

Smoothed Analysis of Social Choice, Revisited

Bailey Flanigan
Carnegie Mellon University
bflaniga@andrew.cmu.edu

Daniel Halpern
Harvard University
dhalpern@g.harvard.edu

Alexandros Psomas
Purdue University
apsomas@cs.purdue.edu

Abstract

A canonical problem in social choice is how to aggregate ranked votes: that is, given n voters’ rankings over m candidates, what *voting rule* f should we use to aggregate these votes and select a single winner? One standard method for comparing voting rules is by their satisfaction of *axioms* — properties that we want a “reasonable” rule to satisfy. This approach, unfortunately, leads to several impossibilities: no voting rule can simultaneously satisfy all the properties we would want, at least in the worst case over all possible inputs.

Motivated by this, we consider a relaxation of this worst case requirement. We analyze this through a “smoothed” model of social choice, where votes are (independently) perturbed with small amounts of noise. If no matter which input profile we start with, the probability (post-noise) of an axiom being satisfied is large, we will view it as nearly as good as satisfied — called “smoothed-satisfied” — even if it may be violated in the worst case.

Our model is a mild restriction of Lirong Xia’s, and corresponds closely to that in Spielman and Teng’s original work on smoothed analysis. Much work has been done so far in several papers by Xia on which axioms can and cannot be satisfied under such noise. In our paper, we aim to give a more cohesive overview on when smoothed analysis of social choice is useful.

Within our model, we give simple sufficient conditions for smoothed-satisfaction or smoothed-violation of several axioms and paradoxes, including most of those studied by Xia as well as some previously unstudied. We then observe that, in a practically important subclass of noise models, although convergence eventually occurs, known rates may require an extremely large number of voters. Motivated by this, we prove bounds specifically within a canonical noise model from this subclass — the Mallows model. Here, we find a more nuanced picture on exactly when smoothed analysis can help.

1 Introduction

One of the most canonical problems in social choice is how to aggregate votes: that is, given a *preference profile* consisting of n voters’ rankings over m alternatives, what *voting rule* f should we use to aggregate this preference profile into a single winner? A predominant way of comparing voting rules is by their satisfaction of *axioms* — logical statements that describe basic, natural properties that voting rules should satisfy.

Traditionally, it is said that a voting rule *satisfies* an axiom if the logical statement specified by the axiom holds on *all possible* preference profiles; if there exists even one profile on which it does not hold (a *counterexample*), the rule is said to *violate* the axiom. Unfortunately, under this worst-case notion of satisfaction, *all* voting rules violate at least some common-sense axioms.¹ There is hope,

¹E.g., the Gibbard-Satterthwaite theorem says that any onto, single-winner, non-dictatorial voting rule with $m > 2$ is not strategyproof [12]. For more examples, see the [Comparison of electoral systems](#) Wikipedia page.

however, because this worst-case approach may cover up an important distinction between voting rules: while some rules may fail to satisfy an axiom on large swaths of preference profiles, others may fail only on precisely contrived instances. This distinction is of practical significance because voting rules of the latter type should usually satisfy the axiom in practice, where the lack of perfect correlation between people’s preferences makes extremely contrived counterexamples unlikely to occur.

In 2020, a paper by Lirong Xia aimed to capture this intuition by modeling profiles as *semi-random*—mostly adversarial, but with random perturbations [24]. More precisely, Xia’s model assumes that “noisy” (but otherwise adversarial) profiles arise in the following way: first, let an adversary choose voters’ “types,” with each type corresponding to a distribution over rankings. The ultimate profile is then sampled by selecting each voter’s ranking independently from their type distribution.

In this paper, we capture the same intuition using a slightly restricted version of Xia’s semi-random model. Our model produces “noisy” profiles in the following way: first, let an adversary fix a *starting profile*. Then, apply a small amount of noise independently to each individual ranking in the profile—that is, for each ranking, draw a new ranking from some distribution (this distribution can be arbitrary, up to regularity conditions). We will refer to this as the *smoothed* model, as it is directly analogous to Spielman and Teng’s “smoothed” model in the distinct context of linear programming. We discuss this connection and the precise relationship between our model and Xia’s in Section 1.2. The key takeaway is that ours is a slight restriction, assumed primarily for ease of exposition.

Because the smoothed model goes so minimally beyond the worst case, resolving axioms in this way is quite meaningful: the small amount of noise it assumes should be sufficient to escape counterexamples only if they are truly isolated among other profiles, i.e., contrived. The model is also well-motivated practically, as people’s preferences are established to be susceptible to small shocks, in daily life and even in the preference elicitation process [7, 14, 15]. The existing work in this area gives some hope of such positive results: it shows that under all standard voting rules considered, the axioms Group-strategyproofness [28], Resolvability [26], and Participation [25] are satisfied post-noise with high probability as n grows large. In contrast, the axiom Condorcet consistency remains violated with high probability, even after semi-random noise, by a popular class of rules [25]. Here we see heterogeneity across axioms: for some, smoothed noise reliably circumvents impossibilities, whereas, for others, it seems insufficient. This motivates our first question:

Question 1: *What properties of axiomatic impossibilities make them resolvable by smoothed noise?*

This question remains to be fully addressed, in part because existing work has mostly analyzed axioms one by one, using technically intricate arguments to get precise asymptotic convergence rates in n . While this approach has yielded detailed insights about specific axioms and voting rules, it remains to distill higher-level patterns across rules and axioms.

Our second research question is a refinement of the first and deals with understanding when semi-random noise can circumvent impossibilities, *absent a key assumption made in the related work so far*. This assumption is called *positivity* in the related work, and it amounts to assuming that when a ranking is perturbed, there is at least some fixed probability, MIN-PROB, of *any other ranking* resulting from this perturbation. Although it is not explicitly written, existing upper bounds on rates of convergence for the axioms we consider implicitly depend multiplicatively on $1/\text{MIN-PROB}^{3/2}$. Although when MIN-PROB is large, this polynomial dependence is not too de-

manding, when MIN-PROB is small, enormous numbers of voters may be necessary to guarantee a reasonable probability of axiom satisfaction.

Although assuming that MIN-PROB is large enough to avoid such issues may seem innocuous, it is anything but. This is because noise models in which MIN-PROB is very small represent perhaps the most realistic subclass of noise models: they encompass any noise model in which small perturbations are unlikely to *completely reverse* someone’s ranking (or make similarly drastic changes). Given that small real-world shocks that may perturb people’s preferences are unlikely to qualitatively change their opinions (to, e.g., the extent that their rankings are reversed), a more realistic noise model might be one in which “less” noise corresponds to lower probability of drawing a perturbed ranking that reflects extreme opinion change. This motivates our second research question:

Question 2: *What properties of axiomatic impossibilities make them resolvable by smoothed noise, for noise models in which rankings reflecting extreme opinion change are drawn with low (or zero) probability?*

1.1 Results and Contributions

Question 1. In Section 3, we study the general smoothed model. We distill two patterns: one across axioms where smoothed noise *does* help, and another across axioms where it does not. In the process, we do the first smoothed (or semi-random) analysis of several core axioms in social choice. We also re-analyze some axioms studied in past work, but we prove these results anew, with derivations that are standardized across axioms and yield bounds that explicitly depend on MIN-PROB, which will be useful in Question 2. We delineate which analyses are new and which are re-proven in Section 1.2.

As summarized in Table 1 (located in Section 3.1), we first show negative results for the axioms CONDORCET CONSISTENCY, INDEPENDENCE OF IRRELEVANT ALTERNATIVES, CONSISTENCY, MAJORITY. Across all voting rules we study, we find that these axioms are all smoothed-violated whenever they are violated, i.e., smoothed noise is insufficient to circumvent impossibilities for these axioms. Moreover, across axioms and rules, this smoothed violation always holds for the same reason: there exist large contiguous segments of the profile space consisting of counterexamples. As shown in the left-hand pane of Figure 1, if such regions exist, then our starting profile can be chosen to be inside one, in which case small perturbations are insufficient to produce profiles outside it. We formalize this underlying pattern into a generic sufficient condition for smoothed-violation in Lemma 7.

As summarized in Table 2 (located in Section 3.2), we next show positive results for the axioms RESOLVABILITY and group notions of STRATEGYPROOFNESS, PARTICIPATION, and MONOTONICITY. In the spirit of generalizing across axioms, we are able to analyze the latter three at once by proving the smoothed-satisfaction of a more general axiom, called $\rho(n)$ -GROUP-STABILITY, which requires that the behavior of $\rho(n)$ voters should not be able to affect the outcome of the election

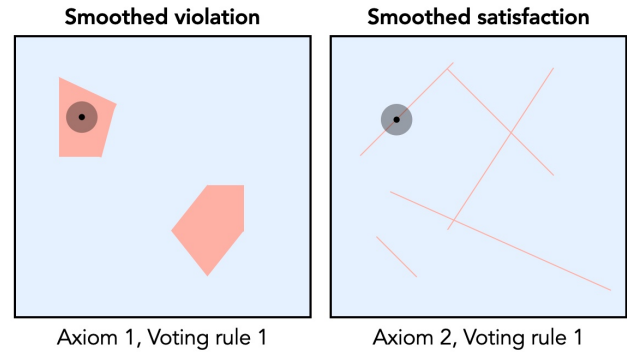


Figure 1: Both panes represent (abstractly) the space of all possible profiles. Red profiles are counterexamples in which the voting rule fails to satisfy the axiom. The black dot is the starting profile, and surrounding region is where the resulting profile is likely to end up after applying noise.

(essentially framing *Margin of Victory* as a binary property [22, 28]). The common feature of these axioms, which permits their smoothed-satisfaction, is that their counterexamples must occur on or near hyperplanes. As depicted in the right-hand pane of Figure 1, this means that any “well-behaved” noise, applied to any starting profile, will place very little probability mass on the sliver of counterexamples contained within (potentially some nonzero but shrinking distance of) the measure-zero hyperplane. We formalize this underlying pattern into a generic sufficient condition for smoothed-satisfaction in Lemma 10.

Question 2. In Section 4, we pursue the first convergence rates to smoothed satisfaction that are not parameterized by MIN-PROB, and which do not rely on the minimum probability being nonzero. Although the core ideas we use naturally extend to broader classes of noise models, for concreteness, we pursue these bounds specifically in the *Mallows model*—perhaps the most canonical model of noisy rankings. This model has the realistic property we want: the probability of drawing a ranking decreases in its *Kendall tau distance*² from the original ranking. More precisely, for a noise parameter $\phi \in [0, 1]$, the probability of drawing a ranking at Kendall tau distance d from the original ranking is proportional to ϕ^d . As a result, MIN-PROB for this model is exponentially small: the probability of reversing a ranking due to Mallows noise is $\phi^{\Omega(m^2)}$.

The Mallows model falls within our smoothed model, so our impossibilities from Section 3.1 carry over. From among the remaining axioms of interest (those in Section 3.2), we characterize precise convergence rates to smoothed satisfaction of RESOLVABILITY and $\rho(n)$ -GROUP-STRATEGYPROOFNESS in the Mallows model.

While MIN-PROB-parameterized upper bounds yielded heterogeneity across axioms, they were almost homogeneous across voting rules per axiom, i.e., for each axiom, either nearly all rules satisfied, or nearly all rules violated them. Interestingly, these new bounds reveal diversity across voting rules as well. As summarized in Table 3, while PLURALITY and BORDA COUNT are smoothed-satisfied at a rate polynomial in ϕ and m , we can lower bound the convergence rates of MAXIMIN and VETO to include a $1/\phi^{\Omega(m)}$ term. This essentially means the number of voters required to guarantee satisfaction is exponential in these parameters, making this property less appealing. Across these voting rules, we can distill a pattern: voting rules whose outcomes are sensitive to changes throughout ranking positions are helped substantially by Mallows noise, while those whose outcomes remain fixed despite potentially many swaps are helped much less. Zooming out, this conceptual finding extends beyond the Mallows model and the studied axioms, suggesting that for reasonable noise models, only voting rules that are truly sensitive to small changes in voters’ preferences will have brittle impossibilities.

Bonus Contribution: smoothed analysis of Arrow’s theorem. As an extension, in Section 5 we do the first smoothed (or semi-random) analysis of *Arrow’s Theorem* [1]—perhaps one of the most influential impossibilities in social choice. Surprisingly, we show that this impossibility is resolved with high probability under the smoothed model. However, as we show, this resolution is for a trivial reason related to Arrow’s definition of the axiom *Non-dictatorship*. We therefore identify a slight strengthening of Arrow’s theorem, and pose the open question of whether smoothed noise is sufficient to resolve it.

²The Kendall tau distance between two rankings π, π' is the number of pairs of candidates on which they disagree.

1.2 Related work

Our model is designed to be a close analog to Spielman and Teng’s celebrated “smoothed” model, introduced in their analysis of the Simplex algorithm [23]. We give a detailed description of the parallels between our model and theirs in Appendix A.1. Within social choice, the model most closely related to ours is Lirong Xia’s semi-random model [24],³ of which our model is a mild restriction. To see how Xia’s model generalizes ours, in our model we always have $m!$ “types”, with each type corresponding to a possible starting ranking. Each type’s associated distribution is then the noise added to that ranking. The key restriction we make is that, while Xia’s model allows types with potentially different “shapes” of noise distributions, we assume that all rankings are perturbed by a common noise distribution, just centered at different rankings. This restriction is technically mild, as the assumption critical to analyses in both models is that noise is *independent* across rankings, and many of our results could be expressed in the more general model but in a less concise way. Because of that, we chose to stick to a slightly more restricted model for ease of exposition.

Beyond Xia’s model, axiom satisfaction has also been studied in other, less-adversarial randomized models. One popular such model is *Impartial Culture*, where profiles are drawn uniformly and i.i.d. [11, 13, 18]. Slightly more generally, the axiom Strategyproofness was studied in a *non-uniform* i.i.d model by Mossel et al. [19]. Further afield, there is empirical work on the frequency of axiom violations on simulated inputs and real elections [21, 10].

Now we situate our results among past work on axiom satisfaction in the semi-random model. First, our results in Section 4 are completely distinct from this existing line of work, pursuing bounds that do not rely on the positivity assumption core to the semi-random model. In contrast, our results in Section 3 are for generic noise distributions that *are* subject to the positivity assumption, along with all other regularity conditions imposed in Xia’s semi-random model. In the general smoothed model, we present the first smoothed (or semi-random) analysis of the axioms CONSISTENCY, INDEPENDENCE OF IRRELEVANT ALTERNATIVES, and MAJORITY, along with Arrow’s impossibility. We also introduce and analyze a new axiom, $\rho(n)$ -GROUP-STABILITY, which we show generalizes multiple standard axioms. In the interest of identifying patterns, we also re-analyze several axiomatic impossibilities for which there already exist similar bounds,⁴ though we show them via different proofs and give bounds that are explicit in their dependence on MIN-PROB. For clarity, we distinguish axioms and impossibilities analyzed for the first time in this paper in Table 1, Table 2, and Table 3 with *. The set of voting rules we study overlaps—but not perfectly—with the existing work on these axioms. In Appendix A.2, we offer a detailed comparison between the techniques we use to prove our bounds versus those used in past work.

2 Model

Rankings. There are m candidates M and n voters N . Voters express their preferences over the candidates as complete rankings. Formally, a ranking π is a bijection from indices to candidates $[m] \rightarrow M$, where $\pi(j)$ represents the candidate ranked j -th in π . Let $a \succ_{\pi} b$ to denote

³To avoid confusion with our model, we clarify: Xia’s model is called “smoothed” in the original paper, but is renamed “semi-random” in most subsequent work.

⁴These axiomatic results include RESOLVABILITY [26]; STRATEGYPROOFNESS [28]; CONDORCET and PARTICIPATION [25], plus Moulin’s impossibility of satisfying them together [27]; and Condorcet’s voting paradox and the ANR impossibility theorem [24].

that candidate a is preferred to b in ranking π , or formally, $\pi^{-1}(a) < \pi^{-1}(b)$. $\mathcal{L}(M)$ is the set of $m!$ possible rankings, or just \mathcal{L} when M is clear from context. Letting the rankings in \mathcal{L} be implicitly ordered, the last ($m!$ -th) ranking in \mathcal{L} is π_{-1} , and the set of rankings without this element as $\mathcal{L}_{-1} = \mathcal{L} \setminus \{\pi_{-1}\}$. For reasons to be clarified next, we will often work with \mathcal{L}_{-1} instead of \mathcal{L} .

Profiles and profile histograms. A profile $\pi = (\pi_i | i \in N)$ is a vector of n rankings, where π_i is voter i 's ranking. Let $\Pi_n = \prod_{i=1}^n \mathcal{L}$ be the set of all profiles on n voters, and let $\Pi = \bigcup_{n \in \mathbb{Z}^+} \Pi_n$ be the set of all profiles overall. We define addition over profiles in the natural way: $(\pi + \pi') = (\pi_i | i \in N) \parallel (\pi'_i | i \in N')$.⁵

Instead of working directly with profiles, we will work primarily with their *histograms*. A *histogram* \mathbf{h} is an $|\mathcal{L}_{-1}|$ -length vector whose π -th entry h_π is the proportion of voters with ranking π in a profile (\mathbf{h} is indexed only up to \mathcal{L}_{-1} because the histogram must add to 1, so the $m!$ -th index is redundant). The histogram associated with a particular profile π is \mathbf{h}^π , with π -th entry $\mathbf{h}_\pi^\pi := 1/n |\{i : \pi_i = \pi\}|$. We define the simplex of all possible histograms as $\Delta_{\mathbf{h}} := \{\mathbf{h} \in [0, 1]^{|\mathcal{L}_{-1}|} : \sum_{\pi \in \mathcal{L}_{-1}} h_\pi \leq 1\}$. Of course, H includes vectors with irrational entries that could never correspond to a well-defined profile; nonetheless, it will be useful to consider the completed space. In order to talk about only the histograms that are realizable from well-defined profiles, we also define $H^{\mathbb{Q}} = H \cap \mathbb{Q}^{|\mathcal{L}_{-1}|}$ to be the subset of H of vectors with rational components.

