2	3.3	2.9	3.8	3.8
KIDNEY-2D	0.5	0.02	3.4	0.8	4.1	1.4	4.9	2.8	14	14
MMACHINE-2D	2.4	0.7	1.5	1.1	0.3	0.2	2.6	1.9	10	9.6
KIDNEYSTUDY1-2D	3.8	3.7	3.8	3.8	3.8	3.8	3.8	3.7	3.8	3.8
KIDNEYSTUDY2-2D	6.8	6.2	6.7	6.3	6.8	6.2	6.8	6.5	6.8	6.6

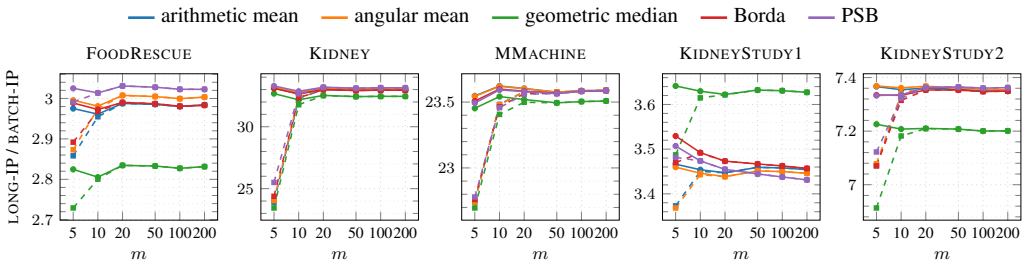


Figure 14: Random-weights variant of Figure 3. LONG-IP (solid) and BATCH-IP (dashed) on the full feature sets of all five datasets, as a function of batch size m, with voter weights drawn once per dataset from a Dirichlet(1) distribution. The qualitative picture is unchanged from the uniform-weights case: the per-batch gap closes rapidly as m grows and the five rules remain nearly indistinguishable on the full-feature electorates.

on the 2D variants, the arithmetic mean falls below the entitlement on KIDNEY-2D and the geometric median on MMACHINE-2D, while the angular mean and PSB satisfy it. Absolute IP values do shift – because the entitlement denominator $\alpha^{(i)}\binom{m}{2}$ changes with the weight draw. In Figure 14, the per-batch gap LONG-IP $-$ BATCH-IP closes at the same rate as in Figure 13.

E.4.2 Non-Uniform Item Distribution

The proportionality guarantee for the angular mean (Theorem 3.1) holds under spherically symmetric batch distributions $\mathcal{D}_\circ$. In this section we probe what happens when this assumption is dropped, by drawing items from an *angular central Gaussian* (ACG) distribution [47] on S^{d-1} parameterized by a concentration axis $v \in S^{d-1}$ and an anisotropy parameter $\lambda \in (0, 1]$:

$$z \sim \mathcal{N}\big(0,\ \lambda \cdot I + (1-\lambda)\, vv^\top\big), \qquad x = z/\|z\|.$$

The covariance has eigenvalue 1 along v and λ along every direction orthogonal to v; $\lambda = 1$ recovers the isotropic Gaussian (and hence the uniform distribution on S^{d-1}), while $\lambda \to 0$ collapses items onto $\{+v, -v\}$. We probe two settings: a synthetic two-voter setup and the real-world 2D electorates.

Synthetic two-voter ACG sweep. Figure 15 extends Figure 7 to non-spherical sampling. We fix $\alpha^{(1)} = 0.7$ and sweep five angles $\phi \in \{45°, 90°, 135°, 150°, 175°\}$ between the two voters' scoring vectors. Each panel column corresponds to one of five concentration axes v placed at geodesic positions $t \in \{0, 0.25, 0.5, 0.75, 1\}$ between $\theta^{(1)}$ and $\theta^{(2)}$, and each row to one of the five aggregation rules. Within each panel, IP_i for voter type 1 (solid) and voter type 2 (dashed) is plotted as a function of λ.

