



# Optimal social choice functions: A utilitarian view <sup>☆</sup>



Craig Boutilier <sup>a,1</sup>, Ioannis Caragiannis <sup>b</sup>, Simi Haber <sup>c</sup>, Tyler Lu <sup>a,2</sup>,  
Ariel D. Procaccia <sup>d,\*</sup>, Or Sheffet <sup>e</sup>

<sup>a</sup> Dept. of Computer Science, University of Toronto, Canada

<sup>b</sup> Computer Technology Institute "Diophantus" and Dept. of Computer Engineering and Informatics, University of Patras, Greece

<sup>c</sup> Dept. of Mathematics, Bar-Ilan University, Israel

<sup>d</sup> Computer Science Dept., Carnegie Mellon University, United States

<sup>e</sup> Center for Research on Computation and Society, Harvard SEAS, United States

## ARTICLE INFO

### Article history:

Received 12 April 2014

Received in revised form 11 June 2015

Accepted 14 June 2015

Available online 25 June 2015

### Keywords:

Computational social choice

## ABSTRACT

We adopt a utilitarian perspective on social choice, assuming that agents have (possibly latent) utility functions over some space of alternatives. For many reasons one might consider mechanisms, or *social choice functions*, that only have access to the ordinal rankings of alternatives by the individual agents rather than their utility functions. In this context, one possible objective for a social choice function is the maximization of (expected) social welfare relative to the information contained in these rankings. We study such *optimal* social choice functions under three different models, and underscore the important role played by *scoring functions*. In our worst-case model, no assumptions are made about the underlying distribution and we analyze the worst-case *distortion*—or degree to which the selected alternative does not maximize social welfare—of optimal (randomized) social choice functions. In our average-case model, we derive optimal functions under neutral (or impartial culture) probabilistic models. Finally, a very general learning-theoretic model allows for the computation of optimal social choice functions (i.e., ones that maximize expected social welfare) under arbitrary, sampleable distributions. In the latter case, we provide both algorithms and sample complexity results for the class of scoring functions, and further validate the approach empirically.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Classic models in social choice theory assume that the preferences of a set of *agents* over a set of *alternatives* are represented as linear orders; a *social choice function*, given these preferences as input, outputs a single socially desirable alternative. A host of clever social choice functions have been designed to satisfy various *normative* criteria. Most work in *computational social choice* studies computational aspects of these models, addressing questions such as the complexity of computing social choice functions [5,17] or manipulating them (see the survey by Faliszewski and Procaccia [13]).

<sup>☆</sup> A preliminary version of this paper appeared in the proceedings of EC'12.

<sup>\*</sup> Corresponding author.

E-mail addresses: [cebly@cs.toronto.edu](mailto:cebly@cs.toronto.edu) (C. Boutilier), [caragian@ceid.upatras.gr](mailto:caragian@ceid.upatras.gr) (I. Caragiannis), [simi@math.biu.ac.il](mailto:simi@math.biu.ac.il) (S. Haber), [tl@cs.toronto.edu](mailto:tl@cs.toronto.edu) (T. Lu), [arielp@cs.cmu.edu](mailto:arielp@cs.cmu.edu) (A.D. Procaccia), [osheffet@seas.harvard.edu](mailto:osheffet@seas.harvard.edu) (O. Sheffet).

<sup>1</sup> Currently on leave at Google, Inc., Mountain View, CA.

<sup>2</sup> Currently at Google, Inc., Mountain View, CA.

Under ordinal preferences, an axiomatic approach to obtaining a socially desirable outcome seems—on the face of it—necessary, absent concrete measures of the quality of an alternative. In contrast, some work in economics assumes *cardinal* preferences and takes a *utilitarian* approach. This viewpoint dates to the work of Bentham at the end of the 18th century, who argued that “it is the greatest happiness of the greatest number that is the measure of right and wrong.” This axiom suggests that happiness can be quantified, and indeed, having coined the term *utility*, Bentham proposed that the goal of government is to maximize the sum of individual utilities—the *social welfare* (defying contemporary wisdom that the goal of government is to enrich the coffers of the ruler). The utilitarian approach is prevalent, for example, in mechanism design, and perhaps even more so in *algorithmic* mechanism design [25].

In this paper we view the social choice problem through this utilitarian lens. Our premise is that agents have (possibly implicit) utility functions, and the goal of a social choice function is to maximize the (utilitarian) social welfare<sup>3</sup>—i.e., (possibly weighted) sum of agent utilities—of the selected alternative. The utilitarian perspective is not appropriate for all social choice problems (a point we discuss further below). However, the methods of social choice—especially voting systems—are finding increasing application in recommender systems, web search, product design, and many more practical domains, in which the primary aim is often, as in much of mechanism design, to aggregate preferences so that utility or efficiency is maximized. Indeed, one motivation for our work is the development of group recommendation systems for a variety of domains, including low-stakes consumer applications and higher profile public policy and corporate decisions. Our work can be viewed as a step toward supporting groups of users making decisions using social choice functions that are automatically optimized for their needs. In these settings, a utilitarian perspective is often called for.

If we could directly access the utilities of agents, the socially desirable alternative could be easily identified. However, such access is often not feasible for a variety of reasons. As a result, we use agent preference orders as a *proxy* for their utility functions; and the social choice function, taking preference orders as input, should perform well with respect to the underlying utilities. From this point of view, a social choice function is *optimal* if it maximizes social welfare given the available information. Using a preference order as proxy for utility in this fashion serves several purposes. First, behavioral economists have argued that people find it difficult to construct utilities for alternatives. Second, the cognitive and communication burden of articulating precise utilities has long been recognized within decision analysis, behavioral economics, and psychology. By contrast, simply comparing and ordering alternatives is considerably easier for most people, which makes soliciting preference orders more practical than eliciting utilities. Furthermore, choice behavior among alternatives can often be interpreted as revealing ordinal (rather than cardinal) preference information, providing ready access to (sometimes incomplete) orders in many of the domains described above. Hence we content ourselves with orders as inputs.

### 1.1. Our results

Our study of optimal social choice functions incorporates three distinct but related models, each with its own assumptions regarding available information and therefore its own notion of optimality. One common thread is that the family of *scoring functions*—social choice functions that score alternatives based only on their position in each agent’s preference order—plays a key role in optimizing social welfare.

In Section 3 we study a model where no information about agents’ utility functions is available when constructing the social choice function. A *worst-case* analysis is thus called for. We believe that the study of this model is of theoretical interest, but it is certainly the least practical of our three models. Specifically, given a collection of agents’ preferences—a *preference profile*—there are many *consistent* collections of utility functions—*utility profiles*—that induce this preference profile in the natural way (by ranking alternatives with higher utility closer to the top). The *distortion* of a social choice function on a preference profile is the worst-case ratio (over feasible utility profiles) of the social welfare of the best alternative to the social welfare of the alternative that is selected by the function. A *worst-case optimal* social choice function minimizes the distortion on every preference profile.

We first derive upper and lower bounds on the least distortion that one can hope for, focusing on *randomized* social choice functions. We show that there exists a preference profile where every randomized social choice function must have distortion at least  $\Omega(\sqrt{m})$ , where  $m$  is the number of alternatives. We complement this result with a randomized social choice function whose distortion on *every* preference profile is  $\mathcal{O}(\sqrt{m} \log^* m)$ . A slightly weaker upper bound is obtained via a randomized variation of a natural scoring function that we call the *harmonic scoring function* (a new canonical scoring function that may be of independent interest). Finally, we establish that the worst-case optimal social choice function (which achieves minimum distortion on every profile) is polynomial-time computable. The proof is based on linear programming, and (roughly speaking) relies on embedding the dual of a sub-problem within a carefully constructed larger LP, in order to avoid quadratic constraints.

In Section 4 we study an *average-case model*, assuming a known distribution  $D$  over utility functions. We assume that the utility function of each agent is drawn independently from  $D$ . Given reported agent preferences, one can compute the expected utility any agent has for an alternative with respect to  $D$ . An *average-case optimal* social choice function selects an alternative that maximizes expected social welfare given the reported profile. We show that when  $D$  is *neutral*, i.e., symmetric with respect to alternatives, the average-case optimal social choice function must be a scoring function. The

<sup>3</sup> Hereinafter, we simply write “social welfare” to refer to “utilitarian social welfare”.

proof leverages Young's [36] characterization of the family of scoring functions. As a corollary, we show that when  $D$  is uniform over an interval, the average-case optimal social choice function is the famous scoring function known as the Borda count.

In Section 5 we develop and analyze a *learning-theoretic model*. Rather than assuming a known distribution  $D$  over utility profiles, we have access only to sampled utility profiles from  $D$ . We use these profiles to compute *sample-optimal* social choice functions. The quality of a sample-optimal function is measured by comparing its expected social welfare to that of the (truly) optimal social choice function for  $D$ . We address two natural questions. First we derive sample complexity results for two classes of social choice functions,  $k$ -approval functions and more general scoring functions; specifically, we derive necessary and sufficient bounds on the number of samples such that the sample-optimal function in this class will have social welfare that is within a small tolerance of the optimal choice function with high probability. Second, we show that computing the sample-optimal scoring function is  $\mathcal{APX}$ -hard, but describe a mixed integer programming formulation of this problem that solves it in practice. Empirical results on a random utility model and a real data set suggest that sample-optimal scoring functions (as well as several more stylized functions, including Borda count) have very low expected distortion.

## 1.2. Perspective and related work

While the utilitarian perspective on social choice—especially the goal of optimizing the (possibly weighted) sum of individual utilities—has been overshadowed by the more axiomatic perspective to a great extent, its foundations are nonetheless firm [15], and it does have its advocates. Our work adopts this utilitarian perspective, and assumes that social welfare is measured using the sum of individual agent utilities in the classic “Benthamite” fashion. Naturally, this position requires making a number of assumptions about the problem domain including: the existence of agent (cardinal) utility functions; the validity of interpersonal comparison of utilities; and having as one's goal the maximization of the sum of individual utilities.

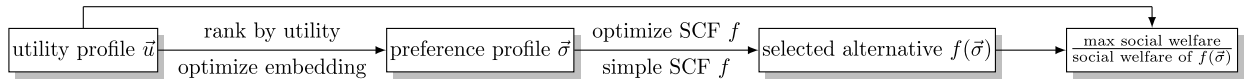
None of these assumptions is valid in all social choice settings. The foundations of von Neumann and Morgenstern [32] expected utility theory treat the strength of preference for alternatives expressed by a utility function as representing an individual's (ordinal) preferences over lotteries or gambles involving those alternatives. While this theory can be operationalized to (roughly) determine an individual's utility function (e.g., using standard gamble queries, as is common in decision analysis), it provides little foundation for a satisfactory account of interpersonal utility comparison. Furthermore, even if one accepts that such interpersonal comparisons are meaningful, many social choice functions and voting schemes studied in the social choice literature cannot, in any sense, be interpreted as maximizing the sum of individual utilities, or as assuming that individual utilities even exist.