Finally, we will use *histogram operator* $\mathcal{H}(\cdot)$ to transform more general, profile-based objects into their histogram-based analogs. For example, $\mathcal{H}(\pi) = \mathbf{h}^\pi$, and $\mathcal{H}(\Pi) = H^{\mathbb{Q}}$. For a single ranking, $\mathcal{H}(\pi)$ is a $|\mathcal{L}_{-1}|$ -length basis vector with a 1 at the π -th index ($\mathcal{H}(\pi_{-1})$ is the 0s vector). We will also apply this operator to *distributions over* profiles and rankings in the natural way, first drawing a profile or ranking from the distribution, and then considering its corresponding histogram.

Voting Rules. A *voting rule* is a function $R : \Pi \mapsto 2^M$ mapping a given profile to a set of winning candidates. Then, $R(\pi)$ is the set of winners chosen by the voting rule on π . If a voting rule by its standard definition results in a tie, rather than specifying a tie-breaking rule, we assume it returns all such winners (this assumption is for ease of exposition only). Let \mathcal{R} be the set of all voting rules. We will study several specific rules, defined colloquially below and formally in Appendix B.1.

Positional Scoring Rules (PSRs) are represented by m -length vectors of weakly decreasing scores $s_1 \geq s_2 \geq \dots \geq s_m$, where $s_1 = 1$ and $s_m = 0$. Candidate a receives s_i points for each voter that ranks it i -th; the candidate with the most points wins. We consider three specific PSRs, defined by their score vectors: PLURALITY: $(1, 0, \dots, 0, 0)$, BORDA: $(1, 1 - \frac{1}{m-1}, \dots, \frac{1}{m-1}, 0)$, and VETO: $(1, 1, \dots, 1, 0)$. Beyond PSRs, we study the rules MINIMAX, KEMENY-YOUNG, and COPELAND. MINIMAX selects the candidate with the smallest *maximum pairwise domination*.⁶ KEMENY-YOUNG selects the candidate ranked first in the ranking with the minimal sum of *Kendall tau* distances from voters' rankings.⁷ COPELAND selects the alternative with the most points, giving a 1 point for each candidate b it pairwise dominates, and $1/2$ point for each b with which it pairwise ties. Finally, in our analysis, we will often study a general class of voting rules called *hyperplane rules*. This class is known to be equivalent to *generalized scoring rules* [29] and encompasses essentially every standard voting rule, including all those listed above.

⁵Here \parallel represents concatenation, and thus the resulting profile $\pi + \pi'$ contains $|N| + |N'|$ voters. We also extend addition to permit positive integer multiples of profiles: $z\pi$ means adding a profile together z times.

⁶Useful definitions: a *pairwise-dominates* b in π when over half of voters rank a ahead of b : $|\{\pi_i | a \succ_{\pi_i} b\}| > n/2$. a and b *pairwise tie* when $|\{\pi_i | a \succ_{\pi_i} b\}| = n/2$. a 's *maximum pairwise domination* is equal to $\max_{b \neq a} |\{\pi_i | a \succ_{\pi_i} b\}|$.

⁷The Kendall tau distance between rankings π, π' is the total number of swaps required to transform π into π' .

DEFINITION 1 (Hyperplane rules [19]). Note that given a set of ℓ affine-hyperplanes, these hyperplanes partition the space of histograms into at most 3^ℓ regions, as every point is either on a hyperplane or on one of two sides. We say that R is a *hyperplane rule* if there exists a finite set of affine hyperplanes H_1, \dots, H_ℓ such that R is constant on each such region.

We will sometimes subdivide this class of rules further: *decisive* hyperplane rules are those which output a single winner on profiles that are not on any hyperplane; *non-decisive* hyperplane rules are all others.

Axioms. Formally, an *axiom* is a function $A : \mathcal{R} \rightarrow (\Pi \rightarrow \{\text{True}, \text{False}\})$, mapping a voting rule R to another mapping describing whether A is satisfied by R on any given profile. We can think of $A(R)$ as representing the true/false statement “ R is consistent with A on π ”. We will study several standard axioms, defined formally in Appendix B.2 and described colloquially below.

R satisfies RESOLVABILITY on π if it selects a single winner. R satisfies CONDORCET CONSISTENCY (abbreviated as CONDORCET) on π if it selects the *Condorcet winner* (the candidate that pairwise dominates all other candidates), or by default if π has no Condorcet winner. R satisfies MAJORITY on π if it selects the *majority winner* (the candidate ranked first by a majority of voters), or by default if π does not have a majority winner. R satisfies CONSISTENCY on π if there is no partition of π into subprofiles such that a unique candidate b , which is *not* the winner in π , is chosen as the winner on all subprofiles in that partition. R satisfies INDEPENDENCE OF IRRELEVANT ALTERNATIVES (IIA) on π if the winner, a , cannot be made to lose to b by adjusting votes in a way that does not change the relative positions of a and b .

We also use $\rho(n)$ -GROUP-STABILITY, which colloquially requires that the outcome of voting rules be *stable* to a change in the behavior of up to $\rho(n)$ voters. This is essentially framing *margin of victory* as a binary condition.

DEFINITION 2 ($\rho(n)$ -GROUP-STABILITY). For a given rule R , $\rho(n)$ -GROUP-STABILITY(R) is satisfied if, for every pair of profiles π, π' that differ on at most $\rho(n)$ of voters’ rankings, $R(\pi) = R(\pi')$.

This axiom implies strong, $\rho(n)$ -parameterized group-level versions of three axioms of common interest: $\rho(n)$ -GROUP-STRATEGYPROOFNESS—no group of up to $\rho(n)$ voters can strategically misreport their preferences in concert and cause R to output an alternative they all weakly prefer, and at least one of them strongly prefers; $\rho(n)$ -GROUP-PARTICIPATION—no group of up to $\rho(n)$ voters can leave the election and cause R to output an alternative they all weakly prefer, and at least one of them strongly prefers); and $\rho(n)$ -GROUP-MONOTONICITY—no group of up to $\rho(n)$ voters can weakly decrease the position of an alternative c , which is not currently a winner by R , in their rankings, and cause c to become a winner by R .⁸ We formally define these axioms in Appendix B.2.

2.1 Smoothed model of profiles

Noise distributions over rankings. A *noise distribution* is effectively a distribution over rankings; formally, it is a distribution over permutations $\sigma : [m] \rightarrow [m]$. When we “apply noise” to a ranking π via a noise distribution \mathcal{S}_ϕ (with parameter ϕ to be defined later), we are drawing a

⁸The fact that $\rho(n)$ -GROUP-STABILITY implies the first and third of these axioms is by definition. For the second, $\rho(n)$ -GROUP-STABILITY technically implies $\frac{\rho(n)}{1-\rho(n)}$ -GROUP-PARTICIPATION, since this axiom involves $\rho(n)$ voters leaving the electorate; this does not change any of the asymptotic results, and we handle it in our proofs.

random permutation $\sigma \sim \mathcal{S}_\phi$, and then permuting π according to σ . Abusing notation slightly, we represent the distribution over rankings induced by perturbing π with noise model \mathcal{S}_ϕ as $\mathcal{S}_\phi(\pi)$.⁹

In this paper, we consider the class of noise distributions \mathcal{S} , encompassing any distribution over rankings satisfying the minimal regularity conditions in Assumptions 1 and 2 below. All $\mathcal{S}_\phi \in \mathcal{S}$ are parameterized by a dispersion parameter $\phi \in [0, 1]$, which captures the “level” of noise applied by \mathcal{S}_ϕ . We will use ϕ to implement multiple different such measures throughout our results. Assumptions 1 and 2 impose the minimal requirements to ensure that ϕ reasonably measures the amount of noise, and that the noise distribution is practical to work with. As we implement specific uses of ϕ , we will impose additional assumptions as needed in later sections.

ASSUMPTION 1 (Extremal values). The distribution \mathcal{S}_0 is the point mass on the identity (so that $\phi = 0$ corresponds to no noise added). The distribution \mathcal{S}_1 is uniform over all permutations (so that $\phi = 1$ corresponds maximum noise). Note that this implies that for all profiles $\pi \in \Pi$, $\mathcal{S}_1(\pi)$ is equivalent to the impartial culture model (see, e.g., [8]).

ASSUMPTION 2 (Continuity). For all σ , the probability \mathcal{S}_ϕ places on σ is continuous in ϕ .

Sampling noisy profiles. Applying noise via \mathcal{S}_ϕ to an entire *profile* means applying it to every ranking within it *independently* (an assumption also made in the semi-random model [24]). Formally, if π is our starting profile, then our noisy profile is drawn from the distribution $\prod_{i=1}^n \mathcal{S}_\phi(\pi_i)$, where each $\mathcal{S}_\phi(\pi_i)$ is independent. The resulting noisy profile is denoted $\mathcal{S}_\phi(\pi)$. We will treat \mathcal{S}_ϕ , $\mathcal{S}_\phi(\pi)$, and $\mathcal{S}_\phi(\pi)$ as distributions and random variables interchangeably.

Since we will work with profiles in histogram form, so we will usually reason about distributions over rankings and profiles *projected into histograms space*. We express the distribution over ranking histograms induced by noise distribution $\mathcal{S}_\phi(\pi)$ as $\mathcal{H}(\mathcal{S}_\phi(\pi))$. Where the former is a distribution over rankings, the latter is a distribution over basis vectors. Then, the corresponding noise distribution over profile histograms is, naturally, $\mathcal{H}(\mathcal{S}_\phi(\pi)) = 1/n \sum_{i=1}^n \mathcal{H}(\mathcal{S}_\phi(\pi_i))$.

2.2 Smoothed-satisfaction and smoothed-violation of axioms

First, we formally define worst-case notions of axiom satisfaction and violation. Recall that $A(R) : \Pi \rightarrow \{\text{True}, \text{False}\}$ intakes a given profile and outputs whether rule R satisfies axiom A on that profile. Then, we say that π is a *counterexample* to $A(R)$ iff $(A(R))(\pi) = \text{False}$. Let $\Pi^{\neg(A(R))}$ be the set of all profiles that are counterexamples to $A(R)$. Then, in the worst-case sense, R *satisfies* A if $\Pi^{\neg(A(R))} = \emptyset$ (i.e., no counterexample exists); otherwise, R *violates* A .

Now, we define what it means for R to satisfy or violate A in the smoothed model. Conceptually, R *smoothed-satisfies* A if the probability that R satisfies A , after applying a noise distribution, converges to 1 as n grows large. In contrast, R *smoothed-violates* A if probability of R satisfying A being satisfied converges to 0 as n grows large. Note that worst-case satisfaction implies smoothed-satisfaction, and smoothed-violation implies violation.

DEFINITION 3 (smoothed-satisfied). Voting rule R *\mathcal{S} -smoothed-satisfies* axiom A at a rate $f(n, \phi)$ if, for all $n \in \mathbb{Z}^+$ and $\phi \in (0, 1]$,

$$\sup_{\phi' \in [\phi, 1]} \sup_{\pi \in \Pi_n} \Pr \left[\mathcal{S}_{\phi'}(\pi) \in \Pi^{\neg A(R)} \right] \leq f(n, \phi).$$

⁹More formally, $\mathcal{S}_\phi(\pi) = \pi + \sigma$ with $\sigma \sim \mathcal{S}_\phi$. Here, the $+$ operator represents composition: if $\sigma(i) = j$, then the i -th ranked candidate in the perturbed ranking will be $\pi(j)$.

DEFINITION 4 (smoothed-violated). A voting rule R \mathcal{S} -smoothed-violates axiom A at a rate of $f(n)$ if there exists $\phi \in (0, 1]$ and a profile π of size n such that for all $z \in \mathbb{Z}^+$,

$$\inf_{\phi' \in [0, \phi]} \Pr \left[\mathcal{S}_{\phi'}(z\pi) \in \Pi^{\neg A(R)} \right] \geq 1 - f(zn).$$

Per Definition 3, convergence to smoothed-satisfaction occurs eventually for *all* $\phi \in (0, 1]$, although the rate depends on ϕ . When we show smoothed-violation, in contrast, we are saying there exists a constant amount of noise ϕ such that this amount, or any less, will not be enough to ensure satisfaction of A by R as n grows large. Note the gap between these two definitions: smoothed-violated is not the negation of smoothed-satisfied, and a claim could therefore be “in-between” these definitions, satisfying neither. This will not end up being the case for any of the voting rules or criteria we study.

3 Patterns across smoothed-violated, smoothed-satisfied axioms

3.1 Condorcet, Majority, Consistency, IIA

This section is dedicated to axioms about which we prove negative results. As summarized in Table 1, we find that for all the voting rules we study, smoothed noise is insufficient to prevent existing violations of any of these axioms. Across these axioms, the reason for this insufficiency is the same: as formalized in our generic sufficient condition in Lemma 7, the histogram space contains contiguous regions of counterexamples.

Voting Rules	Axioms			
	CONDORCET	MAJORITY*	CONSISTENCY*	IIA*
PLURALITY	s-violated	satisfied	satisfied	s-violated
PSRs \ PLURALITY	s-violated	s-violated	satisfied	s-violated
MINIMAX	satisfied	satisfied	s-violated	s-violated
KEMENY-YOUNG	satisfied	satisfied	s-violated	s-violated
COPELAND	satisfied	satisfied	s-violated	s-violated

Table 1: Summary of results stated in Theorem 5. s-violated = smoothed-violated.

THEOREM 5. For all $A \in \{\text{CONDORCET}, \text{MAJORITY}, \text{CONSISTENCY}, \text{IIA}\}$ and all $R \in \text{PSRs} \cup \{\text{MINIMAX}, \text{KEMENY-YOUNG}, \text{COPELAND}\}$, if R violates A , then R smoothed-violates A .

The proof of this theorem is found in Appendix C.1; we explain the intuition here. The core idea is that for small enough ϕ , $\mathcal{H}(\mathcal{S}_\phi(\pi))$ concentrates at an exponential rate very near the starting histogram \mathbf{h}^π (Lemma 6). This lemma, proven in Appendix C.2, follows from the fact that noise is applied independently across rankings, allowing the application of a simple Hoeffding bound.

LEMMA 6. Let \mathcal{S} be a noise model. For all $\varepsilon > 0$, there exists a $\phi \in (0, 1]$ such that for all $\phi' \in [0, \phi]$ and profiles $\pi \in \Pi_n$ on n voters,

$$\Pr \left[\|\mathcal{H}(\mathcal{S}_{\phi'}(\pi)) - \mathbf{h}^\pi\|_1 < \varepsilon \right] > 1 - \exp(-\varepsilon^2 n / 2).$$

From this concentration follows a general sufficient condition for smoothed-violation (Lemma 7): that there exists a counterexample π such that \mathbf{h}^π is contained within a ball of counterexamples. Then, as the noise distribution concentrates around \mathbf{h}^π , most of its probability mass is contained in the ball, and the probability of a counterexample converges to 1.

LEMMA 7. Fix a noise model \mathcal{S} , axiom A , and rule R . Suppose there exists a profile π and radius $r > 0$ such that for all profiles $\pi' \in \Pi$ satisfying (1) $|\pi'| = z|\pi|$ for some $z \in \mathbb{Z}^+$ and (2) $\|\mathbf{h}^{\pi'} - \mathbf{h}^\pi\|_1 < r$, it is the case that $\pi' \in \Pi^{\neg A(R)}$. Then, R **smoothed-violates** A at a rate of

$$f(n) = \exp(-r^2 n / 2).$$

Proof. Fix a noise model \mathcal{S} , an axiom A , and a voting rule R . Fix a profile π and radius $r > 0$ satisfying the preconditions of the theorem. Choose $\varepsilon = r$ and let ϕ be the one from Lemma 6 corresponding to ε and \mathcal{S} . Fix an arbitrary $z \in \mathbb{Z}^+$ and $\phi' \in [0, \phi]$. Notice that $\mathbf{h}^\pi = \mathbf{h}^{z\pi}$. Hence, Lemma 6 guarantees that the post-noise histogram lies in an r -radius ball around the original histogram \mathbf{h}^π with high probability—that is, $\mathcal{H}(\mathcal{S}_{\phi'}(z\pi)) \in B_r^{L_1}(\mathbf{h}^\pi)$ with probability at least $1 - f(z|\pi|) = 1 - \exp(-r^2 z |\pi| / 2)$. Note also that every profile in the support of $\mathcal{S}_{\phi'}(z\pi)$ is guaranteed to have $z|\pi|$ voters. These two facts, taken with both preconditions of the theorem, imply that the post-noise profile histogram is a counterexample with high probability—that is, $\mathcal{S}_{\phi'}(z\pi) \in \Pi^{\neg A(R)}$ with probability at least $1 - f(z|\pi|)$. It then follows, by the definition of smoothed violation (Definition 4), that R smoothed-violates A at a rate of $f(z|\pi|)$, as needed. \square

Finally, we conclude Theorem 5 by finding counterexamples for each A, R that exist within balls of other counterexamples, and then applying Lemma 7. Across rules and axioms, simple counterexamples suffice, supporting the intuition that the insufficiency of smoothed noise is not a quirk of these rules and axioms. Rather, smoothed noise may be insufficient across many interpretable rules and axioms, because—as a natural consequence of their interpretability—they behave similarly on similar profiles.

3.2 Resolvability, Strategyproofness, Participation, and Monotonicity

This section is dedicated to axioms about which we prove (mostly) positive results, summarized in Table 2. Across these axioms, our proofs use a common property that makes smoothed-noise *sufficient* for circumventing impossibilities: that across voting rules, their counterexamples are restricted to regions on or near a limited number of hyperplanes.

