

Improving Randomized Metric Distortion to 2.3282

Nisarg Shah
University of Toronto
nisarg@cs.toronto.edu

Abstract

In metric social choice, each voter ranks a set of m candidates by her distance to them in an unknown metric space. The cost of a candidate is its average distance to the voters. A randomized voting rule must use only the rankings to choose a lottery over candidates. Its distortion is the worst-case ratio between the expected cost under the lottery it returns and the cost of the best candidate. Charikar, Ramakrishnan, Wang, and Wu [CRWW24] prove an upper bound of 2.753, establishing a constant separation from deterministic rules, for which the best achievable distortion is 3. Independently, Frank [Fra26] and Ye [Ye26] improve the bound to 2.5, using an equal mixture of maximal lottery, a Nash equilibrium of the zero-sum game in which a random voter compares two candidates, and Integrated Veto, an adaptation of a well-established rule, in which every voter continuously vetoes her current least-preferred candidate. The existing arguments do not yield a better bound with any mixture of these rules.

We break this barrier with a new ingredient, a *random-size stable lottery*. Let D be a random variable over the domain of positive integers. A random-size stable lottery RSL_D guarantees that the probability of a random voter preferring any fixed candidate c to her favorite of D i.i.d. draws from RSL_D is at most $\mathbb{E}[1/(D+1)]$, where the probability also averages over D . When $D = 1$ deterministically, this is precisely a maximal lottery. More generally, when $D = k$ deterministically, this reduces to the stable k -lottery of Charikar, Ramakrishnan, Tan, and Wang [CRTW25], which they prove exists for every fixed positive integer k . Their minimax argument for a fixed k easily generalizes to a random D . Our main contribution is to show how stability with respect to a random D can be used to bound distortion. Specifically, by mixing a suitably chosen random-size stable lottery with Integrated Veto, we get distortion at most $11641/5000 = 2.3282$. The proof combines infinite-dimensional conic linear-programming duality, heuristic nonlinear optimization, and exact rational verification via nonnegativity testing of polynomials using the Bernstein basis.

All the proofs in this document were obtained using GPT-5.6-Sol with guidance from the author. The author verified the proofs for correctness, and significantly expanded on the exposition and simplified arguments, with the aid of GPT-5.6-Sol and Claude Opus 5.

Contents

1	Introduction	3
1.1	Our Results	4
1.2	Technical Overview	4
1.3	Related Work	7
2	Preliminaries	8
2.1	Analysis via Candidate Sets	8
2.2	Integrated Veto and Maximal Lotteries	9
3	Random-Size Stable Lotteries	11
4	From Random-Size Stability to Loss Ratio	13
4.1	Abstracting the Profile Into an Infinite-Dimensional LP	13
4.2	Simplified PRIMAL	16
4.3	Bounding Loss via DUAL	17
5	A Rule with Distortion at Most 2.3282	18
6	Discussion	20
A	Further Properties of Random-Size Stable Lotteries	26
B	Exact Verification and Numerical Discovery of the Distortion Bound	31

1 Introduction

Social choice theory studies how to choose one of several alternative decisions (candidates) by aggregating the conflicting preferences of multiple parties (voters). The norm is to elicit ranked preferences from voters over the candidates because exact utilities or costs are often difficult for voters to identify and report. Consequently, much of the centuries-old literature has sought voting rules that satisfy qualitative axioms defined on this ordinal data. However, no single axiom has been convincing on its own and seeking multiple axioms together has often led to impossibility results rather than a choice of the “optimal voting rule” [Arr51; Gib73; Sat75]. The rise of approximation algorithms in computer science has offered a new perspective: even when voters report ordinal preferences, one can aim to optimize welfare defined on the underlying numerical utilities or costs, with success measured by *distortion*—the approximation ratio between the achieved welfare and the optimum in the worst case over the unknown preference intensities. This has reignited the search for optimal voting rules, well-defined in this distortion framework.

Specifically, in the metric social choice model [ABEPS18], n voters and m candidates are points in the same underlying metric space. A voter’s distance to a candidate is her cost for that candidate, and she ranks candidates from smallest to largest distance. The social cost of a candidate is its average distance to the voters. A deterministic (resp., randomized) voting rule must choose a candidate (resp., a lottery over candidates) based only on the rankings, and its *distortion* is the ratio of the (expected) social cost of its choice to the smallest social cost of any candidate, in the worst case over all metrics consistent with the observed rankings.

Gkatzelis, Halpern, and Shah [GHS20] prove that the best distortion achievable by deterministic rules is 3. To see why deterministic rules cannot do any better and why randomization *may* help, consider a simple example with two voters (1, 2) and two candidates (a, b). Voter 1 prefers a to b while voter 2 prefers b to a . Let the underlying metric have distance d and $\Delta := d(a, b)$. Only two consistent possibilities are (provably) of interest:

- either voter 1 is located at a and voter 2 is located at the midpoint of a and b , which makes the social costs of a and b equal to $\Delta/4$ and $3\Delta/4$, respectively,
- or voter 1 is located at the midpoint and voter 2 is located at b , which makes the social costs of a and b equal to $3\Delta/4$ and $\Delta/4$, respectively.

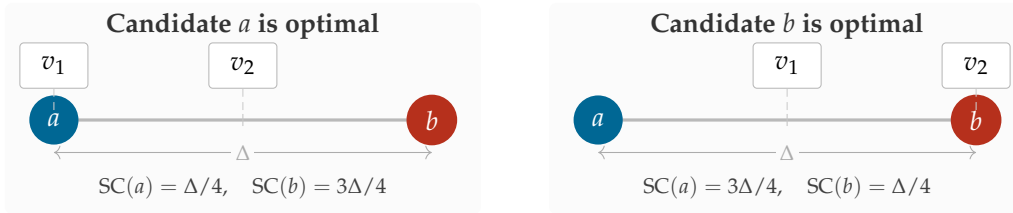


Figure 1: Two metrics consistent with the reported rankings, which shows a lower bound of 3 for deterministic rules and 2 for randomized rules.

Thus, as Figure 1 makes visible, either candidate can be optimal while the other is 3 times worse, making 3 a barrier that deterministic rules cannot cross. However, choosing between a and b uniformly at random achieves an expected social cost of $\Delta/2$, which is only 2 times the minimum

possible cost and that is the best possible distortion on this instance. It was even conjectured that 2 may be the optimal metric distortion of randomized rules, but Pulyassary and Swamy [PS21] and Charikar and Ramakrishnan [CR22] disprove it by establishing a lower bound strictly greater than 2, converging to approximately 2.1126 as $m \rightarrow \infty$.

On the upper-bound side, Charikar, Ramakrishnan, Wang, and Wu [CRWW24] prove a distortion guarantee of 2.753 by combining *maximal lottery* (ML), the Nash equilibrium of a natural zero-sum game dating back to the 1960s [Kre65; Bra17], and RaDiUS, a novel rule they design. Their work, which received the SODA 2024 Best Paper Award, establishes that randomized rules can in fact achieve a constant distortion strictly lower than the best guarantee of 3 for deterministic rules. Independently, Frank [Fra26] and Ye [Ye26] improve this to 2.5 with a rule of striking simplicity: the equal mixture of a maximal lottery and *Integrated Veto* (IV), an algorithmic procedure in which every voter continuously vetoes her current least-preferred surviving candidate. This still leaves a fundamental question open:

What is the smallest distortion that a randomized voting rule can achieve using only voters' preference rankings?

1.1 Our Results

The main result of our work improves the upper bound, closing about 44% of the gap between the upper bound of 2.5 and the lower bound of 2.1126 left by prior work.

Theorem 1. *There exists a randomized voting rule with metric distortion at most $11641/5000 = 2.3282$.*

To accomplish this, we consider a generalization of maximal lotteries to stable k -lotteries (SL_k) by Charikar, Ramakrishnan, Tan, and Wang [CRTW25], who use them to design so-called tournament voting rules with low metric distortion; stable 1-lotteries are precisely maximal lotteries. We design a further generalization of this to *random-size stable lotteries*: given a random variable D , RSL_D effectively provides the guarantee of SL_k but with k drawn from D . Yet, we show that RSL_D cannot be obtained by simply mixing SL_k for different values of k ; we prove its existence for any D directly via a minimax argument similar to what Charikar, Ramakrishnan, Tan, and Wang [CRTW25] used for SL_k , and give a multiplicative-weights approximation algorithm with an explicit running time when D has finite support. The main technical contribution of our work is to connect the guarantee provided by RSL_D to metric distortion via a much more intricate analysis than that of Charikar, Ramakrishnan, Tan, and Wang [CRTW25] for SL_k . This allows us to conduct a computer-assisted search for a suitable distribution D^* , which, when mixed with Integrated Veto, yields the desired distortion bound. While we were not able to analytically derive the optimal distribution, we were able to extract rational coefficients—probabilities under D^* and mixture weights—which exactly yield the rational distortion bound in Theorem 1. Table 1 shows our bound in the context of lower and upper bounds from prior work.

1.2 Technical Overview

The 2.5 barrier. The distortion bound of 2.5 proved independently by Frank [Fra26] and Ye [Ye26] uses the equal mixture of Integrated Veto (IV) and maximal lottery (ML). The analysis is organized around two quantities. Let C denote the candidate set, let I be any nonempty proper subset of C , and let P be a lottery. Informally, $q(I)$ measures how far I is from containing a candidate that

<i>Upper bounds</i>		
3	optimal deterministic rule	[GHS20]
2.753	a mixture of maximal lottery and RaDiUS	[CRWW24]
2.5	the equal mixture of maximal lottery and Integrated Veto	[Fra26; Ye26]
2.3282	a mixture of random-size stable lottery and Integrated Veto	this paper
<i>Lower bounds</i>		
2.0631	an example with $m = 7$	[PS21]
2.1126	a construction for any $m \geq 3$, bound on the left as $m \rightarrow \infty$	[CR22]

Table 1: Randomized metric distortion bounds from our work and prior work.

almost every voter ranks above every candidate outside I . The loss $\ell(P, I)$ measures how much probability P assigns outside I , charging an outside candidate in proportion to the voters who rank every member of I above it. These quantities are defined formally in [Section 2.1](#). The following implication of Charikar, Ramakrishnan, Wang, and Wu [[CRWW24](#)], in the form stated by Frank [[Fra26](#)], lets us work entirely with rankings.

Proposition 1 ([[CRWW24](#); [Fra26](#)]). *Let $\kappa \geq 0$. If a lottery P satisfies $\ell(P, I) \leq \kappa q(I)$ for every nonempty proper $I \subsetneq C$, then $\text{dist}(P) \leq 1 + 2\kappa$.*

Thus, it suffices to bound the ratio $\ell(P, I)/q(I)$ for every set with $q(I) > 0$; when $q(I) = 0$, the required loss is zero. The prior analysis gives

$$\frac{\ell(\text{IV}, I)}{q(I)} \leq q(I) \quad \text{and} \quad \frac{\ell(\text{ML}, I)}{q(I)} \leq \min \left\{ 1, \frac{1}{2q(I)} \right\}.$$

[Figure 2](#) shows these upper bounds. Their equal mixture reaches $3/4$ at both $q = 1/2$ and $q = 1$, and shifting weight toward either rule raises one of these two values. Thus the existing arguments do not yield a better bound with any mixture of these rules. Ye [[Ye26](#)] proves a stronger obstruction: even if the mixture is chosen from the profile and the veto component uses arbitrary, possibly adaptive, time weights along the Simultaneous Plurality Veto trajectory, some profile still forces distortion at least $5/2$.

Random-size stable lotteries. Our new ingredient starts from stable k -lotteries (SL_k) [[CRTW25](#)], which generalize maximal lotteries and satisfy

$$\Pr_v[c \succ_v (\text{SL}_k)^k] \leq \frac{1}{k+1} \quad \text{for every } c \in C.$$

Here $(\text{SL}_k)^k$ consists of k independent draws from SL_k , v is a uniformly random voter, and the probability also includes the draws and the prescribed tie-breaking. The case $k = 1$ is precisely a maximal lottery. We let the number of draws itself be random. For a positive-integer-valued random variable D , a random-size stable lottery RSL_D satisfies

$$\mathbb{E}_D \Pr_v[c \succ_v (\text{RSL}_D)^D] \leq \mathbb{E} \left[\frac{1}{D+1} \right] \quad \text{for every } c \in C.$$

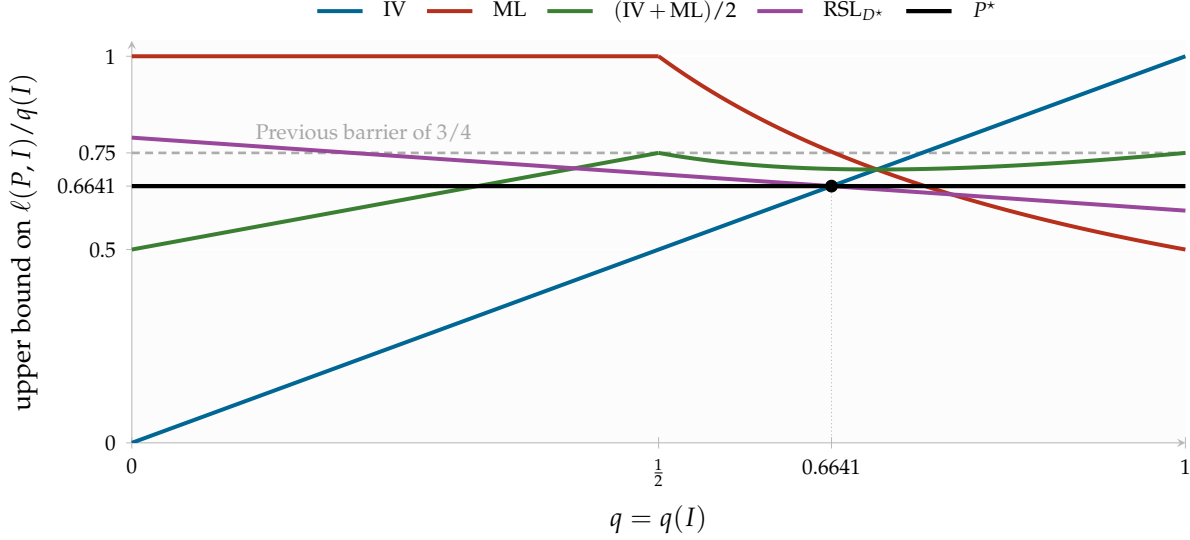


Figure 2: Established upper bounds on the loss ratio $\ell(P, I)/q(I)$ as a function of $q(I)$. The curves are bounds, not exact losses. Under these bounds, reweighting IV and ML cannot cross the horizontal threshold $3/4$. Our rule mixes IV with RSL_{D^*} so that their increasing and decreasing bounds cancel at 0.6641, yielding distortion 2.3282 through [Proposition 1](#).

When $D = k$ deterministically, this recovers SL_k . Crucially, one lottery must be chosen before D is realized: independently choosing an SL_k for each k and then mixing those lotteries does not work because repeated draws from the mixture create uncontrolled cross terms. A convex minimax argument nevertheless proves that RSL_D exists for every D . The definition and proof appear in [Section 3](#); an approximation algorithm is deferred to [Appendix A](#).

Why a random number of draws helps. Fix a voter and a reference candidate, and suppose that RSL_D places mass t on candidates strictly below that candidate and no mass on the candidate itself. The probability that the reference candidate beats all D draws is

$$\phi(t) := \mathbb{E}[t^D] = \sum_{d \geq 1} \Pr[D = d] t^d.$$

It is tempting to use the same formula without the zero-mass assumption, but this is incorrect because draws equal to the reference candidate create ties. Under the random-label tie-breaking defined in [Section 3](#), if b is the mass strictly below the reference candidate and z is the mass on it, then conditioning on its label $s \in [0, 1]$ gives the corrected probability $\int_0^1 \phi(b + sz) ds$. Thus the tie-breaking is an essential part of the comparison rule and of the convexity argument establishing existence.