Despite this, the three key assumptions above hold (at least approximately) in many settings, including those of interest in computational economics, algorithmic mechanism design, and e-commerce. Most work in mechanism design assumes that agents possess real-valued utility or *valuation* functions over alternatives, and while arbitrary social choice functions may be considered, one of the most common is social welfare maximization, which is, for example, the social choice function implemented by the celebrated VCG mechanism [25], which requires monetary transfers (in order to achieve truthfulness). Other papers also assume cardinal utilities and deal with truthfulness and social welfare maximization but in settings without money. For example, in a recent paper, Azrieli and Kim [4] characterize social choice functions that maximize social welfare subject to truthfulness when there are two alternatives—they are weighted majority rules with appropriate weights. In this light, our work can be viewed as providing the means to approximately maximize social welfare, while reducing the elicitation burden of classic mechanisms by having agents rank alternatives rather than specify valuations.

While many of our results on the optimality of scoring rules in the worst-case and average-case models depend on using the sum of utilities as our social choice function, our learning-theoretic model and corresponding empirical optimization framework could, in principle, be adapted to other measures of social welfare (including the “Rawlsian” maximin and other measures) that take as input the utility functions of a collection of agents. In this sense, our framework does not require a commitment to maximizing the sum of individual utilities.

Some researchers argue that agents should express their preferences by *explicitly* reporting utilities. While very common in decision analysis, this perspective is also sometimes adopted in social choice. For example, *utilitarian voting* [18] (or *range voting*) allows voters to express utilities for alternatives in some predefined range, e.g.,  $\{1, \dots, 10\}$ ,  $\{-1, 0, 1\}$ , or  $\{0, 1\}$  (the last coincides with *approval voting* [8]). While utilitarian in approach, such work differs from ours, as we take the (prevalent) view that human voters are far more comfortable expressing ordinal preferences—we seek to optimize the choice of alternative with respect to *implicit* utility functions.

The worst-case model in Section 3 is closely related to work by Procaccia and Rosenschein [27]. Their work shares the premise that ordinal utilities are a proxy for underlying cardinal utilities. They too argue that a social choice function should maximize social welfare, and introduce the notion of distortion to quantify the gap between optimum social welfare and the total utility of the social choice based on the induced preference orders. The main difference from our approach is that they consider deterministic social choice functions, whereas we focus on randomized functions. Deterministic functions inevitably have trivially high distortion, which Procaccia and Rosenschein mitigate by focusing attention on highly structured utility functions. In contrast, our study provides rich theoretical results under a very mild assumption on utility functions.



**Fig. 1.** A comparison of our worst-case model (Section 3) with Caragiannis and Procaccia [10]. The text above (resp., below) the arrows describes our (resp., their) work. SCF stands for social choice function.

A recent paper by Caragiannis and Procaccia [10] builds on [27], and is also closely related to our worst-case results. Although they also aim to optimize social welfare, their work is fundamentally different on a conceptual level: they consider settings where agents are software agents that can easily compute exact utilities for alternatives, and the need for voting arises because of communication restrictions. Hence they focus on simple, fixed social choice functions with low communication requirements, and optimize the *embedding* by which agents translate their utility functions into reported votes. For example, an agent can map its utility function to a plurality vote—which designates a single, favorite alternative, and therefore requires only  $\log_2 m$  bits—by voting for each alternative with probability proportional to that alternative’s utility. While such embeddings are well-motivated in cooperative multiagent systems, in our setting, agents may be people whose utility functions are translated into preference orders in the natural way; thus we optimize the social choice function. Fig. 1 illustrates the two different optimization processes.

Several papers study models that are similar to our average-case model in Section 4. For example, Pivato [26] shows that when cardinal utilities are drawn from a distribution satisfying certain assumptions, a Condorcet winner exists and maximizes the social welfare with high probability. And early empirical work by Bordley [7] and Merrill [24] shows that plurality achieves lower social welfare than other social choice functions.

The unpublished work of Weber [34] is even more closely related. His motivation is similar to ours, but his model and results differ in several important ways. First, he optimizes a measure different from ours. Second, he restricts attention to (a slight generalization of) the family of scoring functions (whereas we identify *optimal* social choice functions, which just happen to be scoring functions). Third, he assumes that the utility of each agent for each alternative is independently and uniformly distributed on an interval, while our assumptions are less restrictive. Weber’s main result is that Borda count is asymptotically optimal (w.r.t. his measure) among scoring functions. Interestingly, under his more restrictive assumptions we show that Borda count is average-case (exactly) optimal (w.r.t. our measure, expected social welfare of the winner) among *all* social choice functions.

Section 4 is also closely related to independent<sup>4</sup> work by Apesteguia et al. [3], who prove a result that is very similar to our Theorem 4.2. Our result is somewhat stronger, in that we assume that the utility function of each agent is drawn i.i.d. from a *neutral* distribution  $D$ , whereas they assume that each agent’s utility *for each alternative* is drawn i.i.d. Moreover, our proof—which builds on a result by Young [36]—is completely different.

The learning-theoretic model in Section 5 is related to a study of the learnability of social choice functions by Procaccia et al. [28]. They consider the reconstruction of a scoring function based on examples, where an example is a preference profile and a winning alternative for that profile. In contrast, in our learning-theoretic setting we optimize expected social welfare, and examples are *utility profiles*. On a conceptual level, their motivation is fundamentally different; on a technical level, we require new tools, but leverage some of their results to derive novel results in our setting.

Finally, we are seeing increasingly more work in computational social choice that views the social choice problem as an optimization problem [21,12]. One such approach views social choice functions as maximum likelihood estimators [11]. This line of work, dating to the 18th century, was revived by Young [37], who studied “optimal” voting rules, but his notion of optimality is very different from ours. Specifically, the maximum likelihood perspective assumes that agents order alternatives reflecting their personal assessment of the relative *likelihood* that particular alternatives are “objectively best”. Voting is intended to determine the alternative (or ranking) with maximum likelihood of being the best given these assessments, assuming that each agent is more likely to rank any pair of alternatives correctly than incorrectly. Young’s view of optimality (and the MLE perspective more broadly) is thus purely statistical and does not address issues of social welfare or utility maximization.<sup>5</sup>

## 2. Preliminaries

Let  $N = \{1, \dots, n\}$  be a set of *agents* and  $A = \{a_1, \dots, a_m\}$  a set of *alternatives*. Each agent has a *preference order* over  $A$ , which is a strict total order. Letting  $[k] = \{1, \dots, k\}$ , we can equivalently view a preference order as a bijection  $\sigma : A \rightarrow [m]$  mapping each alternative to its rank. Let  $S_m$  be the set of such bijections. The alternative ranked in position  $k$  under ranking  $\sigma$  is given by  $\sigma^{-1}(k)$ .

For each  $i \in N$ , let  $\sigma_i$  be the preference order of agent  $i$ . The vector of agent preferences  $\vec{\sigma} = (\sigma_1, \dots, \sigma_n) \in (S_m)^n$  is a *preference profile*. A *social choice function*  $f : (S_m)^n \rightarrow A$  maps preference profiles to alternatives. We draw special attention to a class of social choice functions known as *scoring functions*. A scoring function is defined by a vector  $\vec{s} = (s_1, \dots, s_m)$ .

<sup>4</sup> The conference version of this paper was published in 2012, shortly after the publication of the paper of Apesteguia et al. [3].

<sup>5</sup> Even in his discussion of compromise among *preference* orderings using Kemeny’s rule, Young appeals to a statistical justification, namely, the *median* relative to Kemeny’s distance metric.

Given preference profile  $\vec{\sigma}$ , the score of  $a \in A$  is  $\sum_{i \in N} s_{\sigma_i(a)}$ , i.e.,  $a$  is awarded  $s_k$  points for each agent who ranks it in position  $k$ . The *scoring function*  $f_{\vec{s}}$  induced by  $\vec{s}$  selects some  $a \in A$  with maximum score with ties broken in some fashion (we revisit tie breaking as it becomes relevant). The well-known *Borda* scoring function (or count) is induced by the vector  $(m-1, m-2, \dots, 0)$ .

Unlike classical social choice models, we assume that agents have *utility functions* over alternatives. As discussed above, however, these are not reported or used by the social choice function. Let  $u : A \rightarrow \mathbb{R}_+$  be a utility function. We say a ranking  $\sigma$  is *consistent* with  $u$  if  $u(a) > u(a')$  implies  $\sigma(a) < \sigma(a')$ ; i.e., alternatives with higher utility must be ranked higher than those with lower utility.

Let  $p(u)$  be the set of rankings consistent with (or induced by)  $u$ ;  $p(u)$  is a set to account for ties in utility. We occasionally presume agents use some (randomized) method for selecting a specific ranking  $\sigma \in p(u)$  when they possess utility function  $u$ ; in such a case, we use  $\sigma(u)$  to denote the corresponding random variable (with domain  $p(u)$ ). Abusing notation slightly, let  $p^{-1}(\sigma)$  be the set of utility functions  $u$  such that  $\sigma \in p(u)$ , i.e., the set of utility functions *consistent* with  $\sigma$ . The vector  $\vec{u} = (u_1, \dots, u_n)$  of agent utility functions is a *utility profile*. Let  $p(\vec{u}) = p(u_1) \times \dots \times p(u_n)$  be the set of preference profiles consistent with  $\vec{u}$ . Similarly, let  $\vec{\sigma}(\vec{u})$  denote the random variable over  $p(\vec{u})$  representing the (joint) choice of rankings, and  $p^{-1}(\vec{\sigma})$  denote the set of utility profiles consistent with preference profile  $\vec{\sigma}$ .

Positing a utility model allows one to quantify the social welfare of an alternative. For utility profile  $\vec{u}$ , let  $\text{sw}(a, \vec{u}) = \sum_{i \in N} u_i(a)$  be the (utilitarian) *social welfare* of  $a$ .