Voting Rules		Axioms	
		RESOLVABILITY	$o(\sqrt{n})$ -GROUP STABILITY*
(Decisive hyperplane rules)	PLURALITY	s-satisfied	s-satisfied
	PSRs \setminus PLURALITY	s-satisfied	s-satisfied
	MINIMAX	s-satisfied	s-satisfied
	KEMENY-YOUNG	s-satisfied	s-satisfied
	COPELAND	s-violated	s-satisfied

Table 2: Summary of results stated in Theorem 8 and Theorem 11. s-satisfied = smoothed-satisfied. Positive results for $o(\sqrt{n})$ -GROUP-STABILITY directly imply positive results for $o(\sqrt{n})$ -GROUP-STRATEGYPROOFNESS, $o(\sqrt{n})$ -GROUP-PARTICIPATION, and $o(\sqrt{n})$ -GROUP-MONOTONICITY.

As are upper bounds in Xia’s semi-random model [24], our upper bounds here will be parameterized by $\text{MIN-PROB}(\mathcal{S}_\phi) := \min_\sigma \Pr[\mathcal{S}_\phi = \sigma]$, the smallest probability \mathcal{S}_ϕ assigns to any permutation. This parameterization motivates two additional assumptions: ASSUMPTION 3 ensures that $1/\text{MIN-PROB}(\mathcal{S}_\phi)$ is well-defined, and ASSUMPTION 4 describes how MIN-PROB implements ϕ as a measurement of the noisiness of \mathcal{S}_ϕ .

ASSUMPTION 3 (Positivity). For all $\phi \in (0, 1]$, $\text{MIN-PROB}(\mathcal{S}_\phi) > 0$. That is, for any nonzero amount of noise, the resulting distribution assigns positive probability to all permutations.

ASSUMPTION 4 (Weak Monotonicity). The value $\text{MIN-PROB}(\mathcal{S}_\phi)$ is non-decreasing in ϕ .¹⁰

First, Theorem 8 shows that RESOLVABILITY is smoothed-satisfied for all decisive hyperplane rules. Notably, these rules exclude the known rule COPELAND, due to its lack of sensitivity to changes in rankings in regions surrounding ties.

THEOREM 8. All decisive hyperplane rules smoothed-satisfy RESOLVABILITY at a rate $O_m(1/\sqrt{n} \cdot \text{MIN-PROB}(\mathcal{S}_\phi)^{3/2})$. All non-decisive hyperplane rules smoothed-violate RESOLVABILITY.

The formal proof is found in Appendix C.3. This result corely relies on the fact that, essentially regardless of the noise distribution over *rankings* \mathcal{S}_ϕ , the resulting noise distribution over *entire profile histograms* $\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))$ must converge uniformly¹¹ to a multi-dimensional Gaussian distribution at a rate of $O(1/\sqrt{n})$ (Lemma 9):

LEMMA 9. Let \mathcal{S} be a noise model, $\phi \in [0, 1]$ a parameter, and $\boldsymbol{\pi} \in \Pi_n$ a profile on n voters. Then, for all convex sets $X \subseteq \mathbb{R}^{m!-1}$,

$$|\Pr[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi})) \in X] - \Pr[\mathcal{N}(\mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))], \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))]) \in X]| \leq \frac{O((m!)^{7/4})}{\sqrt{n} \cdot \text{MIN-PROB}(\mathcal{S}_\phi)^{3/2}}.$$

Intuitively, $\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))$ converges to a Gaussian because it is akin to the sum of independent indicators (literally, independently-drawn rankings represented in histogram space as basis vectors). We prove this convergence in Appendix C.4 using a general form of the Berry Esseen bound [3].

This convergence allows us to prove a general sufficient condition for the smoothed-satisfaction: that the set of counterexamples are contained with a measure-zero subset of histograms (Lemma 10).

LEMMA 10. Fix a noise model \mathcal{S} , voting rule R , and axiom A . If there exists some set X such that (1) $X = \bigcup_{j=1}^\ell X_j$ where each $X_j \subseteq \Delta_{\mathbf{h}}$ is convex, (2) $\mathcal{H}(\Pi^{\neg A(R)}) \subseteq X$, and (3) X is measure zero (noting that X is necessarily measurable), then R **smoothed-satisfies** A at a rate of

$$f(n, \phi) = \frac{\ell \cdot O((m!)^{7/4})}{\sqrt{n} \cdot \text{MIN-PROB}(\mathcal{S}_\phi)^{3/2}}.$$

The formal proof is found in Appendix C.5, and uses that the Gaussian places zero probability mass over measure-zero regions of its support. This concludes our analysis of RESOLVABILITY.

¹⁰We note that our results don’t centrally depend on Assumption 4—Assumptions 2 and 3 are sufficient to give the same high-level results. We include Assumption 4 because it is not prohibitive, it simplifies the exposition, and allows us to give more useful parameterized bounds.

¹¹“Uniform” (over profiles) convergence means that for all profiles $\boldsymbol{\pi}$ of any fixed n , $\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))$ is at most $O(1/\sqrt{n})$ “distance away” from the the Gaussian distribution with expectation and variance corresponding to that of $\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))$.

Now, we show that $\rho(n)$ -GROUP-STABILITY (Definition 2) is smoothed-satisfied by all hyperplane rules for $\rho(n) \in o(\sqrt{n})$. Because $\rho(n)$ -GROUP-STABILITY implies $\rho(n)$ -GROUP-STRATEGYPROOFNESS, $\rho(n)$ -GROUP-PARTICIPATION, and $\rho(n)$ -GROUP-MONOTONICITY, these axioms must also be smoothed-satisfied by all hyperplane rules.

THEOREM 11. $o(\sqrt{n})$ -GROUP-STABILITY is smoothed-satisfied by all hyperplane rules at a rate of $O_m(1/\sqrt{n} \cdot \text{MIN-PROB}(S_\phi)^{3/2}) + o(1)$.

We will now prove Theorem 11 via the following anti-concentration lemma, which states that for any π , $\mathcal{H}(\mathcal{S}_\phi(\pi))$ places limited probability mass within distance $\delta(n)$ of any specific hyperplane as n grows large, so long as $\delta(n)$ is decreasing sufficiently quickly in n .

LEMMA 12. Let \mathcal{G} be the set of all hyperplanes in $\mathbb{R}^{m!-1}$. For all noise models \mathcal{S} , parameters $\phi \in [0, 1]$, and $\delta(n) \in o(1/\sqrt{n})$, we have the following, where d is the L_1 distance.

$$\sup_{G \in \mathcal{G}} \sup_{\phi' \in [\phi, 1]} \sup_{\pi \in \Pi_n} \Pr [d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi)), G) \leq \delta(n)] \in O \left(\frac{\delta(n)\sqrt{n}}{\sqrt{\text{MIN-PROB}(S_\phi)}} + \frac{1}{\text{MIN-PROB}(S_\phi)^{3/2}\sqrt{n}} \right) \in o(1).$$

This lemma, proven in Appendix C.6, shows that even if $\mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ falls within $\delta(n)$ distance of the hyperplane (we can think of this as it falling within a “thick” hyperplane), the width of this thick hyperplane is shrinking faster than the distribution over histograms concentrates as n grows large. To make the convergence rate in Lemma 12 $o(1)$, $\delta(n)$ must be in $o(1/\sqrt{n})$.

To apply this lemma to show the smoothed-satisfaction of $\rho(n)$ -GROUP-STABILITY, first observe that for a group of size $\rho(n)$ to be able to impact the outcome of the election in π (i.e., for π to be a counterexample), that group must be *pivotal* in π —that is, \mathbf{h}^π must lie within some $\rho(n)$ -dependent distance from a profile on which the winner changes. Because the set of such profiles are defined by finitely-many hyperplanes (by definition of hyperplane rules), counterexamples are then restricted to “thick” hyperplanes. To apply Lemma 12 for each such hyperplane, we need their width to be $o(1/\sqrt{n})$ in histogram space, corresponding to a coalition of size $\rho(n) \in o(\sqrt{n})$. Then, we conclude Theorem 11 by simply union bounding over the finite number of hyperplanes referred to in Definition 1. \square

4 Beyond dependence on the minimum probability

All convergence rates to smoothed-satisfaction proven so far — both in Section 3.2 and in the related work [24] — depend on $1/\sqrt{n \cdot \text{MIN-PROB}^3}$. As a result, when MIN-PROB is very small, extremely large n is required to get reasonably low probability of an axiom violation: examining the relative n and MIN-PROB dependency above, if we decrease MIN-PROB by factor ℓ , we need to increase the number of voters n by a factor of ℓ^3 in order to recover the same probability bound.¹²

Motivated by the lack of good bounds for the important class of noise models with small MIN-PROB, we now pursue the first MIN-PROB-independent bounds on convergence rates. For concreteness, we specifically pursue these bounds within the well-established *Mallows model*, as defined

¹²One may wonder if different techniques could potentially improve this dependence. Although we do not have a tight cubic lower bound and the exact bound should depend on the exact voting rule/axiom, we can at least get a linear one. Indeed, consider the noise distribution that switches to any ranking other than the starting one with probability ε , but stays on the starting one with probability $1 - (m! - 1)\varepsilon$. The minimum probability here is ε , but we need $n \in \omega(\varepsilon/m!)$ to ensure that the profile changes at all with high probability, a necessary condition to smoothed-satisfy any axiom that fails on even a single profile.

below; however, as we will illustrate in detail throughout this section, our results rely only weakly on the properties of this model, and should apply to more general models as well.

DEFINITION 13 (Mallows noise model [17]). Let $d : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{N}$ be the Kendall tau distance, and let $\phi \in [0, 1]$. Then, the Mallows model $\mathcal{S}_\phi^{Mallows}$ is defined as follows, where $Z = \sum_{\pi' \in \mathcal{L}} \phi^{d(\pi', \pi)}$ is a normalizing term.

$$\Pr [\mathcal{S}_\phi^{Mallows}(\pi) = \pi'] = \frac{1}{Z} \phi^{d(\pi, \pi')}.$$

The Mallows model is an attractive case study for proving MIN-PROB-independent bounds for two reasons. First, it precisely captures the intuition motivating this analysis: that long-range swaps should be rare under less noise, making MIN-PROB low. Longer range swaps are *so* rare in this model, in fact, that $\text{MIN-PROB}(\mathcal{S}_\phi^{Mallows})$ approaches zero exponentially fast as m grows: the maximum Kendall tau distance between rankings is $\binom{m}{2}$, so $\text{MIN-PROB}(\mathcal{S}_\phi^{Mallows}) \in \phi^{\Omega(m^2)}$. This motivates our second reason: that despite its importance in social choice, existing convergence rates grow poorer at an exponential rate for Mallows noise as m gets large and/or ϕ gets small.

Within the Mallows model, we characterize the precise m, ϕ, n -dependent rates at which RESOLVABILITY and $o(\sqrt{n})$ -STRATEGY-PROOFNESS are smoothed-satisfied by four diverse voting rules: PLURALITY, BORDA, VETO, and MINIMAX. Recall that these voting rules are already known to smoothed-satisfy both axioms by our results in Section 3.2. The key difference in this analysis is that, while before n was being treated as the only variable (with m and the noise level treated as constants), we now consider more closely the convergence depends on m and ϕ . In particular, we are interested in how large n must be (as a function of m and ϕ) for the rate to be $o(1)$ (i.e., satisfaction occurring with high probability). We summarize our results in Table 3, framed to directly answer this question. The formal statements and proofs are below, in Section 4.1.

Voting Rules	Axioms		
	RESOLVABILITY	n^p -GROUP-STRATEGYPROOF	
PLURALITY	$\omega(m^5/\phi)$ is sufficient	$\omega((m^5/\phi)^{1/(1-p)})$ is sufficient	(Prop. 17)
BORDA	$\omega(m^6/\phi)$ is sufficient	$\omega((m^8/\phi)^{1/(1-p)})$ is sufficient	(Prop. 18)
VETO	$\Omega(1/\phi^{m-2})$ is necessary	$\Omega(1/\phi^{m-2})$ is necessary	(Prop. 19)
MINIMAX	$\Omega(1/m\phi^{\lfloor m/2 \rfloor})$ is necessary	$\Omega(1/m\phi^{\lfloor m/4 \rfloor})$ is necessary	(Prop. 20)

Table 3:

What is striking in these results is a clear separation between voting rules, which was not visible in our MIN-PROB-parameterized bounds. The convergence rates of PLURALITY and BORDA do not get dramatically worse as ϕ (and thus MIN-PROB) gets small; put another way, as ϕ scales down, n must scale up proportionally to maintain roughly the same probability of satisfaction. In contrast, MINIMAX and VETO require exponentially (in m) large n .¹³

We can, as before, distill a pattern explaining this gap: the voting rules that achieve the best rates (PLURALITY, BORDA) are those whose outcomes are more sensitive to local swaps across the support (or, in critical areas of the support, in the case of PLURALITY). The outcomes of VETO and

¹³To put this in perspective, suppose $m = 6$ and we decrease ϕ from $1/5$ to $1/10$. To maintain a similar probability of violating RESOLVABILITY, if we are using PLURALITY or BORDA it is sufficient to double the number of voters; if we are using VETO, one needs at least $2^4 = 32$ times as many voters, and for MINIMAX, one needs at least $2^3 = 8$ times as many voters. This gap only gets steeper as m gets larger.

MAXIMIN, in contrast, are fairly insensitive to local swaps, allowing us to show that local swaps are not enough to overcome impossibilities. While this pattern is perhaps unsurprising in retrospect, it may be important in informing the choice of voting rules.

4.1 Formal statements and proofs

We include in the body the proofs for one upper bound (PLURALITY) and one lower bound (VETO), and defer the proofs for BORDA and MINIMAX to Appendices D.2 and D.3, respectively. Our arguments will rely on only very weak properties of the Mallows model, essentially requiring just that the noise rarely induces swaps between distantly-ranked candidates. To emphasize the kinds of noise models to which our arguments can generalize, we now recap the precise properties of the noise model required for the proof corresponding to each voting rule. For PLURALITY, the key property of the noise distribution we use is that no alternative is ranked first post-noise with probability near 1. For BORDA, the key property is that there is no ℓ for which two candidates will be exactly ℓ positions apart post-noise with probability near 1. For VETO, the key property is that there is an exponentially small probability of a given voter moving one of their top two candidates to last place. For MINIMAX, the key property is that it is unlikely for two candidates a distance of $m/2$ apart to swap. Before presenting our results, we establish a few useful properties of Mallows noise.

LEMMA 14 (First- and last-place probability [2]). From starting ranking π , the probability that $\pi(j)$ is ranked first in a ranking drawn from $\mathcal{S}_\phi^{\text{Mallows}}(\pi)$ is proportional to ϕ^j , i.e., $\phi^j / \sum_{j'=1}^m \phi^{j'}$. Symmetrically, the probability $\pi(j)$ is ranked last is proportional to ϕ^{m-j} .

From [17], if two candidates i, j are k positions apart in π (i.e., $\pi(i)$ and $\pi(j)$ with $i - j = k$), the probability that they retain their relative order post-Mallows noise is increasing in k , and is always at least $1/2$. This was refined by [5] to an exact value of their swap probability:

LEMMA 15 (swap probability [5]). For candidates i, j such that $\pi(i)$ and $\pi(j)$ with $i - j = k$, their probability of swapping post-Mallows noise is

$$q(k) := \frac{1}{1 - \phi^{k+1}} \left(1 - \frac{(1 - \phi)k\phi^k}{1 - \phi^k} \right).$$

We will often parameterize our bounds by the probability of the opposite event—that i and j at distance k *do* swap, which we denote by $\bar{q}(k) := 1 - q(k)$. We will use the following upper bounds on $\bar{q}(k)$, derived in Appendix D.1.

LEMMA 16. For all $k \in [m]$ and $\phi \in [0, 1]$, $\bar{q}(k) \leq \min \left\{ k\phi^k, \phi^{k/2}, \frac{\phi}{1+\phi} \right\}$.

PROPOSITION 17 (PLURALITY). PLURALITY smoothed-satisfies RESOLVABILITY at a rate of at most

$$O \left(m^{5/2} / \sqrt{n\phi} + m \exp(-n/6m) \right)$$

and, as long as $\rho(n) \leq \frac{n}{6m}$, smoothed-satisfies $\rho(n)$ -GROUP-STABILITY at a rate of

$$O \left(\rho(n)m^{5/2} / \sqrt{n\phi} + m \exp(-n/9m) \right).$$

Proof. We begin with resolvability and show how to extend it to $\rho(n)$ -GROUP-STABILITY after. Fix a starting profile π . For a candidate c , let X_i^c be the indicator that voter i ranks c first, and let $p_i^c = \Pr[X_i^c = 1]$. Then, let $S^c = \sum_i X_i^c$ be the random variable representing plurality score of c post-noise. It follows that $\mathbb{E}[S^c] = \sum_{i \in [n]} p_i^c$. We will partition the candidates into two sets based on these expectations: L (for “low”) is defined as $L = \{c | \mathbb{E}[S^c] < \frac{n}{2m}\}$, and H (for “high”) is defined as $H = \{c | \mathbb{E}[S^c] \geq \frac{n}{2m}\}$. We will first show *Claim 1*: with high probability, the winner will be from H . Then, in *Claim 2*, we will show that the probability any two candidates in H have the same plurality score is small. Union bounding over these two events will upper bound the probability of an unresolvable outcome, i.e., a tie in plurality scores.

