The Metric Distortion of Randomized Social Choice Functions: C1 Maximal Lottery Rules and Simulations

Fabian Frank
Technical University of Munich
Munich, Germany
fabian_w.frank@tum.de

Patrick Lederer
University of New South Wales
Sydney, Australia
p.lederer@unsw.edu.au

ABSTRACT

The metric distortion of a randomized social choice function (RSCF) quantifies its worst-case approximation ratio to the optimal social cost when the voters' costs for alternatives are given by distances in a metric space. This notion has recently attracted significant attention as numerous RSCFs that aim to minimize the metric distortion have been suggested. Since such tailored voting rules have, however, little normative appeal other than their low metric distortion, we will study the metric distortion of well-established RSCFs. Specifically, we first show that C1 maximal lottery rules, a well-known class of RSCFs, have a metric distortion of 4, which is optimal within the class of majoritarian RSCFs. Secondly, we conduct extensive computer experiments on the metric distortion of RSCFs to obtain insights into their average-case performance. These computer experiments are based on a new linear program for computing the metric distortion of a lottery and reveal that the average-case metric distortion of some classical RSCFs is often only slightly worse than that of RSCFs tailored to minimize the metric distortion. Finally, we also analytically study the expected metric distortion of RSCFs for the impartial culture distribution. Specifically, we show that, under this distribution, every reasonable RSCF has an expected metric distortion close to 2 when the number of voters is large.

KEYWORDS

Randomized Social Choice Functions; Metric Distortion; Simulations; Average-case Analysis; Maximal Lottery Rules

ACM Reference Format:
Fabian Frank and Patrick Lederer. 2025. The Metric Distortion of Randomized Social Choice Functions: C1 Maximal Lottery Rules and Simulations. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 9 pages.

1 INTRODUCTION

An important challenge in multi-agent systems is collective decision-making: given the possibly conflicting preferences of a group of agents over some alternatives, a joint decision has to be made. To address this problem, researchers in the field of social choice theory try to identify desirable mechanisms to aggregate the agents' preferences. In more detail, social choice theory is mainly concerned with *social choice functions (SCFs)* and *randomized social choice functions (RSCFs)*, which formalize deterministic and randomized voting rules: an SCF maps the voters' preferences (expressed as linear rankings of the alternatives) to a single winner, and an RSCF returns a probability distribution over the alternatives from which the final winner will eventually be chosen [6, 13].

In an attempt to quantitatively measure the quality of SCFs and RSCFs, Procaccia and Rosenschein [47] introduced the distortion of voting rules. The idea of this notion is that voters have latent cardinal utilities over the alternatives and that voting rules should try to select alternatives with high social welfare. However, (R)SCFs do not have access to the voters' utilities, and the *distortion* of a voting rule thus quantifies the worst-case ratio between the (expected) social welfare of the selected alternative and that of the optimal alternative. A prominent variant of this problem has been suggested by Anshelevich et al. [3]: in the metric distortion setting, voters and alternatives are located in a metric space and the distance between a voter and an alternative specifies the cost incurred to a voter when an alternative is elected. Voting rules should then try to select an alternative with low social cost but, since voters only report ordinal preferences, they can only approximate the optimal social cost. The *metric distortion* of an SCF (resp. RSCF) is hence the worst-case ratio between the (expected) social cost of the selected alternative and of the optimal alternative, where the worst-case is taken over all preference profiles and all metric spaces that are consistent with the given profile.

The metric distortion of SCFs and RSCFs has recently attained significant attention [see, e.g., 4]. In particular, after Anshelevich et al. [3] and Anshelevich and Postl [5] have shown that no SCF (resp. RSCF) has a metric distortion of less than 3 (resp. 2), numerous authors tried to find voting rules with minimal metric distortion [e.g., 2, 16, 36–38]. However, many of the suggested voting rules are specifically tailored to minimize the metric distortion and have otherwise little normative appeal. For example, the recently proposed Plurality-Veto rule [37] is not even anonymous and its latest variant called Simultaneous-Veto [38] fails Pareto-optimality. We thus find it noteworthy that some well-established RSCFs also have a low metric distortion. For instance, the uniform random dictatorship and C2 maximal lottery (C2ML) rules, two of the most prominent RSCFs in the literature, both have a metric distortion of 3 [5, 16, 22]. Since such established RSCFs satisfy numerous desirable properties, we will study their metric distortion in more detail, even though voting rules with lower metric distortion are known.

Our Contribution. The goal of this paper is to enhance the understanding of the metric distortion of established RSCFs. We will contribute to this end in three ways. Firstly, we investigate the metric distortion of C1 maximal lottery (C1ML) rules, a class of RSCFs that is well-known for being robust to small changes in the voters'

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

preferences [12, 34, 39]. C1ML rules intuitively choose randomized Condorcet winners: these rules return a lottery p such that, for every lottery q, it is at least as likely that a majority of the voters prefers an outcome drawn from p to an outcome drawn from q than vice versa. As our first result, we show that every C1ML rule has a metric distortion of at most 4 and give a lower bound on the metric distortion of all majoritarian RSCFs (which only depend on the majority relation) that converges to 4 as the number of alternatives increases. Since C1ML rules are majoritarian, this proves that they minimize the metric distortion within this class of RSCFs.

Secondly, we conduct extensive computer experiments on the metric distortion of five RSCFs: the uniform random dictatorship, C1 and C2 maximal lottery rules, a randomized variant of the Plurality-Veto rule [37], and the CRWW rules suggested by Charikar et al. [16], which have the best currently known metric distortion. In more detail, we sample preference profiles from numerous distributions, compute the lotteries chosen by our RSCFs, and then compute the worst-case metric distortion for the given lotteries and profiles. Moreover, we conduct an analogous experiment also with real-world data taken from PrefLib [42]. Our simulations show that the average metric distortion of all RSCFs is rather similar and significantly better than their worst-case guarantees. In particular, for many "structured" distributions C1ML and C2ML rules are only slightly worse than CRWW rules, which typically have the best metric distortion in our experiments. In light of their normative appeal, this gives a strong argument for using a C1ML or C2ML rule instead of an RSCF designed to minimize the metric distortion.

Our computer experiments rely on a new linear program for computing the metric distortion of a lottery for a given profile, which we believe to be of independent interest. Specifically, our LP has only $O(nm^2)$ constraints, where n is the number of voters and m the number of alternatives, and thus allows us to efficiently compute the metric distortion of a lottery even for large profiles.

Finally, we complement our simulations with an analytical study of the expected metric distortion of RSCFs when preference profiles are sampled from the impartial culture distribution. For this setting, we show that the expected metric distortion of every reasonable RSCF converges to a value between 2 and $2 + \frac{1}{m-1}$ (where m is the number of alternatives) when the number of voters goes to infinity. This result aligns with our simulations for the impartial culture model and shows that, at least under the simplistic impartial culture distribution, the choice of the voting rule has surprisingly little effect on the expected metric distortion.

Related Work. We will next review the most relevant related works and refer to the survey by Anshelevich et al. [4] for more details. An overview of the upper and lower bounds for the metric distortion of various classes of voting rules is given in Table 1. The study of metric distortion was initiated by Anshelevich et al. [3] who have, e.g., shown that no deterministic SCF has a metric distortion of less than 3. Inspired by this work, numerous researchers tried to find rules with a metric distortion of 3 [1, 2, 31, 48], but it was only in a recent line of work that such SCFs have been designed [30, 36–38, 44]. In particular, these works culminated in the Plurality-Veto rule, a simple SCF with a metric distortion of 3 [37]. Interestingly, Kizilkaya and Kempe [38] recently aimed to design a normatively more appealing SCF with optimal metric distortion.

Table 1: Overview of the best known upper and lower bounds on the metric distortion in various classes of voting rules. Each row together with the labels "RSCF" and "SCF" determines a class of voting rules. The columns labeled "LB" and "UB" show the best known lower and upper bounds for the metric distortion of rules within the given class when there is an unbounded number of alternatives. The bold numbers are proven in this paper.