### 3. The worst-case model

We begin our study of optimal social choice functions with a *worst-case model*. A social choice function has access only to a preference profile, but this preference profile is induced by some unknown utility profile. To quantify the *quality* of a social choice function, we use the notion of *distortion* [27,10], which reflects the degree to which the social choice can become distorted when cardinal preferences are mapped to ordinal preferences. More precisely, the distortion of social choice function  $f$  on a preference profile  $\vec{\sigma}$  is given by

$$\text{dist}(f, \vec{\sigma}) = \sup_{\vec{u} \in p^{-1}(\vec{\sigma})} \frac{\max_{a \in A} \text{sw}(a, \vec{u})}{\text{sw}(f(\vec{\sigma}), \vec{u})}.$$

In other words, distortion is the worst-case ratio (over consistent utility profiles) of the social welfare of the optimal alternative to that of the alternative selected by  $f$ .

As observed by Procaccia and Rosenschein [27], deterministic social choice functions must have high distortion. For example, consider a preference profile where  $n/2$  agents rank  $a$  first, and  $n/2$  agents rank  $b$  first. Assume (w.l.o.g.) a social choice function selects  $a$ . Suppose the agents that rank  $b$  first have utility 1 for  $b$  and 0 for other alternatives, while agents that rank  $a$  first have utility  $1/m$  for all alternatives. The ratio between the social welfare of  $b$  and  $a$  is  $\Omega(m)$ . To reduce potential distortion, Procaccia and Rosenschein [27] adopt an extremely restrictive assumption on utility functions (specifically, that utilities are Borda scores). We instead turn to randomization.

We consider *randomized social choice functions*  $f : (S_m)^n \rightarrow \Delta(A)$ , in which  $f(\vec{\sigma})$  is a distribution (or random variable) over  $A$ . We extend the definition of distortion to randomized functions in the natural way:

$$\text{dist}(f, \vec{\sigma}) = \sup_{\vec{u} \in p^{-1}(\vec{\sigma})} \frac{\max_{a \in A} \text{sw}(a, \vec{u})}{\mathbb{E}[\text{sw}(f(\vec{\sigma}), \vec{u})]}.$$

In general, even randomized social choice functions cannot achieve a distortion lower than  $\Omega(m)$ . Consider a preference profile where each  $a \in A$  is ranked first at least once. Given a randomized social choice function, there is some alternative  $a^* \in A$  that is selected with probability at most  $1/m$  given this preference profile. However, this profile is induced by the utility profile where one agent gives arbitrarily high utility to  $a^*$ , and all other utilities are arbitrarily low. The ratio between the social welfare of  $a^*$  and the function's expected social welfare would therefore be  $\Omega(m)$ .

To avoid this, we make the following assumption in this section:

**Assumption 3.1.** (Only in Section 3.) For each agent  $i \in N$ ,  $\sum_{a \in A} u_i(a) = 1$ .

This ensures that agents have equal “weights,” or equal pools of “utility points” to distribute among the alternatives. Otherwise, if, say, agent 1 has utility 1 for  $a$  and 0 for the rest, and agent 2 has utility  $1/2$  for  $b$  and 0 for the rest, then agent 1 has twice as much influence as agent 2 in determining the socially optimal alternative. The same assumption is also made by Caragiannis and Procaccia [10].

Our first result establishes a lower bound on the distortion of randomized social choice functions under [Assumption 3.1](#) (which is almost tight, see below).

**Theorem 3.2.** Assume that  $n \geq \sqrt{m}$ . Then there exists a  $\vec{\sigma} \in (S_m)^n$  such that for any randomized social choice function  $f$ ,  $\text{dist}(f, \vec{\sigma}) = \Omega(\sqrt{m})$ .

**Proof.** For ease of exposition assume that  $\sqrt{m}$  divides  $n$ . Partition the agents into  $\sqrt{m}$  equal subsets  $N_1, \dots, N_{\sqrt{m}}$ . Consider the preference profile  $\vec{\sigma}$  where  $\sigma_i(a_k) = 1$ , for all  $i \in N_k$ , and the remaining alternatives are ranked arbitrarily.

For any randomized  $f$  there must be a  $k^* \in \{1, \dots, \sqrt{m}\}$  such that  $\Pr[f(\vec{\sigma}) = a_{k^*}] \leq 1/\sqrt{m}$ . Let  $\vec{u}$  be a utility profile such that for all  $i \in N_{k^*}$ ,  $v_i(a_{k^*}) = 1$  and  $u_i(a) = 0$  for all  $a \in A \setminus \{a_{k^*}\}$ . For all  $i \notin N_{k^*}$  and  $a \in A$ ,  $v_i(a) = 1/m$ . It holds that

$$\frac{n}{\sqrt{m}} \leq \text{sw}(a_{k^*}, \vec{u}) \leq \frac{2n}{\sqrt{m}},$$

and for all  $a \in A \setminus \{a_{k^*}\}$ ,  $\text{sw}(a, \vec{u}) \leq n/m$ . Therefore:

$$\text{dist}(f, \vec{\sigma}) \geq \frac{\frac{n}{\sqrt{m}}}{\frac{1}{\sqrt{m}} \cdot \frac{2n}{\sqrt{m}} + \frac{\sqrt{m}-1}{\sqrt{m}} \cdot \frac{n}{m}} \geq \frac{\sqrt{m}}{3}. \quad \square$$

We next establish the existence of a randomized social choice function that nearly achieves this lower bound on every preference profile, leaving a tiny gap of only  $\log^* m$  (iterated logarithm of  $m$ ).

**Theorem 3.3.** *There exists a randomized social choice function  $f$  such that for every  $\vec{\sigma} \in (\mathcal{S}_m)^n$ ,  $\text{dist}(f, \vec{\sigma}) = \mathcal{O}(\sqrt{m} \cdot \log^* m)$ .*

The rather intricate proof of this theorem is provided in [Appendix A](#). Here we present a much simpler proof of a weaker upper bound of  $\mathcal{O}(\sqrt{m \log m})$ . This latter proof uses the novel *harmonic scoring function*, given by score vector  $(h_1, \dots, h_m)$ , where  $h_k = 1/k$ .

**Proof of weaker upper bound of  $\mathcal{O}(\sqrt{m \log m})$ .** Let  $\text{sc}(a, \vec{\sigma})$  be the score of  $a$  under  $\vec{\sigma}$  using the harmonic scoring function. It holds that  $u_i(a) \leq 1/\sigma_i(a)$  for all  $i \in N$  and  $a \in A$ , because if  $i \in N$  ranks  $a \in A$  in position  $k$  and gives it utility  $u_i(a)$ , each of the  $k-1$  alternatives ranked above  $a$  must have utility at least  $u_i(a)$ , but the sum of utilities is one. Therefore, for any  $\vec{u} \in p^{-1}(\vec{\sigma})$  and any  $a$ ,

$$\text{sw}(a, \vec{u}) \leq \text{sc}(a, \vec{\sigma}). \quad (1)$$

In addition, note that for any  $\vec{\sigma}$ ,

$$\sum_{a \in A} \text{sc}(a, \vec{\sigma}) = n \cdot \sum_{k=1}^m \frac{1}{k} \leq n(\ln m + 1). \quad (2)$$

Consider the randomized  $f$  that chooses one of the following two schemes (each with probability 1/2): (i) select an alternative uniformly at random, and (ii) select an alternative with probability  $\text{sc}(a, \vec{\sigma}) / (\sum_{a' \in A} \text{sc}(a', \vec{\sigma}))$  (i.e., proportional to  $\text{sc}(a, \vec{\sigma})$ ). Let  $\vec{\sigma} \in (\mathcal{S}_m)^n$ ,  $\vec{u} \in p^{-1}(\vec{\sigma})$ , and  $a \in A$ . It is sufficient to show that

$$\frac{\text{sw}(a, \vec{u})}{\mathbb{E}[\text{sw}(f(\vec{\sigma}), \vec{u})]} \leq 2\sqrt{m(\ln m + 1)}.$$

We consider two cases. First, assume that  $\text{sc}(a, \vec{\sigma}) \geq n\sqrt{(\ln m + 1)/m}$ . With probability 1/2, a winner is selected proportionally to its score. Using Eq. (2), the probability that  $a$  is selected is at least

$$\frac{1}{2} \cdot \frac{n \cdot \sqrt{\frac{\ln m + 1}{m}}}{n(\ln m + 1)} = \frac{1}{2\sqrt{m(\ln m + 1)}}.$$

It follows that

$$\mathbb{E}[\text{sw}(f(\vec{\sigma}), \vec{u})] \geq \Pr[f(\vec{\sigma}) = a] \cdot \text{sw}(a, \vec{u}) \geq \frac{1}{2\sqrt{m(\ln m + 1)}} \cdot \text{sw}(a, \vec{u}).$$

Second, assume that  $\text{sc}(a, \vec{\sigma}) < n\sqrt{(\ln m + 1)/m}$ . From Eq. (1) it follows that  $\text{sw}(a, \vec{u}) < n\sqrt{(\ln m + 1)/m}$ . With probability 1/2, a winner is selected uniformly at random. We have that

$$\mathbb{E}[\text{sw}(f(\vec{\sigma}), \vec{u}) \mid \text{uniform selection}] = \frac{\sum_{i \in N} \sum_{a \in A} u_i(a)}{m} = \frac{n}{m},$$

and therefore  $\mathbb{E}[\text{sw}(f(\vec{\sigma}), \vec{u})] \geq n/(2m)$ . We conclude that

$$\frac{\text{sw}(a, \vec{u})}{\mathbb{E}[\text{sw}(f(\vec{\sigma}), \vec{u})]} \leq \frac{n \cdot \sqrt{\frac{\ln m + 1}{m}}}{\frac{n}{2m}} = 2\sqrt{m(\ln m + 1)}. \quad \square$$



An interesting aspect of this proof is its use of the harmonic scoring function. Despite a large body of (especially computational) work on scoring functions (see, e.g., [16,35,28]), only three scoring functions are considered canonical: *Borda count*; *plurality*, defined by vector  $(1, 0, \dots, 0)$ ; and *veto* (or *anti-plurality*), defined by vector  $(1, \dots, 1, 0)$ . We hope that the harmonic function, with natural parameters and attractive theoretical properties, may in time be accepted into this exclusive club.

While Theorem 3.3 offers attractive theoretical guarantees, its randomized social choice function need not be *optimal*. While there are preference profiles where distortion must be at least  $\Omega(\sqrt{m})$ , there may be many profiles where low distortion is achievable but this function nevertheless yields relatively high distortion. We are thus most interested in *worst-case optimal* (randomized) social choice functions. By this, we simply mean that for every  $\vec{\sigma} \in (\mathcal{S}_m)^n$ , the function  $f$  has minimum possible distortion on  $\vec{\sigma}$ . We can show that such a social choice function is polynomial-time computable via linear programming duality.