Proof of Claim 1. Fix a candidate $a \in L$. Note that a necessary condition for a to be a plurality winner is for their score to be at least n/m . Standard Chernoff bounds says that $\Pr[S^a \geq (1 + \delta)\mu] \leq \exp(-\delta^2\mu/(2 + \delta))$ for all $\delta \geq 0$, where $\mu = \mathbb{E}[S^a]$. Choose δ such that $(1 + \delta)\mu = n/m$. Note that since $a \in L$, $\mu \leq \frac{n}{2m}$, so $\delta \geq 1$. Further, this implies that $\delta\mu \geq \frac{n}{2m}$ and $\frac{\delta}{2+\delta} \leq \frac{1}{3}$. Plugging in both of these bounds, we get an upper bound of $\exp(-\frac{n}{6m})$ on the probability that a wins. Via a union bound, the probability any candidate in L wins is at most $|L| \exp(-\frac{n}{6m})$.

Proof of Claim 2. Fix two candidates $a, b \in H$. Let $Y_i = X_i^a - X_i^b$; then, we can upper-bound $\Pr[S^a = S^b]$ by $\Pr[\sum_i Y_i = 0]$. Because Y_i converges to be Gaussian-distributed, and since the Gaussian places 0 mass on any point, we can upper bound $\Pr[\sum_i Y_i = 0]$ by the convergence rate of $\sum_i Y_i$ to the Gaussian, derived via Berry-Esseen:

$$\Pr \left[\sum_i Y_i = 0 \right] \leq O \left(1/\sqrt{\text{Var}[Y_i]} \cdot \max_i (\rho_i / \text{Var}[Y_i]) \right) \quad (1)$$

where $\rho_i = \mathbb{E}[|Y_i - \mathbb{E}[Y_i]|^3]$. We now bound each of these terms. First, since each $Y_i \in [-1, 1]$, $|Y_i - \mathbb{E}[Y_i]| \leq 2$, so $\max_i (\rho_i / \text{Var}[Y_i]) \leq 2$. Next, we expand out $\text{Var}[Y_i]$ to get the following, where the last step uses that $2E[X_i^a X_i^b] = 0$ since only one of these indicators can be 1 for a given i .

$$\text{Var}[Y_i] = \mathbb{E}[Y_i^2] - \mathbb{E}[Y_i]^2 = E[(X_i^a)^2] + E[(X_i^b)^2] - 2E[X_i^a X_i^b] - E[X_i^a - X_i^b]^2 = p_i^a + p_i^b - (p_i^a - p_i^b)^2,$$

Now, Lemma 14, the probability that *any* candidate $\pi_i(j)$ ends up in the first position is proportional to ϕ^j . Therefore, (assuming $m \geq 2$) each p_i^c is upper-bounded by $\frac{\phi}{\sum_{j=1}^m \phi^j} \leq \frac{\phi}{\phi + \phi^2} = \frac{1}{1+\phi}$. We use this to upper-bound $(p_i^a - p_i^b)^2$ as follows:

$$(p_i^a - p_i^b)^2 \leq \max(p_i^a, p_i^b)(p_i^a + p_i^b) \leq 1/(1+\phi) \cdot (p_i^a + p_i^b).$$

Using this, we conclude that $\text{Var}[Y_i] \geq p_i^a + p_i^b - \frac{1}{1+\phi}(p_i^a + p_i^b) = \frac{\phi}{1+\phi}(p_i^a + p_i^b)$. It follows that $\sum_i \text{Var}[Y_i] \geq \frac{\phi}{1+\phi}(\mathbb{E}[S^a] + \mathbb{E}[S^b]) \geq \frac{\phi}{1+\phi}n/m \geq \frac{\phi}{2} \cdot n/m$. Plugging our bounds into Equation (1),

$$\Pr \left[\sum_i Y_i = 0 \right] \leq O \left(\sqrt{m/(n\phi)} \right).$$

Union bounding over the at most $|H|^2 \leq m^2$ pairs of candidates in H , the probability such a tie occurs is at most $O \left(m^2 \sqrt{m/(n\phi)} \right)$. Finally, union bounding again over the event that the winner was from L (where $|L| \leq m$), the overall probability of a tie is at most the stated bound.

For $\rho(n)$ -GROUP-STABILITY, we upper bound the probability that any non-winning candidate has plurality score within $2\rho(n)$ of the maximal score. make the threshold to be in L having $\mathbb{E}[X^c] \leq \frac{n}{3m}$. For Claim 1, we use a threshold of $\frac{n}{m} - 2\rho(n) \geq \frac{2n}{3m}$. The same argument goes through, we just plug in $\frac{n}{3m}$ for $\delta\mu$ instead, replacing the 6 in the bound with a 9. For Claim 2, rather than just proving it is unlikely for the scores to be equal, we can do the same analysis for all differences in the range $[-2\rho(n), 2\rho(n)]$. Union bounding over all $4\rho(n) + 1$ possible score values, this adds an additional $O(\rho(n))$ factor. \square

PROPOSITION 18 (BORDA). BORDA smoothed-satisfies RESOLVABILITY at a rate of $O(m^3/\sqrt{n\phi})$, and smoothed-satisfies $\rho(n)$ -GROUP-STABILITY at a rate of $O(\rho(n)m^4/\sqrt{n\phi})$.

PROPOSITION 19 (VETO). The rate VETO smoothed-satisfies RESOLVABILITY is lower bounded by $1 - n\phi^{m-2}$, and, as long as $n \geq m$, the rate 1-STRATEGY-PROOFNESS is smoothed-satisfied is lower bounded by $1 - n\phi^{m-2} - m\phi^{\lfloor \frac{n}{m} \rfloor}$.

Proof. Consider a profile where all voters rank two candidates a and b in their top two positions. We upper bound the probability that either a or b is ranked last by any voter; if this never occurs, then a and b will be tied. To this end, recall that by Lemma 14, a voter's first-choice alternative will be ranked last post-noise with probability $\phi^m / \sum_{j=1}^m \phi^j$, and their second-choice alternative will be ranked last post-noise with probability $\phi^{m-1} / \sum_{j=1}^m \phi^j$. Hence, neither of these candidates will ranked last with probability at least

$$1 - \frac{\phi^{m-1} + \phi^m}{\sum_{j=1}^m \phi^j} = 1 - \frac{\phi^{m-2} + \phi^{m-1}}{\sum_{j=0}^1 \phi^j} = 1 - \frac{(\phi^{m-2} + \phi^{m-1})(1 - \phi)}{1 - \phi^m} = 1 - \frac{\phi^{m-2} - \phi^m}{1 - \phi^m} \geq 1 - \phi^{m-2}.$$

Union bounding over all voters, we get that no voter will place a or b last with probability at least $1 - n\phi^{m-2}$. This lower bounds the rate at which VETO can smoothed-satisfy RESOLVABILITY.

For 1-GROUP-STABILITY, we will have the same construction, except each candidate other than a or b will be ranked last by at least $\lfloor \frac{n}{m} \rfloor$ voters. It is still the case that a and b will never be ranked last with probability at least $1 - n\phi^{m-2}$. Note that for any candidate $c \neq a, b$, if a voter ranks them last in the starting profile, they will continue to do so with probability $\frac{1}{\sum_{j=0}^m \phi^j} = \frac{1-\phi}{1-\phi^m} \geq 1 - \phi$. This implies they will continue to be ranked last by at least one of these $\lfloor \frac{n}{m} \rfloor$ voters with probability at least $1 - \phi^{\lfloor \frac{n}{m} \rfloor}$. Union bounding over these $m - 2 \leq m$ candidates, in the sampled profile, with probability $1 - n\phi^{m-2} - m\phi^{\lfloor \frac{n}{m} \rfloor}$, candidates a and b will never be ranked last, while all other candidates will be ranked last at least once. Hence, a and b are tied as veto winners. Further, as there are at least $m > m - 2$ voters, at least one candidate c must be ranked last at least twice. Consider a voter i that ranked c last, and suppose without loss of generality they prefer a to b . In this case, they can move b to the bottom of their ranking. Now a will be the unique candidate never ranked last, and will hence be the unique winner. This is an improvement for the voter, meaning that the resulting profile does not satisfy 1-STRATEGY-PROOFNESS. \square

PROPOSITION 20 (MINIMAX). The rate at which MINIMAX smoothed-satisfies RESOLVABILITY is lower bounded by

$$\Omega\left(\frac{1}{\sqrt{n\bar{q}(\lfloor m/2 \rfloor)}} - m \exp\left(-\frac{n(1 - 2\bar{q}(\lfloor m/2 \rfloor))^2}{32}\right)\right),$$

at least for n divisible by 4, and the rate for 1-GROUP-STRATEGYPROOFNESS is lower bounded by

$$1 - 3nq(\lfloor m/4 \rfloor) - O\left(m \exp\left(\frac{-(n-2)(1-2\bar{q}(\lfloor m/2 \rfloor))^2}{32}\right)\right).$$

Note that using the fact that $\bar{q}(\lfloor m/2 \rfloor) \in O(1/\log m)$, a necessary condition for a high probability bound is $n \in \omega(\bar{q}(\lfloor m/2 \rfloor))$. We can plug in our various upper bounds from Lemma 16 on $\bar{q}(\lfloor m/2 \rfloor)$. For the bound in Table 3, the second of these implies that $n \in \omega(1/(m\phi^{\lfloor m/2 \rfloor}))$ is necessary.

5 Discussion

Because the motivation for the smoothed model is fundamentally practical, we begin by outlining some key practical takeaways from our work. These takeaways help address the question: *how much can a bit of smoothed noise give us, when, and why?*

Although the related work so far has been largely optimistic about smoothed analysis in social choice, our analysis in Section 3.1 illustrates that for many rules and axioms, smoothed noise may be insufficient to circumvent impossibilities. Our sufficient condition for smoothed violation, moreover, suggests that the insufficiency of smoothed noise may be tied fundamentally to axioms’ and voting rules’ simplicity and interpretability, suggesting that these negative results are difficult to get around without making other compromises.

On the other hand, our results in Section 3.2 show that for certain kinds of axioms — i.e., those whose counterexamples must lie near hyperplanes or other small-measure structures — a small amount of noise *is* enough to circumvent impossibilities. However, our results in Section 4 paint a more nuanced picture: if we believe that in practice, any perturbations are truly unlikely to cause drastic opinion changes, then our choice of voting rule is far more important than past work suggests. In particular, while past MIN-PROB-parameterized upper bounds yield similar convergence rates across voting rules (for any given axiom), our results in the Mallows model — for which existing bounds do not usefully apply — show that actually, certain voting rules converge to smoothed-satisfaction far faster than others. The voting rules that converge faster are, perhaps unsurprisingly, those whose outcomes are more sensitive to local swaps.

With these takeaways in hand, we now explore some extensions and future work.

Extension and future work: smoothed analysis of Arrow’s Theorem Our results from Section 3.1 and 3.2, which address the smoothed-satisfaction of individual axioms, also have implications for *multi-axiom impossibilities* of the form, “there exists no voting rule that simultaneously satisfies all axioms in $A \in \mathcal{A}$ ” where \mathcal{A} is some collection of axioms. We can now ask the same question replacing “satisfies” with “smoothed-satisfies.” Past work in the semi-random model has studied several multi-axiom impossibilities, and our results from Section 3 immediately recover several of these existing results.¹⁴ However, perhaps the most famous multi-axiom impossibility, *Arrow’s Theorem* [1], has yet to be analyzed under semi-random noise. This impossibility states

¹⁴Our results imply the smoothed-resolution Gibbard-Satterthwaite [12] (Corollary 30), the ANR impossibility theorem (e.g., see [24]) (Corollary 31), and the impossibility of simultaneously satisfying CONDORCET and PARTICIPATION as identified by Moulin [20] (Corollary 32). Our results also imply that the existence of Condorcet cycles (i.e., *Condorcet’s Paradox*) is *not* smoothed-resolved — that is, existing Condorcet cycles remain with high probability in the presence of smoothed noise (Proposition 33).

that there is no voting rule that (when $m \geq 3$) can simultaneously satisfy IIA, UNANIMITY, and NON-DICTATORSHIP.¹⁵

In light of our findings in Section 3.1 that IIA tends to be smoothed-violated, one might guess that Arrow’s Impossibility is smoothed-impossible; i.e., it still holds under smoothed noise. However, due to how Arrow defines the axiom NON-DICTATORSHIP [1], it is not hard to show that Arrow’s Theorem is in fact *smoothed-resolved*. Indeed, a voting rule is said to be a *dictatorship* if there exists some voter such that the outcome of the voting rule is always that voter’s first choice. Now define an *almost-dictatorship* voting rule, that chooses voter 1’s favorite alternative on every profile except one arbitrary profile. The existence of this single exceptional profile means that our rule satisfies NON-DICTATORSHIP. However, since the rule only differs from a dictatorship on this one profile, it can easily be checked that it smoothed-satisfies IIA and UNANIMITY.

This positive result for Arrow’s theorem is not very conceptually satisfying, as it does not get at the heart of what makes IIA, dictatorship, and unanimity inconsistent. In some sense, this is because satisfying non-dictatorship, as it is defined, is too easy. To say something more meaningful, we propose a strengthening of NON-DICTATORSHIP, called LOCAL NON-DICTATORSHIP, that may be more interesting in the smoothed context: Fix a profile π and define voter i ’s neighborhood around π , $N_i(\pi) \subseteq \Pi$, to be the set of all profiles reachable by switching i ’s ranking for another ranking. Then, we say rule R satisfies LOCAL NON-DICTATORSHIP on profile π , if for all voters i , there is a $\pi' \in N_i(\pi)$ such that i ’s first choice in π' doesn’t win. In a sense, then, satisfying this axiom disallows a voter from being a dictator in any *local* area of profile space.

Because satisfaction of LOCAL NON-DICTATORSHIP implies the satisfaction of NON-DICTATORSHIP, Arrow’s impossibility also implies inconsistency between IIA, UNANIMITY, and LOCAL NON-DICTATORSHIP. We refer to this new, stronger impossibility as *Strengthened Arrow’s Theorem*. As we hoped, Strengthened Arrow’s Theorem is not susceptible to the simple fix discussed above: our *almost-dictatorship* voting rule fails to smoothed-satisfy LOCAL NON-DICTATORSHIP, since voter 1 is the dictator across almost every neighborhood. Then, the question remains open: *can this impossibility be smoothed-resolved?*

Future work: more general models limiting drastic opinion change Our analyses in Section 4 are restricted to the Mallows model, which is perhaps the poster child for a much broader, practically-motivated assumption: *noise-inducing shocks should not, in practice, induce drastic opinion change*. Already, our results reveal that this is a practically meaningful restriction, uncovering previously hidden heterogeneity among voting rules’ rates of convergence, due to their varying sensitivity to local swaps.

Given the practicality of the above assumption — and the different patterns of convergence that emerge when we apply it — we see investigating the broader space of noise models satisfying this assumption as an important frontier. There are many possible such noise models, and many will have MIN-PROB equal to or very near zero, warranting the pursuit of new bounds. Such further analyses may enable more nuanced distinctions between axioms and voting rules, and stronger conclusions over more flexible noise models. Going even further, one exciting potential opportunity in this direction is to potentially trade this restriction for a relaxation of the *independence* assumption (that noise is applied independently across rankings). This strong assumption — relied upon centrally in both our analyses and those in the related work — is difficult to relax when reasoning about

¹⁵UNANIMITY requires that if an alternative is ranked first by all voters, it is the winner. NON-DICTATORSHIP means that there exists no voter such that the outcome of the voting rule is always their first choice.

almost arbitrary noise distributions. However, there may be hope of relaxing it when considering noise distributions restricted to more closely reflect practice.

References

- [1] Kenneth Arrow. Individual values and social choice. *Nueva York: Wiley*, 24, 1951.
- [2] Pranjal Awasthi, Avrim Blum, Or Sheffet, and Aravindan Vijayaraghavan. Learning mixtures of ranking models. In *Proceedings of the 27th Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 2609—2617, 2014.
- [3] Vidmantas Bentkus. A lyapunov-type bound in rd. *Theory of Probability & Its Applications*, 49(2):311–323, 2005.
- [4] Andrew C. Berry. The accuracy of the gaussian approximation to the sum of independent variates. *Transactions of the american mathematical society*, 49(1):122–136, 1941.
- [5] Niclas Boehmer, Piotr Faliszewski, and Sonja Kraicz. Properties of the mallows model depending on the number of alternatives: A warning for an experimentalist. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, 2023.
- [6] Weiwei Cheng, Krzysztof Dembczynski, and Eyke Hüllermeier. Label ranking methods based on the plackett-luce model. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, pages 215–222, 2010.
- [7] Leah Melani Dillman, Don A .and Christian. Survey mode as a source of instability in responses across surveys. *Field methods*, 17(1):30–52, 2005.
- [8] Ömer Eğecioğlu and Ayça E. Giritligil. The impartial, anonymous, and neutral culture model: a probability model for sampling public preference structures. *The Journal of mathematical sociology*, 37(4):203–222, 2013.
- [9] Carl-Gustav Esseen. On the liapunov limit error in the theory of probability. *Arkiv för Matematik, Astronomi och Fysik*, 28:1–19, 1942.
- [10] Dan S. Felsenthal, Zeev Maoz, and Amnon Rapoport. An empirical evaluation of six voting procedures: do they really make any difference? *British Journal of Political Science*, 23(1): 1–27, 1993.
- [11] William V. Gehrlein. Condorcet’s paradox and the likelihood of its occurrence: different perspectives on balanced preferences. *Theory and decision*, 52(2):171–199, 2002.
- [12] Allan Gibbard. Manipulation of voting schemes. *Econometrica*, 41:587–602, 1973.
- [13] James Green-Armytage, T Nicolaus Tideman, and Rafael Cosman. Statistical evaluation of voting rules. *Social Choice and Welfare*, 46(1):183–212, 2016.
- [14] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific, 2013.
- [15] Leonard Lee, On Amir, and Dan Ariely. In search of homo economicus: Cognitive noise and the role of emotion in preference consistency. *Journal of consumer research*, 36(2):173–187, 2009.