	RSCF		SCF	
	LB	UB	LB	UB
All	2.112	2.753	3	3
Tops-only	3	3	∞	∞
Pairwise	3	3	3	$2+\sqrt{5}$
Majoritarian	**4**	**4**	5	5

As an alternative approach to minimize the metric distortion, researchers also studied RSCFs. In particular, Anshelevich and Postl [5] have shown that no RSCF has a metric distortion of less than 2 and that the uniform random dictatorship has a metric distortion of 3. Moreover, Gross et al. [33] have proven that all tops-only RSCFs (i.e., RSCFs that can only access the voters' favorite alternatives) have a metric distortion of at least $3 - \frac{2}{m}$ when there are m alternatives. Similarly, Charikar et al. [16] have shown that C2 maximal lottery rules have a metric distortion of 3 and it is known that all pairwise RSCFs (i.e., RSCFs that only depend on the numbers of voters that prefer x to y for all alternatives x, y) have a metric distortion of at least $3 - \frac{2}{m}$ [31]. Thus, when the number of alternatives is unbounded, the uniform random dictatorship minimizes the metric distortion within the class of tops-only RSCFs and C2 maximal lottery rules within the class of pairwise RSCFs. Moreover, several RSCFs have been designed with the goal to minimize the metric distortion [21, 30, 33], but none of them guarantees a metric distortion of less than 3. It was hence only recently that both the upper and lower bound of the metric distortion of RSCFs has been improved: Charikar and Ramakrishnan [15] have shown that every RSCF has a metric distortion of at least 2.112 and Charikar et al. [16] designed the CRWW rules with a metric distortion of at most 2.753.

Finally, our work is related to several papers [14, 17, 18, 32] that analyze the expected distortion of voting rules when the voters' utilities are drawn from a distribution. By contrast, we study the expected metric distortion of voting rules for the worst-case metrics of randomly drawn profiles, i.e., we consider realistic profiles without imposing any additional structure on the voters' utilities. Furthermore, we note that Ebadian et al. [20] suggested an improved linear program for computing the non-metric distortion of voting rules, which can be seen as a mathematically unrelated analog of our new linear program for computing the metric distortion of RSCFs.

2 MODEL

Let $V_n = \{v_1, \dots, v_n\}$ denote a finite set of $n \geq 1$ voters and $X_m = \{x_1, \dots, x_m\}$ a finite set of $m \geq 1$ alternatives. We suppose that every voter $v \in V_n$ reports a *preference relation* $\succ_v$, which is formally a complete, transitive, and anti-symmetric binary relation over X_m. The set of all preference relations over X_m is denoted by $\mathcal{R}(X_m)$. A *preference profile* R is the collection of the preference relations of all voters in V_n. The set of all preference profiles over an electorate V_n and a set of alternatives X_m is given by $\mathcal{R}(X_m)^{V_n}$. In this paper, we will allow for both varying sets of voters and alternatives. The set of all preference profiles is hence given by $\mathcal{R}^* = \bigcup_{n,m\in\mathbb{N}} R(X_m)^{V_n}$. Moreover, $\mathcal{R}^*_m$ is the set of all profiles on m alternatives, i.e., $\mathcal{R}^*_m = \bigcup_{n\in\mathbb{N}} R(X_m)^{V_n}$. Given a profile R, we will denote by V_R and X_R the sets of voters and alternatives that are present in the profile R, and by n_R and m_R the sizes of these sets.

Next, we introduce additional notation for preference profiles. In particular, we define $t_R(x) = |\{v \in N_R \colon \forall y \in X_R \setminus \{x\} \colon x \succ_v y\}|$ as the number of voters that top-rank alternative x in the profile R. Furthermore, we let the *support* $n_{xy}(R) = |\{v \in V_R \colon x \succ_v y\}|$ for x against y denote the number of voters who prefer x to y in R. Finally, the *majority relation* $\succsim_R$ of a profile R is defined by $x \succsim_R y$ if and only if $n_{xy}(R) \geq n_{yx}(R)$. That is, $x \succsim_R y$ if at least as many voters prefer x to y than vice versa. Following the literature, $\succ_R$ denotes the strict part of $\succsim_R$ (i.e., $x \succ_R y$ iff $x \succsim_R y$ and not $y \succsim_R x$) and $\sim_R$ the indifference part (i.e., $x \sim_R y$ iff $x \succsim_R y$ and $y \succsim_R x$).

2.1 Randomized Social Choice Functions

The study objects of this paper are randomized social choice functions which are voting rules that may use chance to determine the winner of the election. To formalize this, we define *lotteries* as probability distributions over the set of alternatives X_R: a lottery is a function $p : X_R \to [0, 1]$ such that $\sum_{x\in X_R} p(x) = 1$. We furthermore denote by $\Delta(X_R)$ the set of all lotteries over X_R. A *randomized social choice function (RSCF)* f is then a function that maps every preference profile $R \in \mathcal{R}^*$ to a lottery $p \in \Delta(X_R)$. We denote by $f(R, x)$ the probability that f assigns to alternative x in the profile R and next introduce five (classes of) RSCFs:

Uniform random dictatorship. The uniform random dictatorship f_{RD} picks a voter $v \in V_R$ uniformly at random and implements his favorite alternative as the winner of the election. More formally, the probability that an alternative x is selected in a profile R by the uniform random dictatorship is $f_{RD}(R, x) = \frac{t_R(x)}{n_R}$.

Randomized Plurality-Veto. Kizilkaya and Kempe [38] suggested the Plurality-Veto rule as a deterministic SCF with the optimal metric distortion of 3. For this rule, we first fix a sequence of the voters $(v_1, \dots, v_n)$ and assign a score $s(x)$ to each alternative that is initially equal to $t_R(x)$. Then, we iterate through the voters according to the given sequence, ask each voter for his worst alternative with positive score, and reduce the score of this alternative by 1. Finally, the winner of this rule is the last alternative with positive score. Since the winner of Plurality-Veto rule depends on the order over the voters, we denote by $PV(R)$ the set of alternatives that can be chosen for some order. Furthermore, we define the *randomized Plurality-Veto rule* f_{RPV} as the RSCF that picks an alternative from $PV(R)$ uniformly at random. The set $PV(R)$ and hence f_{RPV} can be efficiently computed by solving m matching problems [37, 38].

C2ML rules. C2 maximal lottery (C2ML) rules, which have been suggested by Fishburn [23] and recently promoted by, e.g., Brandl et al. [11], compute a randomized Condorcet winner: these rules select a lottery p such that, for all lotteries q, the expected number of voters that prefer the outcome chosen from p to the outcome chosen from q is at least as large as the expected number of voters that prefer the outcome chosen from q to the outcome chosen from p. To formalize this, we extend the support $n_{xy}(R)$ to lotteries p, q by defining $n_{pq}(R) = \sum_{x,y\in A} p(x)q(y)n_{xy}(R)$. Then, the set of C2 maximal lotteries is given by $C2ML(R) = \{p \in \Delta(X_R) \colon \forall q \in \Delta(X_R) \colon n_{pq}(R) \geq n_{qp}(R)\}$. The set of C2 maximal lotteries is always non-empty by the minimax theorem and almost always a singleton [40, 41]. Finally, an RSCF is a C2ML rule if $f(R) \in C2ML(R)$ for every profile $R \in \mathcal{R}^*$.

C1ML rules. C1 maximal lottery (C1ML) rules, which go back to Fishburn [23], also choose a randomized Condorcet winner but in a different sense: C1ML rules select a lottery p such that, for all lotteries q, it is at least as likely that a majority of the voters prefers the outcome chosen from p to the outcome chosen from q than vice versa. To formalize this, we extend the majority relation to lotteries p, q by defining that $p \succsim_R q$ if and only if $\sum_{x,y\in A \colon x\succ_R y} p(x)q(y) \geq \sum_{x,y\in A \colon x\succ_R y} p(y)q(x)$. The set of C1 maximal lotteries is then $C1ML(R) = \{p \in \Delta(X_R) \colon \forall q \in \Delta(X_R) \colon p \succsim_R q\}$. Just as for C2 maximal lotteries, this set is always non-empty and almost always a singleton. In particular, if the number of voters is odd, there are unique C1 and C2 maximal lotteries. An RSCF is a C1ML rule if $f(R) \in C1ML(R)$ for all profiles $R \in \mathcal{R}^*$.