This is where the distribution of D enters the proof. For a candidate set I , random-size stability can be applied both to the candidate in I that is usually ranked above everything outside I and, on average, to the candidates receiving probability from RSL_D . Together, these two applications constrain how much loss RSL_D can place outside I . The constraints involve $\phi(t)$ at the voter-level values of t , so they retain the entire shape of ϕ , not only the scalar $\mathbb{E}[1/(D+1)] = \int_0^1 \phi(t) dt$.

This extra freedom is exactly what the figure exploits. The distribution D^* , supported on 1, 12, and 13, gives a linearly decreasing upper bound on $\ell(\text{RSL}_{D^*}, I)/q(I)$. Integrated Veto gives the linearly increasing upper bound $q(I)$. Mixing the two cancels these slopes and keeps the loss ratio at most 0.6641 for every I . [Proposition 1](#) then gives distortion at most $1 + 2(0.6641) = 2.3282$.

1.3 Related Work

Metric distortion. The metric model was introduced by Anshelevich, Bhardwaj, and Postl [[ABP15](#)] and Anshelevich, Bhardwaj, Elkind, Postl, and Skowron [[ABEPS18](#)], who proved a lower bound of 3 for every deterministic rule and an upper bound of 5 for the Copeland rule. Goel, Krishnaswamy, and Munagala [[GKM17](#)] showed that Ranked Pairs can have distortion at least 5 and, more broadly, that even randomization cannot take tournament rules below distortion 3. Munagala and Wang [[MW19](#)] reduced the upper bound to $2 + \sqrt{5} \approx 4.236$ using generalized uncovered sets and proposed a candidate-selection condition whose universal existence would imply distortion 3; Kempe [[Kem20a](#)] subsequently developed an LP-duality framework that unified these analyses and led to closely related sufficient conditions. Gkatzelis, Halpern, and Shah [[GHS20](#)] finally matched the lower bound with Plurality Matching, settling the optimal deterministic distortion at 3. Subsequent work made this guarantee considerably more transparent. Plurality Veto initializes each candidate’s score at its plurality share and has voters, in an arbitrary order, veto their least-preferred candidate with positive score until one candidate survives [[KK22](#)]. Simultaneous Plurality Veto removes the dependence on voter order by letting all voters veto continuously and simultaneously, while preserving distortion 3 [[KK23](#)]. Integrated Veto, introduced independently by Frank [[Fra26](#)] and Ye [[Ye26](#)] (under the name *Veto Lottery* in the latter), turns this deterministic process into a lottery by assigning each candidate probability equal to twice the area under its score trajectory. Thus, the Integrated Veto component of our rule descends directly from the sequence of deterministic rules that attained the optimal distortion of 3.

Utilitarian distortion. In the utilitarian model initiated by Procaccia and Rosenschein [[PR06](#)], voters again report only rankings, but the hidden cardinal values are otherwise unrestricted and the objective is to maximize their sum. Here normalization is essential: rankings are unchanged if one rescales each voter’s utilities independently, so rank data alone provide no way to place different voters’ utilities on a common scale. The standard unit-sum model normalizes each voter’s utilities to sum to one, giving every voter equal weight while retaining the relative intensity of her preferences. Metric distortion needs no analogous normalization because a common underlying metric already places all voters’ costs on the same scale. Under unit-sum utilities, the optimal deterministic distortion is $\Theta(m^2)$ [[CP11](#); [CNPS17](#)]; Boutilier, Caragiannis, Haber, Lu, Procaccia, and Sheffet [[BCHLPS15](#)] proved an $\Omega(\sqrt{m})$ lower bound for randomized rules, and Ebadian, Kahng, Peters, and Shah [[EKPS24](#)] matched it with an $O(\sqrt{m})$ rule, closing the randomized case asymptotically as well. Although the metric and utilitarian models are each motivated by the same loss of cardinal information, rules optimized for one can perform poorly in the other. Gkatzelis, Latifian, and Shah [[GLS23](#)] nevertheless design deterministic and randomized rules that simultaneously achieve near-optimal guarantees in both models, establishing best-of-both-worlds distortion.

Information–distortion tradeoffs. Full rankings are only one point on a broader elicitation spectrum. At one end, top- t ballots ask each voter to rank only her t favorite candidates; work

under both utilitarian and metric objectives quantifies what can still be guaranteed as this reported prefix shrinks [Kem20b; BHLS22]. At the other end, a value query asks a voter for her exact utility for one candidate in addition to her ranking. A few such queries can sharply reduce utilitarian distortion [ABFV21]; for any constant λ , Ebadian and Shah [ES25] achieve $O(\min\{n, m\}^{1/\lambda})$ distortion using λ value queries per voter, matching the known dependence on m when $n = \Omega(m)$. More generally, one can optimize the elicitation format itself subject only to a bit budget. Mandal, Procaccia, Shah, and Woodruff [MPSW19] and Mandal, Shah, and Woodruff [MSW20] characterize this communication–distortion frontier up to logarithmic factors: achieving utilitarian distortion d requires $\tilde{\Theta}(m/d)$ bits per voter with deterministic elicitation and $\tilde{\Theta}(m/d^3)$ bits with randomized elicitation. Hence, the format of the elicited information can matter as much as its quantity: a carefully designed short message may yield lower distortion than a substantially longer ranked ballot.

2 Preliminaries

Let C be a finite set of m candidates and V a finite set of n voters. For a positive integer r , write $[r] = \{1, \dots, r\}$. Each voter $v \in V$ reports a strict ranking $\succ_v$ of C , and the *profile* is the collection $\sigma = (\succ_v)_{v \in V}$. Unless stated otherwise, v is drawn uniformly at random from V in every expectation or probability carrying the subscript v .

A *metric instance* places the voters and candidates in a common metric space with distance d .¹ It is *consistent* with σ if every voter’s ranking agrees with her distances, that is, $a \succ_v b$ implies $d(v, a) \leq d(v, b)$ for all $v \in V$ and $a, b \in C$. The social cost of a candidate c and the expected social cost of a lottery $P \in \Delta(C)$, the set of probability distributions on C , are

$$\text{SC}_d(c) = \mathbb{E}_v[d(v, c)] \quad \text{and} \quad \text{SC}_d(P) = \mathbb{E}_{c \sim P}[\text{SC}_d(c)].$$

Definition 1 (Distortion). For a profile σ and a lottery P , define

$$\text{dist}_\sigma(P) = \sup_{d \text{ consistent with } \sigma} \frac{\text{SC}_d(P)}{\min_{c \in C} \text{SC}_d(c)},$$

with the usual conventions $0/0 = 1$ and $z/0 = +\infty$ for $z > 0$. We drop the subscript when the profile is fixed. A randomized social choice rule $\mathcal{R}$ maps each profile σ to a lottery, and its *distortion* is $\sup_\sigma \text{dist}_\sigma(\mathcal{R}(\sigma))$. We write $D_{\text{rand}}(m)$ for the smallest distortion achievable with m candidates and $\sup_m D_{\text{rand}}(m)$ for the unrestricted optimum. Note that the rule is unaware of the underlying consistent metric.

2.1 Analysis via Candidate Sets

Analyzing the worst case over the vast space of metrics consistent with the reported rankings is cumbersome. Charikar, Ramakrishnan, Wang, and Wu [CRWW24] simplify the analysis by showing that it suffices to reason about a set of inequalities involving only the reported rankings. We use the following reformulation of their approach, made explicit by Frank [Fra26].

¹Distinct points are allowed to be at distance zero, so technically this is a pseudometric.

For a nonempty proper subset of candidates $I \subsetneq C$, let $J := C \setminus I$ and define

$$q(I) = \min_{i \in I} \Pr_v[j \succ_v i \text{ for some } j \in J], \quad (2.1)$$

$$\ell(P, I) = \sum_{j \in J} P(j) \Pr_v[i \succ_v j \text{ for every } i \in I]. \quad (2.2)$$

We will sometimes refer to a candidate $z \in I$ attaining the minimum in Equation (2.1). The meanings of these two quantities are simple. Some candidate $z \in I$ is ranked above every candidate outside I by $1 - q(I)$ fraction of voters. When $q(I)$ is very small, almost all voters rank z above every candidate outside I . If a lottery P were to then place significant probability mass outside I , z could lead to P having high distortion. Hence, the loss $\ell(P, I)$ charges any probability that P places outside I , but probability placed on $j \notin I$ is charged only on voters who rank all members of I above j . It turns out that keeping $\ell(P, I)$ small whenever $q(I)$ is small is a guaranteed route to achieving low distortion.

Proposition 1 ([CRWW24; Fra26]). *Let $\kappa \geq 0$. If a lottery P satisfies $\ell(P, I) \leq \kappa q(I)$ for every nonempty proper $I \subsetneq C$, then $\text{dist}(P) \leq 1 + 2\kappa$.*

Frank [Fra26] finds a lottery which achieves $\kappa = 3/4$. Our goal is to find a lottery with $\kappa < 3/4$ — more precisely, $\kappa = 0.6641$, which yields the desired distortion bound of 2.3282. Let us understand these quantities better through an example.

Example 1 (Understanding $q(I)$, $\ell(P, I)$, and κ). Let $C = \{a, b, c\}$ and $|V| = 20$, with the preference profile shown on the right. Take $I = \{a, b\}$, so $J = \{c\}$. Candidate a fails to beat c for a $3/5$ fraction of the voters, whereas candidate b fails to beat c for a $1/4$ fraction of the voters. Hence

#Voters	Ranking
8	$a \succ b \succ c$
7	$b \succ c \succ a$
5	$c \succ a \succ b$

$$q(I) = \min \left\{ \frac{3}{5}, \frac{1}{4} \right\} = \frac{1}{4}, \quad z = b.$$

Only the first group of voters ranks both a and b above c , so $\ell(P, I) = \frac{2}{5}P(c)$ for every lottery P . A rule can therefore satisfy $\ell(P, I) \leq \kappa q(I)$ for this particular I only by keeping $P(c)$ at most $5\kappa/8$. To yield the desired distortion bound, however, it must satisfy the inequality for $I = \{a\}, \{b\}, \{c\}, \{b, c\}, \{a, c\}$ as well. We revisit this example in Example 2.

2.2 Integrated Veto and Maximal Lotteries

We combine two kinds of lotteries. The first follows the progression described in Section 1.3: Integrated Veto was introduced independently by Frank [Fra26] and Ye [Ye26]; the latter calls it Veto Lottery. It runs the Simultaneous Plurality Veto process of Kizilkaya and Kempe [KK23]. Each candidate starts with a score equal to the fraction of voters ranking it first. While positive scores remain, every voter continuously decreases, at rate $1/n$, the score of her least-preferred candidate among those with positive score. Integrated Veto gives each candidate probability equal to twice the area under its score trajectory.

The factor of two and the fact that these probabilities sum to one have the same two-line explanation. The scores start summing to one and, as long as any candidate is active, the total decreases at rate exactly one, so the total score is $1 - t$ at time t , until it reaches zero at time $t = 1$. The area under the total-score curve is therefore the area of a triangle with legs 1 and 1, namely $1/2$, and doubling turns areas into probabilities.

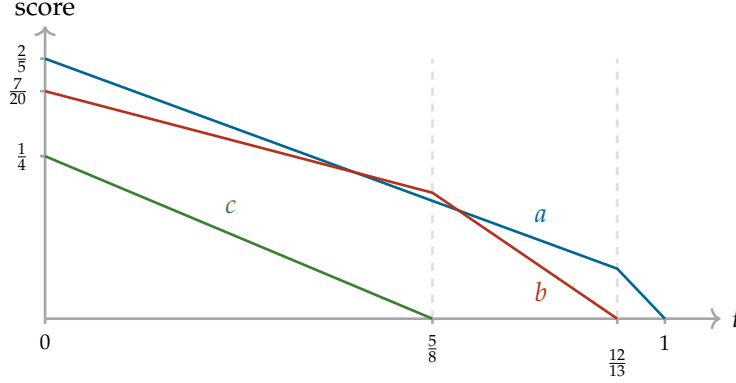


Figure 3: Integrated Veto score trajectories for [Example 2](#). Each candidate’s probability is twice the area under its trajectory. The three areas sum to $1/2$ because the total score is $1 - t$ throughout, which is also why the factor of two is the right normalization. A trajectory changes slope only when some candidate is exhausted and the voters who were vetoing it switch to their next-least-preferred active candidate.

Example 2 (Integrated Veto on the profile of [Example 1](#)). Take the profile of [Example 1](#). The initial scores are the plurality shares $(\frac{2}{5}, \frac{7}{20}, \frac{1}{4})$ for (a, b, c) . Initially the first group vetoes c , the second vetoes a , and the third vetoes b , so the three scores fall at rates of $\frac{7}{20}$, $\frac{1}{4}$, and $\frac{2}{5}$, respectively. Candidate c is exhausted at $t = \frac{5}{8}$; at that moment the first group joins the third in vetoing b , whose score then falls at a rate of $\frac{2}{5} + \frac{1}{4} = \frac{13}{20}$. Candidate b is exhausted at $t = \frac{12}{13}$, after which all voters veto a , which is exhausted at $t = 1$. [Figure 3](#) plots the three trajectories. Doubling the areas gives

$$\text{IV}(a) = \frac{29}{65}, \quad \text{IV}(b) = \frac{827}{2080}, \quad \text{IV}(c) = \frac{5}{32},$$

which sum to one. For the set $I = \{a, b\}$ of [Example 1](#), the loss is $\ell(\text{IV}, I) = \frac{2}{5} \cdot \frac{5}{32} = \frac{1}{16} = q(I)^2$. Frank [[Fra26](#)] more generally proves $\ell(\text{IV}, I) \leq q(I)^2$, stated in [Lemma 1](#) below, which holds with equality in this example.

The second lottery is a maximal lottery, which is a Nash equilibrium of the following two-player zero-sum game. Each player chooses a lottery over candidates. Nature draws one candidate from each lottery and one random voter. The voter awards a point to the player whose candidate she prefers; identical draws split the point equally. The game is symmetric, so its value is $1/2$, and a maximin strategy is a maximal lottery. We write ML for an arbitrary maximal lottery. According to Brandt [[Bra17](#)], this rule was first considered by Kreweras [[Kre65](#)], but it is so fundamental that it has been rediscovered repeatedly [[Fis84](#); [LLL93](#); [FR95](#); [FM92](#); [RS10](#)].

The following bounds are the ingredients of the previous 5/2 result.

Lemma 1 ([[Fra26](#)]). *For every nonempty proper $I \subsetneq C$, with $q = q(I)$,*

$$\ell(\text{IV}, I) \leq q^2 \quad \text{and} \quad \ell(\text{ML}, I) \leq \min\{q, 1/2\}.$$

Dividing by q and averaging the two bounds shows that the loss ratio $\ell((\text{IV} + \text{ML})/2, I)/q$ is at most $\frac{1+q}{2}$ for $q \leq 1/2$ and at most $\frac{q}{2} + \frac{1}{4q}$ for $q \geq 1/2$. Both expressions are at most $3/4$, so

Proposition 1 gives distortion at most $5/2$. As noted in [Section 1.2](#), **Proposition 1** does not yield a better bound even under any unequal mixture of IV and ML; Ye [\[Ye26\]](#) explicitly proves that a much larger family of rules in fact has distortion no better than $5/2$.