- [16] Tyler Lu and Craig Boutilier. Learning mallows models with pairwise preferences. In *Proceedings of the 28th International Conference on Machine Learning (ICML)*, pages 145–152, 2011.
- [17] Colin L. Mallows. Non-null ranking models. i. *Biometrika*, 44(1/2):114–130, 1957.
- [18] Elchanan Mossel. Probabilistic view of voting, paradoxes, and manipulation. *Bulletin of the American Mathematical Society*, 59(3):297–330, 2022.
- [19] Elchanan Mossel, Ariel D. Procaccia, and Miklós Z Rácz. A smooth transition from powerlessness to absolute power. *Journal of Artificial Intelligence Research*, 48:923–951, 2013.
- [20] Hervé Moulin. Condorcet’s principle implies the no show paradox. *Journal of Economic Theory*, 45(1):53–64, 1988.
- [21] Florenz Plassmann and T. Nicolaus Tideman. How frequently do different voting rules encounter voting paradoxes in three-candidate elections? *Social Choice and Welfare*, 42(1): 31–75, 2014.
- [22] Geoffrey Pritchard and Arkadii Slinko. On the average minimum size of a manipulating coalition. *Social Choice and Welfare*, 27:263–277, 2006.
- [23] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004.
- [24] Lirong Xia. The smoothed possibility of social choice. In *Proceedings of the 33rd Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 11044–11055, 2020.
- [25] Lirong Xia. The semi-random satisfaction of voting axioms. In *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 6075–6086, 2021.
- [26] Lirong Xia. How likely are large elections tied? In *Proceedings of the 22nd ACM Conference on Economics and Computation (EC)*, pages 884–885, 2021.
- [27] Lirong Xia. Semi-random impossibilities of condorcet criterion. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5867–5875, 2023.
- [28] Lirong Xia. The impact of a coalition: Assessing the likelihood of voter influence in large elections. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC)*, 2023. Forthcoming.
- [29] Lirong Xia and Vincent Conitzer. Generalized scoring rules and the frequency of coalitional manipulability. In *Proceedings of the 9th ACM Conference on Economics and Computation (EC)*, pages 109–118, 2008.

A Supplemental Materials from Section 1

A.1 Comparison to Spielman and Teng’s smoothed model

The smoothed analysis framework was originally proposed by Spielman and Teng to provide theoretical justification for the Simplex algorithm’s fast runtime in real-world instances, despite its exponential worst-case complexity [23]. In their analysis, they go beyond the worst-case by adding Gaussian noise independently to each entry of the real-valued constraint matrix that is the input to Simplex algorithm. Then, they bound the expected run time of Simplex in the worst-case over inputs, where these guarantees are parameterized by the variance of the Gaussian noise added. Stated as a more general framework, the idea of smoothed analysis is to fix an arbitrary instance, add noise from a parameterized distribution, and then measure the quality of the expected outcome on the worst input (or, whether a property is satisfied with high probability—an alternative formulation proposed by Spielman and Teng that is closer to ours).

This is precisely how our model works, but rather than its input being a constraint matrix of real numbers, it is a *base profile* of complete rankings over m alternatives. This base profile is perturbed by applying generically-structured noise independently across each of its rankings (we refer to this as the *independence* assumption). When evaluating the probability of a criterion being satisfied post-perturbation, we assume that the base profile is chosen adversarially, i.e., to minimize this probability. The noise distribution we apply to each ranking is parameterized by a value $\phi \in [0, 1]$ to measure the quantity of noise added, analogous to Spielman and Teng’s variance parameter. Our main departure from the original smoothed model is, where Spielman and Teng specify Gaussian noise, we do not assume a specific noise distribution, instead allowing *any* ϕ -parameterized distribution that is *neutral* over alternatives.¹⁶ *Neutrality* means that for a given ϕ , if we permute a ranking and add noise, this is equivalent to adding noise and then permuting the output. In that sense, our noise model over profiles can be specified by a single distribution over one ranking, permuted to be applied to ranking permutations. On this distribution over rankings, we also sometimes assume *positivity*—that when $\phi > 0$, this distribution assigns positive probability to all rankings. Under these assumptions, Our class of noise models generalizes the popular Mallows noise model (e.g., [16]) as well as one-dimensional parameterizations of the Plackett-Luce model (e.g., see [6]).¹⁷

A.2 Comparison of techniques with Xia 2020 [24]

At a high level, Xia’s work and our work in the general smoothed model (Section 3) take a similar technical approach: both show that their noise models become well-behaved as the number of voters n grows large, and then use these convergence results to upper and lower bound how much probability mass is placed on “bad” profiles. In both models, the key assumption enabling this type of analysis is the independence of noise across rankings. As a result of this assumption, as n grows large, the resulting distribution over profiles will converge to distributions we understand, analogous to how sums of i.i.d. random variables converge to Gaussians via the Central Limit Theorem, or concentrate around their mean via Hoeffding’s inequality. We prove the convergence

¹⁶We do not commit to a noise distribution because there is no single well-established distribution that is obvious to apply (unlike in the real-valued setting, where the Gaussian is standard).

¹⁷Traditionally, the Plackett-Luce model takes a single m real-valued parameters, one per ranking position. Our model generalizes a variation where each parameter is expressed as a function of ϕ .

of our model in Lemmas 6 and 9, corresponding to Xia’s Lemma 1 of [24].¹⁸ Our convergence rates across the respective lemmas match asymptotically except when roughly, the set of “bad” profiles for a criterion is extremely small; this case tends to only come into effect for very specific kinds of criterion, none of which we formally analyze in this paper. In the limited cases where such differences do exist, they stem from our proof relying on a multi-dimensional version of the well-known Berry-Esseen theorem, while Xia’s relies on faster convergence of Poisson Multinomial variables.

¹⁸Notice that when comparing formal statements, both models consider the space of histograms (vectors representing the fraction of the profile composed of each ranking) rather than profiles directly, and, while we tend to analyze the set of histograms corresponding to “bad” profiles (e.g., those in which an axiom is not satisfied) directly, Xia considers sets which are solutions to a system of linear equations and inequalities. However, in practice, these approaches end up being quite similar.

B Supplementary Materials from Section 2

B.1 Definitions of Voting Rules

We define the voting rules we study as functions of profile histograms \mathbf{h} (rather than a profile π). Further, we define them in the same form: first, we express how they assign candidates' scores, and then express via an $\arg\max_c$ that the winner or set of winners constitutes the candidate(s) with the highest (or in one case, lowest) score.

Positional Scoring Rules (PSRs). For fixed m , a positional scoring rule is characterized by a vector of weights of length m , $\alpha = (\alpha_c | c \in [m])$, where $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_m$. Without loss of generality, we let these weights be translated and scaled such that $\alpha_1 = 1$ and $\alpha_m = 0$. The winner(s) by a PSR R , characterized by α ,

$$R(\mathbf{h}) = \arg\max_c \sum_{j \in [m]} \alpha_j \sum_{\substack{\pi \in \mathcal{L} \\ \pi(j)=c}} h_\pi$$

Minimax. The MINIMAX winner is the candidate whose greatest pairwise defeat is the smallest:

$$\text{MINIMAX}(\mathbf{h}) = \arg\min_c \max_{c' \neq c} \sum_{\substack{\pi \in \mathcal{L} \\ c' \succ_\pi c}} h_\pi$$

Kemeny-Young. We define a candidate c 's Kemeny-Young score to be, at a high level, the level of agreement with voters' rankings of the most-agreeing ranking that ranks c first. Then, the set of winners contains the candidate(s) with the highest Kemeny-Young score.

$$\text{KEMENY-YOUNG}(\mathbf{h}) = \arg\max_c \max_{\pi: \pi(1)=c} \sum_{\substack{c', c'' \in C \\ c' \succ_\pi c''}} \sum_{\substack{\pi' \in \mathcal{L} \\ c' \succ_{\pi'} c''}} h_{\pi'}$$

Copeland.

$$\text{COPELAND}(\mathbf{h}) = \arg\max_c \sum_{c' \neq c} \mathbb{I}[c \succ_{\mathbf{h}} c'] + 1/2 \cdot \mathbb{I}[c \sim_{\mathbf{h}} c']$$

B.2 Definitions of Axioms

We use the following notation. For a ranking $\pi \in \mathcal{L}$, we let $\pi(j)$ for an index $j \in [m]$ be the candidate in the j 'th position of π . For a ranking $\pi \in \mathcal{L}$ and distinct candidate $c, c' \in M$, we use $c \succ_\pi c'$ to denote that c is ranked higher than c' in π . $c \succ_\pi c'$ when $|\{i \in [n] : c \succ_{\pi_i} c'\}| > |\{i \in [n] : c' \succ_{\pi_i} c\}|$.

RESOLVABILITY. A voting rule R satisfies RESOLVABILITY on π iff $|R(\pi)| = 1$ (i.e., there are no ties).

CONDORCET. A *Condorcet winner* is a candidate that would win in a pairwise election against every other candidate. That is, c is a Condorcet winner in π if $c \succ_\pi c'$ for all $c' \neq c$. A voting rule

R satisfies Condorcet Consistency on a given profile π if one of two conditions hold: (1) there is no Condorcet winner in π , or (2) there is a Condorcet winner c , and $R(\pi) = \{c\}$.

MAJORITY. A *majority winner* is a candidate that is ranked first by a majority of agents. That is, c is a majority winner in π if

$$|\{i \in [n] : \pi_i(c) = 1\}| > n/2$$

A voting rule R satisfies MAJORITY on a profile π if it satisfies one of two conditions: (1) there is no majority winner in π , or (2) if there is a majority winner a , then $R(\pi) = \{a\}$.

CONSISTENCY. A voting rule R satisfies CONSISTENCY on a profile π if the following holds: for all partitions of π into sub-profiles, (π^1, \dots, π^t) , if $R(\pi^j)$ for all $j \in [t]$ is the same set of winners W , then $R(\pi) = W$.

INDEPENDENCE OF IRRELEVANT ALTERNATIVES (IIA). A voting rule R satisfies INDEPENDENCE OF IRRELEVANT ALTERNATIVES (IIA) on profile π if the following holds. Suppose $R(\pi) = a$. Then, for all candidates $b \neq a$, if π' is such that $a \succ_{\pi_i} b$ if and only if $a \succ_{\pi'_i} b$ for all voters i , then $R(\pi') \neq b$.

The axioms we study are defined formally as follows:

DEFINITION 21 (**$\rho(n)$ -group-strategyproofness**). A voting rule R satisfies $\rho(n)$ -GROUP-STRATEGY-PROOFNESS on a profile π if there exists no group of agents of size at most $\rho(n)$ such that if they change their votes, resulting in some profile π' with outcome $R(\pi')$, they are all at least as well off and at least one is strictly better off. An agent is at least as well (resp. strictly better) off if their favorite candidate in the set $R(\pi')$ is weakly (resp. strictly) preferred to their favorite candidate in $R(\pi)$.

DEFINITION 22 (**$\rho(n)$ -group-monotonicity**). A voting rule R satisfies $\rho(n)$ -GROUP-MONOTONICITY on a profile π if there exists no candidate a and no group of agents of size at most $\rho(n)$ such that if they change their votes *without decreasing the position of a in any of their rankings*, producing a new profile π' , then it cannot be that $a \in R(\pi)$ and $a \notin R(\pi')$.

DEFINITION 23 (**$\rho(n)$ -group-participation**). A voting rule R satisfies $\rho(n)$ -GROUP-PARTICIPATION on a profile π if there exists no group of agents of size at most $\rho(n)$ such that if they collectively leave the election, producing a new profile π' , they are all at least as well off and at least one is strictly better off with the outcome $R(\pi')$ than $R(\pi)$.

C Supplementary material from Section 3

C.1 Proof of Theorem 5

We prove this result for each axiom separately. For each axiom, we define a sufficient condition for a counterexample π to be “strict”, i.e., all profiles π' with histograms nearby to \mathbf{h}^π will also be counterexamples. We then show that the existence of a strict counterexample implies smoothed-violation via Lemma 7. We later give (or point to existing) strict counterexamples for all relevant pairs of rules and axioms. In the following arguments, we say a profile π is *robust* with respect to a voting rule R if there is an $\varepsilon > 0$ such that all profiles π' with $\|\mathbf{h}^\pi - \mathbf{h}^{\pi'}\|_1 < \varepsilon$, $R(\pi') = R(\pi)$. Notice that for hyperplane rules, all profiles that do not fall on hyperplanes are robust.

Condorcet: We say a counterexample π is a strict counterexample to R satisfying CONDORCET if π is robust with respect to R , π has a strict Condorcet winner a (i.e., a candidate that beats every other candidate on *strictly more* than half of the voters), and $R(\pi) \neq a$. If such a strict counterexample π exists, if it is robust with value ε_1 and a wins with at least a $1/2 + \varepsilon_2$ fraction against each candidate, all profiles whose histogram falls within $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ of \mathbf{h}^π have the same Condorcet winner and same output under R , and thus are also counterexamples. Hence, Lemma 7 implies smoothed-violation.

Majority: We say a counterexample π is a strict counterexample to R satisfying MAJORITY if π is robust and π has a strict Majority winner a (i.e., a candidate that is ranked first by *strictly more* than half of the voters), and $R(\pi) \neq a$. If such a strict counterexample π exists, if it is robust with value ε_1 and a is ranked first by at least a $1/2 + \varepsilon_2$ fraction of the voters, all profiles whose histogram falls within $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ of \mathbf{h}^π have the same Condorcet winner and same output under R , and thus are also counterexamples. Hence, Lemma 7 implies smoothed-violation.

Consistency: We say a counterexample π is a strict counterexample to R satisfying CONSISTENCY if $\pi = \pi^1 \cup \dots \cup \pi^t$ such that $R(\pi^1) = \dots = R(\pi^t) \neq R(\pi)$ and all of π^1, \dots, π^t and π are robust with respect to R . Suppose such a strict counterexample $\pi = \pi^1 \cup \dots \cup \pi^t$ exists. Let ε^{\min} be an amount by which all the relevant profiles are robust. Suppose each profile π^j has n_j voters and let $n = n_1 + \dots + n_t$ be the number of voters in π . Let $p = \min_j n_j/n$. Notice that any π' on zn voters within $\varepsilon = p\varepsilon^{\min}$ of π can be decomposed into $\pi^{1'}, \dots, \pi^{t'}$ such that each $\pi^{j'}$ has zn_j voters and $\mathbf{h}^{\pi^{j'}}$ is at most ε^{\min} away from \mathbf{h}^{π^j} . Hence, $R(\pi^{1'}) = \dots = R(\pi^{t'}) \neq R(\pi')$, so π' is a counterexample to CONSISTENCY. Hence, Lemma 7 implies smoothed-violation.

IIA: We say a counterexample π^1 is a strict counterexample to R satisfying IIA if there is another profile π^2 such that both π^1 and π^2 are robust with respect to R , and there are candidates a, b such that the relative ranking of a and b are the same under π_i^1 and π_i^2 for all voters i , yet $R(\pi^1) = a$ and $R(\pi^2) = b$. Suppose such a strict counterexample π^1 and π^2 that are both robust by at least ε . Consider any profile $\pi^{1'}$ where $\mathbf{h}^{\pi^{1'}}$ is within ε of \mathbf{h}^{π^1} . Notice that there must be a profile $\pi^{2'}$ such that $\mathbf{h}^{\pi^{2'}}$ is within ε of \mathbf{h}^{π^2} that matches $\pi^{1'}$ in terms of all voter’s relative rankings of a and b , and by robustness $R(\pi^{1'}) = a$ while $R(\pi^{2'}) = b$. Hence, Lemma 7 implies smoothed-violation.

Counterexamples: The following table points to strict counterexamples that can be used to show the claims above. Most can be found on Wikipedia pages. The missing ones are presented afterwards.

Voting Rules	Axioms			
	CONDORCET	MAJORITY	CONSISTENCY	IIA
PLURALITY	[a]	satisfied	satisfied	[c]
(non-PLURALITY) PSRs	Example 24	Example 24	satisfied	Example 25
MINIMAX	satisfied	satisfied	[b]	[c]
KEMENY-YOUNG	satisfied	satisfied	[b]	[c]
COPELAND	satisfied	satisfied	[b]	[c]

Table 4: Externally-referenced examples come from the following Wikipedia pages [a]: *Condorcet Criterion*, [b]: *Consistency Criterion*, [c]: *Independence of Irrelevant Alternatives*.