CRWW rules. Finally, we introduce the RSCFs suggested by Charikar et al. [16], which we refer to as CRWW rules. As a subroutine, these rules rely on another RSCF called $f_{\beta-radius}$. To define this RSCF, we say x β-covers y in a profile R for some $\beta \in [0, 1]$ if $n_{xy}(R) \geq \beta n_R$ and $n_{zx}(R) \geq \beta n_R$ implies $n_{zy}(R) \geq \beta n_R$ for all $z \in X_R$. Moreover, we define $U_\beta(R)$ as the set of alternatives that are not β-covered in R and $R|_{U_\beta(R)}$ as the profile that arises from R by removing all alternatives not in $U_\beta(R)$. Then, $f_{\beta-radius}$ computes the uniform random dictatorship on $R|_{U_\beta(R)}$, i.e., $f_{\beta-radius}(R) = f_{RD}(R|_{U_\beta(R)})$. Based on this subroutine, constants $B = 0.876353$, $p = \frac{1}{1+\int_{0.5}^{B} \frac{1}{1-x^2} dx} \approx 0.552327$, and the distribution $\rho(\beta) = \frac{p}{(1-p)(1-\beta^2)}$ on the interval $(\frac{1}{2}, B)$, CRWW rules are defined as follows: with probability p, we execute a C2ML rule and with probability $1 - p$, we sample a value $\beta \in (0.5, B)$ from the distribution $\rho(\beta)$ and return $f_{\beta-radius}(R)$. Hence, an RSCF f is a CRWW rule if there is a C2ML rule f' such that

$$f(R) = pf'(R) + (1 - p) \int_{0.5}^{B} \rho(\beta) f_{\beta-radius}(R) d\beta.$$

The uniform random dictatorship f_{RD}, C2ML rules, and C1ML rules are well-established in the literature. For example, f_{RD} is known to be strategyproof [29], whereas both C2ML rules and C1ML rules satisfy, e.g., Condorcet-consistency and composition-consistency [11]. By contrast, the randomized Plurality-Veto rule and the CRWW rules are designed to minimize the metric distortion and only known to satisfy basic further axioms. Moreover, we note that the uniform random dictatorship f_{RD}, C2ML rules, and C1ML rules belong to important classes of RSCFs: f_{RD} is a *tops-only* RSCF as it only accesses the voters' favorite alternatives, C2ML rules are *pairwise* as they only access the supports $n_{xy}(R)$ for all $x, y \in X_R$, and C1ML rules are *majoritarian* as they only depend on the majority relation $\succsim_R$. In more detail, an RSCF f is majoritarian if $f(R) = f(R')$ for all profiles $R, R' \in \mathcal{R}^*$ with $\succsim_R = \succsim_{R'}$.

2.2 Metric Distortion

In order to assess the quality of RSCFs, we analyze their metric distortion. The idea of this approach is that voters and alternatives are embedded in a metric space and that the distance between a voter v and an alternative x specifies the cost that v experiences when x is selected. Following the utilitarian approach, the optimal alternative is then the one that minimizes the total distance to all voters. However, since voters only report ordinal preferences over the alternatives instead of their cardinal costs, we cannot determine the best alternative. The goal of metric distortion is hence to select a lottery that approximates the optimal social cost well for every metric space that is consistent with the voters' preferences.

To formalize this, we call a function $d : (V_R \cup X_R)^2 \to \mathbb{R}_{\geq 0}$ a *metric* if it satisfies for all $x, y, z \in V_R \cup X_R$ that *i)* $d(x, x) = 0$, *ii)* $d(x, y) = d(y, x)$, and *iii)* $d(x, z) \leq d(x, y) + d(y, z)$. We note that some definitions of metrics also require that $d(x, y) > 0$ if $x \neq y$, but the literature on metric distortion typically omits this condition since it does not affect the results. The distance $d(v, x)$ states the cost incurred to voter v when alternative x is selected. The *social cost* of an alternative x is thus $sc(x, d) = \sum_{v \in V_R} d(v, x)$ and the social cost of lottery p is $sc(p, d) = \sum_{x \in X_R} p(x) sc(x, d)$. Finally, a metric d is *consistent* with a profile R if $x \succ_v y$ implies $d(v, x) \leq d(v, y)$ for all voters $v \in V_R$ and alternatives $x, y \in X_R$. We denote by $D(R)$ the set of metrics that are consistent with R.

Given a profile R, the goal of metric distortion is to find a lottery whose social cost is close to the optimal social cost for all metrics that are consistent with R. We thus define the metric distortion of a lottery p in a profile R as $dist(p, R) = \sup_{d \in D(R)} \frac{sc(p,d)}{\min_{x \in X_R} sc(x,d)}$. Note that $\min_{x \in X_R} sc(x, d)$ might be 0; we hence define $\frac{0}{0} = 1$ and $\frac{z}{0} = \infty$ for $z > 0$. For the ease of presentation, we will use in our results that $\infty > x$ for all $x \in \mathbb{R}$ and $y + z \cdot \infty = \infty$ for all $y \in \mathbb{R}$, $z \in \mathbb{R}_{>0}$. Next, the *metric distortion* $dist(f)$ of an RSCF f is its worst-case metric distortion over all possible profiles, i.e., $dist(f) = \sup_{R \in \mathcal{R}^*} dist(f(R), R)$. To allow for a more fine-grained analysis, we further define $dist_m(f) = \sup_{R \in \mathcal{R}^*_m} dist(f(R), R)$ as the metric distortion of f when only profiles on m alternatives are considered. We note that $dist(f) = \infty$ and $dist_m(f) = \infty$ if the respective suprema are unbounded.

We recall here that the uniform random dictatorship f_{RD}, the randomized Plurality-Veto rule f_{RPV}, C2ML rules f_{C2ML}, and CRWW rules f_{CRWW} have a metric distortion of $dist(f_{RD}) = 3$, $dist(f_{RPV}) = 3$, $dist(f_{C2ML}) = 3$, and $dist(f_{CRWW}) \leq 2.753$, respectively. By contrast, the metric distortion of C1ML rules is unknown.

3 ANALYSIS OF C1ML RULES

As our first contribution, we will show that C1ML rules have a metric distortion of 4 and that no other majoritarian RSCF has a lower metric distortion when the number of alternatives is unbounded. Thus, our results show that C1ML rules minimize the metric distortion among majoritarian RSCFs. Our analysis of the C1ML rule is further motivated by the fact that maximal lottery rules have been repeatedly recommended for practical usage [11, 12]. All missing proofs can be found in the full version of this paper [25].

To prove our results, we first show a strong relation between the metric distortion of majoritarian RSCFs and distances in the majority relation. To this end, we define the *majority distance* $md(x, y, \succsim_R)$ as the length of the shortest path from x to y in the majority relation $\succsim_R$. In particular, $md(x, x, \succsim_R) = 0$, $md(x, y, \succsim_R) = 1$ if $x \succsim_R y$, and $md(x, y, \succsim_R) = \infty$ if there is no path from x to y in $\succsim_R$. We extend this notion also to lotteries p by defining $md(p, y, \succsim_R) = \sum_{x \in X_R} p(x) md(x, y, \succsim_R)$ and note that $md(p, y, \succsim_R) = \infty$ if and only if there is an alternative x with $p(x) > 0$ and $md(x, y, \succsim_R) = \infty$.

Proposition 1. *It holds for all majoritarian RSCFs f and preference profiles R on $m \geq 3$ alternatives that*
(1) $dist(f(R), R) \leq 1 + 2 \max_{x \in X_R} md(f(R), x, \succsim_R)$.
(2) $dist_m(f) \geq 1 + 2 \max_{x \in X_R} md(f(R), x, \succsim_R)$.

Proof sketch. For Claim (1), we first note that there is nothing to show if $\max_{x \in X_R} md(f(R), x, \succsim_R) = \infty$ and we hence suppose that $md(f(R), x, \succsim_R) < \infty$ for all $x \in X_R$. We then prove that $sc(x, d) \leq (1 + 2md(x, y, \succsim_R)) sc(y, d)$ for all $x, y \in X_R$ and $d \in D(R)$ by an induction on the majority distance between x and y. This insight implies Claim (1) as $dist(f(R), R) = \sup_{d \in D(R)} \frac{\sum_{x \in X_R} f(R,x) sc(x,d)}{\min_{y \in X_R} sc(y,d)}$. For Claim (2), we show that there is for every $\epsilon > 0$ a preference profile R^ϵ and a metric space $d \in D(R^\epsilon)$ such that $\succsim_{R^\epsilon} = \succsim_R$ and $\frac{sc(f(R),d)}{\min_{y \in X_R} sc(y,d)} \geq 1 + 2 \max_{x \in X_R} md(f(R), x, \succsim_R) - \epsilon$. Since $f(R^\epsilon) = f(R)$ as f is majoritarian, we then infer Claim (2) by letting ϵ go to 0. □

Claims related to Proposition 1 have been shown by Anshelevich et al. [2, Lemma 6] and Kempe [36, Corollary 5.1], but these results lack the lower bound given in (2). Based on our proposition, we will next compute the metric distortion of C1ML rules. In particular, our subsequent theorem shows that C1ML rules have a metric distortion of at most 4 and that no majoritarian RSCF has a lower metric distortion if the number of alternatives m is unbounded.