3 Random-Size Stable Lotteries

We now formally define our key novel ingredient: random-size stable lotteries. We prove their existence by extending the minimax argument of Charikar, Ramakrishnan, Tan, and Wang [\[CRTW25\]](#) from a fixed number of draws k to a random number of draws D . In [Appendix A](#), we show that this also yields a standard multiplicative-weights algorithm to compute such lotteries approximately and establish other properties of interest that we will not use in our work.

Comparing multisets. First, we need to lift each voter’s comparison of individual candidates to a comparison between multisets of candidates. Fix a voter v and two nonempty multisets S and T . It is natural to compare S and T by the candidates that v ranks highest in them. If these candidates are distinct, v ’s ranking determines the comparison; the only ambiguity arises when the same candidate is highest-ranked in both multisets, possibly with different multiplicities. To resolve this ambiguity, we use the random-label tie-breaking scheme of Charikar, Ramakrishnan, Tan, and Wang [\[CRTW25\]](#): give every duplicate an independent label drawn uniformly from $[0, 1]$, and let v compare duplicates by their labels. Thus, if voter v ’s highest-ranked candidate in $S \cup T$ has multiplicity r in S and s in T , then $\Pr[S \succ_v T] = r/(r + s)$. By choosing S or T to be singleton, this also yields comparisons between individual candidates and multisets. For a lottery P , a positive integer d , and a positive-integer-valued random variable D , we will be interested in the multisets P^d and P^D obtained by taking d and D independent draws from P , respectively.

The random sample size. For a random number of draws D , let

$$\phi(t) = \mathbb{E}[t^D] = \sum_{d \geq 1} \Pr[D = d] t^d$$

be the probability-generating function of D . Note that ϕ is increasing and convex on $[0, 1]$, with $\phi(0) = 0$, $\phi(1) = 1$, and $\int_0^1 \phi(t) dt = \mathbb{E}[1/(D + 1)]$. Fix a voter v . If a lottery P places mass t on candidates that v ranks strictly below a candidate c and $P(c) = 0$, then

$$\Pr[c \succ_v P^D] = \sum_{d \geq 1} \Pr[D = d] \cdot \Pr[c \succ_v P^d] = \sum_{d \geq 1} \Pr[D = d] \cdot t^d = \phi(t).$$

Without the assumption $P(c) = 0$, this simple formula is incorrect because a draw equal to c ties the challenger. Let b be the P -mass strictly below c and let $z = P(c)$. Conditional on the challenger copy of c receiving label $s \in [0, 1]$, a single draw from P loses to it with probability $b + sz$. Random-label tie-breaking therefore gives the corrected formula

$$\Pr[c \succ_v P^D] = \int_0^1 \phi(b + sz) ds.$$

When $z = 0$, this reduces to $\phi(b)$.

Definition 2 (Random-size stable lottery). We write RSL_D for an arbitrary lottery that is *random-size stable* for D , meaning that, for every challenger lottery A ,

$$\mathbb{E}_D \Pr_v[A \succ_v (\text{RSL}_D)^D] \leq \mathbb{E} \left[\frac{1}{D+1} \right].$$

Here, $A = A^1$ denotes a single random sample from A . When $D = 1$ deterministically, **Definition 2** is precisely a maximal lottery; when $D = k$ deterministically, it is a stable k -lottery defined by Charikar, Ramakrishnan, Tan, and Wang [CRTW25]. Cheng, Jiang, Munagala, and Wang [CJMW20] and Jiang, Munagala, and Wang [JMW20] define a related notion of stable lotteries over committees, which are more restrictive in that the committee does not necessarily come from independent draws from the same lottery, but more general in other dimensions, as discussed in **Appendix A**.

$\mathbb{E}_D \Pr_v[A \succ_v (\text{RSL}_D)^D]$ is linear in A , so it is enough to consider a single challenger candidate when maximizing or minimizing it. Because the comparison, with random-label tie-breaking, is designed to yield exactly one probabilistic winner, the equation in **Definition 2** is also equivalent to

$$\mathbb{E}_D \Pr_v[(\text{RSL}_D)^D \succ_v A] \geq 1 - \mathbb{E} \left[\frac{1}{D+1} \right].$$

Theorem 2 (Existence of random-size stable lotteries). *For every finite preference profile and every distribution of a positive-integer-valued random variable D , there exists a random-size stable lottery RSL_D such that*

$$\mathbb{E}_D \Pr_v[c \succ_v (\text{RSL}_D)^D] \leq \mathbb{E} \left[\frac{1}{D+1} \right] \quad \text{for every } c \in C. \quad (3.1)$$

No smaller right-hand side works for every profile. Every candidate in the support of RSL_D satisfies the displayed inequality with equality.

Proof. For this proof, abbreviate $p_D = \mathbb{E}[1/(D+1)]$. For a challenger lottery A and a lottery P , define

$$F(A, P) = \mathbb{E}_D \Pr_v[A \succ_v P^D].$$

The function is linear in A . To see that it is continuous and convex in P , first take A to be a fixed candidate c and fix a voter. Let b be the P -mass strictly below c and let $z = P(c)$. Conditional on the challenger copy of c having label $s \in [0, 1]$, this voter contributes $\phi(b + sz)$. Both b and z are linear in P , while ϕ is continuous and convex; integrating over s , the voter, and A preserves these properties.

The convex minimax theorem therefore gives

$$\min_P \max_A F(A, P) = \max_A \min_P F(A, P).$$

For every lottery P , exchangeability among $d+1$ independent labeled draws gives $\Pr_v[P \succ_v P^d] = 1/(d+1)$, and hence $F(P, P) = p_D$. Choosing $P = A$ shows that the right-hand side of the minimax equality is at most p_D , while $\max_A F(A, P) \geq F(P, P) = p_D$ for every P . The minimax value is therefore exactly p_D . Because F is linear in A , an optimal P , denoted RSL_D , satisfies **Equation (3.1)** for every candidate.

Finally,

$$p_D = F(\text{RSL}_D, \text{RSL}_D) = \sum_{c \in C} \text{RSL}_D(c) \mathbb{E}_D \Pr_v[c \succ_v (\text{RSL}_D)^D].$$

Every term in this weighted average is at most p_D , so equality holds for every candidate in the support. The same identity shows that no smaller universal right-hand side is possible. $\square$

The cases of deterministic D recover maximal lotteries and stable k -lotteries directly. Random-size existence cannot instead be obtained by averaging separately chosen stable k -lotteries: independent draws from that average contain cross terms not controlled by fixed-size stability. [Appendix A.1](#) gives a two-candidate example.

The same convexity argument also gives a standard multiplicative-weights algorithm for approximating RSL_D ; the precise guarantee and proof appear in [Proposition 2](#).

An exact identity for reversing which side receives the D draws, together with the nonnegative convexity term it produces, is given in [Lemma 7](#). Its outside-mass consequence is incorporated into the profile-independent loss optimization below.

4 From Random-Size Stability to Loss Ratio

To use [Proposition 1](#), we would like to upper bound the loss $\ell(\text{RSL}_D, I)$ of a random-size stable lottery RSL_D in terms of only D and $q(I)$, abstracting away any other relevant information from the full profile. The three lemmas in this section successively simplify this problem.

4.1 Abstracting the Profile Into an Infinite-Dimensional LP

The first lemma, [Lemma 2](#), is one of the most significant technical results enabling our approach. It condenses the worst-case over all profiles into an optimization program over a scalar x and two probability measures η and ξ , establishing the necessary link between random-size stability and the loss $\ell(\text{RSL}_D, I)$.

Later, [Lemma 3](#) applies a useful transformation to this program and [Lemma 4](#) paves a pathway for upper bounding its optimal objective value.

Lemma 2. Fix a positive-integer-valued random variable D . Write $\phi_D(t) = \mathbb{E}[t^D]$ and $p_D = \mathbb{E}[\frac{1}{D+1}]$. For $q \in [0, 1]$, define

$$\begin{aligned}
g_D(q) &= \sup_{x, \eta, \xi} x \left((1-q) \int_{[0,1]^2} b_0 d\eta(b_0, y) + q \int_{[0,1]} b_1 d\xi(b_1) \right) \\
\text{subject to } & (1-q) \int_{[0,1]^2} \phi_D(x + (1-x)y) d\eta(b_0, y) + q \int_{[0,1]} \phi_D(xb_1) d\xi(b_1) \leq p_D, \quad (C1) \\
& (1-q) \int_{[0,1]^2} \left(\int_0^y \phi_D((1-x)u + xb_0) du + \int_y^1 \phi_D((1-x)u + x) du \right) d\eta(b_0, y) \\
& \quad + q \int_{[0,1]} \left(\int_0^1 \phi_D((1-x)u + xb_1) du \right) d\xi(b_1) \leq p_D, \quad (C2) \\
& x \left((1-q) \int_{[0,1]^2} b_0 d\eta(b_0, y) + q \int_{[0,1]} b_1 d\xi(b_1) \right) \leq (1-p_D)x, \quad (C3) \\
& \eta([0,1]^2) = 1, \quad \xi([0,1]) = 1, \quad x \in [0, 1], \quad \eta, \xi \geq 0.
\end{aligned}$$

Note that η and ξ range over nonnegative finite Borel measures on $[0, 1]^2$ and $[0, 1]$, respectively. Then, for every profile, every random-size stable lottery RSL_D , and every nonempty proper $I \subsetneq C$,

$$\ell(\text{RSL}_D, I) \leq g_D(q(I)).$$

We acknowledge that the statement of this lemma may be jarring for a reader at this point in the paper as the complicated optimization program has no prior foundation. So, before diving into a rigorous proof, let us provide some insights on the high-level approach and where some of these terms come from. Some subtleties are omitted in this exposition, but discussed in the rigorous proof that follows.

Insights behind Lemma 2. Fix a nonempty proper $I \subsetneq C$, put $J = C \setminus I$, and choose $z \in I$ attaining the minimum in the definition of $q(I)$. Recall that

$$\ell(\text{RSL}_D, I) = \sum_{j \in J} \text{RSL}_D(j) \cdot \Pr_v[i \succ_v j \text{ for every } i \in I].$$

To separate the ingredients of this expression, write $\text{RSL}_D = (1 - x)P_I + xP_J$, where x is its probability mass on J , and P_I, P_J are RSL_D restricted (and renormalized) on I, J , respectively. Let $V_0 = \{v \in V : z \succ_v j \text{ for every } j \in J\}$ and $V_1 = V \setminus V_0$. By the definition of $q(I)$, V_0 contains a $1 - q(I)$ fraction of the voters. For $s \in \{0, 1\}$ and $v \in V_s$, let $b_s(v)$ be the P_J -mass on candidates that v ranks below every member of I . For $v \in V_0$, let $y(v)$ be the P_I -mass on candidates ranked below z by v . Let η be the joint distribution of $(b_0(v), y(v))$ with v drawn uniformly from V_0 , and let ξ be the distribution of $b_1(v)$ with v drawn uniformly from V_1 . Then, the loss decomposes as

$$\ell(\text{RSL}_D, I) = x \left((1 - q(I)) \int_{[0,1]^2} b_0 \, d\eta(b_0, y) + q(I) \int_{[0,1]} b_1 \, d\xi(b_1) \right).$$

Notice that this matches the objective of the optimization program in Lemma 2. The proof then identifies three necessary constraints that random-size stability imposes on x, η , and ξ : (C1) comes from challenging RSL_D with z , (C2) from challenging it with P_I , and (C3) from reversing the latter comparison. We then forget the originating profile and maximize the displayed loss expression over all possible choices of x, η, ξ that satisfy (C1)–(C3). While some possible choices may not arise from any profile, this supremum upper bounds the loss on any feasible profile. The next two lemmas convert this profile-independent optimization into the pointwise scalar inequalities used in the proof of our main bound.

Proof of Lemma 2. Fix a nonempty proper set $I \subsetneq C$, put $J = C \setminus I$ and $q = q(I)$, and choose $z \in I$ attaining the minimum in Equation (2.1). Define

$$V_0 = \{v \in V : z \succ_v j \text{ for every } j \in J\} \quad \text{and} \quad V_1 = V \setminus V_0.$$

By the definition of $q(I)$, $|V_0|/|V| = 1 - q$. Let $x = \text{RSL}_D(J)$. If $x = 0$, then the loss is zero. Moreover, taking η and ξ to be point masses at $(0, 0)$ and 0 , respectively, is feasible in the definition of $g_D(q)$: the left-hand sides of (C1) and (C3) are zero, while that of (C2) is $\int_0^1 \phi_D(u) \, du = p_D$. Hence, suppose $x > 0$. Write $P_J = \text{RSL}_D \mid J$ for the conditional lottery outside I . When $x < 1$, write $P_I = \text{RSL}_D \mid I$; when $x = 1$, let P_I be any lottery supported on I .

Encoding the loss. For $s \in \{0, 1\}$ and each voter $v \in V_s$, define

$$b_s(v) = \sum_{j \in J} P_J(j) \cdot \mathbf{1}\{i \succ_v j \text{ for every } i \in I\}.$$

Thus $b_s(v)$ is the fraction of the outside conditional lottery that this voter ranks below every member of I . For each $v \in V_0$, give z an independent random label and let $y(v)$ be the fraction of P_I lying below this labeled copy of z . Let η have the joint law of $(b_0(v), y(v))$ when v is drawn uniformly from V_0 , including the randomness of the label, and let ξ have the law of $b_1(v)$ when v is drawn uniformly from V_1 . If either set is empty, choose the corresponding measure arbitrarily. Since $\text{RSL}_D(j) = xP_I(j)$ for every $j \in J$, [Equation \(2.2\)](#) gives

$$\begin{aligned} \ell(\text{RSL}_D, I) &= x \sum_{j \in I} P_I(j) \cdot \Pr_v[i \succ_v j \text{ for every } i \in I] \\ &= x \mathbb{E}_v \left[\sum_{j \in I} P_I(j) \cdot \mathbf{1} \{i \succ_v j \text{ for every } i \in I\} \right] \\ &= x \left((1-q) \int_{[0,1]^2} b_0 \, d\eta(b_0, y) + q \int_{[0,1]} b_1 \, d\xi(b_1) \right). \end{aligned}$$

The final equality is the law of total expectation over the partition V_0 and V_1 . It remains to show that x, η, ξ satisfy (C1)–(C3).

Deriving (C1) by applying stability to z . For $v \in V_0$, the lottery RSL_D places mass at least $x + (1-x)y(v)$ below the labeled copy of z ; for $v \in V_1$, it places mass at least $xb_1(v)$ below it. Since $\phi_D(t)$ is the probability that all D draws fall within a mass of t , monotonicity of ϕ_D and random-size stability for z give

$$(1-q) \int_{[0,1]^2} \phi_D(x + (1-x)y) \, d\eta(b_0, y) + q \int_{[0,1]} \phi_D(xb_1) \, d\xi(b_1) \leq \mathbb{E}_D \Pr_v[z \succ_v (\text{RSL}_D)^D] \leq p_D.$$

This proves (C1).