Example 24. Let R be the positional scoring rule represented by the weights vector $(1, \alpha, \dots, 0)$ without loss of generality. We assume $\alpha > 0$ is separated from 0 (otherwise, R is just Plurality). Let

$$\pi : \begin{array}{l} n \left(\frac{1}{2} - \frac{\alpha}{4(2-\alpha)} \right) \text{ voters} \\ n \left(\frac{1}{2} + \frac{\alpha}{4(2-\alpha)} \right) \text{ voters} \end{array} \left| \begin{array}{l} \mathbf{a}_1 \succ a_3 \succ \dots \succ a_m \succ \mathbf{a}_2 \\ \mathbf{a}_2 \succ \mathbf{a}_1 \succ a_3 \succ \dots \succ a_m \end{array} \right.$$

Example 25. Let R be the positional scoring rule represented by the weights vector $(1, \alpha, \dots, 0)$ without loss of generality. We assume $\alpha > 0$ is separated from 0 (otherwise, R is just Plurality).

$$\pi^1 : \begin{array}{l} n/2 \text{ voters} \\ n/4 \text{ voters} \\ n/4 \text{ voters} \end{array} \left| \begin{array}{l} \mathbf{a}_1 \succ a_3 \succ \dots \succ a_m \succ \mathbf{a}_2 \\ \mathbf{a}_2 \succ a_3 \succ \dots \succ a_m \succ \mathbf{a}_1 \\ \mathbf{a}_2 \succ \mathbf{a}_1 \succ a_3 \succ \dots \succ a_m \end{array} \right. \quad \pi^2 : \begin{array}{l} n/2 \text{ voters} \\ n/4 \text{ voters} \\ n/4 \text{ voters} \end{array} \left| \begin{array}{l} \mathbf{a}_1 \succ \mathbf{a}_2 \succ a_3 \succ \dots \succ a_m \\ \mathbf{a}_2 \succ a_3 \succ \dots \succ a_m \succ \mathbf{a}_1 \\ \mathbf{a}_2 \succ \mathbf{a}_1 \succ a_3 \succ \dots \succ a_m \end{array} \right.$$

This concludes the proof. \square

C.2 Proof of Lemma 6

Fix a noise model \mathcal{S} and $\varepsilon > 0$. By continuity of \mathcal{S} , there exists a $\phi > 0$ such that for all $\phi' \in [0, \phi]$, $\Pr[\mathcal{S}_{\phi'}(\pi) = \pi] > 1 - \varepsilon/2$. Choose this to be our ϕ .

Fix such a ϕ' and a profile $\pi \in \Pi_n$. Notice that after applying \mathcal{S} , each ranking π_i will stay the same with probability at least $1 - \varepsilon/2$. As this is independent accross voters, a straightforward application of Hoeffding's inequality tells us that at least a $1 - \varepsilon$ fraction of rankings will not change with probability at least $1 - \exp(-\varepsilon^2 n/2)$, as needed. \square

C.3 Proof of Theorem 8

For a decisive hyperplane rule, note that all profiles that fail RESOLVABILITY must fall on one of the ℓ hyperplanes (where by definition ℓ is finite). Further, note that these hyperplanes are convex and measure-zero. Hence, we can immediately apply Lemma 10 using the hyperplanes as convex sets.

For non-decisive hyperplane rules, there must exist a profile π not lying on any hyperplane for which RESOLVABILITY fails. Further, \mathbf{h}^π is at least some L^1 distance $\varepsilon > 0$ from all hyperplanes.

All such profiles with histograms in this ball have the same outcome as π and hence do not satisfy RESOLVABILITY. We can then directly apply Lemma 7 using this profile π and $r = \varepsilon$.

C.4 Proof of Lemma 9

To prove this, we first prove the following technical lemma. One takeaway from this lemma is that $\mathcal{H}(\mathcal{S}_\phi(\pi))$'s covariance matrix is quite easy to work with: its inverse not only exists but has a simple closed-form, and its eigenvalues are lower-bounded by a constant.

LEMMA 26. For all noise models \mathcal{S} , parameters $\phi \in (0, 1]$, and rankings $\pi \in \Pi_n$, the covariance matrix $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ is invertible and has all positive real eigenvalues lower bounded by $\text{MIN-PROB}(\mathcal{S}_\phi)/(m!n)$.

Proof. We first express the expectation and variance of $\mathcal{H}(\mathcal{S}_\phi(\pi))$ in terms of the analogous values for the $\mathcal{H}(\mathcal{S}_\phi(\pi_i))$. The relationships between these quantities are shown below, derived by applying simple properties of the expectation and variance in conjunction with the fact that $\mathcal{H}(\mathcal{S}_\phi(\pi)) = 1/n \sum_{i=1}^n \mathcal{H}(\mathcal{S}_\phi(\pi_i))$:

$$\mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi))] = 1/n \sum_{i=1}^n \mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))], \quad \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))] = 1/n^2 \sum_{i=1}^n \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$$

We now use these relationships to find closed forms for each of these objects. Note that $\mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$ is an $|\mathcal{L}_{-1}|$ -length vector whose π -th component is simply $\Pr[\mathcal{S}_\phi(\pi_i) = \pi]$. $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$ is a $|\mathcal{L}_{-1}| \times |\mathcal{L}_{-1}|$ matrix whose entries each correspond to a pair of rankings π, π' , such that the π, π' -th entry is equal to the covariance between the random variables $\mathcal{H}(\mathcal{S}_\phi(\pi_i))_\pi$ and $\mathcal{H}(\mathcal{S}_\phi(\pi_i))_{\pi'}$.

The covariance matrix $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$ is a $|\mathcal{L}_{-1}| \times |\mathcal{L}_{-1}|$ matrix whose entries each correspond to a pair of rankings π, π' , such that the π, π' -th entry is equal to the covariance between the random variables $\mathcal{H}(\mathcal{S}_\phi(\pi_i))_\pi$ and $\mathcal{H}(\mathcal{S}_\phi(\pi_i))_{\pi'}$. Given that $\mathcal{H}(\mathcal{S}_\phi(\pi_i))$ can take on the values of only basis vectors, the values of these random variables are either 0 or 1. To characterize these entries, we will use the fact that for general X and Y , $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

For distinct rankings $\pi \neq \pi'$, at most one of $\mathcal{H}(\mathcal{S}_\phi(\pi_i))_\pi$ and $\mathcal{H}(\mathcal{S}_\phi(\pi_i))_{\pi'}$ can be nonzero, so the expectation their product must be 0. Then,

$$\text{Cov}(\mathcal{H}(\mathcal{S}_\phi(\pi_i))_\pi, \mathcal{H}(\mathcal{S}_\phi(\pi_i))_{\pi'}) = 0 - \Pr[\mathcal{S}_\phi(\pi_i) = \pi] \cdot \Pr[\mathcal{S}_\phi(\pi_i) = \pi'].$$

For diagonal entries where $\pi = \pi'$, since the values of our random variables are always 0 or 1, we have

$$\text{Cov}(\mathcal{H}(\mathcal{S}_\phi(\pi_i))_\pi, \mathcal{H}(\mathcal{S}_\phi(\pi_i))_\pi) = \Pr[\mathcal{S}_\phi(\pi_i) = \pi] - \Pr[\mathcal{S}_\phi(\pi_i) = \pi]^2.$$

With these in hand, we now prove the lemma statement. Fix \mathcal{S}, ϕ , and π . Recall that $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))] = 1/n^2 \sum_{i=1}^n \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$. Hence, we first consider $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ for individual rankings π .

Fix an arbitrary ranking π . First, we will prove that all eigenvalues of $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ are positive and the minimum is at least $\text{MIN-PROB}(\mathcal{S}_\phi)/m!$.

To simplify notation in the subsequent computations, for the j 'th ranking π' , let $q_j = \Pr[\mathcal{S}_\phi(\pi) = \pi']$. Recall that in $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$, the (j, k) -th entry when $j = k$ (a diagonal entry) has value

$q_j(1 - q_j)$ and for $j \neq k$, the entry has value $-q_j \cdot q_k$. We can then write the covariance matrix as

$$\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))] = \begin{pmatrix} q_1(1 - q_1) & -q_1 \cdot q_2 & \cdots & -q_1 \cdot q_{m!-1} \\ -q_2 \cdot q_1 & q_2(1 - q_2) & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ -q_{m!-1} \cdot q_1 & \cdots & \cdots & q_{m!-1}(1 - q_{m!-1}) \end{pmatrix}.$$

Note that $q_{m!}$ is the probability of the “missing” ranking, and that $\sum_{j=1}^{m!} q_j = 1$. We first demonstrate an inverse of $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$. Consider the matrix:

$$M^{inv} = \begin{pmatrix} \frac{1}{q_1} + \frac{1}{q_{m!}} & \frac{1}{q_{m!}} & \cdots & \frac{1}{q_{m!}} \\ \frac{1}{q_{m!}} & \frac{1}{q_2} + \frac{1}{q_{m!}} & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \frac{1}{q_{m!}} & \cdots & \cdots & \frac{1}{q_{m!-1}} + \frac{1}{q_{m!}} \end{pmatrix}.$$

More formally, the j th diagonal entry is $1/q_j - 1/q_{m!}$ and all off-diagonal entries are simply $1/q_{m!}$.

We now show that $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))] \cdot M^{inv} = I_{m!-1}$ where $I_{m!-1}$ is the identity matrix, that is, M^{inv} is in fact the inverse of $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$. To that end, let us consider the i th diagonal entry of the product. It is precisely

$$\begin{aligned} (\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))] \cdot M^{inv})_{jj} &= \sum_{k=1, k \neq j}^{m!-1} -\frac{q_j q_k}{q_{m!}} + q_j(1 - q_j) \left(\frac{1}{q_j} + \frac{1}{q_{m!}} \right) \\ &= \sum_{k=1, k \neq j}^{m!-1} -\frac{q_j q_k}{q_{m!}} + (1 - q_j) + \frac{q_j}{q_{m!}}(1 - q_j) \\ &= \sum_{k=1, k \neq j}^{m!-1} -\frac{q_j q_k}{q_{m!}} + 1 - q_j + \frac{q_j}{q_{m!}} - \frac{q_j \cdot q_j}{q_{m!}} \\ &= \sum_{k=1}^{m!-1} -\frac{q_j q_k}{q_{m!}} + 1 - q_j + \frac{q_j}{q_{m!}} \\ &= \frac{-q_j(1 - q_{m!})}{q_{m!}} + 1 - q_j + \frac{q_j}{q_{m!}} \\ &= \frac{-q_j + q_j \cdot q_{m!}}{q_{m!}} + 1 - \frac{q_j \cdot q_{m!}}{q_{m!}} + \frac{q_j}{q_{m!}} \\ &= 1. \end{aligned}$$

For a non-diagonal entry j, k with $j \neq k$, we have

$$\begin{aligned} (\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))] \cdot M^{inv})_{jk} &= \sum_{\ell=1, \ell \neq j, k}^{m!} -\frac{q_j q_\ell}{q_{m!}} + \frac{q_j(1 - q_j)}{q_{m!}} - q_j q_k \cdot \left(\frac{1}{q_k} + \frac{1}{q_{m!}} \right) \\ &= \sum_{\ell=1, \ell \neq j, k}^{m!} -\frac{q_j q_\ell}{q_{m!}} + \frac{q_j}{q_{m!}} - \frac{q_j q_j}{q_{m!}} - q_j - \frac{q_j q_k}{q_{m!}} \end{aligned}$$

$$\begin{aligned}
&= \sum_{\ell=1}^{m!} -\frac{q_j q_\ell}{q_{m!}} + \frac{q_j}{q_{m!}} - q_j \\
&= \frac{q_j(1 - q_{m!})}{q_{m!}} + \frac{q_j}{q_{m!}} - \frac{q_j q_{m!}}{q_{m!}} \\
&= 0.
\end{aligned}$$

We now consider the eigenvalues of $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$. Since it is a covariance matrix, it is symmetric, and therefore positive semi-definite. Since we now know it is invertible, it is in fact positive definite. This implies all of its eigenvalues exist and are positive. Further, since the eigenvalues of $M^{inv} = \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]^{-1}$ are the reciprocals of the eigenvalues of $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$, we can lower bound the eigenvalues of $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ by upper bounding the the eigenvalues of M^{inv} .

To upper bound the maximum eigenvalue of M^{inv} , we can upperbound the maximum absolute row sum. Note that the sum of row j is

$$\frac{1}{q_j} + \frac{m! - 1}{q_{m!}} \leq \frac{m!}{\text{MIN-PROB}(\mathcal{S}_\phi)}.$$

This lower bounds the minimum eigenvalue of $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ by $\frac{\text{MIN-PROB}(\mathcal{S}_\phi)}{m!}$, as needed.

For our fixed profile π , we now have that each $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$ has minimum eigenvalue at least $\frac{\text{MIN-PROB}(\mathcal{S}_\phi)}{m!}$. Since the minimum eigenvalue of the sum of matrices is at least the sum of the minimum eigenvalues of each matrix, the minimum eigenvalue of $\sum_{i=1}^n \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$ is at least $n \cdot \frac{\text{MIN-PROB}(\mathcal{S}_\phi)}{m!}$. Finally, to get $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$, we scale this sum down by n^2 which scales the minimum eigenvalues equivalently, yielding a minimum eigenvalue of at least $\frac{\text{MIN-PROB}(\mathcal{S}_\phi)}{nm!}$. Note that the minimum eigenvalue here is positive, meaning $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ has no zero eigenvalues, meaning it is invertible. \square

Proof of Lemma 9. Fix \mathcal{S}, ϕ , and $\pi \in \Pi_n$. Since $\mathcal{H}(\mathcal{S}_\phi(\pi)) = 1/n \sum_{i=1}^n \mathcal{H}(\mathcal{S}_\phi(\pi_i))$ where each of these summands is independent, our goal will be to apply a version of the Berry-Esseen bound, as stated below:

LEMMA 27 (Restatement of Berry-Esseen as in [3]). Let Y_1, \dots, Y_n be independent, mean-zero, $\mathbb{R}^{m!-1}$ -valued random variables. Let $S = Y_1 + \dots + Y_n$, and let C^2 be the covariance matrix of S , assumed invertible. Let $\mathcal{N}(0, C^2)$ be a $m! - 1$ -dimensional Gaussian with mean zero and covariance C^2 . Then for any convex subset $X \subseteq \mathbb{R}^{m!-1}$,

$$|\Pr[S \in X] - \Pr[\mathcal{N}(0, C^2) \in X]| \leq O((m! - 1)^{1/4}) \cdot \left(\sum_{i=1}^n \mathbb{E}[|C^{-1}Y_i|^3] \right).$$

In order to apply the Berry-Esseen bound, we use the properties of the covariance matrix $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ from Lemma 26.

By Lemma 26, $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ is invertible, so we can apply the Berry-Esseen bound. For consistency with the form of the stated bound, we first translate both our distribution, $\mathcal{H}(\mathcal{S}_\phi(\pi))$, and the Gaussian to which we want to show it converges, to make both mean-zero. That is, we will show the equivalent statement about $\mathcal{H}(\mathcal{S}_\phi(\pi)) - \mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi))]$ approaching $\mathcal{N}(0, \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))])$. Note that subtracting the expectations of both $\mathcal{H}(\mathcal{S}_\phi(\pi))$ and $\mathcal{N}(\mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi))], \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi))])$ translates the the distributions identically, and note that the convexity of the quantified sets X is invariant under

translation. Therefore, proving the claim on the translated version of our distribution implies the claim on our original distribution.

By linearity of expectation, we can express our translated distribution as the sum of n independent random variables:

$$\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi})) - \mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))] = \sum_{i=1}^n \frac{1}{n} (\mathcal{H}(\mathcal{S}_\phi(\pi_i)) - \mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]).$$

We let Y_i be the random variable distributed as a single term of the above sum, $\frac{1}{n}(\mathcal{H}(\mathcal{S}_\phi(\pi_i)) - \mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))])$. Note that Y_i has mean zero, and covariance $\frac{1}{n^2} \cdot \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$.

Now, all that remains to show is that the following bound on the convergence rate holds:

$$O((m! - 1)^{1/4}) \cdot \left(\sum_{i=1}^n \mathbb{E} \left[\left| \text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))]^{-1/2} Y_i \right|^3 \right] \right) \leq \frac{O((m! - 1)^{1/4})}{(\lambda^{\min, \mathcal{S}, \phi})^{3/2}} \cdot \frac{1}{\sqrt{n}}.$$

We note that the exponentiated $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))]^{-1/2}$ in this expression is well defined because the matrix is symmetric.

We first show that $|Y_i| \leq 2/n$ for all i , where $|Y_i|$ denotes the L_2 norm of Y_i . Recall that $\mathcal{H}(\mathcal{S}_\phi(\pi_i))$ is always either a basis vector or the all 0s vector, and $\mathbb{E}[\mathcal{H}(\mathcal{S}_\phi(\pi_i))]$ is the vector whose entry corresponding to π' is $\Pr[\mathcal{S}_\phi(\pi) = \pi']$. Then, after subtracting the second vector from the first, the negative entries in the resulting vector can sum in magnitude to at most the sum of these probabilities, which is at most 1. Similarly, the positive entries can also sum to at most 1. hence, the L_1 norm before scaling by $1/n$ is at most 2. Using the fact that L_2 norms are at most L_1 norms, we get that this continues to hold for the L_2 norm. After dividing by n , we get that $|Y_i| \leq 2/n$, as needed.

Now, per Lemma 26, $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))]$ has minimum eigenvalue at least $\frac{\text{MIN-PROB}(S)}{m!n}$. It therefore holds that $\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))]^{-1/2}$ has maximum eigenvalue at most

$$\left(\frac{\text{MIN-PROB}(S)}{m!n} \right)^{-1/2} = \sqrt{\frac{m!n}{\text{MIN-PROB}(S)}}.$$

Multiplying a vector by a matrix can scale the norm by at most the matrix's maximum eigenvalue. Thus, combined with our observation that $|Y_i| \leq 2/n$, the following bound will always hold:

$$|\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))]^{-1/2} Y_i|^3 \leq \left(\sqrt{\frac{m!n}{\text{MIN-PROB}(S)}} \cdot \frac{2}{n} \right)^3 = \frac{8 \cdot (m!)^{3/2}}{n^{3/2} \cdot \text{MIN-PROB}(\mathcal{S}_\phi)^{3/2}}$$

Because this bound holds deterministically on the term above, it must hold also for the expectation of the term above. Thus, by summing over all i , we get that

$$\sum_{i=1}^n \mathbb{E}[|\text{Cov}[\mathcal{H}(\mathcal{S}_\phi(\boldsymbol{\pi}))]^{-1/2} Y_i|^3] \leq \frac{8(m!)^{3/2}}{\sqrt{n} \cdot \text{MIN-PROB}(\mathcal{S}_\phi)^{3/2}}.$$

Since $O((m! - 1)^{1/4}) \cdot 8m!^{3/2} \in O((m!)^{7/4})$, multiplying by the Berry-Esseen constant yields the lemma statement. \square

C.5 Proof of Lemma 10

Fix \mathcal{S} , C , and $X = \bigcup_{j=1}^{\ell} X_j$ where each X_j is convex, X is measure 0, and $\mathcal{H}(\Pi^{-C}) \subseteq X$. Since X is measure 0, this implies each X_j has measure 0. Fix $\phi \in (0, 1]$, a profile $\pi \in \Pi_n$, and $\phi' \in [\phi, 1]$.