**Theorem 3.4.** *The worst-case optimal randomized social choice function is polynomial-time computable.*

**Proof.** For an alternative  $a^* \in A$  and  $\vec{\sigma} \in (\mathcal{S}_m)^n$ , let

$$U(\vec{\sigma}, a^*) = \left\{ \vec{u} \in p^{-1}(\vec{\sigma}) : a^* \in \operatorname{argmax}_{a \in A} \{ \text{sw}(a, \vec{u}) \} \right\}.$$

Given a preference profile  $\vec{\sigma} \in (\mathcal{S}_m)^n$ , an alternative  $a^* \in A$ , a vector of non-negative values  $\{p_a\}_{a \in A}$ , and non-negative  $\beta$ , define the set of inequalities  $\text{INEQ}(\{p_a\}_{a \in A}, \beta, \vec{\sigma}, a^*)$  as follows:

$$\begin{aligned} \sum_{j \in N} y(j, a^*) &\geq 0 \\ \beta + \sum_{a \in A} x(a, a^*) &\leq 0 \\ \forall j \in N, \forall a \in A : \sigma_j(a) = 1, & p_a + x(a, a^*) - y(j, a^*) - z(\sigma_j(a), j, a^*) \geq 0 \\ \forall j \in N, \forall a \in A : 2 \leq \sigma_j(a) \leq m-1, & p_a + x(a, a^*) - y(j, a^*) - z(\sigma_j(a), j, a^*) + z(\sigma_j(a) - 1, j, a^*) \geq 0 \\ \forall j \in N, \forall a \in A : \sigma_j(a) = m, & p_a + x(a, a^*) - y(j, a^*) + z(\sigma_j(a) - 1, j, a^*) \geq 0 \\ \forall a \in A \setminus \{a^*\}, & x(a, a^*) \geq 0 \\ x(a^*, a^*) &\in (-\infty, +\infty) \\ \forall j \in N, & y(j, a^*) \in (-\infty, +\infty) \\ \forall k \in [m-1], j \in N, & z(k, j, a^*) \geq 0 \end{aligned} \tag{3}$$

The connection of this set of inequalities to randomized social choice functions is revealed in the following statement.

**Lemma 3.5.** *A randomized social choice function  $f$  that uses probability distribution  $\{p_a = \Pr[f(\vec{\sigma}) = a]\}_{a \in A}$  for profile  $\vec{\sigma}$  has distortion at most  $\beta^{-1}$  when  $\beta \in (0, 1]$  with respect to utility profiles in  $U(\vec{\sigma}, a^*)$  if and only if the set of inequalities  $\text{INEQ}(\{p_a(\vec{\sigma})\}_{a \in A}, \beta, \vec{\sigma}, a^*)$  is satisfied.*

**Proof.** The fact that the distortion is at most  $\beta^{-1}$  is equivalent to the following linear program having a non-negative objective value.

$$\begin{aligned} \text{minimize} \quad & \sum_{j \in N} \sum_{a \in A} p_a u_j(a) - \beta q \\ \text{subject to} \quad & \sum_{j \in N} u_j(a^*) = q \\ & \forall a \in A \setminus \{a^*\}, \sum_{j \in N} u_j(a) \leq q \\ & \forall j \in N, \sum_{a \in A} u_j(a) = 1 \\ & \forall j \in N, k \in [m-1], u_j(\sigma_j^{-1}(k)) \geq u_j(\sigma_j^{-1}(k+1)) \\ & \forall j \in N, a \in A, u_j(a) \geq 0 \\ & q \geq 0 \end{aligned}$$

In the above LP, the variable  $u_j(a)$  denotes the utility of agent  $j \in N$  for alternative  $a \in A$  while the variable  $q$  is the maximum social welfare among all alternatives. The constraints guarantee that  $\vec{u} \in U(\vec{\sigma}, a^*)$ .

Next, consider the dual linear program using the variable  $x(a^*, a^*)$  for the first constraint, the variables  $x(a, a^*)$  for the second set of constraints, the variables  $y(j, a^*)$  for the third set of constraints, and the variables  $z(k, j, a^*)$  for the fourth set of constraints. This program maximizes  $\sum_{j \in N} y(j, a^*)$ , subject to the constraints in  $\text{INEQ}(\{p_a\}_{a \in A}, \beta, \vec{\sigma}, a^*)$  except Equation (3). By LP duality, the maximum objective value of the dual LP should be non-negative as well. Observe that this is exactly the requirement that  $\text{INEQ}(\{p_a(\vec{\sigma})\}_{a \in A}, \beta, \vec{\sigma}, a^*)$  is satisfied.  $\square$

We are now ready to define the worst-case optimal randomized social choice function  $f^*$ , i.e., the one that achieves optimal distortion  $\beta(\vec{\sigma})^{-1}$  with respect to every profile  $\vec{\sigma} \in (\mathcal{S}_m)^n$ . To do so, we consider the probability distribution  $\{p_a(\vec{\sigma}) = \Pr[f^*(\vec{\sigma}) = a]\}_{a \in A}$  as a set of variables and  $\beta(\vec{\sigma})$  as our objective that has to be maximized. We employ an LP that is defined as follows.

$$\begin{aligned} & \text{maximize} && \beta(\vec{\sigma}) \\ & \text{subject to} && \forall a^* \in A, \text{INEQ}(\{p_a(\vec{\sigma})\}_{a \in A}, \beta(\vec{\sigma}), \vec{\sigma}, a^*) \text{ is satisfied} \\ & && \sum_{a \in A} p_a(\vec{\sigma}) = 1 \\ & && \forall a \in A, p_a(\vec{\sigma}) \geq 0 \end{aligned}$$

Using Lemma 3.5 we find that the probability distribution obtained as (part of) the solution to this LP induces the lowest possible distortion  $\beta(\vec{\sigma})^{-1}$  with respect to  $\vec{\sigma}$ .

We conclude that in order to compute the worst-case optimal social choice function  $f^*$  with respect to any preference profile, it suffices to solve a linear program with  $\mathcal{O}(nm^2)$  variables and  $\mathcal{O}(nm^2)$  constraints. We observe that leveraging LP duality is crucial to the proof, as naively embedding the primal LP presented in Lemma 3.5 into the above LP would result in a quadratic program.  $\square$

Interestingly, even though we can concisely describe the optimal function, we do not know whether its distortion on every profile is at most  $\mathcal{O}(\sqrt{m})$ . Of course, by Theorem 3.3, we do know that its distortion on any profile can only be slightly larger: at most  $\mathcal{O}(\sqrt{m} \log^* m)$ .

#### 4. The average-case model

We now consider a model in which agent utility functions are drawn from a probability distribution  $D$ . As discussed in Section 1, this model is closely related to independent work by Apestegui et al. [3]. We *do not* assume that utilities are normalized (as in Section 3), but we do assume (in this section only) that each agent's utility function is drawn independently from the same distribution.

**Assumption 4.1.** (Only in Section 4.) Agent utility functions  $u_1, \dots, u_n$  are drawn i.i.d. from  $D$ .

This assumption, while admittedly restrictive, enables us to prove strong results; it will not be used when we move to a more general learning-theoretic model in Section 5.

This model gives rise to the product distribution  $D^n$  over utility profiles. As above, utility profiles induce preference profiles in the natural way, but since we will need to reason about the induced distribution over preference profiles, we make the specific, but mild, assumption that ties in utility are broken uniformly at random; that is, if  $u(a) = u(b)$  then  $\Pr[(\sigma(u))(a) < (\sigma(u))(b)] = \Pr[(\sigma(u))(b) < (\sigma(u))(a)] = 1/2$ . This assumption is essentially without loss of generality under non-atomic distributions (since ties occur with probability zero).

The notion of optimality takes a slightly different meaning in this setting: instead of maximizing the ratio to the optimal social welfare, a social choice function should perform as well as possible on average. We say that a social choice function  $f$  is *average-case optimal* if for every preference profile  $\vec{\sigma}$  it maximizes expected social welfare  $\mathbb{E}[\text{sw}(f(\vec{\sigma}(\vec{u})), \vec{u}) \mid \vec{\sigma}(\vec{u})]$ . Note that expectation is conditional on the reported preference profile  $\vec{\sigma}(\vec{u})$ .

In this section, we consider distributions  $D$  that possess a special structure. Distribution  $D$  is *neutral*<sup>6</sup> if for any measurable  $U \subseteq \mathbb{R}_+^m$  and any permutation  $\pi \in \mathcal{S}_m$ , we have  $D(U) = D(U \circ \pi)$ , where  $U \circ \pi = \{u \circ \pi : u \in U\}$  (here  $u \circ \pi$  denotes a permutation of utility function  $u$ ). Informally, a neutral distribution is symmetric with respect to alternatives. A neutral distribution induces a distribution over preference profiles where each agent draws a ranking  $\sigma$  independently and uniformly at random; this is exactly the *impartial culture assumption*, a model that plays an important role in social choice theory [31, 30]. We now show that scoring functions play a crucial role in the average-case model, underscoring even more deeply the importance of this family in the study of optimal social choice functions.

**Theorem 4.2.** Assume a neutral distribution  $D$  over utility functions. Then the average-case optimal social choice function is a scoring function.

Apestegui et al. [3] independently prove a similar theorem, but make the stronger assumption that each agent's utility for each alternative is drawn i.i.d. Importantly, they use a direct proof, whereas we provide a completely different, potentially more broadly useful proof by exploiting machinery developed by Young [36].

To this end, we define a *social choice correspondence* as a function from preference profiles to nonempty subsets of  $A$ . A scoring correspondence is defined by a vector  $\vec{s}$  as before, but selects *all* alternatives with maximum score. An anonymous

<sup>6</sup> Neutrality is also known as *exchangeability* in this context.



social choice correspondence operates on *anonymous preference profiles*, i.e., vectors  $\vec{x} \in \mathbb{N}^m!$  that count the number of agents holding each of the  $m!$  possible rankings of  $A$  in the preference profile (i.e., without regard for *which* agent holds what preference). Let  $x_\sigma$  denote the number of agents holding ranking  $\sigma$  in  $\vec{x}$ .

An anonymous social choice correspondence  $f$  is: *consistent* if  $f(\vec{x} + \vec{y}) = f(\vec{x}) \cap f(\vec{y})$  when  $f(\vec{x}) \cap f(\vec{y}) \neq \emptyset$ ; *continuous* if whenever  $f(\vec{x}) = \{a\}$  then for any anonymous profile  $\vec{y}$  there is  $T \in \mathbb{N}$  such that  $f(\vec{y} + t\vec{x}) = \{a\}$  for every  $t \geq T$ ; and *neutral* if  $f \circ \sigma = \sigma \circ f$  for every  $\sigma \in S_m$ . Denote by  $p^*(\vec{u})$  the set of anonymous preference profiles consistent with  $\vec{u}$ .