Deriving (C2) by averaging stability over the inside lottery. Conditional on a voter v , draw an independently labeled candidate from P_I , and let $U_v \in [0, 1]$ be its labeled percentile within P_I , measured from worst to best in v 's ranking. Conditional on v , random labeling makes U_v uniform on $[0, 1]$ and independent of the label used to define $y(v)$ for $v \in V_0$. For a realization $U_v = u$ and $v \in V_s$, the challenger beats at least the $(1-x)u$ mass of P_I below it and the $xb_s(v)$ mass of P_J lying below every member of I . For $v \in V_0$, if $u > y(v)$, it also lies above z and hence above all of P_J , so it beats mass at least $(1-x)u + x$. Averaging the resulting lower bound on the probability that all D draws lie below this candidate gives

$$\begin{aligned} (1-q) \int_{[0,1]^2} \left(\int_0^y \phi_D((1-x)u + xb_0) \, du + \int_y^1 \phi_D((1-x)u + x) \, du \right) d\eta(b_0, y) \\ + q \int_{[0,1]} \left(\int_0^1 \phi_D((1-x)u + xb_1) \, du \right) d\xi(b_1) \leq \mathbb{E}_D \Pr_v[P_I \succ_v (\text{RSL}_D)^D] \leq p_D. \end{aligned}$$

This proves (C2).

Deriving (C3) using the reverse comparison identity. Random-size stability with challenger lottery P_I gives $\mathbb{E}_D \Pr_v[P_I \succ_v (\text{RSL}_D)^D] \leq p_D$, so [Lemma 7](#) gives $\mathbb{E}_D \Pr_v[P_I^D \succ_v \text{RSL}_D] \leq 1 - p_D$. Decomposing the one draw from $\text{RSL}_D = (1-x)P_I + xP_J$ and using exchangeability gives

$$\mathbb{E}_D \Pr_v[P_I^D \succ_v \text{RSL}_D] = (1-x)(1-p_D) + x \mathbb{E}_D \Pr_v[P_I^D \succ_v P_J].$$

Whenever every member of I beats $j \in J$, every nonempty multiset drawn from P_I beats j . Therefore,

$$\begin{aligned} \mathbb{E}_D \Pr_v[P_I^D \succ_v P_J] &\geq \sum_{j \in J} P_J(j) \cdot \Pr_v[i \succ_v j \text{ for every } i \in I] \\ &= (1-q) \int_{[0,1]^2} b_0 d\eta(b_0, y) + q \int_{[0,1]} b_1 d\zeta(b_1) \\ &= \frac{\ell(\text{RSL}_D, I)}{x}. \end{aligned}$$

Combining the last two equations gives $\ell(\text{RSL}_D, I) \leq (1-p_D)x$, which is precisely (C3). Thus every profile produces a feasible triple (x, η, ζ) whose objective equals its loss. Since the definition of $g_D(q)$ takes the supremum over all such x and all probability measures η and ζ , including choices that need not arise from a profile, it upper bounds the loss and proves the lemma. $\square$

4.2 Simplified PRIMAL

Note that the program in [Lemma 2](#) is not jointly linear in x , η , and ζ : its objective multiplies x by linear functionals of η and ζ , and its constraints depend nonlinearly on x through ϕ_D . To obtain a linear program, we fix x and search over x separately. When D , q , and x are fixed, the objective and constraints are linear in the decision measures η and ζ .²

(C3) directly bounds the objective by $(1-p_D)x$. Hence, instead of keeping (C3) in the program, we remove it—this can only increase the optimal objective value and hence keep it a valid upper bound on $g_D(q)$ —and instead constrain $g_D(q)$ directly using the upper bound provided by (C3).

To succinctly write the resulting LP, we define some additional notation. For fixed D , abbreviate

$$f(t) = \phi_D(t), \quad p = p_D = \int_0^1 f(t) dt, \quad r = 1 - x.$$

For $x > 0$, define

$$K_x = \int_0^1 f(ru) du, \quad W_x(y) = f(x + ry), \quad w_x(b_1) = f(xb_1), \quad (4.1)$$

$$\begin{aligned} G_x(b_0, y) &= \frac{\int_0^y f(ru + xb_0) du + \int_y^1 f(ru + x) du - K_x}{x}, \\ F_x(b_1) &= \frac{\int_0^1 f(ru + xb_1) du - K_x}{x}, \quad R_x = \frac{p - K_x}{x}. \end{aligned} \quad (4.2)$$

For any optimization program $\mathcal{P}$, we will write $\text{val}(\mathcal{P})$ to denote its optimal objective value.

Lemma 3. *For fixed D , $q \in [0, 1]$, and $x \in (0, 1]$, let η and ζ range over probability measures on $[0, 1]^2$ and $[0, 1]$, respectively, and consider*

²For any fixed function h and scalars α, β ,

$$\int_S h(t) d(\alpha\eta_1 + \beta\eta_2)(t) = \alpha \int_S h(t) d\eta_1(t) + \beta \int_S h(t) d\eta_2(t).$$

Thus the integral is linear in the measures regardless of the form of h . For a reader unfamiliar with programs involving measures, the finite-support analogue is $\sum_{z=1}^{10} h_z \eta_z$, which is linear in the probability masses $\eta_1, \dots, \eta_{10}$ even if the coefficients h_z come from a nonlinear function.

PRIMAL(D, q, x)

$$\begin{aligned} & \sup_{\eta, \xi \geq 0} (1-q) \int_{[0,1]^2} b_0 d\eta(b_0, y) + q \int_{[0,1]} b_1 d\xi(b_1) \\ \text{subject to } & (1-q) \int_{[0,1]^2} W_x(y) d\eta(b_0, y) + q \int_{[0,1]} w_x(b_1) d\xi(b_1) \leq p, \\ & (1-q) \int_{[0,1]^2} G_x(b_0, y) d\eta(b_0, y) + q \int_{[0,1]} F_x(b_1) d\xi(b_1) \leq R_x. \end{aligned}$$

Then

$$g_D(q) \leq \max \left\{ 0, \sup_{\substack{x \in (0,1]: \\ \text{PRIMAL}(D, q, x) \text{ is feasible}}} \min \{ x \text{val}(\text{PRIMAL}(D, q, x)), (1-p)x \} \right\}. \quad (4.3)$$

Proof. The first constraint of PRIMAL(D, q, x) is exactly (C1). In (C2), the function integrated against η is $K_x + xG_x$, and the function integrated against ξ is $K_x + xF_x$. Since both measures have mass one and $(1-q) + q = 1$, subtracting the common term K_x from (C2) and dividing by x gives the second constraint of PRIMAL(D, q, x). Dividing the objective in Lemma 2 by x gives the objective of PRIMAL(D, q, x).

We leave (C3) outside PRIMAL because it already gives the independent upper bound $(1-p)x$ on the original, unnormalized objective. Consequently, every feasible choice in Lemma 2 with $x > 0$ has objective at most

$$\min \{ x \text{val}(\text{PRIMAL}(D, q, x)), (1-p)x \}.$$

Taking the supremum over the feasible values of x proves Equation (4.3). At $x = 0$, the original objective is zero, which accounts for the first term in the maximum. Note that optimizing over all measures thus constrained provides only an upper bound; we do not claim that every pair of measures feasible for PRIMAL comes from a preference profile. $\square$

4.3 Bounding Loss via DUAL

The last step removes the distribution-valued variables by taking the dual of PRIMAL; since PRIMAL involves measures η and ξ over $[0, 1]^2$ and $[0, 1]$, respectively, as variables, this requires using duality theory of generalized moment problems [Sha01]. The dual has only four scalar variables, but infinitely many constraints: one for each $(b_0, y) \in [0, 1]^2$ and one constraint for each $b_1 \in [0, 1]$. Hence, the dual is a semi-infinite linear program.

Lemma 4. For fixed $D, q \in [0, 1]$, and $x \in (0, 1]$, the dual of PRIMAL(D, q, x) is

DUAL(D, q, x)

$$\begin{aligned} & \inf_{\substack{\lambda, \mu \geq 0 \\ u, v \in \mathbb{R}}} \lambda p + \mu R_x + (1-q)u + qv \\ \text{subject to } & u + \lambda W_x(y) + \mu G_x(b_0, y) \geq b_0 \quad \text{for every } (b_0, y) \in [0, 1]^2, \\ & v + \lambda w_x(b_1) + \mu F_x(b_1) \geq b_1 \quad \text{for every } b_1 \in [0, 1]. \end{aligned}$$

Every feasible solution (λ, μ, u, v) of $\text{DUAL}(D, q, x)$, whether optimal or not, satisfies

$$\text{val}(\text{PRIMAL}(D, q, x)) \leq \lambda p + \mu R_x + (1 - q)u + qv. \quad (4.4)$$

Consequently, given $T : [0, 1] \rightarrow [0, \infty)$, in order to prove that $g_D(q) \leq T(q)$ for all $q \in [0, 1]$, it suffices to prove that for each $q \in [0, 1]$ and $x \in (0, 1]$, at least one of the following holds.

1. $\text{PRIMAL}(D, q, x)$ is infeasible.
2. $(1 - p)x \leq T(q)$.
3. $\text{DUAL}(D, q, x)$ has a feasible solution (λ, μ, u, v) with $x(\lambda p + \mu R_x + (1 - q)u + qv) \leq T(q)$.

Proof. Since all objective and constraint functions are continuous, $\text{PRIMAL}(D, q, x)$ is a so-called generalized moment problem, and its dual is precisely the semi-infinite linear program provided in the lemma. For an interested reader, [Appendix B.1](#) derives the displayed dual via the Lagrangian.

Moreover, whenever $\text{PRIMAL}(D, q, x)$ is feasible, generalized-moment duality [[Sha01](#), Corollary 3.1] gives $\text{val}(\text{PRIMAL}(D, q, x)) = \text{val}(\text{DUAL}(D, q, x))$. The strong duality (equality of optimal values) requires checking a number of conditions on the program, but is stronger than what we need. We only need weak duality (which follows directly from first principles, as explained in [Appendix B.1](#)): the objective value of $\text{DUAL}(D, q, x)$ at every feasible solution (λ, μ, u, v) is at least $\text{val}(\text{PRIMAL}(D, q, x))$, which proves [Equation \(4.4\)](#). Combining this inequality with [Equation \(4.3\)](#) proves the final claim: for each feasible $x > 0$, the minimum in [Equation \(4.3\)](#) is at most $T(q)$ by one of the two stated alternatives, while $x = 0$ contributes zero. $\square$

5 A Rule with Distortion at Most 2.3282

By [Lemma 4](#), it is enough to exhibit (a) a draw distribution D and a target $T(q)$, and (b) for every $(q, x) \in [0, 1]^2$, which of its three conditions are met, and in case of the third condition, via which dual feasible solution (λ, μ, u, v) .

Table 2: Rule and target for [Lemma 5](#)

Let D^* be supported on $\{1, 12, 13\}$ with

$$\Pr[D^* = 1] = \frac{423239}{10^6}, \quad \Pr[D^* = 12] = \frac{471105}{10^6}, \quad \Pr[D^* = 13] = \frac{105656}{10^6}.$$

Its probability-generating function and stability threshold are

$$\phi_{D^*}(t) = \frac{423239t + 471105t^{12} + 105656t^{13}}{10^6}, \quad p_{D^*} = \int_0^1 \phi_{D^*}(t) dt = \frac{46483747}{182000000}.$$

Put

$$a = \frac{33205}{42059}, \quad b = \frac{7941}{42059}, \quad T(q) = q(a - bq).$$

[Table 2](#) provides (a). For (b), [Table 3](#) in [Appendix B.2](#) gives 49 parametric dual tuples $\mathbf{z}_j = (\lambda_j, \mu_j, u_j, v_j)$ for $j \in [49]$, and [Table 4](#) in [Appendix B.2](#) lists the 50 candidate bounds obtained from the (C3) bound and these 49 tuples. [Appendix B.3](#) proves that each dual tuple is feasible on its

stated range of x , while [Appendix B.4](#) proves that at least one eligible candidate bound applies at every relevant (q, x) . The case $x = 0$ is immediate, while [Appendix B.2](#) proves primal infeasibility at $x = 1$ for $q < 1 - p_{D^*}$.

Verifying a dual solution requires plugging the dual solution, a function of x , into the objective and constraints of $\text{DUAL}(D^*, q, x)$ and verifying throughout its stated domain that (a) all the constraints are met and (b) the objective is upper bounded by $T(q)$. Both parts reduce to verifying that a number of polynomials in one or two variables are nonnegative over their respective domains. This is verified using exact rational arithmetic by the algorithm of Garloff [[Gar93](#)], which rewrites the polynomials in their so-called Bernstein bases and uses the derived Bernstein coefficients to establish nonnegativity throughout the desired domain.

To verify that at least one of the 50 candidate bounds applies at every relevant point, we apply the algorithm of Garloff [[Gar93](#)] to the resulting disjunction of polynomial inequalities.³ When no single candidate can be verified throughout the current domain, the algorithm automatically subdivides the domain and repeats the test on the resulting subdivisions. This process is not guaranteed to terminate because nonnegative Bernstein coefficients are sufficient but not necessary for polynomial nonnegativity. However, whenever it terminates, the resulting subdivisions cover the entire domain and at least one required inequality holds throughout each subdivision, so the verification is sound. It terminates in our case using exact rational arithmetic. The verification code is available via our GitHub repository⁴ and instructions for running it as well as the expected output are listed in [Appendix B.7](#). This verification concludes in the following lemma, for which a complete proof is provided in [Appendix B.4](#).

Lemma 5. *Let D^* , a , and b be as specified in [Table 2](#). Then, $g_{D^*}(q) \leq q(a - bq)$ for all $q \in [0, 1]$. Consequently, for every profile and every random-size stable lottery RSL_{D^*} , we have $\ell(\text{RSL}_{D^*}, I) \leq q(I) \cdot (a - b \cdot q(I))$ for every nonempty proper $I \subsetneq C$.*

We are now ready for our main result.

Theorem 1. *There exists a randomized voting rule with metric distortion at most $11641/5000 = 2.3282$.*

Proof. Consider the rule

$$P^* = \frac{7941}{50000} \text{IV} + \frac{42059}{50000} \text{RSL}_{D^*},$$

and fix a nonempty proper $I \subsetneq C$. Write $q = q(I)$. By [Lemmas 1](#) and [5](#) and the linearity of $\ell(P, I)$ in P ,

$$\ell(P^*, I) \leq \frac{7941}{50000} q^2 + \frac{42059}{50000} q \left(\frac{33205}{42059} - \frac{7941}{42059} q \right) = \frac{6641}{10000} q. \quad (5.1)$$

Therefore, [Proposition 1](#) gives $\text{dist}(P^*) \leq 1 + 2 \cdot \frac{6641}{10000} = \frac{11641}{5000} = 2.3282$. $\square$

[Equation \(5.1\)](#) shows why we sought a loss upper bound of the form $T(q) = q(a - bq)$. This is so that the loss ratio is upper bounded as $\ell(\text{RSL}_{D^*}, I)/q(I) \leq a - bq(I)$, a linearly decreasing function of q . When mixed appropriately with the linearly increasing upper bound of $\ell(\text{IV}, I)/q(I) \leq q(I)$

³Garloff [[Gar93](#)] describe the subdivision test for one polynomial. We apply the same test to every eligible polynomial on each subdivision and accept that subdivision when any one test succeeds.

⁴<https://github.com/nisarg89/metric-distortion>

for IV (Lemma 1), the variation cancels out, giving us the constant upper bound of $\frac{a}{b+1} = \frac{6641}{10000}$ on κ and allowing us to apply Proposition 1.

Figure 2 shows the upper bounds on the loss ratios of IV (linearly increasing), RSL_{D^*} (linearly decreasing), and the mixture P^* (constant at $\frac{6641}{10000}$).

A note on the search process. While the search process used to identify the candidate solutions presented in Tables 2 and 3 is irrelevant from the perspective of proving the claimed distortion bound, let us describe its stages at a high level to provide some intuition; Appendices B.5 and B.6 describe the full process, and Appendix B.7 provides instructions for an interested reader to execute our code and perform an independent search.