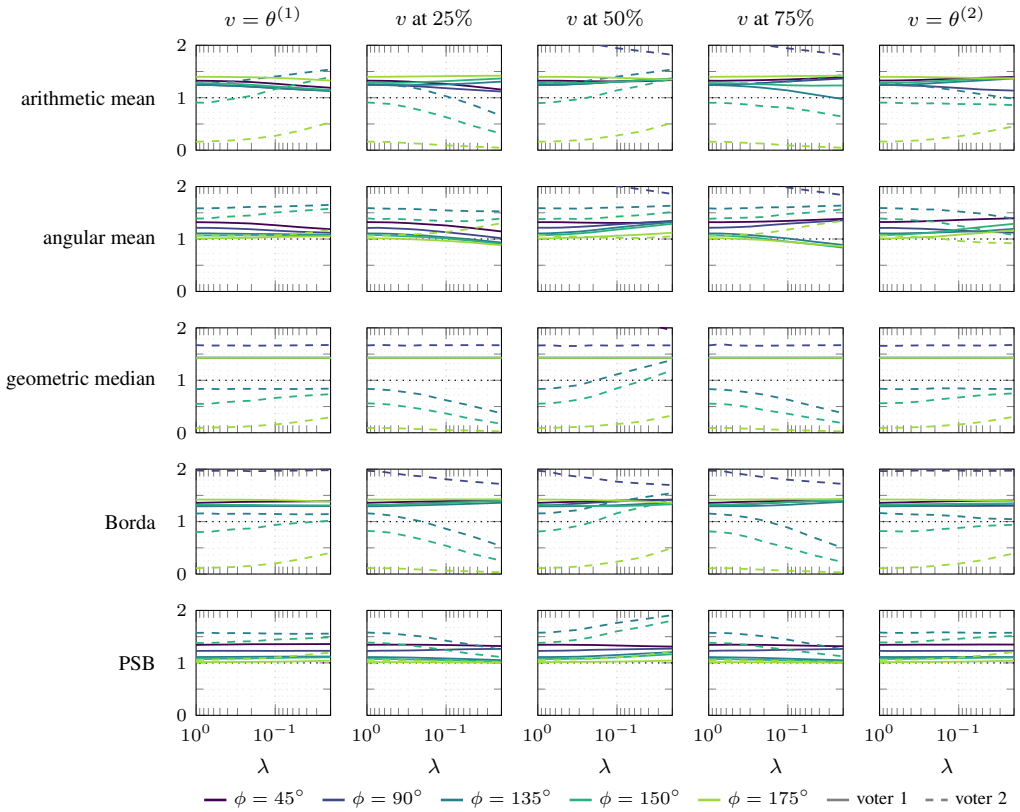


Figure 15: Per-voter IP_i on the two-voter profile with items drawn from an angular central Gaussian (ACG) batch distribution. Anisotropy $\lambda \in [0.02, 1]$ on the x-axis (log scale, decreasing left-to-right); concentration axis v placed at geodesic positions $t \in \{0, 0.25, 0.5, 0.75, 1\}$ between $\theta^{(1)}$ and $\theta^{(2)}$ (columns); rows are mechanisms. Solid lines: voter 1 ($\alpha^{(1)} = 0.7$); dashed: voter 2. Colors: $\phi \in \{45°, 90°, 135°, 150°, 175°\}$. $\lambda = 1$ recovers spherical symmetry; $\lambda \to 0$ collapses items onto $\{+v, -v\}$.

Real-world ACG sweep. We replicate the same sweep on the 2D variants of all five real-world datasets. Because each electorate now contains many voters in specific positions on S^1 (rather than a stylized two-voter placement), the natural reference axes for v are no longer the voters themselves. We use four fixed compass directions $v \in \{0°, 45°, 90°, 135°\}$ on S^1, which cover the half-circle without redundancy under the $v \sim -v$ symmetry of the ACG. We show the results in Figure 16.

We take this as additional empirical evidence that the angular mean's proportionality is robust to substantial departures from spherical symmetry on the real-world electorates studied in Section 5.