**Lemma 4.3.** (See Young [36].) *An anonymous social choice correspondence is a scoring correspondence if and only if it is neutral, consistent, and continuous.*

**Proof of Theorem 4.2.** An optimal social choice function is clearly anonymous and neutral because agent utilities are i.i.d. and  $D$  is neutral. Thus, we restrict our attention to functions that receive anonymous preference profiles as input.

Let  $f^*$  be the social choice correspondence that, given an anonymous preference profile  $\vec{x}$ , returns all  $a \in A$  that maximize  $\mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{x} \in p^*(\vec{u})]$ , i.e.,

$$f^*(\vec{x}) = \text{argmax}_{a \in A} \mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{x} \in p^*(\vec{u})].$$

It is sufficient to show that  $f^*$  is a scoring correspondence. Indeed, if  $f^*$  is a scoring correspondence then any choice from  $f^*$  (i.e., a choice from  $f^*(\vec{\sigma})$  for every preference profile  $\vec{\sigma}$ ) is a scoring function. Moreover, the set of choices from  $f^*$  is exactly the set of optimal choice functions.

To show that  $f^*$  is a scoring correspondence, it suffices, by Lemma 4.3, to demonstrate that  $f^*$  is consistent and continuous. To see that  $f^*$  is consistent, let  $\vec{x}$  and  $\vec{y}$  be two anonymous profiles such that  $f^*(\vec{x}) \cap f^*(\vec{y}) \neq \emptyset$ , and let  $a, a' \in A$  such that  $a \in f^*(\vec{x}) \cap f^*(\vec{y})$  and  $a' \notin f^*(\vec{x}) \cap f^*(\vec{y})$ . Then

$$\mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{x} \in p^*(\vec{u})] \geq \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{x} \in p^*(\vec{u})]$$

and

$$\mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{y} \in p^*(\vec{u})] \geq \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{y} \in p^*(\vec{u})],$$

where one of the inequalities is strict. Moreover, for any  $b \in A$ ,

$$\begin{aligned} \mathbb{E}[\text{sw}(b, \vec{u}) \mid \vec{x} + \vec{y} \in p^*(\vec{u})] &= \sum_{\sigma \in S_m} (x_\sigma + y_\sigma) \cdot \mathbb{E}[u(b) \mid \sigma \in p(u)] \\ &= \sum_{\sigma \in S_m} x_\sigma \cdot \mathbb{E}[u(b) \mid \sigma \in p(u)] + \sum_{\sigma \in S_m} y_\sigma \cdot \mathbb{E}[u(b) \mid \sigma \in p(u)] \\ &= \mathbb{E}[\text{sw}(b, \vec{u}) \mid \vec{x} \in p^*(\vec{u})] + \mathbb{E}[\text{sw}(b, \vec{u}) \mid \vec{y} \in p^*(\vec{u})]. \end{aligned}$$

This last equality simply says that the overall expected social welfare of an alternative with respect to two separate electorates—whose preferences are represented as  $\vec{x}$  and  $\vec{y}$  on the right hand side—is equal to the expected social welfare of that alternative when the two electorates are put together. We conclude that

$$\mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{x} + \vec{y} \in p^*(\vec{u})] > \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{x} + \vec{y} \in p^*(\vec{u})].$$

This shows that  $a \in f^*(\vec{x} + \vec{y})$  and  $a' \notin f^*(\vec{x} + \vec{y})$ , proving that  $f(\vec{x} + \vec{y}) = f(\vec{x}) \cap f(\vec{y})$ .

To prove continuity, assume  $f^*(\vec{x}) = \{a\}$ . Then there exists an  $\epsilon > 0$  such that

$$\mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{x} \in p^*(\vec{u})] - \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{x} \in p^*(\vec{u})] \geq \epsilon$$

for every  $a' \in A \setminus \{a\}$ . Let  $\vec{y}$  and let  $T > (\mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{y} \in p^*(\vec{u})])/\epsilon$  for every  $a' \in A$ . Then for every  $t \geq T$  and every  $a' \in A \setminus \{a\}$ ,

$$\begin{aligned} &\mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{y} + t \cdot \vec{x} \in p^*(\vec{u})] - \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{y} + t \cdot \vec{x} \in p^*(\vec{u})] \\ &= \mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{y} \in p^*(\vec{u})] - \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{y} \in p^*(\vec{u})] \\ &\quad + t \cdot (\mathbb{E}[\text{sw}(a, \vec{u}) \mid \vec{x} \in p^*(\vec{u})] - \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{x} \in p^*(\vec{u})]) \\ &\geq T \cdot \epsilon - \mathbb{E}[\text{sw}(a', \vec{u}) \mid \vec{y} \in p^*(\vec{u})] > 0 \end{aligned}$$

It follows that  $f^*(\vec{y} + t\vec{x}) = \{a\}$  for every  $t \geq T$ , and therefore continuity is satisfied.  $\square$

The proof implies that the optimal social choice function scores alternatives based only on their position in each agent's preference order. This observation allows us to construct the optimal scoring function given the distribution  $D$ .

**Corollary 4.4.** Assume a neutral distribution  $D$  over utility functions. For each  $k = 1, \dots, m$ , let  $s_k^* = \mathbb{E}[u(a) \mid (\sigma(u))(a) = k]$  for some arbitrary  $a \in A$ . Then the average-case optimal social choice function is a scoring function with parameters  $(s_1^*, \dots, s_m^*)$ .

The optimal scoring function has an especially natural form: the score  $s_k^*$  for position  $k$  is simply the expected utility of any alternative  $a$  for any agent conditional on  $a$  being ranked  $k$ th by that agent. (Notice that the arbitrary choice of  $a$  is justified by the neutrality of  $D$ .) We now consider the special case where agent utilities for each alternative are drawn uniformly from some interval (w.l.o.g., take this to be  $[0, 1]$ ).

**Corollary 4.5.** Let  $D$  be the uniform distribution over  $[0, 1]^m$  (i.e., the utility for each alternative is drawn independently and uniformly from  $[0, 1]$ ). Then the average-case optimal social choice function is the Borda count.

**Proof.** It suffices to compute  $s_k^*$  for  $k = 1, \dots, m$ . A folk theorem about the expectation of  $k$ -order statistics immediately implies that  $s_k^* = (m + 1 - k)/(m + 1)$ ; we provide an informal proof for completeness. Consider the random variables  $X_1, \dots, X_m$ , where  $X_k$  is the utility of the alternative ranked in position  $k$ . The lengths of the  $m + 1$  intervals  $[0, X_m], [X_m, X_{m-1}], \dots, [X_1, 1]$  are identically distributed (to see this, choose  $m + 1$  points on a circle uniformly at random—their distances are identically distributed—and then cut the circle at the first point, which becomes both 0 and 1), and the sum of their lengths is 1. Thus the expected length of each interval is  $1/(m + 1)$ .

Now, clearly the scoring functions defined by the vectors  $\vec{s}$  and  $c \cdot \vec{s}$ , or  $\vec{s}$  and  $\vec{s} + (c, \dots, c)$ , are identical (up to tie breaking). The optimal scoring function defined by the vector  $(m/(m + 1), \dots, 1/(m + 1))$  is therefore equivalent to the Borda count.  $\square$

## 5. The learning-theoretic model

We now consider a *learning-theoretic model* for computing optimal social choice functions that is likely to have the greatest practical impact of our three models. Similarly to the average case model in the previous section, we assume some (possibly unknown) distribution  $D$  over *utility profiles* (rather than utility functions, as in Section 4). However, strong analytical results were made possible in the average case model only by accepting strong assumptions about the distribution, essentially equivalent to the impartial culture assumption. This model is unrealistic for a variety of reasons (e.g., see critiques by Regenwetter et al. [29]).

Instead, we devise techniques to compute approximately optimal social choice functions—specifically, *optimal scoring functions*—for arbitrary distributions  $D$  over utility profiles, without assuming a specific parameterized or stylized form, or independence of agent preferences. Most realistic distributions are likely to be analytically intractable, so we develop a *sample-based optimization framework* for this purpose. We assume access only to a set of sampled profiles from  $D$ —or the ability to generate such samples from a known distribution. With sufficiently many samples, the optimal scoring function with respect to these samples will be approximately optimal for  $D$ .

Because we rely only on samples from  $D$ , the model can be interpreted as *learning* an optimal social choice function. We first address the question of sample complexity by deriving bounds on the number of samples needed to compute approximately optimal scoring functions (as well as the more restricted class of  $k$ -approval functions). We then consider the problem of computing an optimal scoring function for a given sample set. We show that this problem (i.e., computing scores to optimize social welfare) is  $\mathcal{APX}$ -hard, but develop a mixed integer program (MIP) for its optimization. While we discuss the model in learning-theoretic terms, we emphasize that the approach is equally valid when  $D$  is known: sample-based optimization offers a viable and very general computational model in this case.<sup>7</sup>

### 5.1. Requisite concepts

To quantify sample complexity, we rely on two well-known measures of the complexity of a class of functions. Let  $\mathcal{F}$  be some class of functions of the form  $f : \mathcal{X} \rightarrow A$  for some set  $A$ . We say a sample  $x_1, \dots, x_d \in \mathcal{X}$  is *shattered* by  $\mathcal{F}$  if there exist  $f, g \in \mathcal{F}$  such that  $f(x_i) \neq g(x_i)$  for each  $i \leq d$ , and for every boolean vector  $(b_1, \dots, b_d) \in \{0, 1\}^d$  there is an  $h \in \mathcal{F}$  such that  $h(x_i) = f(x_i)$  if  $b_i = 1$  and  $h(x_i) = g(x_i)$  if  $b_i = 0$ . The *generalized dimension*  $D_G(\mathcal{F})$  of  $\mathcal{F}$  is the maximum  $d$  such that some sample  $x_1, \dots, x_d \in \mathcal{X}$  is shattered by  $\mathcal{F}$ .<sup>8</sup> The *pseudo-dimension* is a variation of this concept defined for real-valued functions. If  $\mathcal{F}$  is a class of functions of the form  $f : \mathcal{X} \rightarrow \mathbb{R}$ , the pseudo-dimension  $D_P(\mathcal{F})$  of  $\mathcal{F}$  is the maximum  $d$  such that there are some  $x_1, \dots, x_d \in \mathcal{X}$  and thresholds  $t_1, \dots, t_d \in \mathbb{R}$  such that, for every  $(b_1, \dots, b_d) \in \{0, 1\}^d$ , there exists an  $h \in \mathcal{F}$  such that  $h(x_i) \geq t_i$  if  $b_i = 1$  and  $h(x_i) < t_i$  if  $b_i = 0$ .