Theorem 1. *The following claims are true:*
(1) It holds for all C1ML rules f and $m \geq 3$ that $dist_m(f) \leq 4$ and $dist_m(f) \geq 4 - (\frac{1}{3})^{\lfloor \frac{m-3}{2} \rfloor}$.
(2) It holds for all majoritarian RSCFs f that $dist_m(f) \geq 4 - \frac{3}{m}$ if $m \geq 3$ is odd and $dist_m(f) \geq 4 - \frac{3}{m-1}$ if $m \geq 3$ is even.

Proof. We will only prove Claim (1) here and give a proof sketch for Claim (2). The full proof of Claim (2) can be found in [25].

Claim (1), upper bound: Let f denote a C1ML rule, let R denote a profile, and define $p = f(R)$. It follows from a result by Dutta and Laslier [19] that $p(x) > 0$ implies $md(x, y, \succsim_R) \leq 2$ for all $x, y \in X_R$. Based on this insight, we will show that $md(p, z, \succsim_R) \leq \frac{3}{2}$ for all $z \in X_R$ as Claim (1) of Proposition 1 then proves that $dist(p, R) \leq 4$. We thus fix an alternative $z \in X_R$ and let q denote the lottery with $q(z) = 1$. Further, we define $X^+ = \{x \in X_R : x \succ_R z\}$ and $X^- = \{x \in X_R : z \succ_R x\}$. By the definition of C1ML rules, it holds that $p \succsim_R q$, which implies that $\sum_{x \in X^+} p(x) \geq \sum_{x \in X^-} p(x)$. This means that $\sum_{x \in X^-} p(x) \leq \frac{1}{2}$. Next, it holds for all $x \in X_R$ with $p(x) > 0$ that $md(x, z, \succsim_R) = 1$ if $x \succsim_R z$ and $md(x, z, \succsim_R) = 2$ if $z \succ_R x$ due to our previous observation. Hence, we infer that $md(p, z, \succsim_R) \leq \sum_{x \in X_R : x \succsim_R z} p(x) + 2 \sum_{x \in X_R : z \succ_R x} p(x) = 1 + \sum_{x \in X^-} p(x) \leq \frac{3}{2}$. Finally, Claim (1) of Proposition 1 shows that $dist(p, R) \leq 4$.

Claim (1), lower bound: For proving our lower bound, we recall that C1ML rules are majoritarian and that $|C1ML(R)| = 1$ if the majority relation of R is strict [40]. Moreover, by McGarvey's construction [43], there is for every complete binary relation $\succsim$ on X_m a profile R with $\succsim_R = \succsim$. Due to Claim (2) of Proposition 1, we can hence show the lower bound by constructing a complete and anti-symmetric binary relation $\succsim^*$ for every X_m with $m \geq 3$ such that $\max_{x \in X_R} md(p, x, \succsim^*) = \frac{3}{2} - \frac{1}{2} \cdot (\frac{1}{3})^{\lfloor \frac{m-3}{2} \rfloor}$, where p is the unique C1 maximal lottery of a profile R with $\succsim_R = \succsim^*$. We first suppose that $m \geq 3$ is odd and consider the following relation $\succsim^*$ on X_m: for all odd $k < m$ and all j with $k+2 \leq j \leq m$, it holds that $x_{k+1} \succ^* x_k$, $x_k \succ^* x_j$, and $x_j \succ^* x_{k+1}$. It can be checked that the unique C1 maximal lottery p for this relation is defined by $p(x_k) = p(x_{k+1}) = (\frac{1}{3})^{\frac{k+1}{2}}$ for all odd $k < m$ and $p(x_m) = (\frac{1}{3})^{\frac{m-1}{2}}$. This means that $\sum_{x \in X^o} p(x) = \sum_{x \in X^e} p(x_k) = \frac{1}{2} - \frac{1}{2} p(x_m)$ for the sets $X^o = \{x_1, x_3, \dots, x_{m-2}\}$ and $X^e = \{x_2, x_4, \dots, x_{m-1}\}$. Next, by definition of $\succsim^*$, it holds for all odd $k < m$ that $md(x_k, x_m, \succsim^*) = 1$ and $md(x_{k+1}, x_m, \succsim^*) = 2$. Hence, $md(p, x_m, \succsim^*) = \sum_{x \in X^o} p(x) + 2 \sum_{x \in X^e} p(x) = 3\left(\frac{1}{2} - \frac{1}{2} p(x_m)\right) = \frac{3}{2} - \frac{1}{2} \cdot (\frac{1}{3})^{\frac{m-3}{2}}$. Proposition 1 then shows that $dist_m(f) \geq 4 - (\frac{1}{3})^{\frac{m-3}{2}}$. Finally, to extend this result to even m, we add a new alternative to $\succsim^*$ that loses all majority comparisons. Every C1ML rule will assign probability 0 to this alternative and it does hence not affect our analysis.

Claim (2): In this proof sketch, we assume that $m \geq 3$ is odd. To prove the theorem in this case, we will again use Claim (2) of Proposition 1 and hence construct a profile R such that $\max_{x \in X_R} md(p, x, \succsim_R) \geq \frac{3}{2} - \frac{3}{2m}$ for every lottery p. Next, McGarvey's theorem [43] allows us to focus on complete binary relations on X_m. The theorem then follows by proving that $\max_{x \in X_R} md(p, x, \succsim) \geq \frac{3}{2} - \frac{3}{2m}$ for all lotteries p and the "cyclic" relation $\succsim$ given by $x_i \succ x_{i+_m k}$ for all $i \in \{1, \dots, m\}$, $k \in \{1, \dots, \frac{m-1}{2}\}$ (where $i +_m k = i + k$ if $i + k \leq m$ and $i +_m k = i + k - m$ else). □

Remark 1. The upper bound in Claim (1) of Theorem 1 is tight as there are C1ML rules f with $dist(f) = 4$. To see this, consider the lottery p given by $p(a) = p(c) = \frac{1}{2}$ and a profile R with $X_R = \{a, b, c\}$, $a \succ_R b$, $b \succ_R c$, and $c \sim_R a$. Since p is C1 maximal in R and $md(p, b, \succsim_R) = \frac{3}{2}$, Proposition 1 shows that $dist(f) = 4$ for all C1ML rules f with $f(R) = p$. By contrast, the lower bound for C1ML rules is not tight. It can be shown that every C1ML rule has a metric distortion of at least $4 - 3\gamma_m$, where γ_m denotes the minimal non-zero probability that a C1ML rule assigns to an alternative in a profile with m alternatives and an odd number of voters. However, the probabilities γ_m are not well-understood [26], so we cannot use them to improve our lower bound for C1ML rules.

Remark 2. Proposition 1 allows us to identify the majoritarian RSCF that minimizes $dist_m(f)$ for a fixed number of alternatives m: this RSCF f^* chooses for each profile R a lottery p that minimizes $\max_{x \in X_R} md(p, x, \succsim_R)$. Based on a computer-aided approach, we have shown that $dist_m(f^*) = 4 - \frac{3}{m}$ for all odd $m \leq 9$, which proves that Claim (2) of Theorem 1 is tight in these cases.

Remark 3. Proposition 1 recovers known bounds on the metric distortion of majoritarian SCFs. For instance, this proposition implies that every alternative in the uncovered set has a metric distortion of 5 because the uncovered set is the set of alternatives that can reach every other alternative in at most two steps. This result has been first shown by Anshelevich et al. [2].

4 SIMULATIONS

As our second contribution, we conduct extensive computer experiments to gain insights into the average-case metric distortion of the RSCFs in Section 2.1. To this end, we first derive a linear program that efficiently computes the metric distortion of a lottery for a profile (Section 4.1), and then explain the setup and results of our experiments (Sections 4.2 and 4.3). The code for our experiments is publicly available on Zenodo [24].