Fix an arbitrary X_j . Note that since X_j has measure zero, the probability mass placed on X_j by a Gaussian with invertible covariance matrix is 0. Hence, Lemma 9 immediately imply that $\Pr[\mathcal{H}(\mathcal{S}_{\phi'}(\pi)) \in X_j] - 0 \leq \frac{O((m!)^{7/4})}{\sqrt{n} \cdot \text{MIN-PROB}(\mathcal{S}_{\phi'})^{3/2}}$. Using the monotonicity of MIN-PROB (Assumption 4), we get that this is at most $\frac{O((m!)^{7/4})}{\sqrt{n} \cdot \text{MIN-PROB}(\mathcal{S}_{\phi})^{3/2}}$ (with ϕ' replaced with ϕ). Union bounding over all ℓ sets X_j tells us that

$$\Pr[\mathcal{H}(\mathcal{S}_{\phi'}(\pi)) \in X] \leq \frac{\ell \cdot O((m!)^{7/4})}{\sqrt{n} \cdot \text{MIN-PROB}(\mathcal{S}_{\phi})^{3/2}},$$

as needed. \square

C.6 Proof of Theorem 12

Proof. Fix \mathcal{S} , ϕ , and $\delta(n)$. Next, fix a number n , a hyperplane $G \in \mathcal{G}$, a $\phi' \in [\phi, 1]$, and a $\pi^* \in \Pi_n$. Here, G is a hyperplane in $m! - 1$ dimensions, meaning it is defined by a linear equation whose variables are the profile proportions for rankings in \mathcal{L}_{-1} , each weighted by a coefficient, a_{π} , along with a constant, b . In other words, it is the set of proportions \mathbf{h} satisfying

$$b = \sum_{\pi \in \mathcal{L}_{-1}} a_{\pi} \cdot h_{\pi}. \quad (2)$$

Formally, we need to consider only the intersection of G with our convex hull H , but for our results, it is irrelevant whether we formally make this restriction.

We will upper bound $\Pr[d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G) \leq \delta(n)]$ by an $o(1)$ function of n (allowing this convergence rate to depend on ϕ , m , and \mathcal{S} , but not G , ϕ' , or π^*).

Recall that the hyperplane G has a coefficient for each ranking except the “missing” ranking, π_{-1} . We choose $\pi^{max} \in \mathcal{L}$ such that π^{max} is the ranking corresponding to a largest-magnitude coefficient in the definition of G , i.e., $\pi^{max} \in \arg\max_{\pi \in \mathcal{L}_{-1}} |a_{\pi}|$. Now, we will define two types of events.

First, let $\mathcal{E}^{few} = \{\pi \in \Pi_n : |\{i \in N : \pi_i \in \{\pi^{max}, \pi_{-1}\}\}| < np\}$ — that is, \mathcal{E}^{few} is the set of profiles in which fewer than np agents end up voting either π^{max} or π_{-1} in $\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*))$.

Let $\mathcal{V}^{\geq np} = \{V \subseteq N : |V| \geq np\}$ be the collection of all sets of agents of size at least np . For a set of such agents $V \in \mathcal{V}^{\geq np}$, we denote the complement of this set of agents as $\bar{V} = N \setminus V$. Now, slightly abusing notation, define $\tilde{\Pi}^{\bar{V}} = \{(\pi_i)_{i \in \bar{V}} : \pi_i \in \mathcal{L} \setminus \{\pi^{max}, \pi_{-1}\}\}$ to be the set of all *partial* profiles in which all agents in \bar{V} have a ranking other than π^{max} or π_{-1} . For $V \in \mathcal{V}^{\geq np}$ and partial profile $\tilde{\pi} \in \tilde{\Pi}^{\bar{V}}$, let $\mathcal{E}^{V, \tilde{\pi}}$ be the event that agents in V either vote π or π' , and all other agents vote as in $\tilde{\pi}$, that is,

$$\mathcal{E}^{V, \tilde{\pi}} = \{\pi \in \Pi : \pi_i = \tilde{\pi}_i \text{ for } i \in \bar{V} \text{ and } \pi_i \in \{\pi^{max}, \pi_{-1}\} \text{ for } i \in V\}.$$

In the remainder of the proof, we will use that the event \mathcal{E}^{few} , along with the set of all events of the form $\mathcal{E}^{V, \tilde{\pi}}$ (that is, for all subsets of agents $V \in \mathcal{V}^{\geq np}$ and partial profiles $\tilde{\pi} \in \tilde{\Pi}^{\bar{V}}$), form a partition of the full space of profiles. To see how we will use this, recall that we are trying to show that $d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G)$ is likely to be somewhat large. We will show this by showing that,

conditioned on any individual event $\mathcal{E}^{V,\tilde{\pi}}$, the probability of this distance being large is fairly high, and the remaining event, \mathcal{E}^{few} is unlikely to occur.

We will show two claims:

1. There is some $f(n) \in o(1)$ such that

$$\Pr \left[\mathcal{E}^{few} \right] \leq f(n).$$

2. There is some $g(n) \in o(1)$ such that for all $\mathcal{E}^{V,\pi}$,

$$\Pr \left[d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G) \leq \delta(n) \mid \mathcal{E}^{V,\tilde{\pi}} \right] \leq g(n).$$

We first show that together these are sufficient to prove the bound. Indeed, by the law of total probability, as these events form a partition

$$\begin{aligned} \Pr \left[d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G) \leq \delta(n) \right] &= \Pr \left[d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G) \leq \delta(n) \mid \mathcal{E}^{few} \right] \cdot \Pr \left[\mathcal{E}^{few} \right] \\ &\quad + \sum_{\mathcal{E}^{V,\pi}} \Pr \left[d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G) \leq \delta(n) \mid \mathcal{E}^{V,\pi} \right] \cdot \Pr \left[\mathcal{E}^{V,\pi} \right] \\ &\leq \Pr \left[d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G) \leq \delta(n) \mid \mathcal{E}^{few} \right] \cdot f(n) + \sum_{\mathcal{E}^{V,\pi}} g(n) \cdot \Pr \left[\mathcal{E}^{V,\pi} \right] \\ &= \Pr \left[d(\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*)), G) \leq \delta(n) \mid \mathcal{E}^{few} \right] \cdot f(n) + g(n) \sum_{\mathcal{E}^{V,\pi}} \Pr \left[\mathcal{E}^{V,\pi} \right] \\ &\leq 1 \cdot f(n) + g(n) \cdot 1 \\ &= f(n) + g(n) \in o(1). \end{aligned}$$

We now show the claims. The first claim follows from a straightforward Chernoff bound: each of the n agents places probability at least $2p$ on ending up in either π^{max} or π_{-1} , hence the expected number of agents with either π^{max} or π_{-1} is at least $2np$. Then, given that agents' rankings are sampled independently, the probability of the total number of such agents being less than half of this expectation is exponentially small in n .

We now show the second claim. Fix an arbitrary $\mathcal{E}^{V,\tilde{\pi}}$ and let us consider the conditional distribution of $\mathcal{H}(\mathcal{S}_{\phi'}(\pi^*))$ conditioned on this event. Note that all proportions in the support of this distribution match on all entries except that of π^{max} . Let $\tilde{\mathbf{h}}_{-\pi^{max}}$ be the partial proportions over all these matching entries, i.e., all rankings except π^{max} , so that $\mathbf{h}_{-\pi^{max}} = \tilde{\mathbf{h}}_{-\pi^{max}}$ for all $\mathbf{h} \in \mathcal{H}(\mathcal{E}^{V,\tilde{\pi}})$.

Now, since $a_{\pi^{max}} \neq 0$, by Equation (2), there exists exactly one completion of the partial proportions $\mathbf{h}_{-\pi^{max}}$ such that the resulting proportion lies on G . Namely,

$$p^{V,\tilde{\pi}} := \frac{b - \sum_{\pi \in \mathcal{L} \setminus \{\pi^{max}, \pi_{-1}\}} a_{\pi} \cdot \tilde{h}_{\pi}}{a_{\pi^{max}}}.$$

Note that this completion need not be a valid proportion, as the value assigned to the proportion π^{max} , $p^{V,\tilde{\pi}}$, could need to be irrational, negative, or make all proportions add up to strictly more than one.

Now, we consider all completions of $\mathbf{h}_{-\pi^{max}}$ which assign values to the proportion of ranking π^{max} at least $\delta(n)$ larger or smaller than $p^{V,\tilde{\pi}}$. In particular, we will show (1) that all such proportions are at least $\delta(n)$ L_1 distance from all points on G , and (2) that we are likely to draw such a profile from $\mathcal{H}(\mathcal{S}_\phi(\pi^*))$ conditioned on $\mathcal{E}^{V,\tilde{\pi}}$.

Fix such a completion of $\tilde{\mathbf{h}}_{-\pi^{max}}$, call it \mathbf{h} , so that $|h_{\pi^{max}} - p^{V,\tilde{\pi}}| > \delta(n)$. By this assumption, we have the following, where in the first step we replace b according to Equation (2):

$$\begin{aligned} \left| \sum_{\pi \in \mathcal{L}_{-1}} a_\pi h_\pi - b \right| &= \left| \sum_{\pi \in \mathcal{L}_{-1}} a_\pi h_\pi - \left(p^{V,\tilde{\pi}} \cdot a_{\pi^{max}} + \sum_{\pi \in \mathcal{L} \setminus \{\pi^{max}, \pi_{-1}\}} a_\pi h_\pi \right) \right| \\ &= |a_{\pi^{max}} \cdot h_{\pi^{max}} - p^{V,\tilde{\pi}} \cdot a_{\pi^{max}}| \\ &> |\delta(n) \cdot a_{\pi^{max}}| \end{aligned}$$

Since $a_{\pi^{max}}$ is maximal in magnitude over all coefficients in G , the $\delta(n)$ -radius L_1 -ball around \mathbf{h} does not intersect G , because changing any entry of \mathbf{h} by at most $\delta(n)$ cannot change $\sum_{\pi \in \mathcal{L}_{-1}} a_\pi p_\pi$ by more than $\delta(n) \cdot a_{\pi^{max}}$.

We have shown that all completions of $\mathbf{h}_{-\pi^{max}}$ assigning values of π^{max} outside $p^{V,\tilde{\pi}} \pm \delta(n)$ must be more than $\delta(n)$ L_1 distance from G . Next, we show that the probability of drawing such a completion is likely. We will do this by upper-bounding the probability placed on $h_{\pi^{max}}$ falling in the interval $p^{V,\tilde{\pi}} \pm \delta(n)$ by showing its distribution is approximately normal using the Berry-Esseen bound, and showing the normal distribution does not place much mass on this small interval.

To this end, let I_i be the indicator random variable that agent $i \in V$ chooses π^{max} (if they do not choose π^{max} , they choose π_{-1}). Let $S = \sum_{i \in V} I_i$ be the random variable representing the total number of agents in V that vote for π^{max} . Note that the resulting proportion $\mathcal{H}_{\pi^{max}}$ will be in the range $p^{V,\tilde{\pi}} \pm \delta(n)$ iff S is in the range $n(p^{V,\tilde{\pi}} \pm \delta(n))$, an interval of size $2n\delta(n)$. We now show that the probability S is in *any* interval of size $2n\delta(n)$ is $o(1)$.

To do this, we first state the Berry-Esseen bound.

LEMMA 28 (Berry-Esseen Bound [4, 9]). Let X_1, \dots, X_n be independent random variables with $\mathbb{E}[X_i] = \mu_i$, $\mathbb{E}[|X_i^2 - \mu_i|] = \sigma_i^2 > 0$, and $\mathbb{E}[|X_i^3 - \mu_i|] = \rho_i < \infty$. Let $T = X_1 + \dots + X_n$. Let F be the cdf of $\frac{T - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$. Then, there exists an absolute constant C_1 such that

$$\sup_{x \in \mathbb{R}} |F(x) - \Phi(x)| \leq C_1 \cdot \left(\sum_{i=1}^n \sigma_i^2 \right)^{-1/2} \cdot \max_{1 \leq i \leq n} \frac{\rho_i}{\sigma_i^2}.$$

We would like to directly apply Lemma 28 to our S , where each $X_i = I_i$. We first consider the quantity σ_i^2 for each i . Note that $\text{MIN-PROB}(\mathcal{S}'_\phi) \leq \Pr[I_i = 1] = 1 - \Pr[I_i = 0] \leq 1 - \text{MIN-PROB}(\mathcal{S}'_\phi)$. Further, by monotonicity (Assumption 4), $\text{MIN-PROB}(\mathcal{S}'_\phi) \geq \text{MIN-PROB}(\mathcal{S}_\phi)$, so $\text{MIN-PROB}(\mathcal{S}_\phi) \leq \Pr[I_i = 1] \leq 1 - \text{MIN-PROB}(\mathcal{S}_\phi)$. Since each I_i is a Bernoulli, each $\sigma_i^2 \geq \text{MIN-PROB}(\mathcal{S}_\phi) \cdot (1 - \text{MIN-PROB}(\mathcal{S}_\phi)) \geq \text{MIN-PROB}(\mathcal{S}_\phi)/2$. This implies that $\sum_{i=1}^n \sigma_i^2 \geq n \cdot \text{MIN-PROB}(\mathcal{S}_\phi)/2$. Further, note that since each I_i is bounded in $[0, 1]$, each $\rho_i \leq 1$. This implies the error bound of Lemma 28 is at most

$$C_1 \cdot \left(n \cdot \frac{\text{MIN-PROB}(\mathcal{S}_\phi)}{2} \right)^{-1/2} \cdot \frac{2}{\text{MIN-PROB}(\mathcal{S}_\phi)} = O \left(\frac{1}{\text{MIN-PROB}(\mathcal{S}_\phi)^{3/2} \cdot \sqrt{n}} \right).$$

Further, the probability S is in some range of size $2n\delta(n)$ will be equivalent to the probability $\frac{S - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$ is in a specific range of size $\frac{2n\delta(n)}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \leq \frac{2\delta(n)\sqrt{2n}}{\sqrt{\text{MIN-PROB}(S_\phi)}}$. Let F be the CDF of this random variable. The probability this random variable is in a range of this size is at most

$$\begin{aligned}
& \sup_x \left(F \left(x + \frac{2\delta(n)\sqrt{2n}}{\sqrt{\text{MIN-PROB}(S_\phi)}} \right) - F(x) \right) \\
& \leq \sup_x \left(\Phi \left(x + \frac{2\delta(n)\sqrt{2n}}{\sqrt{\text{MIN-PROB}(S_\phi)}} \right) - \Phi(x) \right) + 2O \left(\frac{1}{\text{MIN-PROB}(S_\phi)^{3/2} \cdot \sqrt{n}} \right) \\
& = \sup_x \left(\int_x^{x + \frac{2\delta(n)\sqrt{2n}}{\sqrt{\text{MIN-PROB}(S_\phi)}}} \phi(x) \right) + 2O \left(\frac{1}{\text{MIN-PROB}(S_\phi)^{3/2} \cdot \sqrt{n}} \right) \\
& \leq \sup_x \left(\int_x^{x + \frac{2\delta(n)\sqrt{2n}}{\sqrt{\text{MIN-PROB}(S_\phi)}}} \frac{e}{\sqrt{2\pi}} \right) + 2O \left(\frac{1}{\text{MIN-PROB}(S_\phi)^{3/2} \cdot \sqrt{n}} \right) \\
& = \sup_x \left(\frac{e}{\sqrt{2\pi}} \cdot \frac{2\delta(n)\sqrt{2n}}{\sqrt{\text{MIN-PROB}(S_\phi)}} \right) + 2O \left(\frac{1}{\text{MIN-PROB}(S_\phi)^{3/2} \cdot \sqrt{n}} \right) \\
& = O \left(\frac{1}{\sqrt{\text{MIN-PROB}(S_\phi)}} \cdot \delta(n)\sqrt{n} \right) + O \left(\frac{1}{\text{MIN-PROB}(S_\phi)^{3/2} \cdot \sqrt{n}} \right) \\
& \in o(1).
\end{aligned}$$

This is our desired $g(n)$ which completes the proof of the second claim. □

D Supplemental Materials for Section 4

D.1 Proof of Lemma 16

Proof. We will prove these bounds by lower-bounding $q(k)$. First, we have that

$$q(k) = \frac{1}{1 - \phi^{k+1}} \left(1 - \frac{(1 - \phi)k\phi^k}{1 - \phi^k} \right) \geq 1 - \frac{(1 - \phi)k\phi^k}{1 - \phi^k} \geq 1 - k\phi^k.$$

Second, by the AM-GM inequality, we have that

$$q(k) = \frac{1}{1 - \phi^{k+1}} \left(1 - \frac{(1 - \phi)k\phi^k}{1 - \phi^k} \right) \geq 1 - \frac{(1 - \phi)k\phi^k}{1 - \phi^k}.$$

Simplifying, this is equal to $1 - \frac{k\phi^k}{\sum_{j=0}^{k-1} \phi^j}$, and the second upper-bound follows:

$$1 - \frac{k\phi^k}{\sum_{j=0}^{k-1} \phi^j} = 1 - \frac{1}{\frac{1}{k} \sum_{j=1}^k \frac{1}{\phi^j}} \geq 1 - \frac{1}{\left(\prod_{j=1}^k \frac{1}{\phi^j} \right)^{1/k}} = 1 - \phi^{\frac{k(k+1)}{2} \cdot \frac{1}{k}} = 1 - \phi^{\frac{k+1}{2}} \geq 1 - \phi^{k/2}.$$

Third,

$$q(k) \geq q(1) = \frac{1}{1 - \phi^2} \left(1 - \frac{(1 - \phi)\phi}{1 - \phi} \right) = \frac{1 - \phi}{1 - \phi^2} = \frac{1}{1 + \phi} = 1 - \frac{\phi}{1 + \phi}. \quad \square$$

D.2 Proof of Proposition 18

Proof. We begin with RESOLVABILITY and show $\rho(n)$ -GROUP-STABILITY later. Fix a starting profile π and two candidates, a and b . Our approach will be to first upper bound the probability that a and b have the same Borda score post-noise, and then union bound over all pairs of candidates. For a candidate c , let $X_i^c \in \{0, \dots, m-1\}$ be the random variable representing the number of Borda points given to candidate c by voter i , post-noise. Let $Y_i = X_i^a - X_i^b$ be the difference in Borda points given to a and b by voter i . Our goal is to upper bound $\Pr[\sum_i Y_i = 0]$. As in the proof for PLURALITY, we will do so via use Berry-Esseen using that the distribution of $\sum_i Y_i$ converges to a Gaussian. Since the Gaussian places zero mass on the value zero, we simply upper bound $\Pr[\sum_i Y_i = 0]$ by the convergence rate of its distribution to a Gaussian, given by Berry-Essen exactly as in Equation (1).