We will use bounds on the pseudo-dimension to derive bounds on the sample complexity. We first observe:

**Observation 5.1.** For any finite function class  $\mathcal{F}$ , its (generalized or pseudo-) dimension is no greater than  $\log_2 |\mathcal{F}|$ .

<sup>7</sup> Our sample complexity results make no distributional assumptions. If sampling a *known* distribution  $D$  for computational reasons, much tighter distribution-dependent sample size results should be possible.

<sup>8</sup> The generalized dimension is a generalization of the VC dimension; the latter notion is only defined for Boolean-valued functions.

Let  $\mathcal{F}^{n,m}$  be some class of randomized social choice functions over  $n$  agents and  $m$  alternatives. For any  $f \in \mathcal{F}^{n,m}$ , we can adopt the usual perspective, where  $f : (\mathcal{S}_m)^n \rightarrow \Delta(A)$  maps preference profiles into distributions over alternatives—in this case, we focus on the *generalized dimension* of  $\mathcal{F}^{n,m}$ , by which we refer to the generalized dimension of the correspondence defined by mapping  $\vec{u}$  to the support of  $f(\vec{u})$ . We can take a different perspective by transforming  $f$  as follows: define  $\hat{f}(\vec{u}) = \mathbb{E}[\text{sw}(f(\vec{\sigma}(\vec{u})), \vec{u})]$ , where  $\hat{f}$  maps a utility profile  $\vec{u}$  into the expected social welfare realized by applying  $f$  to the preference profile  $\vec{\sigma}(\vec{u})$  induced by  $\vec{u}$ . Define  $\hat{\mathcal{F}}^{n,m} = \{\hat{f} : f \in \mathcal{F}^{n,m}\}$ . With this view, we focus on the pseudo-dimension of  $\hat{\mathcal{F}}^{n,m}$ . These are not unrelated:

**Lemma 5.2.** For any set of randomized social choice functions  $\mathcal{F}^{n,m}$ ,  $D_G(\mathcal{F}^{n,m}) \leq D_P(\hat{\mathcal{F}}^{n,m})$ .

**Proof.** Assume  $D_G(\mathcal{F}^{n,m}) = d$ . Let  $\vec{\sigma}^1, \dots, \vec{\sigma}^d$  be  $d$  preference profiles that are shattered by  $\mathcal{F}^{n,m}$ , and let  $f, g \in \mathcal{F}^{n,m}$  be two social choice functions that differ on each of these profiles. Define utility profiles  $\vec{u}^1, \dots, \vec{u}^d$  that induce  $\vec{\sigma}^1, \dots, \vec{\sigma}^d$  such that  $\mathbb{E}[\text{sw}(f(\vec{\sigma}^i), \vec{u}^i)] \neq \mathbb{E}[\text{sw}(g(\vec{\sigma}^i), \vec{u}^i)]$  for all  $i$ ; this is always possible using some small perturbation in the utility of a relevant candidate. For each  $i \leq d$ , let  $t_i = (\mathbb{E}[\text{sw}(f(\vec{\sigma}^i), \vec{u}^i)] + \mathbb{E}[\text{sw}(g(\vec{\sigma}^i), \vec{u}^i)])/2$ .

Given any  $(b'_1, \dots, b'_d) \in \{0, 1\}^d$ , we need to show that there is some  $\hat{h} \in \hat{\mathcal{F}}^{n,m}$  such that  $\hat{h}(\vec{u}_i) \geq t_i$  if  $b'_i = 1$  and  $\hat{h}(\vec{u}_i) < t_i$  if  $b'_i = 0$ . Let  $(b_1, \dots, b_d)$  be a binary vector which we set as follows: if  $\mathbb{E}[\text{sw}(f(\vec{\sigma}^i), \vec{u}^i)] \geq t_i$  and  $\mathbb{E}[\text{sw}(g(\vec{\sigma}^i), \vec{u}^i)] < t_i$ , set  $b_i = b'_i$ ; otherwise if  $\mathbb{E}[\text{sw}(f(\vec{\sigma}^i), \vec{u}^i)] < t_i$  and  $\mathbb{E}[\text{sw}(g(\vec{\sigma}^i), \vec{u}^i)] \geq t_i$ , set  $b_i = 1 - b'_i$ . Since  $\mathcal{F}^{n,m}$  shatters  $\vec{\sigma}^1, \dots, \vec{\sigma}^d$ , there is an  $h \in \mathcal{F}^{n,m}$  that serves as a witness for  $\vec{\sigma}^1, \dots, \vec{\sigma}^d$  w.r.t.  $(b_1, \dots, b_d)$ . By construction, its corresponding  $\hat{h} \in \hat{\mathcal{F}}^{n,m}$  serves as a witness (relative to  $\hat{f}, \hat{g}$ ) to the shattering of the same  $d$  profiles given  $(b'_1, \dots, b'_d)$ . Hence  $D_P(\hat{\mathcal{F}}^{n,m})$  is at least  $d$ .  $\square$

## 5.2. Sample-based optimization

Let  $\mathcal{F}^{n,m}$  be some class of social choice functions from which we must select an optimal function  $f^*$  relative to some (possibly unknown) distribution  $D$  over utility profiles. We assume access to  $t$  sampled profiles,  $\vec{u}^1, \dots, \vec{u}^t$ . These may be samples from a population of interest, or drawn randomly from a generative model or known distribution. For each  $\vec{u}^i$ , we also sample, generate, or compute the corresponding (possibly random, due to ties) preference profile  $\vec{\sigma}^i$ . We treat these collectively as our sample:  $T = [(\vec{u}^1, \vec{\sigma}^1), \dots, (\vec{u}^t, \vec{\sigma}^t)]$ . A *sample-optimal social choice function* for sample  $T$  is

$$\hat{f} \in \operatorname{argmax}_{f \in \mathcal{F}^{n,m}} \sum_{i=1}^t \mathbb{E}_{f(\vec{\sigma}^i)}[\text{sw}(f(\vec{\sigma}^i), \vec{u}^i)].$$

In a learning-theoretic sense,  $\hat{f}$  is the *empirical risk minimizer*, while from an optimization standpoint,  $\hat{f}$  is the solution to a sample-based optimization problem.

In a sample-based model, we must content ourselves with approximate optimality. Let  $f^*$  be an optimal social choice function w.r.t. distribution  $D$ . We say a social choice function  $\hat{f}$  is  $\varepsilon$ -optimal for some  $\varepsilon \geq 0$  if, for any utility profile  $\vec{u}$ ,

$$\mathbb{E}[\text{sw}(\hat{f}(\vec{\sigma}(\vec{u})), \vec{u})] \geq \mathbb{E}[\text{sw}(f^*(\vec{\sigma}(\vec{u})), \vec{u})] - \varepsilon.$$

This definition will also be used relative to restricted classes of functions  $\mathcal{F}^{n,m}$ .

## 5.3. Sample complexity of $k$ -approval

We first consider the class of social choice functions known as  *$k$ -approval functions*. For any  $1 \leq k \leq m-1$ , the  $k$ -approval function is the scoring function, denoted  $f_k$ , with score vector  $\vec{s}_k = (1, 1, \dots, 0, 0)$  that has exactly  $k$  ones and  $m-k$  zeros. We assume ties among highest-scoring alternatives are broken uniformly at random.

Given distribution  $D$ , the *optimal  $k$ -approval function*—where our only choice is over the value of  $k$ —maximizes expected social welfare w.r.t.  $D$ . We denote this function by  $f_{k^*}^D$ . With only a collection of  $t$  sample profiles, the best we can attain is approximate optimality with the sample-optimal function  $\hat{f}$ . We determine the required sample complexity  $t$ , that is, the number of samples needed to ensure that  $\hat{f}$  is approximately optimal to some desired degree  $\varepsilon$  with high probability  $1 - \delta$  (for some  $\delta > 0$ ).

Our class of social choice functions is very limited: let  $\mathcal{F}_{\text{app}}^{n,m} = \{f_k : 1 \leq k \leq m-1\}$  be the class of  $k$ -approval functions with  $n$  agents and  $m$  alternatives. Define, as above,  $\hat{\mathcal{F}}_{\text{app}}^{n,m} = \{\hat{f}_k : f_k \in \mathcal{F}_{\text{app}}^{n,m}\}$ . Let  $\text{sc}_k(\vec{\sigma}, a)$  be the  $k$ -approval score of  $a \in A$  under preference profile  $\vec{\sigma}$ . Sample complexity depends on the pseudo-dimension of  $k$ -approval functions; since there are only  $m-1$  such functions, we can provide an immediate upper bound using [Observation 5.1](#):

**Observation 5.3.**  $D_P(\hat{\mathcal{F}}_{\text{app}}^{n,m}) \leq \log_2(m-1)$ .

This bound is asymptotically tight:

**Theorem 5.4.** For all  $m \geq 6$  and  $n \geq 2m - 3$ ,  $D_P(\hat{\mathcal{F}}_{app}^{n,m}) = \Omega(\log m)$ .

**Proof.** Let  $d \geq 1$  be an integer. We will show that the generalized dimension  $D_G(\mathcal{F}_{app}^{n,m})$  of  $k$ -approval functions on preference profiles with  $m \geq 2^{d+1} + 2$  alternatives and  $n \geq 2^{d+2} + 1$  agents is at least  $d$ . We will first construct  $d$  preference profiles  $\vec{\sigma}^1, \vec{\sigma}^2, \dots, \vec{\sigma}^d$  with exactly  $2^{d+1} + 2$  alternatives, including two special ones  $a$  and  $b$ , and exactly  $2^{d+2} + 1$  agents with the following property. For any integers  $k = 2, 3, \dots, 2^d + 1$  and  $t = 1, 2, \dots, d$ ,  $k$ -approval returns as the sole winner in profile  $\vec{\sigma}^t$ :

- alternative  $a$ , if the  $t$ -th bit in the binary representation of  $k - 2$  is 1, and
- alternative  $b$  otherwise.