4.1 Computing the Metric Distortion

The main challenge for our experiments is to compute the metric distortion $dist(p, R)$ for a given lottery p and profile R. To this end, we note that it suffices to compute the term $dist(p, R, x) = \sup_{d \in D(R)} \frac{sc(p,d)}{sc(x,d)}$ for all alternatives $x \in X_R$ because $dist(p, R) = \max_{x \in X_R} dist(p, R, x)$. Moreover, we can assume that $sc(x, d) = 1$ since the term $\frac{sc(p,d)}{sc(x,d)}$ is invariant under scaling d. Hence, we only need to find for every alternative x the metric d_x that maximizes $sc(p, d_x)$ subject to $d_x \in D(R)$ and $sc(x, d_x) = 1$. While this can be done by linear programs that use the distances $d(x, v)$ as variables and encode that $d \in D(R)$ and $sc(x, d) = 1$, this straightforward approach is too slow for our experiments as we need $O((n+m)^3)$ constraints to formalize the triangle inequalities for metrics.

To derive a more efficient method to compute $dist(p, R, x)$, we will use that the metric distortion of a lottery p for a profile R can be computed by only considering the biased metrics of Charikar and Ramakrishnan [15]. To define these metrics, we let $\succeq_v$ denote the relation given by $x \succeq_v y$ if and only if $x \succ_v y$ or $x = y$ for all $x, y \in X_R$. Then, a metric d is *biased* for a profile R if there is an alternative $x^* \in X_R$ and a function $t : X_R \to \mathbb{R}_{\geq 0}$ such that *(i)* $t(x^*) = 0$, *(ii)* $d(x^*, v) = \frac{1}{2} \max_{x, y \in X_R \colon x \succeq_v y} t(x) - t(y)$ for all $v \in V_R$, and *(iii)* $d(x, v) = d(x^*, v) + \min_{y \in X_R \colon x \succeq_v y} t(y)$ for all $v \in V_R$ and all $x \in X_R \setminus \{x^*\}$. Unfortunately, due to the maxima and minima in the definition of these metrics, we cannot directly use them to compute $dist(p, R)$. We thus adapt the idea of biased metrics to construct a linear program that efficiently computes this value. In more detail, we will show that the following LP (called LP 1), which uses variables $d(x, v)$ and $t(x)$ for $x \in X_R$ and $v \in V_R$, computes $dist(p, R, x^*)$ for every lottery p, profile R, and alternative x^*.

$$
\begin{aligned}
\max \ & \sum_{x \in X_R} p(x) \sum_{v \in V_R} d(x, v) \\
\text{s.t. } & t(x^*) = 0 \\
& t(x) \geq 0 && \forall x \in X_R \\
& d(x^*, v) \geq \tfrac{1}{2}(t(x) - t(y)) && \forall v \in V_R,\ x, y \in X_R \colon x \succeq_v y \\
& d(x, v) \leq d(x^*, v) + t(y) && \forall v \in V_R,\ x, y \in X_R \colon x \succeq_v y \\
& d(x, v) + d(x^*, v) \geq t(x) && \forall v \in V_R,\ x \in X_R \\
& \textstyle\sum_{v \in V_R} d(x^*, v) = 1
\end{aligned}
\tag{LP 1}
$$

Proposition 2. *Fix a profile R, a lottery p, and an alternative x^*. If the optimal objective value o^*_{LP} of LP 1 is bounded, then $dist(p, R, x^*) = o^*_{LP}$ and otherwise $dist(p, R, x^*) = \infty$.*

Proof sketch. Let R denote a profile, p a lottery, and x^* an alternative. It can be checked that every biased metric d together with its inducing function t satisfies the conditions of LP 1, so the optimal objective value o^*_{LP} of our LP is lower bounded by $dist(p, R, x^*)$. Specifically, the constraints in the first four lines follow directly from the definition of d and t, the fifth line follows by substituting the definitions of $d(x, v)$ and $d(x^*, v)$, and the constraint that $\sum_{v \in V_R} d(x^*, v) = 1$ can be enforced by scaling d and t without affecting $dist(p, R, x^*)$. Conversely, to show that $dist(p, R, x^*) \geq o^*_{LP}$, we prove that every feasible solution of LP 1 with objective value o_{LP} can be transformed into a metric $d \in D(R)$ such that $\frac{sc(p,d)}{sc(x^*,d)} \geq o_{LP}$. We note that this direction is independent of the work of Charikar et al. [16] as we need to reason about our constraints to turn a feasible solution of our LP into a metric. ◻

Given a profile R on n voters and m alternatives, LP 1 has $O(nm^2)$ constraints and it is thus very fast to solve this LP. For example, based on LP 1, we need in average roughly 20 seconds on a single core of an Apple M1 Ultra chip to compute the metric distortion of a lottery for a profile with 201 voters and 15 alternatives.

4.2 Simulations with Synthetic Data

As our first computer experiment, we conduct extensive simulations based on synthetically generated preference profiles. In more detail, we generate 1000 preference profiles on m alternatives and n voters for 14 distributions over preference profiles and all combinations of $(m, n) \in \{5, 10, 15\} \times \{11, 21, \dots, 201\}$. For each of the sampled preference profiles R, we then compute the lottery $f(R)$ chosen by the five RSCFs discussed in Section 2.1 and their respective metric distortion $dist(f(R), R)$. Finally, for each $m \in \{5, 10, 15\}$ and each distribution, we plot the average metric distortion of each RSCF as a function depending on the number of voters n. Due to space restrictions, we show the results of these simulations only for two exemplary distributions, namely the impartial culture model and the 3-dimensional Euclidean cube model. The plots for the other distributions (the Mallow's model, the Pólya-Eggenberg urn model, and the t-dimensional Euclidean cube and ball models with various parameterizations) as well as further statics can be found in the full version [25]. In particular, our experiments cover all major models used in the "map of elections" [7, 8, 49]. We next define the impartial culture and the Euclidean cube models.

Impartial Culture (IC). In this model, each voter is assigned a preference relation independently and uniformly at random. Hence, for each voter $v \in V_n$ and preference relation $\succ \in \mathcal{R}(X_m)$, the probability that $\succ$ is assigned to v is $\frac{1}{m!}$.

t-Dimensional Euclidean Cube (tEC). In this model, we assign voters and alternatives independently and uniformly at random to points in the t-dimensional cube $[-1, 1]^t$. The voters' preference relations are then given by their distances to the alternatives: a voter v prefers alternative x to alternative y if $|p_v - p_x|_2 < |p_v - p_y|_2$ where p_v, p_x, and p_y denote the points of v, x, and y in the hypercube. In the main body, we use this model with $t = 3$ dimensions.

The results of our simulations for these two models are shown in Figure 1. We first note that, in most experiments, the measured variance is rather lower, with typical values lying between 0.05 to 0.01 (see the full version [25] for details). Moreover, in all experiments, the average metric distortion of the considered RSCFs is significantly smaller than their worst-case metric distortion, thus indicating that such worst-case bounds are too pessimistic for more realistic profiles. In particular, the average metric distortion of all RSCFs is usually in the interval $[2, 2.5]$, which also shows that the choice of a particular rule has only limited effect. This is especially striking when comparing C1ML and C2ML rules, which are almost indistinguishable in our experiments even though the worst-case metric distortion is 3 for C2ML rules and 4 for C1ML rules.

Beyond these general observations, there are several interesting trends in our experiments that can be observed for most of the distributions. We explain these trends for each RSCF individually.

CRWW rule. In most of our simulations and especially when $m \in \{10, 15\}$, the CRWW rule f_{CRWW} has the lowest average metric distortion among the tested rules. In particular, for effectively all distributions and all numbers of voters, the average metric distortion of this rule lies between 2 and 2.15. These results suggest that, when the metric distortion is the central factor for deciding on the RSCF, we should use the CRWW rule as it has both the best worst-case and average-case metric distortion.

Randomized Plurality-Veto. The randomized Plurality-Veto rule f_{RPV} has often a very low average metric distortion if there are only $m = 5$ alternatives, but it becomes worse as m increases. For instance, in the 3-dimensional Euclidean cube model, it has for most values of n an average metric distortion of less than 2, but its average metric distortion increases to over 2.2 when $m = 15$. By contrast, in the impartial culture model, the average metric distortion of f_{RPV} depends significantly on the number of voters and roughly converges against $2 + \frac{1}{m-1}$. We believe the reason for this is that, in our simulations, f_{RPV} randomizes over larger sets of alternatives when m increases. This is beneficial for the metric distortion if preference profiles are sufficiently close to uniform, but detrimental if there is an alternative that every voter appreciates.