- First, for fixed (D, q, x) , we aim to solve $\text{DUAL}(D, q, x)$ via constraint-generation: we start by solving it with only finitely many constraints imposed at some $E \subset [0, 1]^2$ and $E_1 \subset [0, 1]$; call a separation oracle on the returned solution to check if a constraint corresponding to some $(b_0, y) \in [0, 1]^2$ or $b_1 \in [0, 1]$ is violated and add a violated constraint to E or E_1 , respectively; and repeat until the separation oracle fails to find a violated dual constraint. This process is plugged inside a broader grid-search over $(q, x) \in [0, 1]^2$ for each fixed D , and that in turn inside an ad hoc search over D in the probability simplex with support $S \subset [60]$ of size $|S| \leq 4$.
- Note that this search is by no means exact. The use of ϕ in the DUAL coefficients makes the separation oracle nonlinear; SciPy’s `differential_evolution` heuristic is used to implement it, which may fail to find a violated constraint even when one exists. Hence, exact feasibility of any candidate dual feasible solution identified must be verified separately using rational arithmetic. The same issue applies to the numerical LP solver HiGHS from SciPy we use to solve DUAL on a finite set of constraints corresponding to (E, E_1) , which may return a solution that may turn out to be infeasible under an exact feasibility check. Finally, an upper bound $g_D(q) \leq T(q)$ requires taking the supremum over $x \in [0, 1]$ and must hold for all $q \in [0, 1]$, whereas we use only a grid over $(q, x) \in [0, 1]^2$.
- Nonetheless, the grid search produces a candidate (D^*, a, b) with the lowest identified $a/(1+b)$, listed in Table 2, for which we have evidence to believe that $g_{D^*}(q) \leq q(a - bq)$ for all $q \in [0, 1]$ may hold. To certify this, based on the candidate dual solutions found at the grid points, it also identified which constraints are tight in different regions and used that to analytically derive and rationalize the 49 parametrized dual solutions as functions of x listed in Table 3. As we described above, subsequent exact verification was successful on these candidates identified by the search.

6 Discussion

Room for improvement. We see four avenues to improve the bound with differing prospects.

- *Numerical certification.* The numerical search estimated a bound of approximately 2.32818, compared with the certified bound of 2.3282. Hence, more precise rationalization and stronger numerical certification may have little to no room for improvement.

- *The draw distribution.* Even when we searched over draw distributions with support size 4, the best we could find placed positive masses only on 1, 12, and 13. Our search suggests that the optimal distribution seems to converge to bimodality with modes at 1 and p for some positive real number p , with mass placed on p discretized by placing appropriate masses on $\lfloor p \rfloor$ and $\lceil p \rceil$; if this is true, it would suggest that the optimal distribution has support structure very similar to that of our D^* and looking for alternative draw distributions may only yield modest improvement.
- *The analysis.* Replacing the affine loss envelope by a tighter one may help, but we are more optimistic about a joint analysis of IV and RSL_{D^*} that does not simply take a convex combination of their separate worst-case bounds, or about retaining more of the constraints imposed by the profile in our initial optimization program in [Lemma 2](#). Even more fundamentally, our entire analysis hinges on optimizing κ to then apply [Proposition 1](#), but this result itself is only a sufficient condition, not necessary.
- *The rule.* We genuinely do not know whether a mixture of IV, ML, and one or several random-size stable lotteries $\text{RSL}_{D_1}, \text{RSL}_{D_2}, \dots$ might already reach optimal distortion (the lower bound sits at 2.1126...), or whether closing the gap may require a fundamentally new rule.

For $m \geq 4$ candidates, determining the optimal randomized metric distortion remains open, including whether its limit as $m \rightarrow \infty$ equals the asymptotic lower bound 2.1126... of Charikar and Ramakrishnan [[CR22](#)].

A fascinating experience working with GPT-5.6-Sol. The bound of 2.3282 emerged through a rapid series of improvements.

- When we provided the work of Frank [[Fra26](#)], together with detailed instructions on how to effectively make improvements on open problems, GPT-5.6-Sol autonomously identified an appropriate ingredient to add—stable 2-lottery SL_2 [[CRTW25](#)—and achieved a bound of 2.4924 using a mixture of $(\text{IV} + \text{ML})/2$ and SL_2 .
- Then, it explored placing unequal weights on IV and ML, which reduced the bound to 2.4662.
- When we asked it to find a further improvement, GPT-5.6-Sol autonomously came up with the idea of using a lottery P for which $\frac{1}{2} \Pr_v[c \succ_v P] + \frac{1}{2} \Pr_v[c \succ_v P^2] \leq \frac{5}{12}$, proved its existence using a minimax argument, and reduced the bound to 2.42 by mixing it with IV and ML.
- At this point, we generalized its proposed lottery into the random-size stable lottery RSL_D presented in this work and asked it to systematically search over the space of D .
- This led to a rapid series of autonomous improvements, which culminated in the 2.3282 bound presented in this work.

We have not manually verified all these intermediate results, only the final bound and its corresponding mathematical claims presented in this work. Watching this sequence unfold over the short span of 2 days was both a fascinating experience and thought-provoking regarding the nature of research and implications for science in an era in which such rapid autonomous discoveries are possible.

AI Disclosure

All the mathematical proofs were derived by OpenAI Codex (GPT-5.6-Sol at Max effort) based on research directions, literature connections, proof and search strategies, and inspirations supplied by the author. The author has verified all mathematical details, presented simplified arguments and detailed expositions, often with the aid of GPT-5.6-Sol and Claude Opus 5, and retains full responsibility for all the content and any errors.

Acknowledgments

This work was supported by an NSERC Discovery Grant and an NSERC-CSE Research Communities Grant. Researchers funded through the NSERC-CSE Research Communities Grants do not represent the Communications Security Establishment Canada or the Government of Canada. Any research, opinions or positions they produce as part of this initiative do not represent the official views of the Government of Canada.

References

- [ABEPS18] Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. “Approximating Optimal Social Choice under Metric Preferences”. In: *Artificial Intelligence* 264 (2018), pp. 27–51.
- [ABFV21] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. “Peeking Behind the Ordinal Curtain: Improving Distortion via Cardinal Queries”. In: *Artificial Intelligence* 296 (2021), p. 103488.
- [ABP15] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. “Approximating Optimal Social Choice under Metric Preferences”. In: *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*. 2015, pp. 777–783.
- [Arr51] Kenneth J. Arrow. *Social Choice and Individual Values*. Wiley, 1951.
- [BCHLPS15] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. “Optimal Social Choice Functions: A Utilitarian View”. In: *Artificial Intelligence* 227 (2015), pp. 190–213.
- [BHLS22] Allan Borodin, Daniel Halpern, Mohamad Latifian, and Nisarg Shah. “Distortion in Voting with Top- t Preferences”. In: *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*. 2022, pp. 116–122.
- [Bra17] Felix Brandt. “Rolling the Dice: Recent Results in Probabilistic Social Choice”. In: *Trends in Computational Social Choice*. 2017, pp. 3–26.
- [CJMW20] Yu Cheng, Zhihao Jiang, Kamesh Munagala, and Kangning Wang. “Group Fairness in Committee Selection”. In: *ACM Transactions on Economics and Computation* 8.4 (2020), 23:1–23:18.
- [CNPS17] Ioannis Caragiannis, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. “Subset Selection via Implicit Utilitarian Voting”. In: *Journal of Artificial Intelligence Research* 58 (2017), pp. 123–152.

- [CP11] Ioannis Caragiannis and Ariel D. Procaccia. “Voting Almost Maximizes Social Welfare Despite Limited Communication”. In: *Artificial Intelligence* 175.9–10 (2011), pp. 1655–1671.
- [CR22] Moses Charikar and Prasanna Ramakrishnan. “Metric Distortion Bounds for Randomized Social Choice”. In: *Proceedings of the 33rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2022, pp. 2986–3004.
- [CRTW25] Moses Charikar, Prasanna Ramakrishnan, Zihan Tan, and Kangning Wang. “Metric Distortion for Tournament Voting and Beyond”. In: *Proceedings of the 26th ACM Conference on Economics and Computation (EC)*. 2025, pp. 790–818.
- [CRWW24] Moses Charikar, Prasanna Ramakrishnan, Kangning Wang, and Hongxun Wu. “Breaking the Metric Voting Distortion Barrier”. In: *Journal of the ACM* 71.6 (2024), 42:1–42:33.
- [EKPS24] Soroush Ebadian, Anson Kahng, Dominik Peters, and Nisarg Shah. “Optimized Distortion and Proportional Fairness in Voting”. In: *ACM Transactions on Economics and Computation* 12.1 (2024), 3:1–3:39.
- [ES25] Soroush Ebadian and Nisarg Shah. “Every Bit Helps: Achieving the Optimal Distortion with a Few Queries”. In: *Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI)*. 2025, pp. 13788–13795.
- [Fis84] Peter C. Fishburn. “Probabilistic Social Choice Based on Simple Voting Comparisons”. In: *The Review of Economic Studies* 51.4 (1984), pp. 683–692.
- [FM92] Dan S. Felsenthal and Moshé Machover. “After Two Centuries, Should Condorcet’s Voting Procedure Be Implemented?”. In: *Behavioral Science* 37.4 (1992), pp. 250–274.
- [FR95] David C. Fisher and Jennifer Ryan. “Tournament Games and Positive Tournaments”. In: *Journal of Graph Theory* 19.2 (1995), pp. 217–236.
- [Fra26] Fabian Frank. *An Improved Bound for the Randomized Metric Distortion Problem*. arXiv:2608.17863. 2026.
- [Gar93] Jürgen Garloff. “The Bernstein Algorithm”. In: *Interval Computations* 2 (1993), pp. 154–168.
- [GHS20] Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. “Resolving the Optimal Metric Distortion Conjecture”. In: *Proceedings of the 61st IEEE Symposium on Foundations of Computer Science (FOCS)*. 2020, pp. 1427–1438.
- [Gib73] Allan Gibbard. “Manipulation of Voting Schemes: A General Result”. In: *Econometrica* 41.4 (1973), pp. 587–601.
- [GKM17] Ashish Goel, Anilesh Kollagunta Krishnaswamy, and Kamesh Munagala. “Metric Distortion of Social Choice Rules: Lower Bounds and Fairness Properties”. In: *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*. 2017, pp. 287–304.
- [GLS23] Vasilis Gkatzelis, Mohamad Latifian, and Nisarg Shah. “Best of Both Distortion Worlds”. In: *Proceedings of the 24th ACM Conference on Economics and Computation (EC)*. 2023, pp. 738–758.

- [JMW20] Zhihao Jiang, Kamesh Munagala, and Kangning Wang. “Approximately Stable Committee Selection”. In: *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing (STOC)*. 2020, pp. 463–472.
- [Kem20a] David Kempe. “An Analysis Framework for Metric Voting Based on LP Duality”. In: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*. 2020, pp. 2079–2086.
- [Kem20b] David Kempe. “Communication, Distortion, and Randomness in Metric Voting”. In: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*. 2020, pp. 2087–2094.
- [KK22] Fatih Erdem Kizilkaya and David Kempe. “Plurality Veto: A Simple Voting Rule Achieving Optimal Metric Distortion”. In: *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*. 2022, pp. 349–355.
- [KK23] Fatih Erdem Kizilkaya and David Kempe. “Generalized Veto Core and a Practical Voting Rule with Optimal Metric Distortion”. In: *Proceedings of the 24th ACM Conference on Economics and Computation (EC)*. 2023, pp. 913–936.
- [Kre65] Germain Kreweras. “Aggregation of Preference Orderings”. In: *Mathematics and Social Sciences I*. 1965, pp. 73–79.
- [LLL93] Gilbert Laffond, Jean-François Laslier, and Michel Le Breton. “The Bipartisan Set of a Tournament Game”. In: *Games and Economic Behavior* 5.1 (1993), pp. 182–201.
- [Mas90] Pascal Massart. “The Tight Constant in the Dvoretzky–Kiefer–Wolfowitz Inequality”. In: *The Annals of Probability* 18.3 (1990), pp. 1269–1283.
- [MPSW19] Debmalya Mandal, Ariel D. Procaccia, Nisarg Shah, and David P. Woodruff. “Efficient and Thrifty Voting by Any Means Necessary”. In: *Proceedings of the 33rd Annual Conference on Neural Information Processing Systems (NeurIPS)*. 2019, pp. 7178–7189.
- [MSW20] Debmalya Mandal, Nisarg Shah, and David P. Woodruff. “Optimal Communication–Distortion Tradeoff in Voting”. In: *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*. 2020, pp. 795–813.
- [MW19] Kamesh Munagala and Kangning Wang. “Improved Metric Distortion for Deterministic Social Choice Rules”. In: *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*. 2019, pp. 245–262.
- [PR06] Ariel D. Procaccia and Jeffrey S. Rosenschein. “The Distortion of Cardinal Preferences in Voting”. In: *Proceedings of the 10th International Workshop on Cooperative Information Agents (CIA)*. 2006, pp. 317–331.
- [PS21] Haripriya Pulyassary and Chaitanya Swamy. *On the Randomized Metric Distortion Conjecture*. arXiv:2111.08698. 2021.
- [RS10] Ronald L. Rivest and Emily Shen. “An Optimal Single-Winner Preferential Voting System Based on Game Theory”. In: *Proceedings of the 3rd International Workshop on Computational Social Choice (COMSOC)*. 2010, pp. 399–410.
- [Sat75] Mark Allen Satterthwaite. “Strategy-Proofness and Arrow’s Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions”. In: *Journal of Economic Theory* 10.2 (1975), pp. 187–217.

- [Sha01] Alexander Shapiro. “On Duality Theory of Conic Linear Problems”. In: *Semi-Infinite Programming: Recent Advances*. 2001, pp. 135–165.
- [Ye26] Qilin Ye. *Half Veto, Half Maximal Lottery, Five-Halves Distortion*. arXiv:2608.21202. 2026.

Appendix

A Further Properties of Random-Size Stable Lotteries

This appendix records further consequences of random-size stability. It relates the definition to deterministic challenger committees, comparisons using other numbers of samples, and lotteries with small support. These consequences do not subsume stable lotteries over committees [CJMW20; JMW20]: here every committee consists of independent draws from one lottery and a voter compares committees by their favorite members, whereas that literature permits correlated committee lotteries and more general committee preferences.

A.1 Why separately chosen stable lotteries cannot simply be mixed

Suppose masses $3/5$ and $2/5$ of the voters rank $a \succ b$ and $b \succ a$, respectively. The lottery $P_1 = \delta_a$ is stable for one draw, while $P_2 = (4/5)\delta_a + (1/5)\delta_b$ is stable for two draws because

$$\begin{aligned}\Pr_v[a \succ_v P_2^2] &= \frac{3}{5} \int_0^1 \left(\frac{1}{5} + \frac{4}{5}s \right)^2 ds + \frac{2}{5} \int_0^1 \left(\frac{4}{5}s \right)^2 ds = \frac{1}{3}, \\ \Pr_v[b \succ_v P_2^2] &= \frac{3}{5} \int_0^1 \left(\frac{1}{5}s \right)^2 ds + \frac{2}{5} \int_0^1 \left(\frac{4}{5} + \frac{1}{5}s \right)^2 ds = \frac{1}{3}.\end{aligned}$$

If D is uniform on $\{1, 2\}$, their average $P = (P_1 + P_2)/2$ satisfies

$$\Pr_v[a \succ_v P] = \frac{51}{100}, \quad \Pr_v[a \succ_v P^2] = \frac{33}{100},$$

and hence $\mathbb{E}_D \Pr_v[a \succ_v P^D] = 21/50 > 5/12 = \mathbb{E}[1/(D+1)]$. The cross terms created by drawing repeatedly from P are not controlled by the separate guarantees for P_1 and P_2 . Selecting P_D after observing D would satisfy a different averaged statement, not the requirement that one base lottery be chosen before D .