Then, we will show how to easily extend our construction for  $m \geq 2^{d+1} + 2$  alternatives and  $n \geq 2^{d+2} + 1$  agents. As a corollary, we will have obtained that the particular sample of  $d$  profiles is shattered by the class of  $k$ -approval functions. The theorem will then follow by Lemma 5.2 and by the relation between  $d$ ,  $m$ , and  $n$ .

Our construction for preference profile  $\vec{\sigma}^t$  is as follows. The set of alternatives contains the two special alternatives  $a$  and  $b$  as well as two disjoint sets of alternatives  $A_1$  and  $A_2$ , each of size  $2^d$ . Each profile has  $2^{d+2} + 1$  agents partitioned into five disjoint sets: a set of  $2^d$  agents  $N_q$  for  $q = 1, 2, 3, 4$ , and an extra agent. For  $q \in \{1, 2\}$ , the agents in  $N_q$  rank alternatives  $a$  and  $b$  in the first two positions (in arbitrary order), then the alternatives in  $A_q$  (in arbitrary order), and then the alternatives in  $A_{3-q}$  (again, in arbitrary order). All the agents in  $N_3$  rank alternative  $a$  last. For  $j = 0, 1, \dots, 2^{d-t} - 1$ ,  $2^t$  agents of  $N_3$  rank alternative  $b$  in position  $2 + j2^t$ . The remaining top  $2^d + 1$  positions in the ranking of each agent in  $N_3$  are occupied by alternatives in  $A_1$  in arbitrary order; the next  $2^d$  positions are occupied by alternatives in  $A_2$  (again, in arbitrary order). All the agents in  $N_4$  rank alternative  $b$  last. For  $j = 0, 1, \dots, 2^{d-t} - 1$ ,  $2^t$  agents of  $N_4$  rank alternative  $a$  in position  $2 + 2^{t-1} + j2^t$ . The remaining top  $2^d + 1$  positions in the ranking of each agent in  $N_4$  are occupied by alternatives in  $A_2$  in arbitrary order; the next  $2^d$  positions are occupied by alternatives in  $A_1$  (again, in arbitrary order). Finally, the extra agent ranks alternative  $a$  first, alternative  $b$  last, and the alternatives in  $A_1 \cup A_2$  in between in arbitrary order.

Now, note that each alternative in  $A_q$  is ranked in the first  $2^d + 1$  positions by at most  $2^{d+1} + 1$  agents (the agents in  $N_q \cup N_{2+q}$  and, possibly, the extra agent). Hence, the  $k$ -approval score of each alternative in  $A_1 \cup A_2$  is at most  $2^{d+1} + 1$  for every  $k$  in  $\{2, \dots, 2^d + 1\}$ . Also, observe that alternative  $b$  is ranked in one of the top two positions by at least  $2^{d+1} + 2$  agents (the agents in  $N_1 \cup N_2$  as well as  $2^t$  agents of  $N_3$ ). Hence, no alternative from  $A_1 \cup A_2$  is a winner in profile  $\vec{\sigma}^t$  under  $k$ -approval, for  $t = 1, \dots, d$  and  $k = 2, \dots, 2^d + 1$ .

It remains to compute the  $k$ -approval score of alternatives  $a$  and  $b$ . Consider profile  $\vec{\sigma}^t$ . Observe that, for every  $k = 2, \dots, 2^d + 1$ , the number of agents in  $N_3$  that rank alternative  $b$  in the top  $k$  positions is

$$2^t \left( \left\lfloor \frac{k-2}{2^t} \right\rfloor + 1 \right) = 2^t \left\lfloor \frac{k-2+2^t}{2^t} \right\rfloor$$

and similarly the number of agents in  $N_4$  that rank alternative  $a$  in the top  $k$  positions is  $2^t \lfloor \frac{k-2+2^{t-1}}{2^t} \rfloor$ . Hence, the  $k$ -approval scores of  $a$  and  $b$  are  $2^{d+1} + 1 + 2^t \lfloor \frac{k-2+2^{t-1}}{2^t} \rfloor$  and  $2^{d+1} + 2^t \lfloor \frac{k-2+2^t}{2^t} \rfloor$ , respectively. Let  $b_t$  be the  $t$ -th least significant bit in the binary representation of  $k - 2$  and let  $\lambda, \nu$  be unique integers such that  $0 \leq \lambda \leq 2^{t-1} - 1$ ,  $0 \leq \nu \leq 2^{d-t} - 1$ , and  $k - 2 = \lambda + b_t 2^{t-1} + \nu 2^t$ . Then,

$$\begin{aligned} \text{sc}_k(a, \vec{\sigma}^t) - \text{sc}_k(b, \vec{\sigma}^t) &= 1 + 2^t \left\lfloor \frac{k-2+2^{t-1}}{2^t} \right\rfloor - 2^t \left\lfloor \frac{k-2+2^t}{2^t} \right\rfloor \\ &= 1 + 2^t \left\lfloor \frac{\lambda + b_t 2^{t-1} + \nu 2^t + 2^{t-1}}{2^t} \right\rfloor - 2^t \left\lfloor \frac{\lambda + b_t 2^{t-1} + \nu 2^t + 2^t}{2^t} \right\rfloor \\ &= 1 + 2^t \left\lfloor \frac{\lambda + b_t 2^{t-1} + \nu 2^t + 2^{t-1}}{2^t} \right\rfloor - 2^t(\nu + 1) \\ &= 1 + 2^t \left\lfloor \frac{\lambda + b_t 2^{t-1} - 2^{t-1}}{2^t} \right\rfloor \\ &= 1 + 2^t(b_t - 1), \end{aligned}$$

where the third and fifth equalities hold because  $0 \leq \lambda \leq 2^{t-1} - 1$  (and  $b_t \in \{0, 1\}$ ). Using the fact that  $t \geq 1$ , observe that  $\text{sc}_k(a, \vec{\sigma}^t) - \text{sc}_k(b, \vec{\sigma}^t)$  is strictly positive if  $b_t = 1$  and strictly negative if  $b_t = 0$ . Hence, the sole winner under  $k$ -approval is alternative  $a$  if  $b_k = 1$  and alternative  $b$  otherwise.

To complete the proof, it suffices to modify the construction above and capture the more general case of  $m \geq 2^{d+1} + 2$  alternatives and  $n \geq 2^{d+2} + 1$  agents as follows. Each of the  $2^{d+2} + 1$  agents in the above construction ranks the  $m - 2^{d+1} - 2$  additional alternatives in the last positions of its ranking. In addition to these agents, there are  $n - 2^{d+2} - 1$  agents that rank alternatives  $a$  and  $b$  in the first two positions (in arbitrary order) and then the remaining alternatives (again, in arbitrary

order). These modifications do not change the winner of profile  $\vec{\sigma}^t$  under  $k$ -approval for every  $k \in \{2, 3, \dots, 2^d + 1\}$ . This is due to the facts that alternative  $b$  still has strictly higher  $k$ -approval score than any other alternative different than  $a$  (alternative  $b$  is ranked above all alternatives besides  $a$  by the additional agents and is clearly ranked above the additional alternatives by all agents), and the difference in the  $k$ -approval scores between  $a$  and  $b$  is not affected by the additional agents (since  $k \geq 2$ ).  $\square$

**Observation 5.3** and **Theorem 5.4** show that  $D_P(\hat{\mathcal{F}}_{\text{app}}^{n,m}) = \Theta(\log m)$  provided that  $n \geq 2m - 3$ . Standard learning-theoretic results [2] allow us to bound sample complexity for optimizing  $k$ -approval (within a constant factor).

**Theorem 5.5.** *For any  $\varepsilon, \delta > 0$ , there exists a  $C > 0$  such that if  $t \geq C \log(m/\delta)/\varepsilon^2$ , then for any distribution  $D$  over utility profiles, with probability at least  $1 - \delta$  over  $t$  i.i.d. utility profiles, the sample-optimal  $k$ -approval function  $\hat{f}_k$  is  $\varepsilon$ -optimal for  $D$ . Furthermore, for  $n \geq 2m - 3$ , there is a  $C' > 0$  such that no algorithm can construct an  $\varepsilon$ -optimal  $k$ -approval function, with probability at least  $1 - \delta$ , if  $t < C' \log(m/\delta)/\varepsilon^2$ .*

#### 5.4. Sample complexity of scoring functions

The class of  $k$ -approval functions is quite restrictive, so we now consider construction of an approximately optimal scoring function without restricting score vector structure. Limiting attention to scoring functions does not ensure optimality within the class of *arbitrary* functions. However, it is a natural restriction, first, because of the prominence of scoring functions as illustrated above, and second, because of the natural interpretation and appeal of such social choice functions.<sup>9</sup>

Let  $f_{\vec{s}}$  denote the scoring (social choice) function induced by score vector  $\vec{s}$ , and let  $\mathcal{F}_s^{n,m} = \{f_{\vec{s}} : \vec{s} \in \mathbb{R}^m\}$  be the class of all scoring functions with  $n$  agents and  $m$  alternatives. We again assume ties among highest-scoring alternatives are broken uniformly at random. Define  $\hat{\mathcal{F}}_s^{n,m} = \{\hat{f}_{\vec{s}} : f_{\vec{s}} \in \mathcal{F}_s^{n,m}\}$ . We derive the sample complexity for scoring functions, i.e., the number of sampled utility profiles needed to ensure that the sample-optimal  $\hat{f}_{\vec{s}}$  is  $\varepsilon$ -optimal for some desired  $\varepsilon$ , with probability at least  $1 - \delta$ .

We first bound the pseudo-dimension of  $\hat{\mathcal{F}}_s^{n,m}$ . Procaccia et al. [28] prove a lower bound of  $m - 3$  on  $D_G(\mathcal{F}_s^{n,m})$  for all  $n \geq 4$  and  $m \geq 4$ . By **Lemma 5.2**, we obtain the following statement.

**Corollary 5.6.** *For all  $n \geq 4$  and  $m \geq 4$ ,  $D_P(\hat{\mathcal{F}}_s^{n,m}) \geq m - 3$ .*

In the same paper, Procaccia et al. [28] prove that the number of distinct scoring functions is at most  $2^{O(m^2 \log n)}$ . Even though their original result assumes a deterministic tie-breaking rule, their proof can be adapted for randomized tie-breaking. Using this bound together with **Observation 5.1**, we immediately obtain that  $D_P(\hat{\mathcal{F}}_s^{n,m}) = O(m^2 \log n)$ . We can derive a significantly better upper bound that depends only on  $m$ :

**Theorem 5.7.**  $D_P(\hat{\mathcal{F}}_s^{n,m}) = O(m \log m)$ .

We require the following result of Warren [33] (though we use an extension due to Alon [1], where the signum function takes values in  $\{-1, 0, 1\}$ ):

**Lemma 5.8.** (See [33,1].) *Let  $\mathcal{P}$  be a set of  $K$  polynomials of degree  $\tau$  on  $\ell$  real variables. Then, the number of different sign patterns  $\mathcal{P}$  may have is at most  $\left(\frac{8e\tau K}{\ell}\right)^\ell$ .*

In the context of the lemma, a sign pattern is a vector of values in  $\{-1, 0, 1\}$  that is obtained by applying each polynomial in  $\mathcal{P}$  to the same input.