Uniform random dictatorship. The average metric distortion of the uniform random dictatorship f_{RD} becomes smaller as the number of voters increases when using distributions that are close to uniform. For instance, for the impartial culture model and all $m \in \{5, 10, 15\}$, the average metric distortion of f_{RD} converges to a value close to 2 as the number of voters n increases. By contrast, if the voters' preferences are more structured (e.g., in the Euclidean cube model), the average metric distortion is largely independent of the number of voters and significantly worse than that of the other rules. A possible explanation for this is that f_{RD} only considers the voters' favorite alternatives and thus fails to identify strong compromise alternatives. In particular, such strong alternatives are likely to exist for structured distributions, but typically do not exist if n is large and the distribution over profiles is close to uniform.

C1ML and C2ML rules. The average metric distortion of C1ML and C2ML rules is rather high for the impartial culture model with $m = 5$ (close to 2.25 when $n \geq 100$), but it is close to that of the CRWW rule for more structured distributions (e.g., the 3-dimensional Euclidean cube model). Moreover, for most distributions, the average metric distortion of these rules decreases when the number of alternatives increases. Our explanation for this is that C1ML and C2ML rules use very little randomization as they select randomized Condorcet winners. This behavior results in a low metric distortion if there are alternatives that severely beat all other alternatives in a pairwise comparison, but it is detrimental if all alternatives are roughly equally good.

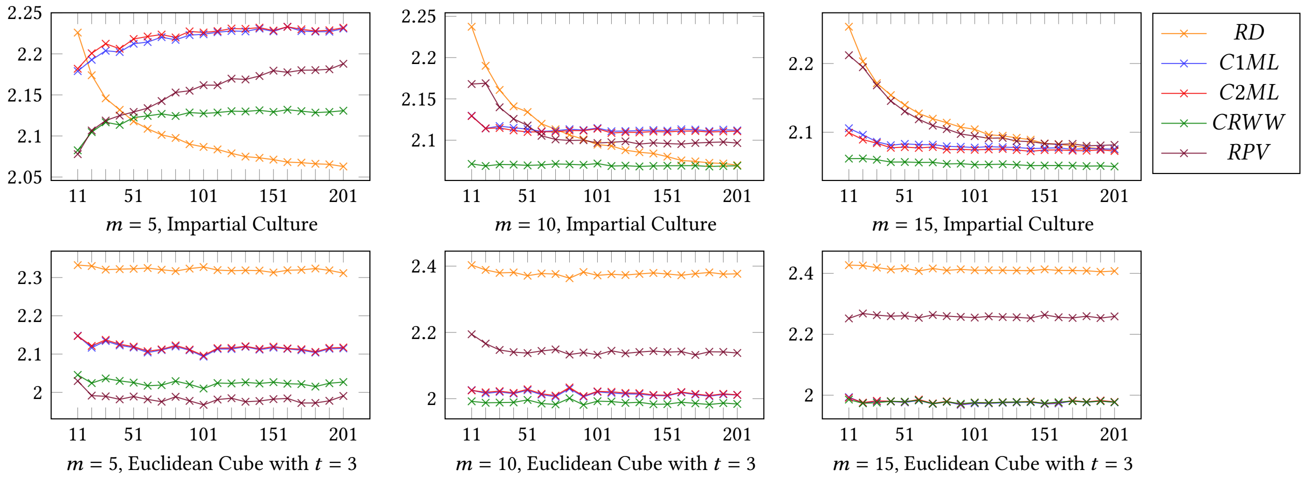

Figure 1: Results of our simulations for the impartial culture and 3-dimensional Euclidean cube models. For both models and $m \in \{5, 10, 15\}$ alternatives, we plot the average metric distortion (y-axis) of the uniform random dictatorship, the C2ML and C1ML rules, the randomized Plurality-Veto rule, and the CRWW rule subject to the number of voters $n \in \{11, 21, \dots, 201\}$ (x-axis).

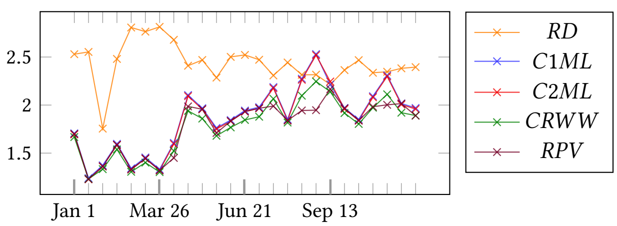

Figure 2: Results of our simulations with the Spotify Daily dataset. Each data point presents the average metric distortion of one of our RSCFs over 14 days (e.g., the first data point averages the metric distortion from January 01 to January 14, the second one from January 15 to January 28, etc.).

4.3 Simulations with Real-world Data

We also conduct computer experiments on the average metric distortion of our RSCFs on real-world data from PrefLib [42]. In more detail, we use the *Spotify Daily dataset* provided by Boehmer and Schaar [9] for our experiments. This dataset contains the rankings of the 200 most listened songs on Spotify for 53 countries and every day in 2017, and the rankings of each day form an election. Since not every country ranks the same songs among their top-200, Boehmer and Schaar [9] have identified maximal complete subelections for each day, which typically contain between 40 and 50 voters and around 20 alternatives. These maximal subelections are very well-suited for our computer experiments because they contain a reasonable number of voters and alternatives and all voters rank all alternatives. Due to these characteristics, we can directly compute our five RSCFs and their metric distortion on these maximal subelections, i.e., we conduct our experiments on the real-world data without any modifications other than those made by Boehmer and Schaar [9]. This also means that it suffices to compute the metric distortion of our RSCFs for each subelection once since no randomization is used in the generation of preference profiles.

The results of our experiments with the Spotify Daily dataset are shown in Figure 2, where we display the average metric distortion of each RSCF in a biweekly rhythm, i.e., each data point is the average of 14 days. Additional statistics can again be found in the full version [25]. We note that the simulations on the Spotify Daily dataset roughly agree with our computer experiments based on, e.g., the Euclidean models or Mallow's model (see [25] for more details). In more detail, Figure 2 shows that the metric distortion of the randomized Plurality-Veto and CRWW rules is typically only slightly better than that of the C1ML and C2ML rules, whereas the uniform random dictatorship often performs significantly worse. There are, however, two central differences between our simulations with synthetic data and the Spotify Daily dataset. Firstly, in the first three month, there is often a very dominant alternative in the elections from Spotify, which results in a very low metric distortion for all tested RSCFs but f_{RD}. Such profiles do typically not appear in our synthetic data, which may hint at the fact that our computer experiments are still too pessimistic. However, after the first three month, this effect vanishes and our RSCFs take similar values as in the simulation with synthetic data. Secondly, the randomized Plurality-Veto rule f_{RPV} performs surprisingly well on the real-world data, in particular in light of the large number of alternatives. The reason for this may be that, while there are many alternatives, only few of them are first-ranked in the preference profiles and only such alternatives have a chance to win under f_{RPV}.

To summarize our computer experiments with both synthetic and real-world data, we believe that they firstly show that that tailored RSCFs, such as the CRWW rule and the randomized Plurality-Veto rule, typically also have the smallest average-case metric distortion. However, especially when preference profiles are sufficiently structured, C1ML and C2ML rules are only slightly worse, thus providing an argument in favor of these rules. Lastly, our simulations show that the uniform random dictatorship is not suitable to minimize the metric distortion in practice, especially when we expect strong alternatives to exist.

5 THEORETICAL AVERAGE-CASE ANALYSIS

Lastly, we will analytically examine the average-case metric distortion of our RSCFs by calculating their expected metric distortion for a randomly drawn profile. In particular, the results in this section can be seen as a rigorous mathematical counterpart to the simulations in Section 4. For mathematical feasibility, we will restrict our attention to the impartial culture model and write $IC(m,n)$ for the respective probability distribution over profiles with m alternatives and n voters. While the impartial culture model is somewhat unrealistic, it is frequently analyzed as it is often seen as a good starting point for average-case analyses [e.g., 10, 27, 28, 35, 45, 46].