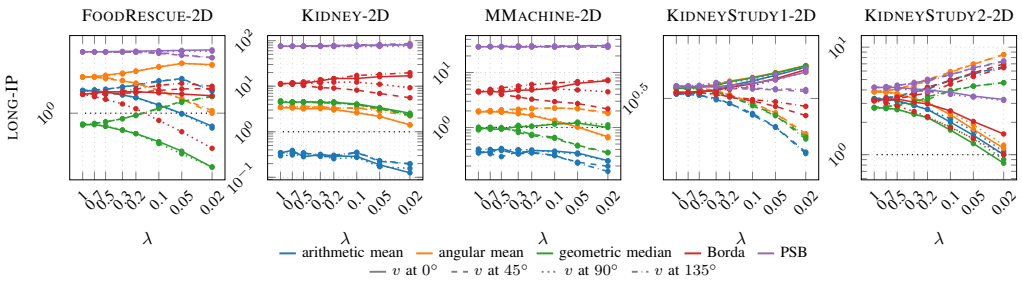


Figure 16: Real-world counterpart of Figure 15. Long-run individual proportionality LONG-IP on the 2D variants of all five real-world datasets, as a function of the ACG anisotropy parameter λ (log scale, decreasing left-to-right) and four concentration axes $v \in \{0°, 45°, 90°, 135°\}$ on S^1. Mechanism = color; axis = linestyle. Voter weights are uniform $\alpha^{(i)} = 1/n$. At $\lambda = 1$ axis direction has no effect and values match the corresponding rows of Table 1. The horizontal dotted line marks the proportional entitlement LONG-IP $= 1$. Note the log y-axis: PSB lives one to two orders of magnitude above the fixed linear rules on KIDNEY-2D and MMACHINE-2D.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction do accurately and honestly reflect both the contributions and scope of the work.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss the limitations of both the theoretical work and the experiments at relevant points in the work. From the theoretical perspective: we highlight distributional assumptions in the model (Section 2), we discuss computational feasibility of the angular mean in the model (Remark 2.1), and we highlight that we are learning an aggregate scoring vector in a very information-rich setting (as opposed to other commonly studied low-information settings) in Section 6. We give extensive information about our real-world and synthetic datasets in the empirical section, reflecting on how homogeneity in some of them impacts our results.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best

judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: All theorems and mathematical claims (lemmas, propositions, and corollaries) are formally stated and rigorously shown (either in the main body or at the referenced location in the appendix).

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: In Appendix E.2, Appendix E.3, and Appendix E.4, extensive detail is given on the methods for producing the experimental results in the paper.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.

(c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: The supplementary material contains a code appendix that contains a script that reproduces all experimental results from the paper.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: The paper does not contain any training and instead focuses on the performance of various mechanisms. Tests are described in full detail.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

Justification: Figure 4 is the only central experiment that exhibits a non-trivial variance. We report inter-quartile ranges.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Yes, at the beginning of Appendix E.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: Yes, the research conforms to the NeurIPS code of ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: This paper is, in fact, heavily motivated by societal impact, and is very explicit in the proportionality ideal that we optimize for and its implications.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: This paper does not pose such risks.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: We use five existing preference datasets and cite the respective papers in Section 5 and Appendix E.3.1. We do not redistribute any of these datasets, but only include the anonymized scoring vectors derived from them in our code appendix. For the

moral machine dataset, data were obtained from the original publication's data appendix available at `https://osf.io/3hvt2/overview`. For the three datasets related to kidney exchange, we obtained access through the author's GitHub repo at `https://github.com/vijaykeswani/Preference-Instability/tree/main`. The FoodRescue dataset was obtained from the authors of the original study and used in accordance with their requested terms.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [Yes]

Justification: The only new asset is our code, which is appropriately documented.

Guidelines:

- The answer [N/A] means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [N/A]

Justification: No crowdsourcing nor experiments with human subjects were conducted for this paper.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.