A.2 Computing a random-size stable lottery

Proposition 2. *Suppose D has finite support of size s , whose largest value is $d_{\max}$, and let $m = |C| \geq 2$. For every positive integer N , one can compute a lottery $\bar{P}$ satisfying*

$$\max_{c \in C} \mathbb{E}_D \Pr_v[c \succ_v \bar{P}^D] \leq \mathbb{E} \left[\frac{1}{D+1} \right] + \sqrt{\frac{\log m}{2N}}.$$

In the real-arithmetic model, the algorithm uses $O(Nnms \log(d_{\max} + 1))$ arithmetic operations and exponential evaluations. Thus additive error at most ε is obtained in $O(nms \log(d_{\max} + 1) \log(m)/\varepsilon^2)$ time.

Proof. Start with the uniform lottery P_1 and set $\eta = \sqrt{8 \log(m)/N}$. At iteration t , put $g_{t,c} = \mathbb{E}_D \Pr_v[c \succ_v P_t^D]$ and update

$$P_{t+1}(c) = \frac{P_t(c) \exp(\eta g_{t,c})}{\sum_{b \in C} P_t(b) \exp(\eta g_{t,b})}.$$

The standard exponential-weights calculation for numbers in $[0, 1]$ gives

$$\max_{c \in C} \frac{1}{N} \sum_{t=1}^N g_{t,c} \leq \frac{1}{N} \sum_{t=1}^N \sum_{c \in C} P_t(c) g_{t,c} + \frac{\log m}{\eta N} + \frac{\eta}{8}.$$

The exchangeability identity proved inside [Theorem 2](#) makes the inner sum $\mathbb{E}[1/(D+1)]$ at every iteration. With the chosen η , the last two terms sum to $\sqrt{\log(m)/(2N)}$. Finally, let $\bar{P} = N^{-1} \sum_{t=1}^N P_t$. The convexity established in the same proof gives $\mathbb{E}_D \Pr_v[c \succ_v \bar{P}^D] \leq N^{-1} \sum_t \mathbb{E}_D \Pr_v[c \succ_v P_t^D]$ for every c .

It remains to account for the running time. For a voter v and candidate c , let $b_{v,c}$ be the P_t -mass strictly below c . Because every $P_t(c)$ is positive, integrating over the label of the challenger copy gives

$$g_{t,c} = \frac{1}{n} \sum_{v \in V} \sum_{d \in \text{supp}(D)} \Pr[D = d] \frac{(b_{v,c} + P_t(c))^{d+1} - b_{v,c}^{d+1}}{(d+1)P_t(c)}.$$

For each voter, all values $b_{v,c}$ are obtained in one scan of her ranking. Repeated squaring then evaluates the displayed powers in $O(nms \log(d_{\max} + 1))$ arithmetic operations per iteration; the multiplicative-weights update costs another $O(m)$ operations and exponential evaluations. This proves the claimed total. $\square$

For an infinite-support distribution, the same iteration guarantee holds provided that the required expectations can be evaluated to sufficient precision, but the running time then depends on the cost of those evaluations.

For a deterministic nonempty committee $A \subseteq C$, the event $A \succ_v (\text{RSL}_D)^D$ means that voter v 's favorite labeled member of A beats her favorite of the D independent draws.

Proposition 3. *Every nonempty committee $A \subseteq C$ satisfies*

$$\mathbb{E}_D \Pr_v[A \succ_v (\text{RSL}_D)^D] \leq \min \left\{ 1, |A| \mathbb{E} \left[\frac{1}{D+1} \right] \right\}.$$

Moreover,

$$\min_{P \in \Delta(C)} \max_{\emptyset \neq A \subseteq C} \frac{\mathbb{E}_D \Pr_v[A \succ_v P^D]}{|A|} = \mathbb{E} \left[\frac{1}{D+1} \right].$$

Proof. The event that A defeats $(\text{RSL}_D)^D$ is the union, over $c \in A$, of the event that the labeled copy of c defeats $(\text{RSL}_D)^D$. The union bound and [Equation \(3.1\)](#) give the first inequality. For any lottery P , the exchangeability identity in the proof of [Theorem 2](#) gives some candidate c with $\mathbb{E}_D \Pr_v[c \succ_v P^D] \geq \mathbb{E}[1/(D+1)]$, which proves the matching lower bound for the normalized maximum. $\square$

Fix a committee A , expose the voter and the labeled copies in A , and let $T \in [0, 1]$ be the probability that one draw from RSL_D lies below the voter's favorite labeled member of A . Then $\Pr_v[A \succ_v (\text{RSL}_D)^D] = \mathbb{E}[T^D]$ and $\mathbb{E}_D \Pr_v[A \succ_v (\text{RSL}_D)^D] = \mathbb{E}[\phi(T)]$.

Because D is positive-integer-valued, ϕ is continuous and strictly increasing from 0 to 1 on $[0, 1]$, so its inverse $\phi^{-1} : [0, 1] \rightarrow [0, 1]$ is well-defined.

Lemma 6. Let $s_A = \min \{1, |A| \mathbb{E}[1/(D+1)]\}$. For an integer $r \geq 1$, define $h_{r,\phi}(u) = (\phi^{-1}(u))^r$, and let $\text{cav } h_{r,\phi}$ be the smallest concave function on $[0, 1]$ lying above $h_{r,\phi}$. Then

$$\Pr_v[A \succ_v (\text{RSL}_D)^r] \leq (\text{cav } h_{r,\phi})(s_A).$$

If $D = d$ deterministically, this bound is $s_A^{r/d}$ for $r \leq d$ and s_A for $r \geq d$.

Proof. By [Proposition 3](#), $\mathbb{E}[\phi(T)] \leq s_A$. The function $\text{cav } h_{r,\phi}$ is nondecreasing, since it is at most the constant concave upper bound 1 and equals 1 at the right endpoint. Because $T^r = h_{r,\phi}(\phi(T))$, the definition of the concave upper bound and Jensen's inequality give

$$\mathbb{E}[T^r] \leq \mathbb{E}[(\text{cav } h_{r,\phi})(\phi(T))] \leq (\text{cav } h_{r,\phi})(\mathbb{E}[\phi(T)]) \leq (\text{cav } h_{r,\phi})(s_A).$$

For $\phi(t) = t^d$, the function $u^{r/d}$ is concave when $r \leq d$; when $r \geq d$, its smallest concave upper bound on $[0, 1]$ is the line u . $\square$

Let B be a lottery, let R be an independent positive integer, and write $\psi(t) = \mathbb{E}[t^R]$.

Theorem 3. Every random-size stable lottery satisfies

$$\Pr_v[B^R \succ_v (\text{RSL}_D)^D] \leq 1 - \psi\left(1 - \mathbb{E}\left[\frac{1}{D+1}\right]\right).$$

For deterministic $R = r$, the bound becomes $1 - (1 - \mathbb{E}[1/(D+1)])^r$.

Finally, random-size stable lotteries can be approximated in support, although the resulting additive error is too coarse for the main distortion proof. Let $\bar{d} = \mathbb{E}[D] = \phi'(1)$.

Proposition 4. For every integer $s \geq 1$, there is a lottery $\hat{P}$ supported on at most s candidates such that

$$\mathbb{E}_D \Pr_v[c \succ_v \hat{P}^D] \leq \mathbb{E}\left[\frac{1}{D+1}\right] + \bar{d} \sqrt{\frac{\pi}{2s}} \quad \text{for every } c \in C.$$

We now prove the last two statements and then record an extension to signed coefficients. The signed extension is mathematically useful but does not describe a random value of D unless the coefficients are nonnegative and sum to one.

A.3 A calculation with cumulative distributions

Draw a uniformly random voter v and order labeled candidates from worst to best according to v . For lotteries A and P , let $F_{A,v}$ and $F_{P,v}$ be their cumulative distribution functions in this order. Draw a labeled outcome ω from A and put

$$U = F_{P,v}(\omega), \quad Y = F_{A,v}(\omega).$$

Conditional on v , Y is uniform on $[0, 1]$. For every integer $d \geq 1$,

$$\Pr_v[A \succ_v P^d] = \mathbb{E}[U^d], \quad \Pr_v[A^d \succ_v P] = d \mathbb{E}[UY^{d-1}]. \quad (\text{A.1})$$

The first identity conditions on the draw from A . For the second, the best of d independent A -draws has density $dF_{A,v}^{d-1} dF_{A,v}$. Summing [Equation \(A.1\)](#) over the distribution of D gives the formulas used in [Lemma 7](#):

$$\mathbb{E}_D \Pr_v[A \succ_v P^D] = \mathbb{E}[\phi(U)], \quad \mathbb{E}_D \Pr_v[A^D \succ_v P] = \mathbb{E}[U\phi'(Y)].$$

Lemma 7. For finitely supported D , define

$$\mathcal{D}_\phi(u, v) = \phi(u) - \phi(v) - \phi'(v)(u - v).$$

Then

$$\mathbb{E}_D \Pr_v[A^D \succ_v P] = \mathbb{E}_D \Pr_v[A \succ_v P^D] + 1 - 2 \mathbb{E} \left[\frac{1}{D+1} \right] - \mathbb{E}[\mathcal{D}_\phi(U, Y)].$$

Consequently, $\mathbb{E}_D \Pr_v[A \succ_v P^D] \leq \mathbb{E}[1/(D+1)]$ implies $\mathbb{E}_D \Pr_v[A^D \succ_v P] \leq 1 - \mathbb{E}[1/(D+1)]$. The probability inequality extends to arbitrary D by truncation.

Proof. Convexity gives $\mathcal{D}_\phi \geq 0$, and rearranging its definition gives

$$U\phi'(Y) = \phi(U) - \mathcal{D}_\phi(U, Y) + Y\phi'(Y) - \phi(Y).$$

Because Y is uniform conditional on every voter,

$$\mathbb{E}[Y\phi'(Y) - \phi(Y)] = \int_0^1 (t\phi'(t) - \phi(t)) dt = 1 - 2 \int_0^1 \phi(t) dt = 1 - 2 \mathbb{E} \left[\frac{1}{D+1} \right].$$

Substitution into the two preceding cumulative-distribution formulas proves the identity, and discarding the nonnegative term proves the consequence. Equality holds when $A = P$ by exchangeability. $\square$

A.4 Several draws from another lottery

For this proof, abbreviate $p_D = \mathbb{E}[1/(D+1)]$. Draw a labeled ω from B . The best of R draws from B has density $\psi'(F_{B,v}(\omega))$ with respect to a draw from B . Conditional on this best draw, all D draws from RSL_D lie below it with probability $\phi(F_{\text{RSL}_D,v}(\omega))$. With $U = F_{\text{RSL}_D,v}(\omega)$ and $Y = F_{B,v}(\omega)$, therefore,

$$\Pr_v[B^R \succ_v (\text{RSL}_D)^D] = \mathbb{E}[\phi(U)\psi'(Y)]. \quad (\text{A.2})$$

By linearity of Equation (3.1) in the candidate drawn from B ,

$$\mathbb{E}[\phi(U)] = \mathbb{E}_D \Pr_v[B \succ_v (\text{RSL}_D)^D] \leq p_D.$$

Also $0 \leq \phi(U) \leq 1$, and Y is uniform on $[0, 1]$. The function ψ' is nondecreasing. We claim that for every random pair (Z, Y) with $0 \leq Z \leq 1$, $\mathbb{E} Z \leq p_D$, and uniform marginal Y ,

$$\mathbb{E}[Z\psi'(Y)] \leq \int_{1-p_D}^1 \psi'(t) dt. \quad (\text{A.3})$$

To prove this directly, put $\lambda = \psi'(1 - p_D)$. Pointwise,

$$Z(\psi'(Y) - \lambda) \leq (\psi'(Y) - \lambda)_+.$$

Taking expectations and using $\mathbb{E} Z \leq p_D$ gives

$$\mathbb{E}[Z\psi'(Y)] \leq \lambda p_D + \int_0^1 (\psi'(t) - \lambda)_+ dt = \int_{1-p_D}^1 \psi'(t) dt,$$

where a flat portion at the threshold does not affect the equality. Applying Equation (A.3) to $Z = \phi(U)$ in Equation (A.2) yields

$$\Pr_v[B^R \succ_v (\text{RSL}_D)^D] \leq \int_{1-p_D}^1 \psi'(t) dt = 1 - \psi(1 - p_D).$$

For deterministic $R = r$, substitute $\psi(t) = t^r$.

A.5 Approximating by a lottery with small support

Draw s candidates independently from RSL_D , and let $\hat{P}$ be their empirical distribution. Fix a voter. Place the candidates in this voter's order and linearly interpolate across the interval corresponding to each candidate; this interpolation is exactly the proportional label convention. Let F and $\hat{F}$ be the resulting cumulative distribution functions for RSL_D and $\hat{P}$. The sharp Dvoretzky–Kiefer–Wolfowitz inequality [Mas90] gives

$$\Pr \left[\sup_t |\hat{F}(t) - F(t)| > \varepsilon \right] \leq 2e^{-2s\varepsilon^2}.$$

Integrating the tail bound gives

$$\mathbb{E} \left[\sup_t |\hat{F}(t) - F(t)| \right] \leq \int_0^\infty 2e^{-2s\varepsilon^2} d\varepsilon = \sqrt{\frac{\pi}{2s}}. \quad (\text{A.4})$$

The same sampled candidates are used for every voter; Equation (A.4) is applied separately and then averaged over voters.

Because $\phi'(t) \leq \phi'(1) = \bar{d}$ on $[0, 1]$, the function ϕ is $\bar{d}$ -Lipschitz. For any pure candidate c , the probability that c defeats the empirical committee differs from the corresponding probability under RSL_D by at most $\bar{d} \sup_t |\hat{F}(t) - F(t)|$ for that voter. Therefore

$$\mathbb{E}_{\hat{P}} \left[\max_{c \in C} (\mathbb{E}_D \Pr_v[c \succ_v \hat{P}^D] - \mathbb{E}_D \Pr_v[c \succ_v (\text{RSL}_D)^D]) \right] \leq \bar{d} \sqrt{\frac{\pi}{2s}}.$$

The maximum causes no factor depending on $|C|$: for each voter, the same uniform error in the cumulative distributions controls every candidate position. Since $\mathbb{E}_D \Pr_v[c \succ_v (\text{RSL}_D)^D] \leq \mathbb{E}[1/(D+1)]$ for every c , some realization of the sample satisfies

$$\mathbb{E}_D \Pr_v[c \succ_v \hat{P}^D] \leq \mathbb{E} \left[\frac{1}{D+1} \right] + \bar{d} \sqrt{\frac{\pi}{2s}}$$

simultaneously for all candidates. Its empirical distribution has support at most s .

A.6 An extension to signed coefficients

Theorem 4. Let $(a_d)_{d \geq 1}$ be an absolutely summable sequence of real numbers, define

$$f(t) = \sum_{d \geq 1} a_d t^d,$$

and let $p_f = \int_0^1 f(t) dt$. Every finite profile admits a lottery P such that

$$\sum_{d \geq 1} a_d \Pr_v[c \succ_v P^d] \leq \int_0^1 f(t) dt$$

for every candidate c . The threshold is exact: for every lottery P , the maximum of the left-hand side over candidates is at least $\int_0^1 f(t) dt$.