As we show next, if the number of voters goes to infinity, the expected metric distortion of every RSCF with bounded metric distortion converges to a value between 2 and $2 + \frac{1}{m-1}$ under the impartial culture model. This means that the choice of the voting rule has only a small effect on the expected metric distortion if there is a large number of voters and alternatives.

Theorem 2. *Let $m \geq 3$. It holds for every RSCF f with $dist_m(f) < \infty$ and $z = \liminf_{n\to\infty} \mathbb{P}_{R\sim IC(m,n)}[\exists x \in X_R : f(R,x) = 0]$ that*

(1) $\limsup_{n\to\infty} \mathbb{E}_{R\sim IC(m,n)}[dist(f(R),R)] \leq 2 + \frac{1}{m-1}$
(2) $\liminf_{n\to\infty} \mathbb{E}_{R\sim IC(m,n)}[dist(f(R),R)] \geq 2 + \frac{z}{m-1}$.

Proof sketch. The basic idea of this proof is that, as n grows larger, a profile drawn from $IC(m,n)$ is with high probability close to the profile R^* where each preference relation is reported by the same amount of voters. We thus start by analyzing the profile R^* and show with the help of LP 1 that the metric distortion of a lottery p on R^* is $2 + \frac{1}{m-1} - \frac{m}{m-1} \cdot \min_{x\in X_m} p(x)$. Next, we prove that the metric distortion of a lottery p for a profile R that contains R^* as a large subprofile can be bounded based on the metric distortion of p for R^*. Based on these two insights, we then fix a number of voters n, set $\alpha = \frac{1}{\sqrt[3]{n}}$, and define T^α as the set of profiles for n voters such that each preference relation is reported by more than $(1-\alpha)\frac{n}{m!}$ voters. Using the law of total probability, we derive that

$$\begin{aligned}\mathbb{E}[dist(f(R),R)] &= \mathbb{P}[R \notin T^\alpha] \cdot \mathbb{E}[dist(f(R),R)|R \notin T^\alpha] \\ &+ \mathbb{P}[R \in T^\alpha] \cdot \mathbb{E}[dist(f(R),R)|R \in T^\alpha]\end{aligned}$$

Finally, for our upper bound, we use standard concentration bounds to show that $\mathbb{P}[R \notin T^\alpha]$ goes to 0 as n increases. Since $\mathbb{E}[dist(f(R),R)|R \notin T^\alpha] \leq dist_m(f) < \infty$, it hence follows that $\mathbb{E}[dist(f(R),R)]$ converges to $\mathbb{E}[dist(f(R),R)|R \in T^\alpha]$ when n goes to infinity, and using our previous insights, we can bound this expectation by $2 + \frac{1}{m-1}$. Similarly, for our lower bound, we use that $\mathbb{E}[dist(f(R),R)] \geq \mathbb{P}[R \in T^\alpha] \cdot \mathbb{E}[dist(f(R),R)|R \in T^\alpha]$ and derive a lower bound on $\mathbb{E}[dist(f(R),R)|R \in T^\alpha]$. □

Based on a similar approach as in Theorem 2, we will next precisely compute the expected metric distortion of the uniform random dictatorship f_{RD} and the randomized Plurality-Veto rule f_{RPV}. In particular, we will show that, in the limit, f_{RD} has an optimal expected metric distortion of 2 under the impartial culture model, whereas the expected metric distortion of f_{RPV} is $2 + \frac{1}{m-1}$.

Theorem 3. *It holds for every $m \geq 3$ that*

(1) $\lim_{n\to\infty} \mathbb{E}_{R\sim IC(m,n)}[dist(f_{RD}(R),R)] = 2$.
(2) $\lim_{n\to\infty} \mathbb{E}_{R\sim IC(m,n)}[dist(f_{RPV}(R),R)] = 2 + \frac{1}{m-1}$.

Proof Sketch. For Claim (1), we show that $f_{RD}(R)$ returns with high probability a lottery close to the uniform one when n is large and R is drawn from $IC(m,n)$. From this insight, we then derive that $\limsup_{n\to\infty} \mathbb{E}_{R\sim IC(m,n)}[dist(f(R),R)] \leq 2$. Combined with the lower bound in Theorem 2, this proves Claim (1). For the claim on f_{RPV}, we prove that this rule only randomizes over all alternatives when each alternative is top-ranked and bottom-ranked by the same number of voters. Since this happens with probability 0 when n goes to infinity, we infer that $\liminf_{n\to\infty} \mathbb{P}_{R\sim IC(m,n)}[\exists x \in X_R : f_{RPV}(R,x) = 0] = 1$, and Claim (2) follows from Theorem 2. □

Remark 4. We leave an analogous result to Theorem 3 for C1ML and C2ML rules open because completely reversed preference relations cancel each other out for these RSCFs. Thus, a small part of the profile may determine the outcome, which severely complicates the analysis of these rules. However, computer experiments by Brandl et al. [12] suggest that the probability $\mathbb{P}_{R\sim IC(m,n)}[\exists x \in X_R : f(R,x) = 0]$ is close to 1 for C1ML and C2ML rules. Hence, Theorem 2 implies that the expected metric distortion of these RSCFs under the IC model is close to $2 + \frac{1}{m-1}$ when n goes to ∞.

Remark 5. Theorem 2 shows that, under the impartial culture distribution, *every* deterministic SCF f with $dist_m(f) < \infty$ has an expected metric distortion of $2 + \frac{1}{m-1}$ when n goes to ∞. This holds because every SCF is an RSCF that always assigns probability 1 to a single alternative, so the value z in Theorem 2 is 1.

6 CONCLUSION

In this paper, we study the metric distortion of randomized social choice functions, with a particular focus on well-established RSCFs such as the uniform random dictatorship, C1ML rules, and C2ML rules. Specifically, we first show that every C1ML rule has a metric distortion of at most 4, and we give a lower bound on the metric distortion of all majoritarian RSCFs that converges to 4 as m increases. This means that C1ML rules minimize the metric distortion within the class of majoritarian RSCFs when the number of alternatives is unbounded. Secondly, we conduct extensive computer experiments on the metric distortion of these three classical RSCFs as well as two RSCFs designed to minimize the metric distortion. These experiments show that, while RSCFs designed to minimize the metric distortion have also the best average-case metric distortion, C1ML and C2ML rules are often only slightly worse. Finally, we also conduct an analytical average-case analysis for the impartial culture model and, surprisingly, derive that the exact choice of voting rule has only a negligible influence on the expected metric distortion if the number of voters is large. In summary, we believe that these results demonstrate that established RSCFs, such as C1ML and C2ML rules, are also appealing when studied through the lens of metric distortion as they have a reasonable worst-case metric distortion and their average-case metric distortion is only slightly worse than that of RSCFs that are tailored to minimize the metric distortion.

ACKNOWLEDGMENTS

We thank three anonymous reviewers for their valuable feedback. This work was supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/11-2 and BR 2312/12-1, and by the NSF-CSIRO grant on "Fair Sequential Collective Decision-Making" (RG230833).

REFERENCES

[1] Ioannis Anagnostides, Dimitris Fotakis, and Panagiotis Patsilinakos. 2022. Dimensionality and Coordination in Voting: The Distortion of STV. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*. 4776–4784.

[2] Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. 2018. Approximating Optimal Social Choice under Metric Preferences. *Aritificial Intelligence* 264 (2018), 27–51.

[3] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. 2015. Approximating Optimal Social Choice under Metric Preferences. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*. 777–783.

[4] Elliot Anshelevich, Aris Filos-Ratsikas, Nisarg Shah, and Alexandros A. Voudouris. 2021. Distortion in social choice problems: The first 15 years and beyond. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*. 4294–4301.

[5] Elliot Anshelevich and John Postl. 2017. Randomized Social Choice Functions Under Metric Preferences. *Journal of Artificial Intelligence Research* 58 (2017), 797–827.

[6] Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura (Eds.). 2011. *Handbook of Social Choice and Welfare*. Vol. 2. North-Holland.

[7] Niclas Boehmer, Robert Bredereck, Piotr Faliszewski, Rolf Niedermeier, and Stanislaw Szufa. 2021. Putting a Compass on the Map of Elections. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*. 59–65.

[8] Niclas Boehmer, Piotr Faliszewski, Łukas Janeczko, Andrzej Kaczmarczyk, Grzegorz Lisowski, Grzegorz Pierczyński, Simon Rey, Dariusz Stolicki, Stanisław Szufa, and Tomasz Wąs. 2024. Guide to Numerical Experiments on Elections in Computational Social Choice. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI)*. 7962–7970.