Again, bounding this upper bound's component terms: Note that $Y_i \in [-(m-1), (m-1)]$, so $|Y_i - \mathbb{E}[Y_i]| \leq 2m$, implying the same bound on $\max_i \frac{\rho_i}{\sigma_i^2}$. To lower bound each σ_i^2 , we consider the distribution of Y_i . In particular, we will upper bound the probability of any specific (integer) value by $\frac{1}{1+\phi}$. That is, for all values ℓ ,

$$\Pr[Y_i = \ell] \leq \frac{1}{1 + \phi} \quad (3)$$

Once we have shown Inequality (3) holds, note that at most one value ℓ can be within $1/2$ of $\mathbb{E}[Y_i]$, so this will imply that with probability at least $1 - \frac{1}{1+\phi} = \frac{\phi}{1+\phi} \geq \frac{\phi}{2}$, $|Y_i - \mathbb{E}[Y_i]| \geq 1/2$. This allows us to lower bound the variance by $\frac{\phi}{2} \cdot (\frac{1}{2})^2 = \frac{\phi}{8}$. This gives us an overall bound of $O(\frac{m}{\sqrt{n\phi}})$, and therefore $O(\frac{m^3}{\sqrt{n\phi}})$ after a union bound.

Finally, we show Inequality (3). Fix a value ℓ . Let $\tau \in \mathcal{L}(M \setminus \{a\})$ be an arbitrary partial ranking of the candidates without ℓ . We will write $\sigma \setminus \{a\} = \tau$ when the partial ranking of σ without a included matches τ . We will show that when σ is sampled from a Mallows model, conditioned on it matching τ , the probability a is ℓ positions above b is at most $\frac{1}{1+\phi}$. Since τ was arbitrary, by law of total probability, this implies the same bound in general. Indeed, note that once τ is fixed, there are m possible rankings after the insertion of a . At most one can correspond to a difference of ℓ . Call this ranking σ . Further, consider inserting a in the neighboring position. We will call this ranking σ' . Note that the Kendall tau distance from the starting ranking σ_i of σ and σ' can differ by at most 1. This implies that $\Pr[\sigma' | \sigma \setminus \{a\} = \tau] / \Pr[\sigma | \sigma \setminus \{a\} = \tau] \geq \phi$. This directly implies that $\Pr[\sigma | \sigma \setminus \{a\} = \tau] \leq \frac{1}{1+\phi}$, as needed.

For $\rho(n)$ -GROUP-STABILITY, note that a single voter can affect the difference in scores of a candidate by at most $2(m-1)$. Hence, as long as the maximal candidate is winning by at least $2(m-1)\rho(n)$, no group of size $\rho(n)$ can change the outcome. Rather than just doing the analysis for $\sum_i Y_i = 0$, we can do the same for all values in $[-2(m-1)\rho(n), 2(m-1)\rho(n)]$, which, via a union bound, adds an additional $O(m\rho(n))$ factor. \square

D.3 Proof of Proposition 20

Proof. We begin with RESOLVABILITY. Fix some n that is divisible by 4 and first, assume m is even (we will describe how to extend it to the odd case later). Consider the following instance. Let a and b be two candidates and let C and D be a partition of the remaining candidates into two sets of equal size, each having $m/2 - 1$ candidates. Let τ^C and τ^D be arbitrary orders of the candidates in C and D respectively, and let $\text{rev}(\tau^C)$ and $\text{rev}(\tau^D)$ be the reversals of these rankings respectively. Finally, assume the voter rankings are as follows, each occurring $n/4$ times:

1. $b \succ \text{rev}(\tau^C) \succ a \succ \tau^D$
2. $b \succ \text{rev}(\tau^D) \succ a \succ \tau^C$
3. $a \succ \text{rev}(\tau^C) \succ b \succ \tau^D$
4. $a \succ \text{rev}(\tau^D) \succ b \succ \tau^C$

We will show that except for with exponentially small probability, a and b still beat every other candidate in a pairwise competition. However, with a reasonably large probability, a and b are pairwise tied. When these events occur, a and b have the same optimal Minimax score of $n/2$, while all others have strictly less, and are hence tied.

For the first, consider an arbitrary candidate $c \in C$. The arguments for if $c \in D$ or with b instead of a will be symmetric. Let X_i be a random variable that is 1 if i ranks a above c in the final ranking and -1 otherwise. We would like to upper bound the probability that $\sum_i X_i \leq 0$. We will do this using an application of Hoeffding's inequality. To do this, we must obtain bounds on $\mathbb{E}[\sum_i X_i]$.

Suppose c is ranked k 'th in τ^C . Note that c is only preferred to a on voters in the first group. For such a voter i , as there is a probability $\bar{q}(k)$ of a swap, so $\mathbb{E}[X_i] = 2\bar{q}(k) - 1$. For voters in the second group, a is preferred to c by the same magnitude, hence, for such a voter i , $\mathbb{E}[X_i] = 1 - 2\bar{q}(k)$, the exact negative of the first. For voters in the third group, since a is preferred to c , they retain their order with probability at least $1/2$, so for such voters i , $\mathbb{E}[X_i] \geq 0$. For voters in the fourth

group, a is preferred to c by $|D| + 1 + k$ positions, hence, $\mathbb{E}[X_i] \geq 2q(|D| + 1 + k) - 1$. Since \bar{q} is decreasing and $|D| + 1 + k \geq m/2$, this is at least $1 - 2\bar{q}(m/2)$. Putting this together, we have that

$$\mathbb{E}[\sum_i X_i] \geq n/4 * (2q(k) - 1) + n/4(1 - 2q(k)) + n/4(1 - 2\bar{q}(m/2)) = n/4(1 - 2\bar{q}(m/2)).$$

Plugging this into Hoeffding's inequality, since the variables take on values in $[-1, 1]$, this implies that

$$\Pr[\sum_i X_i \leq 0] \leq \exp\left(-\frac{n(1 - 2\bar{q}(m/2))^2}{32}\right).$$

Union bounding over all m candidates choices of c and 2 choices of a and b , we get that the probability a and b pairwise beat all other candidates except for with probability $2m \exp\left(-\frac{n(1 - 2\bar{q}(m/2))^2}{32}\right)$.

Next, we consider the relative rankings of a and b . Note that on groups one and two, b is beating a , and on groups three and four, a is beating b . In all cases, the margin of victory is $m/2$ positions. The probability of a swap on any of these rankings is $\bar{q}(m/2)$. Hence, the number of swaps on the first $n/2$ candidates follows a $\text{Bin}(n/2, \bar{q}(m/2))$ distribution, and symmetrically, so does the number of swaps on the second $n/2$. Candidates a and b will be tied in a pairwise competition if the two (independent) draws from these binomials are equal. We show next that this occurs with probability $\Omega(\frac{1}{\sqrt{n\bar{q}(m/2)}})$ which follows directly from the following lemma.

LEMMA 29. The probability two independent draws from a $\text{Bin}(n, p)$ distribution are equal is $\Omega(\frac{1}{\sqrt{np}})$.

Proof. Fix an arbitrary n and p . Note that the mean of a binomial distribution is np . By standard Chernoff bounds, if X follows a $\text{Bin}(n, p)$ distribution, $\Pr[|X - \mu| \geq \delta\mu] \leq 2\exp(-\delta^2\mu/3)$ for all $\delta \in (0, 1)$ where $\mu = np$. In particular, plugging in $\delta = 2/\sqrt{\mu}$, we get that with constant probability ($\geq 1 - 2\exp(-4/3) \geq .47$), X is within $2\sqrt{\mu}$ of its mean. Note that there are at most $4\sqrt{np} + 1$ integer values within $2\sqrt{np}$ of the mean. With constant probability ($\geq .47^2$), both binomials will take on one of these values. Conditioned on this event, the probability they are equal is at least $\frac{1}{4\sqrt{np}+1}$. Indeed, suppose conditioned on being one of these $\frac{1}{4\sqrt{np}+1}$ values, each was taken with probability $p_1, p_2, \dots, p_{4\sqrt{np}+1}$ with $\sum_{i=1}^{4\sqrt{np}+1} p_i = 1$. The probability they are equal is $\sum_{i=1}^{4\sqrt{np}+1} p_i^2 \geq \sum_i \frac{p_i}{4\sqrt{np}+1} = \frac{1}{4\sqrt{np}+1}$. Hence, the probability they are equal is at least $\Omega(\frac{1}{\sqrt{np}})$. \square

Combining the fact that a and b are pairwise tied with probability $\Omega(\frac{1}{\sqrt{n\bar{q}(m/2)}})$ and pairwise beat all others with probability $1 - O\left(\exp(-\frac{n(1 - 2\bar{q}(m/2))^2}{32})\right)$ completes the bound.

To handle m which is odd, we can add a dummy candidate to the end of all the rankings. This candidate wins with even smaller probability, so all of the analysis continues to hold. We simply need to replace $\bar{q}(m/2)$ with $\bar{q}(\lfloor m/2 \rfloor)$.

Next, consider 1-STRATEGY-PROOFNESS. We will have the following starting profile. Let a, b, x, y be candidates, and let S_1, S_2, S_3, S_4 be sets of candidates of size $\lfloor \frac{m}{4} \rfloor$ (as before, we will put remaining candidates at the bottom, but this will not affect the analysis much).

$$\begin{array}{ll}
(n-2)/4 \text{ voters:} & a \succ \text{rev}(\tau_{S_1}) \succ x \succ \text{rev}(\tau_{S_2}) \succ b \succ \tau_{S_3} \succ y \succ \tau_{S_4} \\
(n-2)/4 \text{ voters:} & a \succ \text{rev}(\tau_{S_4}) \succ x \succ \text{rev}(\tau_{S_3}) \succ b \succ \tau_{S_2} \succ y \succ \tau_{S_1} \\
(n-2)/4 \text{ voters:} & b \succ \text{rev}(\tau_{S_1}) \succ y \succ \text{rev}(\tau_{S_2}) \succ a \succ \tau_{S_3} \succ x \succ \tau_{S_4} \\
(n-2)/4 \text{ voters:} & b \succ \text{rev}(\tau_{S_4}) \succ y \succ \text{rev}(\tau_{S_3}) \succ a \succ \tau_{S_2} \succ x \succ \tau_{S_1} \\
1 \text{ voter:} & x \succ \text{rev}(\tau_{S_2}) \succ b \succ \text{rev}(\tau_{S_4}) \succ y \succ \tau_{S_3} \succ a \succ \tau_{S_1} \\
1 \text{ voter:} & y \succ \text{rev}(\tau_{S_1}) \succ a \succ \text{rev}(\tau_{S_3}) \succ b \succ \tau_{S_4} \succ x \succ \tau_{S_2}
\end{array}$$

The analysis will begin similarly to RESOLVABILITY, showing that a and b both beat all candidates outside of $\{a, b, x, y\}$ by a pairwise majority with all but exponentially small probability. Then, we will show that it is reasonably likely that the relative orders of a, b, x , and y remain exactly as they are in the starting profile for all voters. Under this condition, we can check that the winner is in fact b . Indeed, b at minimum pairwise ties with all others, while for all other candidates, we can find a candidate that strictly pairwise beats them, namely a beats x on $n-1$ votes, b beats y on $n-1$ votes, and y beats a on $n/2+1$ votes, and the rest are beaten by both a and b . However, if the last voter bumps up a to the top of their ranking and pushes b to the bottom to instead report $a \succ y \succ \text{rev}(\tau_{S_1}) \succ \text{rev}(\tau_{S_3}) \succ \tau_{S_4} \succ x \succ \tau_{S_2} \succ b$, a will now pairwise beat all other candidates while b will be beaten by x on $n/2+1$ votes. This leads a to be the unique winner instead of b , an improvement according to manipulating voter.

We now compute the relevant probability bounds. One can check as we did with resolvability that for any candidate $j \in S_i$, even if they beat one of a or b on a certain number of votes, these can be matched with an equal number where a and b beat them by the same amount. For example, consider a candidate $j \in S_1$ ranked k 'th in τ_{S_1} . Although they beat a on the third group of voters, this can be matched to the voters in the fourth group that prefer a to j by the exact same margin. Further, although the last voter prefers j to a , the second to last prefers a to j by the same margin. As before, there is an additional $(n-2)/4$ size group that prefers a to j by a margin of at least $\lfloor m/2 \rfloor$. Hence, by the same analysis as for resolvability, we get that a and b strictly pairwise beat all other candidates except for with probability $2m \exp\left(\frac{-(n-2)(1-2\bar{q}(\lfloor m/2 \rfloor))^2}{32}\right)$.

Next, we consider the probability that an individual voter maintains their order of $\{a, x, b, y\}$. Suppose, without loss of generality, they rank $a \succ x \succ b \succ y$. Note that as long as they continue to rank $a \succ x$, $x \succ b$, and $b \succ y$, then by transitivity, the entire partial ranking will remain. Each of these occurs with probability $1 - q(\lfloor m/4 \rfloor)$ as there is always a gap at least this large between the candidates. Hence, union bounding over all voters, we get that all will maintain their orders with probability at least $1 - 3nq(\lfloor m/4 \rfloor)$. Putting this together we have that the overall convergence rate is lower bounded by $1 - 3nq(\lfloor m/4 \rfloor) - O\left(m \exp\left(\frac{-(n-2)(1-2\bar{q}(\lfloor m/2 \rfloor))^2}{32}\right)\right)$. \square

E Supplemental Material from Section 5

COROLLARY 30 (Gibbard-Satterthwaite). Per Table 2, there exist several voting rules that are simultaneously non-dictatorial, permit $m \geq 3$, and smoothed-satisfy STRATEGYPROOFNESS.¹⁹

COROLLARY 31 (ANR). Per Table 1, there exist several voting rules that simultaneously satisfy ANONYMITY and NEUTRALITY and smoothed-satisfy RESOLVABILITY.

¹⁹ \sqrt{n} -GROUP-STRATEGYPROOFNESS \implies 1--GROUP-STRATEGYPROOFNESS, i.e., STRATEGYPROOFNESS.

COROLLARY 32 (Condorcet-Participation). Per Theorem 11, there exist several voting rules that simultaneously satisfy CONDORCET and smoothed-satisfy PARTICIPATION.²⁰

More general criteria. Here; we consider the more general “criterion” that π contains no Condorcet cycle. Note that this is not strictly an axiom, as whether it holds on a profile π depends only on the profile itself—not on any voting rule. To be formal, we think of this as a more generic criterion $C : \Pi \rightarrow \{\text{true}, \text{false}\}$, representing the true/false statement “ C holds on π ”. We can think of $A(R)$ as an example of such a generic criterion.

PROPOSITION 33. Let C be the criterion that π contains no Condorcet Cycle. C is smoothed-violated.

Proof. This proposition is proven using the same procedure as for criterion of the form $A(R)$, as before: we find a profile where all nearby histograms have a Condorcet Cycle, implying the preconditions of Lemma 7. We define this profile to be any consistent with the histogram \mathbf{h} , in which $1/3$ of voters have $a \succ b \succ c$, $1/3$ of voters have $b \succ c \succ a$, and $1/3$ of voters have $c \succ a \succ b$. This profile has a Condorcet cycle: a pairwise-dominates b , b pairwise dominates c , and c pairwise dominates a .

Now, we will establish that \mathbf{h} satisfies the precondition of Lemma 7. Fix any histogram \mathbf{h}' such that $\|\mathbf{h} - \mathbf{h}'\|_1 < 1/6$. Assume for the sake of contradiction that \mathbf{h}' does not contain a Condorcet cycle. It follows that at least one relative pairwise domination flipped from \mathbf{h} to \mathbf{h}' . All pairwise dominations in \mathbf{h} were symmetric, so we wlog consider a ’s pairwise domination of b . In order for b to pairwise dominate a , at least $1/2 - 1/3 = 1/6$ of voters must have reversed their relative ranking of a and b from \mathbf{h} to \mathbf{h}' —i.e., $\|\mathbf{h}' - \mathbf{h}\|_1 \geq 1/6$. This is a contradiction, and we conclude the claim. \square

²⁰ \sqrt{n} -GROUP-PARTICIPATION \implies 1-GROUP-PARTICIPATION, i.e., PARTICIPATION.