Proof. Within this proof, write $G_c(P)$ for the left-hand side. Since $p_f = \sum_{d \geq 1} a_d / (d + 1)$, absolute summability gives uniform convergence of the defining series on $[0, 1]$, hence continuity of every G_c . The exchangeability argument in the proof of [Theorem 2](#) is linear in the coefficients and gives

$$\sum_c P(c) G_c(P) = p_f.$$

For a lottery P , let $M(P) = \arg \max_{c \in \mathcal{C}} G_c(P)$, and let $B(P)$ be the set of lotteries supported on $M(P)$. Each $B(P)$ is nonempty, convex, and compact. Continuity of the functions G_c implies that the correspondence B has a closed graph. Kakutani's fixed-point theorem therefore gives a lottery P with $P \in B(P)$. Every candidate in the support of this lottery maximizes $G_c(P)$, so the displayed identity gives

$$\max_c G_c(P) = \sum_c P(c) G_c(P) = p_f.$$

For every lottery, the maximum of the numbers $G_c(P)$ is at least their P -weighted average p_f , which proves the converse. $\square$

Signed coefficients do not define probabilities and need not make the weighted comparison function convex in P , so the minimax proof of [Theorem 2](#) does not apply. They also do not imply the union bound or the consequence of [Lemma 7](#) used in [Proposition 3](#). We use nonnegative coefficients summing to one everywhere in the distortion results.

B Exact Verification and Numerical Discovery of the Distortion Bound

This appendix first gives the construction of DUAL omitted from the main body, states the 49 exact dual families, and then proves [Lemma 5](#) from these data and the rule and target in [Section 5](#). It gives a uniform, labeled description of every polynomial inequality checked by the program, explains the exact method used to prove those inequalities on continuous domains, and verifies that at least one of the candidate bounds applies at every relevant (q, x) . Only afterward do we describe how finite restrictions of PRIMAL and DUAL suggested the support and masses of D^* , the target coefficients a and b , and the three constructions that produce the dual families. The search is numerical evidence; the proof uses the original programs from [Lemmas 3](#) and [4](#) and exact rational arithmetic.

B.1 Construction of DUAL

Fix D , q , and x as in [Lemma 4](#). Give the two moment constraints of PRIMAL nonnegative multipliers λ and μ , while retaining the requirements $\eta \in \Delta([0, 1]^2)$ and $\xi \in \Delta([0, 1])$. The resulting Lagrangian is

$$\begin{aligned} \mathcal{L}(\eta, \xi; \lambda, \mu) = & \lambda p + \mu R_x \\ & + (1 - q) \int_{[0, 1]^2} (b_0 - \lambda W_x(y) - \mu G_x(b_0, y)) \, d\eta(b_0, y) \\ & + q \int_{[0, 1]} (b_1 - \lambda w_x(b_1) - \mu F_x(b_1)) \, d\xi(b_1). \end{aligned}$$

For every continuous function h on a compact set S ,

$$\sup_{\nu \in \Delta(S)} \int_S h \, d\nu = \max_{s \in S} h(s),$$

because an average cannot exceed the pointwise maximum and a point mass at a maximizer attains it. Thus, for fixed λ and μ , maximizing the Lagrangian over η and ξ replaces its two integrals by the corresponding pointwise maxima. Introducing u and v as upper bounds on these maxima gives

$$\begin{aligned} u &\geq b_0 - \lambda W_x(y) - \mu G_x(b_0, y) && \text{for every } (b_0, y) \in [0, 1]^2, \\ v &\geq b_1 - \lambda w_x(b_1) - \mu F_x(b_1) && \text{for every } b_1 \in [0, 1]. \end{aligned}$$

Minimizing $\lambda p + \mu R_x + (1 - q)u + qv$ subject to these inequalities and $\lambda, \mu \geq 0$ is exactly DUAL.

Equivalently, the pair (η, ξ) can be represented by one measure on the compact disjoint union $[0, 1]^2 \sqcup [0, 1]$, with moment constraints fixing the mass of each component. The objective and all moment functions are continuous, so generalized-moment duality implies that the primal and dual optimal values coincide whenever PRIMAL is feasible [Sha01, Corollary 3.1]. Weak duality, which is all that Lemma 4 uses, also follows immediately by integrating the two pointwise dual inequalities against η and ξ and applying the two moment constraints of PRIMAL.

B.2 Exact Dual Families and Candidate Bounds

Throughout this subsection and the two verification subsections that follow, set $f = \phi_{D^*}$, $p = p_{D^*}$, and $r = 1 - x$; all the functions from Equations (4.1) and (4.2) use this f .

Table 3: The 49 dual families

Dual families 1 and 2. Put $A_x = G_x(1, 0)$, $y_1 = 80529/100000$, and $y_2 = 1$. For $0 < x < 1$ and each $j \in \{1, 2\}$, let $\mathbf{z}_j = (\lambda_j, \mu_j, u_j, v_j)$, where

$$\Delta_j(x) = (f(x) - W_x(y_j)) \partial_y G_x(0, y_j) - (A_x - G_x(0, y_j)) W'_x(y_j)$$

and

$$\lambda_j = \frac{\partial_y G_x(0, y_j)}{\Delta_j(x)}, \quad \mu_j = -\frac{W'_x(y_j)}{\Delta_j(x)}, \quad u_j = -\lambda_j W_x(y_j) - \mu_j G_x(0, y_j), \quad v_j = 0.$$

Dual family 3. Let $\Psi(s) = \int_0^s f(t) \, dt$ and $L_x = p - \Psi(x) - \Psi(r) - xf(r)$. For $0 < x < 1$, let $\mathbf{z}_3 = (\lambda_3, \mu_3, u_3, v_3)$, where

$$\lambda_3 = \frac{L_x}{xf(r)(1 - f(x))}, \quad \mu_3 = \frac{r}{f(r)}, \quad u_3 = -\lambda_3, \quad v_3 = 0.$$

Dual families 4–49. For each $j \in \{4, 5, \dots, 49\}$, put $\tau_j = (j + 11)/100$, $S_j(x) = \lambda_3 x f'(x \tau_j) + \mu_3 F'_x(\tau_j)$, and $\theta_j(x) = 1/S_j(x)$. For $0 < x < 1$, let $\mathbf{z}_j = (\lambda_j, \mu_j, u_j, v_j)$, where

$$\begin{aligned} \lambda_j &= \theta_j \lambda_3, & \mu_j &= \theta_j \mu_3, \\ u_j &= (1 - \theta_j) + \theta_j u_3, & v_j &= \tau_j - \lambda_j f(x \tau_j) - \mu_j F_x(\tau_j). \end{aligned}$$

At $x = 1$, define $\mathbf{z}_j = (\lambda_j, \mu_j, u_j, v_j)$ by its continuous limit, where

$$\lambda_j = \frac{1}{f'(\tau_j)}, \quad \mu_j = 0, \quad u_j = 1 - \frac{1}{f'(\tau_j)}, \quad v_j = \tau_j - \frac{f(\tau_j)}{f'(\tau_j)}.$$

Table 4: Candidate bounds for the exact verification

For any dual tuple $\mathbf{z} = (\lambda, \mu, u, v)$, write

$$U(q, x; \mathbf{z}) = x(\lambda p + \mu R_x + (1 - q)u + qv).$$

The verification considers the following 50 candidate bounds. A dual bound may be used only on the stated range of x , on which [Appendix B.3](#) proves that its tuple is feasible for DUAL.

Candidate	Required bound	Eligible x
(C3)	$(1 - p)x \leq T(q)$	$[0, 1]$
$\mathbf{z}_1$	$U(q, x; \mathbf{z}_1) \leq T(q)$	$(0, 1)$
$\mathbf{z}_2$	$U(q, x; \mathbf{z}_2) \leq T(q)$	$[1/5, 1)$
$\mathbf{z}_3$	$U(q, x; \mathbf{z}_3) \leq T(q)$	$(0, 1)$
$\mathbf{z}_j, 4 \leq j \leq 49$	$U(q, x; \mathbf{z}_j) \leq T(q)$	$(0, 1]$

Primal infeasibility and endpoints. At $x = 0$, the unnormalized loss is zero. At $x = 1$, the first constraint of $\text{PRIMAL}(D^*, q, 1)$ has left-hand side at least $1 - q$. Consequently, this primal is infeasible whenever

$$q < 1 - p = \frac{135516253}{182000000}.$$

It therefore remains to verify that at least one candidate in [Table 4](#) applies for every $(q, x) \in [0, 1] \times (0, 1)$ and for $x = 1$ when $q \in [1 - p, 1]$. On the latter boundary, only the (C3) bound and the tuples $\mathbf{z}_4, \dots, \mathbf{z}_{49}$ are needed; the displayed limits in [Table 3](#) define those tuples at $x = 1$.

B.3 Exact Polynomials for Dual Feasibility

If $f(t) = \sum_d a_d t^d$, the binomial theorem rewrites the normalized functions from [Equation \(4.2\)](#) as the following polynomials:

$$F_x(b_1) = \sum_d a_d \sum_{k=1}^d \binom{d}{k} \frac{x^{k-1} r^{d-k}}{d - k + 1} b_1^k, \quad (\text{B.1})$$

$$G_x(b_0, y) = \sum_d a_d \sum_{k=1}^d \binom{d}{k} \frac{x^{k-1} r^{d-k}}{d - k + 1} \left(b_0^k y^{d-k+1} + 1 - y^{d-k+1} \right), \quad (\text{B.2})$$

$$R_x = \sum_d \frac{a_d}{d + 1} \sum_{j=0}^{d-1} r^j. \quad (\text{B.3})$$

These identities agree with [Equation \(4.2\)](#) for $x > 0$ and give their continuous extensions at $x = 0$. They also show directly that every expression used below is a rational function with rational coefficients.

We use one uniform convention to describe the polynomials checked by the program without expanding their large integer coefficients. Let R be any rational expression below and let $\mathcal{X}$ be its stated domain. After substituting f and [Equations \(B.1\) to \(B.3\)](#), write $R = N/Q$ in lowest terms with coprime integer polynomials and choose the common sign so that Q is positive on the relative interior of $\mathcal{X}$. We write $\text{num}_{\mathcal{X}}(R) = N$; multiplying N and Q by the same positive integer does

not change any sign claim. Whenever an exact factor is displayed beneath $\text{num}_{\mathcal{X}}(R)$, polynomial division by that factor is exact. This convention gives a short but complete implicit specification of every polynomial.

Dual families 1 and 2. Let $\underline{x}_1 = 0$ and $\underline{x}_2 = 1/5$, and define $\mathcal{X}_j = [\underline{x}_j, 1] \times [0, 1]$ in the variables (x, y) . For $j \in \{1, 2\}$, let Δ_j and $\mathbf{z}_j$ be as in [Table 3](#). Put

$$\begin{aligned} I_{j,0}(x, y) &= u_j + \lambda_j W_x(y) + \mu_j G_x(0, y), \\ I_{j,1}(x, y) &= I_{j,0}(x, y) + \mu_j \frac{f(ry)}{r} - 1. \end{aligned}$$

The six polynomial families used for these tuples are

$$\begin{aligned} H_{1,j}^{\text{feas}}(x) &= -\frac{\Delta_j(x)}{x(1-x)^2}, & H_{2,j}^{\text{feas}}(x) &= -\partial_y G_x(0, y_j), \\ H_{3,j}^{\text{feas}}(x) &= W'_x(y_j), \\ H_{4,j}^{\text{feas}}(x) &= \text{num}_{[\underline{x}_j, 1]} \left(\lambda_j a_1 x + \mu_j \frac{f(r)}{r} - 1 \right), \\ H_{5,j}^{\text{feas}}(x, y) &= \frac{\text{num}_{\mathcal{X}_j}(I_{j,0})}{(y - y_j)^2}, & H_{6,j}^{\text{feas}}(x, y) &= \frac{\text{num}_{\mathcal{X}_j}(I_{j,1})}{y}, \end{aligned}$$

where $a_1 = 423239/10^6$. The factors $x(1-x)^2$, $(y - y_j)^2$, and y divide the corresponding numerators exactly. The program proves $H_{1,j}^{\text{feas}} > 0$ on the relative interior of its interval and proves $H_{i,j}^{\text{feas}} \geq 0$ on the stated domain for $i = 2, \dots, 6$.

These signs imply that $\lambda_j, \mu_j \geq 0$. To verify the first pointwise dual inequality, convexity of $G_x(\cdot, y)$ gives, for $r > 0$,

$$G_x(b_0, y) \geq G_x(0, y) + b_0 \frac{f(ry)}{r}. \quad (\text{B.4})$$

The resulting lower bound is affine in b_0 , and the signs of $H_{5,j}^{\text{feas}}$ and $H_{6,j}^{\text{feas}}$ verify its values at $b_0 = 0$ and $b_0 = 1$. For the second pointwise inequality, the difference $v_j + \lambda_j f(xb_1) + \mu_j F_x(b_1) - b_1$ is convex in b_1 , vanishes at $b_1 = 0$, and has nonnegative derivative there by $H_{4,j}^{\text{feas}} \geq 0$. Hence both pointwise inequalities hold throughout the domains on which these tuples are used.

Dual family 3. Recall $\Psi(s) = \int_0^s f(t) dt$ and $L_x = p - \Psi(x) - \Psi(r) - xf(r)$. Define

$$\begin{aligned} I_{3,0}(x, y) &= \mu_3 G_x(0, y) - \lambda_3 (1 - W_x(y)), \\ I_{3,1}(x, y) &= I_{3,0}(x, y) + \mu_3 \frac{f(ry)}{r} - 1. \end{aligned}$$

The three polynomials are

$$H_{1,3}^{\text{feas}}(x) = \frac{L_x}{x(1-x)}, \quad H_{2,3}^{\text{feas}}(x, y) = \frac{\text{num}_{[0,1]^2}(I_{3,0})}{1-y}, \quad H_{3,3}^{\text{feas}}(x, y) = \frac{\text{num}_{[0,1]^2}(I_{3,1})}{y(1-y)}.$$

All displayed divisions are exact. The program proves $H_{1,3}^{\text{feas}} \geq 0$ on $[0, 1]$ and $H_{2,3}^{\text{feas}}, H_{3,3}^{\text{feas}} \geq 0$ on $[0, 1]^2$. Thus $\lambda_3 \geq 0$ and $\mu_3 > 0$, and [Equation \(B.4\)](#) again reduces the first pointwise inequality

to the two endpoint checks. The second pointwise difference is convex, vanishes at $b_1 = 0$, and has nonnegative derivative there because $\mu_3 f(r)/r = 1$ and $\lambda_3 \geq 0$. Hence this tuple is feasible for every $0 < x < 1$.

Dual families 4–49. For $j \in \{4, 5, \dots, 49\}$, define

$$Z_j(x) = L_x f'(x\tau_j) + (1 - f(x)) (f(r + x\tau_j) - f(x\tau_j)).$$

The two polynomial families are

$$H_{1,j}^{\text{feas}}(x) = \frac{Z_j(x)}{1-x}, \quad H_{2,j}^{\text{feas}}(x) = \frac{Z_j(x) - (1 - f(x))f(1-x)}{x(1-x)}.$$

The divisions are exact, and the program proves $H_{1,j}^{\text{feas}} > 0$ and $H_{2,j}^{\text{feas}} \geq 0$ on $[0, 1]$ for every $j \in \{4, 5, \dots, 49\}$. These signs imply $S_j(x) \geq 1$, so $0 < \theta_j(x) \leq 1$. The first pointwise inequality for $\mathbf{z}_j$ is therefore a convex combination of the inequality $1 \geq b_0$ and the first inequality for $\mathbf{z}_3$. The second pointwise difference is convex in b_1 , equals zero at $b_1 = \tau_j$, and has derivative $\theta_j S_j - 1 = 0$ there. It is consequently nonnegative on $[0, 1]$. At $x = 1$, the limiting tuple in [Table 3](#) is feasible because its first inequality is $1 \geq b_0$ and its second is the tangent-line inequality for the convex polynomial f .