[9] Niclas Boehmer and Nathan Schaar. 2023. Collecting, Classifying, Analyzing, and Using Real-World Ranking Data.. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 1706–1715.

[10] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. 2015. Optimal social choice functions: a utilitarian view. *Artificial Intelligence* 227 (2015), 190–213.

[11] Florian Brandl, Felix Brandt, and Hans Georg Seedig. 2016. Consistent Probabilistic Social Choice. *Econometrica* 84, 5 (2016), 1839–1880.

[12] Florian Brandl, Felix Brandt, and Christian Stricker. 2022. An Analytical and Experimental Comparison of Maximal Lottery Schemes. *Social Choice and Welfare* 58, 1 (2022), 5–38.

[13] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. 2016. Introduction to Computational Social Choice. In *Handbook of Computational Social Choice*, Felix Brandt, Vincent Conitzer, Ulle Endriss, J. Lang, and Ariel D. Procaccia (Eds.). Cambridge University Press, Chapter 1.

[14] Ioannis Caragiannis and Karl Fehers. 2024. *Beyond the worst case: Distortion in impartial culture electorates*. Technical Report. https://arxiv.org/pdf/2307.07350.

[15] Moses Charikar and Prasanna Ramakrishnan. 2022. Metric Distortion Bounds for Randomized Social Choice. *Proceedings of the 33th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (2022), 2986–3004.

[16] Moses Charikar, Prasanna Ramakrishnan, Kangning Wang, and Hongxun Wu. 2024. Breaking the Metric Voting Distortion Barrier. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1621–1640.

[17] Yu Cheng, Shaddin Dughmi, and David Kempe. 2017. Of the people: voting is more effective with representative candidates. In *Proceedings of the 18th ACM Conference on Economics and Computation (ACM-EC)*. 305–322.

[18] Yu Cheng, Shaddin Dughmi, and David Kempe. 2018. On the Distortion of Voting with Multiple Representative Candidates. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*. 973–980.

[19] Bhaskar Dutta and Jean-François Laslier. 1999. Comparison Functions and Choice Correspondences. *Social Choice and Welfare* 16, 4 (1999), 513–532.

[20] Soroush Ebadian, Aris Filos-Ratsikas, Mohamad Latifan, and Nisarg Shah. 2024. Computational Aspects of Distortion. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 499–507.

[21] Brandon Fain, Ashish Goel, Kamesh Munagala, and Nina Prabhu. 2019. Random Dictators with a Random Referee: Constant Sample Complexity for Social Choice. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. 1893–1900.

[22] Michael Feldman, Amos Fiat, and Iddan Golomb. 2016. On Voting and Facility Location. In *Proceedings of the 17th ACM Conference on Economics and Computation (ACM-EC)*. 269–286.

[23] Peter C. Fishburn. 1984. Probabilistic Social Choice Based on Simple Voting Comparisons. *Review of Economic Studies* 51, 4 (1984), 683–692.

[24] Fabian Frank and Patrick Lederer. 2025. *Code for the Paper "The Metric Distortion of Randomized Social Choice Functions: C1 Maximal Lottery Rules and Simulations" (v1.0.0). Zenodo.* https://doi.org/10.5281/zenodo.14854198.

[25] Fabian Frank and Patrick Lederer. 2025. *The Metric Distortion of Randomized Social Choice Functions: C1 Maximal Lottery Rules and Simulations*. Technical Report. https://arxiv.org/abs/2403.18340.

[26] David C. Fisher and Jennifer Ryan. 1995. Tournament Games and Condorcet voting. *Linear Algebra Appl.* 217 (1995), 87–100.

[27] William V. Gehrlein and Peter C. Fishburn. 1976. The Probability of the Paradox of Voting: A Computable Solution. *Journal of Economic Theory* 13, 1 (1976), 14–25.

[28] William V. Gehrlein and Peter C. Fishburn. 1978. Probabilities of election outcomes for large electorates. *Journal of Economic Theory* 19, 1 (1978), 38–49.

[29] Allan Gibbard. 1977. Manipulation of schemes that mix voting with chance. *Econometrica* 45, 3 (1977), 665–681.

[30] Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. 2020. Resolving the Optimal Metric Distortion Conjecture. In *Proceedings of the 61st Symposium on Foundations of Computer Science (FOCS)*. 1427–1438.

[31] Ashish Goel, Anilesh K. Krishnaswamy, and Kamesh Munagala. 2017. Metric Distortion of Social Choice Rules: Lower Bounds and Fairness properties. In *Proceedings of the 18th ACM Conference on Economics and Computation (ACM-EC)*. 287–304.

[32] Yannal A. Gonczarowski, Gregory Kehne, Ariel D. Procaccia, Ben Schiffer, and Shirley Zhang. 2023. The Distortion of Binomial Voting Defies Expectation. In *Proceedings of the 37th Annual Conference on Neural Information Processing Systems (NeurIPS)*.

[33] Stephen Gross, Elliot Anshelevich, and Lirong Xia. 2017. Vote Until Two of You Agree: Mechanisms with Small Distortion and Sample Complexity. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*. 544–550.

[34] Lê Nguyên Hoang. 2017. Strategy-Proofness of the Randomized Condorcet Voting System. *Social Choice and Welfare* 48, 3 (2017), 679–701.

[35] Joshua Kavner and Lirong Xia. 2021. Strategic Behavior is Bliss: Iterative Voting Improves Social Welfare. In *Proceedings of the 35th Annual Conference on Neural Information Processing Systems (NeurIPS)*.

[36] David Kempe. 2020. An analysis framework for metric voting based on LP duality. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*. 2079–2086.

[37] Fatih Erdem Kizilkaya and David Kempe. 2022. Plurality Veto: A simple voting rule achieving optimal metric distortion. In *Proceedings of the 31th International Joint Conference on Artificial Intelligence (IJCAI)*. 349–355.

[38] Fatih Erdem Kizilkaya and David Kempe. 2023. Generalized Veto Core and a Practical Voting Rule with Optimal Metric Distortion. In *Proceedings of the 24th ACM Conference on Economics and Computation (ACM-EC)*. 913–936.

[39] Gilbert Laffond, Jean-François Laslier, and Michel Le Breton. 1993. The Bipartisan Set of a Tournament Game. *Games and Economic Behavior* 5, 1 (1993), 182–201.

[40] Gilbert Laffond, Jean-François Laslier, and Michel Le Breton. 1997. A Theorem on Symmetric Two-Player Zero-Sum Games. *Journal of Economic Theory* 72, 2 (1997), 426–431.

[41] Michel Le Breton. 2005. On the Uniqueness of Equilibrium in Symmetric Two-Player Zero-Sum Games with Integer Payoffs. *Économie publique* 17, 2 (2005), 187–195.

[42] N. Mattei and Toby Walsh. 2013. PrefLib: A Library for Preference Data. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT) (Lecture Notes in Computer Science (LNCS), Vol. 8176)*. Springer-Verlag, 259–270.

[43] David C. McGarvey. 1953. A Theorem on the Construction of Voting Paradoxes. *Econometrica* 21, 4 (1953), 608–610.

[44] Kamesh Munagala and Kangning Wang. 2019. Improved Metric Distortion for Deterministic Social Choice Rules. In *Proceedings of the 20th ACM Conference on Economics and Computation (ACM-EC)*. 245–262.

[45] Elisha A. Pazner and Eugene Wesley. 1978. Cheatproofness Properties of the Plurality Rule in Large Elections. *The Review of Economic Studies* 45, 1 (1978), 85–91.

[46] Geoffrey Pritchard and Marc C. Wilson. 2009. Asymptotics of the minimum manipulating coalition size for positional voting rules under impartial culture behaviour. *Mathematical Social Sciences* 58 (2009), 35–57.

[47] Ariel D. Procaccia and Jeffrey S. Rosenschein. 2006. The Distortion of Cardinal Preferences in Voting. In *Cooperative Information Agents X*. Springer, 317–331.

[48] Piotr Skowron and Edith Elkind. 2017. Social choice under metric preferences: scoring rules and STV. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*. 706–712.

[49] Stanislaw Szufa, Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. 2020. Drawing a Map of Elections in the Space of Statistical Cultures. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 1341–1349.