B.4 Polynomial Verification and Proof of [Lemma 5](#)

We now describe the exact outer polynomials and the method used to prove that at least one of the candidate bounds in [Table 4](#) applies at every relevant (q, x) . Recall that $U(q, x; \mathbf{z})$ is the unnormalized objective of a dual tuple $\mathbf{z}$. Define the following polynomial gaps:

$$\begin{aligned} H_0^{\text{out}}(q, x) &= T(q) - (1 - p)x, \\ H_j^{\text{out}}(q, x) &= -\frac{\Delta_j(x)}{x} (T(q) - U(q, x; \mathbf{z}_j)) && \text{for } j \in \{1, 2\}, \\ H_3^{\text{out}}(q, x) &= f(1 - x)(1 - f(x)) (T(q) - U(q, x; \mathbf{z}_3)), \\ H_j^{\text{out}}(q, x) &= \frac{Z_j(x)}{1 - x} (T(q) - U(q, x; \mathbf{z}_j)) && \text{for } j \in \{4, 5, \dots, 49\}. \end{aligned}$$

After substitution, every right-hand side is a polynomial in (q, x) with rational coefficients; the factors shown are precisely positive common denominators on the relative interiors where the corresponding tuples are used, and the divisions by x and $1 - x$ are exact polynomial divisions. Thus nonnegativity of an outer polynomial proves that its associated upper bound is at most $T(q)$.

We use the algorithm of Garloff [[Gar93](#)] to prove polynomial nonnegativity on rational boxes. Map a box $[q_0, q_1] \times [x_0, x_1]$ affinely to $(s, t) \in [0, 1]^2$. If H has bidegree at most (m, n) , write it in the tensor-product Bernstein basis as

$$H(s, t) = \sum_{i=0}^m \sum_{j=0}^n c_{ij} \binom{m}{i} s^i (1-s)^{m-i} \binom{n}{j} t^j (1-t)^{n-j}.$$

Every basis function is nonnegative, and their sum is one. Therefore $H(s, t)$ is a convex combination of the coefficients c_{ij} and

$$\min_{i,j} c_{ij} \leq H(s, t) \leq \max_{i,j} c_{ij}.$$

In particular, nonnegative coefficients prove $H \geq 0$ throughout the box. The program computes these coefficients using exact rational arithmetic.

For a finite disjunction of polynomial inequalities, we apply the same Garloff test to every eligible polynomial on each current rational box. If one candidate has only nonnegative coefficients, that inequality holds throughout the box, which becomes an accepted subdivision. If no candidate succeeds, the program divides the box at rational midpoints and repeats the test on the resulting boxes. This procedure is not guaranteed to terminate: the Bernstein-coefficient test is sufficient but not necessary, and a true pointwise disjunction need not initially have one member that holds throughout a large box. Whenever the procedure does terminate, however, the accepted subdivisions form a finite cover and give a sound proof of the disjunction. It terminates for the polynomials considered here, and no floating-point arithmetic enters any check.

Specifically, for every $(q, x) \in [0, 1] \times (0, 1)$, the program verifies the disjunction

$$H_0^{\text{out}}(q, x) \geq 0 \quad \vee \quad H_1^{\text{out}}(q, x) \geq 0 \quad \vee \quad \left(x \geq \frac{1}{5} \wedge H_2^{\text{out}}(q, x) \geq 0 \right) \quad \vee \quad \bigvee_{j=3}^{49} H_j^{\text{out}}(q, x) \geq 0.$$

For $x = 1$ and $q \in [1 - p, 1]$, it verifies

$$H_0^{\text{out}}(q, 1) \geq 0 \quad \vee \quad \bigvee_{j=4}^{49} H_j^{\text{out}}(q, 1) \geq 0.$$

The condition $x \geq 1/5$ in the first disjunction is a precondition for using $\mathbf{z}_2$, not another candidate bound: a subdivision may be accepted using H_2^{out} only if it lies entirely in $x \geq 1/5$. At $q = 0$, the target is $T(0) = 0$, and the program never divides by q .

Proof of Lemma 5. Fix $q \in [0, 1]$ and an $x \in [0, 1]$ for which $\text{PRIMAL}(D^*, q, x)$ is feasible. If $x = 0$, the unnormalized loss is zero and hence at most $T(q)$, so suppose $x > 0$. If $x = 1$, primal feasibility implies $q \geq 1 - p$ by the argument following Table 4. Thus one of the two verified disjunctions applies. If $H_0^{\text{out}}(q, x) \geq 0$, then $(1 - p)x \leq T(q)$. Otherwise, one of the eligible outer polynomials is nonnegative, and its associated dual tuple $\mathbf{z}_j$ is feasible by Appendix B.3 and satisfies $U(q, x; \mathbf{z}_j) \leq T(q)$. Hence the sufficient condition in Lemma 4 holds, and

$$g_{D^*}(q) \leq T(q) = q \left(\frac{33205}{42059} - \frac{7941}{42059} q \right)$$

for every $q \in [0, 1]$. The loss bound follows from Lemma 2. □

B.5 Discretized Programs Used in the Search

The exact programs in Lemmas 3 and 4 are not finite LPs. For the numerical search, fix arbitrary nonempty finite sets $E \subseteq [0, 1]^2$ and $E_1 \subseteq [0, 1]$, and restrict η and ζ to these sets. The resulting

discretized primal is

$$\begin{aligned}
& \max_{\eta, \xi \geq 0} \quad (1-q) \sum_{(b_0, y) \in E} b_0 \eta(b_0, y) + q \sum_{b_1 \in E_1} b_1 \xi(b_1) \\
& \text{subject to} \quad (1-q) \sum_{(b_0, y) \in E} W_x(y) \eta(b_0, y) + q \sum_{b_1 \in E_1} w_x(b_1) \xi(b_1) \leq p, \\
& \quad (1-q) \sum_{(b_0, y) \in E} G_x(b_0, y) \eta(b_0, y) + q \sum_{b_1 \in E_1} F_x(b_1) \xi(b_1) \leq R_x, \\
& \quad \sum_{(b_0, y) \in E} \eta(b_0, y) = 1, \quad \sum_{b_1 \in E_1} \xi(b_1) = 1.
\end{aligned}$$

Here $\eta(b_0, y)$ and $\xi(b_1)$ are probability masses, not densities. Its ordinary finite LP dual is

$$\begin{aligned}
& \min_{\substack{\lambda, \mu \geq 0 \\ u, v \in \mathbb{R}}} \quad \lambda p + \mu R_x + (1-q)u + qv \\
& \text{subject to} \quad u + \lambda W_x(y) + \mu G_x(b_0, y) \geq b_0 \quad \text{for every } (b_0, y) \in E, \\
& \quad v + \lambda w_x(b_1) + \mu F_x(b_1) \geq b_1 \quad \text{for every } b_1 \in E_1.
\end{aligned}$$

Thus the dual of the discretized primal is precisely what results from imposing the exact dual inequalities only on the finite sets. Finite LP duality equates the two discretized values. Restricting the measures can only decrease the primal optimum, while omitting pointwise dual constraints can only decrease the dual optimum; therefore this common value generally underestimates the exact value. A tuple returned by the discretized dual does not upper bound the original program unless its inequalities are subsequently proved for every $(b_0, y) \in [0, 1]^2$ and every $b_1 \in [0, 1]$.

B.6 Numerical Search and Rationalization

None of the calculations in this subsection is used in the proof of [Lemma 5](#). They explain only how we found the exact data later verified in [Appendix B.4](#).

Fixed (D, q, x) . Start with finite sets E and E_1 , solve the discretized dual, and let (λ, μ, u, v) be the returned tuple. We then numerically search for the largest violation of each of its two types of constraint:

$$\begin{aligned}
\delta_0 &= \max_{(b_0, y) \in [0, 1]^2} \{b_0 - u - \lambda W_x(y) - \mu G_x(b_0, y)\}, \\
\delta_1 &= \max_{b_1 \in [0, 1]} \{b_1 - v - \lambda w_x(b_1) - \mu F_x(b_1)\}.
\end{aligned}$$

When the numerical search finds a positive value, we add a corresponding point to E or E_1 and solve the finite dual again. This is dual constraint generation; from the discretized primal viewpoint, the same step adds a possible point at which a measure may place mass. The two maximizations are nonlinear optimization problems, and the implementation uses SciPy's `differential_evolution`. Failure of this heuristic to find a violation is not a proof of dual feasibility. Numerical LP tolerances can also leave a returned tuple slightly infeasible. By contrast, a genuinely feasible but suboptimal solution of the minimization dual would be completely safe: it would merely give a weaker upper bound.

Finite sets of q and x . The preceding computation was repeated on finite sets of q - and x -values. Let $\hat{L}_D(q, x)$ denote the numerical objective value returned for a searched pair. For each searched q , we formed

$$\hat{g}_D(q) = \max_{x \text{ in the search set}} \min \left\{ x \hat{L}_D(q, x), (1-p)x \right\}.$$

This is an estimate, not an upper bound on $g_D(q)$. A second finite LP chose weights $\alpha, \beta, \gamma \geq 0$ on Integrated Veto, maximal lottery, and the random-size stable lottery by solving

$$\begin{aligned} & \min_{\alpha, \beta, \gamma, C_{\text{cut}}} C_{\text{cut}} \\ \text{subject to} \quad & \alpha q_i^2 + \beta \min \{q_i, 1/2\} + \gamma \hat{g}_D(q_i) \leq C_{\text{cut}} q_i \quad \text{for every searched } q_i, \\ & \alpha + \beta + \gamma = 1, \quad \alpha, \beta, \gamma, C_{\text{cut}} \geq 0. \end{aligned}$$

The implementation also included the analytically derived endpoint constraints $\beta + \gamma \Lambda_D \leq C_{\text{cut}}$ and $\alpha + \beta/2 + \gamma \rho_D \leq C_{\text{cut}}$, where Λ_D is the numerically evaluated $q \downarrow 0$ limit and ρ_D solves $f(\rho_D) = p$. The best candidate found had $\beta = 0$. In that case, the remaining constraints have the form

$$\hat{g}_D(q_i) \leq q_i \left(\frac{C_{\text{cut}}}{\gamma} - \frac{\alpha}{\gamma} q_i \right),$$

which suggests $a = C_{\text{cut}}/\gamma$ and $b = \alpha/\gamma$.

The draw distribution and target. For a selected finite support S of positive integer degrees, we wrote $f(t) = \sum_{d \in S} a_d t^d$, parameterized the masses $a_d > 0$ by unconstrained log ratios followed by normalization, and optimized those real parameters using `differential_evolution`. Thus the degrees were selected discretely, but the probability masses were optimized continuously; there was no grid over the probability simplex. The search tested many supports containing degree 1, including supports of three and four degrees, perturbed promising supports by adding or moving mass to other degrees, and reoptimized the outer LP. It also constructed sparse feasible measures for the discretized primal. Such measures give lower bounds on the true loss and can expose an erroneously low value caused by a violated dual constraint; they never prove an upper bound. The search was not exhaustive and does not prove that D^* is optimal among draw distributions.

From numerical dual solutions to explicit tuples. At each searched pair (q, x) , the finite dual returned numerical values of (λ, μ, u, v) and indicated which pointwise inequalities were nearly tight. In the small- q regime, the observed pattern was that the first pointwise inequality was tight at $(b_0, y) = (0, y_\circ)$ and $(1, 0)$ and stationary in y at $(0, y_\circ)$, while $v = 0$. For fixed x and $y_\circ$, these conditions are

$$\begin{aligned} u + \lambda W_x(y_\circ) + \mu G_x(0, y_\circ) &= 0, \quad \lambda W'_x(y_\circ) + \mu \partial_y G_x(0, y_\circ) = 0, \\ u + \lambda f(x) + \mu A_x &= 1. \end{aligned}$$

Solving these three linear equations with $y_\circ = 80529/100000$ and $y_\circ = 1$ gives $\mathbf{z}_1$ and $\mathbf{z}_2$, respectively. The first rational value was suggested by the numerical active point, which was approximately 0.8052905.

Near the middle numerical peak, the active pattern instead suggested that the first pointwise inequality should be tight at $(0, 1)$, tangent in b_0 there, and tight at $(1, 0)$, again with $v = 0$. Since $W_x(1) = 1$, $G_x(0, 1) = 0$, and $\partial_{b_0} G_x(0, 1) = f(r)/r$, these conditions are

$$u + \lambda = 0, \quad \mu \frac{f(r)}{r} = 1, \quad u + \lambda f(x) + \mu A_x = 1, \quad v = 0.$$

Their solution is $\mathbf{z}_3$; in particular, $\mu_3 = r/f(r)$ and $\lambda_3 = (\mu_3 A_x - 1)/(1 - f(x))$, which is algebraically identical to the formula using L_x in [Table 3](#).

For larger q , the numerical dual solutions placed the active complementary-state constraint at an interior point $b_1 = t$. To retain feasibility of the first pointwise inequality, we mixed the first three coordinates of $\mathbf{z}_3$ with the trivially feasible triple $(0, 0, 1)$:

$$(\lambda, \mu, u) = \theta(\lambda_3, \mu_3, u_3) + (1 - \theta)(0, 0, 1).$$

We then imposed equality and zero derivative in the second pointwise inequality at $b_1 = t$. Equality determines $v = t - \lambda f(xt) - \mu F_x(t)$, while the derivative condition determines the mixing weight. Numerical solutions suggested a moving interior value of t , so we retained the rational values $t = 15/100, 16/100, \dots, 60/100$; these give $\mathbf{z}_4, \mathbf{z}_5, \dots, \mathbf{z}_{49}$, respectively. Only the exact coverage proof in [Appendix B.4](#), not this numerical pattern, establishes that these tuples suffice.

Rationalization. The best numerical candidate used

$$f(t) \approx 0.42323858t + 0.47110490t^{12} + 0.10565652t^{13}$$

and an Integrated Veto weight of approximately 0.1588198. We replaced the probabilities, mixture weight, active value y_1 , and tangency points by the rational values in [Section 5](#) and [Table 3](#). The outer LP relation $a = C_{\text{cut}}/\gamma$ and $b = \alpha/\gamma$ then gives the displayed rational target. These replacements change both the rule and the proposed dual curves: numerical proximity transfers no guarantee. The proof in [Appendix B.4](#) therefore verifies the rational data independently over all states and all $(q, x) \in [0, 1]^2$.

B.7 Reproducing the Computer-Assisted Checks

The GitHub repository cited above contains a README, the exact verifier, the checked-in rational candidate, and the numerical search code. The verifier uses only the Python standard library and reconstructs every polynomial in [Appendices B.3](#) and [B.4](#) using exact rational arithmetic. From the repository root, either run

```
python3 verify_distortion.py
python3 verify_distortion.py candidate.json
```

The first command uses the built-in candidate, while the second supplies the same checked-in data explicitly. Either command produces exactly

All polynomial nonnegativity checks passed.

Any failed check stops the program. Successful completion verifies dual feasibility on the stated domains and that an eligible bound applies at every relevant $(q, x) \in [0, 1]^2$, establishing the computer-assisted claim in [Lemma 5](#). To rerun the numerical discovery process, install `requirements.txt` and run `python3 search_distortion.py`; its floating-point output becomes part of the proof only after exact verification of the rationalized candidate.