**Proof of Theorem 5.7.** Let  $A$  be a set of  $m \geq 3$  alternatives. Consider a sample of  $d$  preference profiles  $\vec{\sigma}^1, \vec{\sigma}^2, \dots, \vec{\sigma}^d$  from  $(\mathcal{S}_m)^n$ , corresponding utility profiles  $\vec{u}^1, \dots, \vec{u}^d$ , and real thresholds  $t_1, t_2, \dots, t_d$  such that for every binary vector  $b$  with  $d$  entries, there exists a scoring function  $f_{\vec{s}_b}$ , with score vector  $\vec{s}_b$ , such that  $\mathbb{E}[\text{sw}(f_{\vec{s}_b}(\vec{\sigma}^i), \vec{u}^i)] \geq t_i$  if  $b_i = 1$  and  $\mathbb{E}[\text{sw}(f_{\vec{s}_b}(\vec{\sigma}^i), \vec{u}^i)] < t_i$  otherwise. Denote by  $\mathcal{F}^* \subseteq \mathcal{F}_s^{n,m}$  the set of these scoring functions. We will show that  $d < 6m \log_2 m$ .

For  $i = 1, \dots, d$ , we partition the set of alternatives  $A$  into the sets  $A_i^+ = \{a \in A : \text{sw}(a, \vec{u}^i) \geq t_i\}$  and  $A_i^- = \{a \in A : \text{sw}(a, \vec{u}^i) < t_i\}$ . By definition of the scoring functions in  $\mathcal{F}^*$ , the sets  $A_i^+$  and  $A_i^-$  are non-empty. In addition, for  $i = 1, \dots, d$  and any pair of alternatives  $a^+ \in A_i^+$  and  $a^- \in A_i^-$ , define the function

<sup>9</sup> Optimization over the class of arbitrary social choice functions may well give results that cannot be communicated without enumerating all possible profiles.

$$L_{a^+, a^-}^i(\vec{s}) = \text{sc}_{\vec{s}}(a^+, \vec{\sigma}^i) - \text{sc}_{\vec{s}}(a^-, \vec{\sigma}^i).$$

This is a linear function on  $m$  variables, the entries of the score vector  $\vec{s}$ . Let  $\text{sgn} : \mathbb{R} \rightarrow \{-1, 0, 1\}$  denote the signum function. The sign of  $L_{a^+, a^-}^i(\vec{s})$  denotes whether the score of alternative  $a^+$  under  $f_{\vec{s}}$  on profile  $\vec{\sigma}^i$  is lower than, equal to, or higher than the score of alternative  $a^-$ , respectively. Let  $L$  denote the collection of the functions above and observe that there are at most  $dm^2/4$  such functions.

We will show that the score vectors of the functions in  $\mathcal{F}^*$  define at least  $2^d$  different sign patterns for  $L$ . Consider two binary vectors  $b^1$  and  $b^0$  that differ in the  $i$ -th coordinate. W.l.o.g., assume that  $b_i^1 = 1$  and  $b_i^0 = 0$ . For  $k \in \{0, 1\}$ , let  $A^k = \text{argmax}_{a \in A} \{\text{sc}_{\vec{s}_{b^k}}(a, \vec{\sigma}^i)\}$  be the set of alternatives with the highest score under  $f_{\vec{s}_{b^k}}(\vec{\sigma}^i)$ . Observe that the sets  $A_i^+ \cap A^1$  and  $A_i^- \cap A^0$  are non-empty since  $\mathbb{E}[\text{sw}(f_{\vec{s}_{b^1}}(\vec{\sigma}^i), \vec{u}^i)] \geq t_i$  and  $\mathbb{E}[\text{sw}(f_{\vec{s}_{b^0}}(\vec{\sigma}^i), \vec{u}^i)] < t_i$ .

We claim that there are two alternatives  $a_1 \in A_i^+ \cap A^1$  and  $a_0 \in A_i^- \cap A^0$  such that either  $a_0 \notin A^1$  or  $a_1 \notin A^0$ . Assume that this is not the case; it follows that  $A_i^+ \cap (A^1 \setminus A^0) = \emptyset$  and  $A_i^- \cap (A^0 \setminus A^1) = \emptyset$ . Hence, for every  $a \in A^0 \setminus A^1$  (if any) it holds that  $a \in A_i^+$  and hence  $\text{sw}(a, \vec{u}^i) \geq t_i$ . Similarly, for every alternative  $a \in A^1 \setminus A^0$  (if any) it holds that  $a \in A_i^-$  and hence  $\text{sw}(a, \vec{u}^i) < t_i$ . Using these observations together with the fact that  $f_{\vec{s}_{b^1}}(\vec{\sigma}^i)$  and  $f_{\vec{s}_{b^0}}(\vec{\sigma}^i)$  select the winning alternative uniformly at random among  $A^1$  and  $A^0$ , respectively, and  $\mathbb{E}[\text{sw}(f_{\vec{s}_{b^1}}(\vec{\sigma}^i), \vec{u}^i)] \geq t_i > \mathbb{E}[\text{sw}(f_{\vec{s}_{b^0}}(\vec{\sigma}^i), \vec{u}^i)]$ , we obtain

$$\begin{aligned} \sum_{a \in A^1 \cup A^0} \text{sw}(a, \vec{u}^i) &= \sum_{a \in A^1} \text{sw}(a, \vec{u}^i) + \sum_{a \in A^0 \setminus A^1} \text{sw}(a, \vec{u}^i) \geq |A^1| \cdot t_i + |A^0 \setminus A^1| \cdot t_i \\ &= |A^0| \cdot t_i + |A^1 \setminus A^0| \cdot t_i > \sum_{a \in A^0} \text{sw}(a, \vec{u}^i) + \sum_{a \in A^1 \setminus A^0} \text{sw}(a, \vec{u}^i) \\ &= \sum_{a \in A^1 \cup A^0} \text{sw}(a, \vec{u}^i), \end{aligned}$$

a contradiction.

If  $a_1 \in A^1 \setminus A^0$ , this means that  $\text{sc}_{\vec{s}_{b^1}}(a_1, \vec{\sigma}^i) \geq \text{sc}_{\vec{s}_{b^1}}(a_0, \vec{\sigma}^i)$  and  $\text{sc}_{\vec{s}_{b^0}}(a_1, \vec{\sigma}^i) < \text{sc}_{\vec{s}_{b^0}}(a_0, \vec{\sigma}^i)$  (i.e.,  $L_{a_1, a_0}^i(\vec{s}_{b^1}) \geq 0$  and  $L_{a_1, a_0}^i(\vec{s}_{b^0}) < 0$ ). Similarly, if  $a_0 \in A^0 \setminus A^1$ , then  $\text{sc}_{\vec{s}_{b^1}}(a_1, \vec{\sigma}^i) > \text{sc}_{\vec{s}_{b^1}}(a_0, \vec{\sigma}^i)$  and  $\text{sc}_{\vec{s}_{b^0}}(a_1, \vec{\sigma}^i) \leq \text{sc}_{\vec{s}_{b^0}}(a_0, \vec{\sigma}^i)$  (i.e.,  $L_{a_1, a_0}^i(\vec{s}_{b^1}) > 0$  and  $L_{a_1, a_0}^i(\vec{s}_{b^0}) \leq 0$ ). Both cases imply that the sign patterns of  $L$  corresponding to  $\vec{s}_{b^1}$  and  $\vec{s}_{b^0}$  are different.

Hence, we have obtained a lower bound of  $2^d$  on the number of different sign patterns of  $L$ . By applying [Lemma 5.8](#) (with  $K \leq dm^2/4$ ,  $\tau = 1$ , and  $\ell = m$ ), we obtain an upper bound of  $(2edm)^m$  on the number of different sign patterns for  $L$ . Hence,  $2^d \leq (2edm)^m$  and, equivalently,  $2^d/d^m \leq (2em)^m$ .

For the sake of contradiction, assume that  $d \geq 6m \log_2 m$  and observe that  $2^d/d^m$  increases with  $d$  in this range. Then,  $2^d/d^m \geq \left(\frac{m^6}{6m \log_2 m}\right)^m$  which, together with the above inequality, yields  $m^4 \leq 12e \log_2 m$ ; a contradiction since  $m \geq 3$ .  $\square$

Again, standard results allow us to bound the sample complexity:

**Theorem 5.9.** For any  $\varepsilon, \delta > 0$ , there exists a  $C > 0$  such that if  $t \geq C[m \log_2 m + \log(1/\delta)]/\varepsilon^2$ , then for any distribution  $D$  over utility profiles, with probability at least  $1 - \delta$  over  $t$  i.i.d. utility profiles, the sample-optimal scoring function  $\hat{f}_{\vec{s}}$  is  $\varepsilon$ -optimal for  $D$ . Furthermore, there is a  $C' > 0$  such that no algorithm can construct an  $\varepsilon$ -optimal scoring function, with probability at least  $1 - \delta$ , if  $t < C'[m + \log(1/\delta)]/\varepsilon^2$ .

### 5.5. Computing optimal scoring functions

We now turn our attention to the question of computing approximately optimal scoring functions. Specifically, given a sample  $T = [(\vec{u}^1, \vec{\sigma}^1), \dots, (\vec{u}^t, \vec{\sigma}^t)]$ , we must compute the scoring vector  $\vec{s}$  corresponding to the sample-optimal scoring function  $\hat{f}_{\vec{s}}$ :

$$\hat{f}_{\vec{s}} \in \text{argmax}_{\vec{s}} \sum_{i=1}^t \mathbb{E}_{f_{\vec{s}}(\vec{\sigma}^i)} [\text{sw}(f_{\vec{s}}(\vec{\sigma}^i), \vec{u}^i)].$$

This problem turns out to be computationally hard.

**Theorem 5.10.** Computing the sample-optimal scoring function is  $\mathcal{APX}$ -hard.

The theorem, whose proof appears in [Appendix B](#), implies that, if  $\mathcal{P} \neq \mathcal{NP}$ , the problem does not even admit a polynomial time approximation scheme.



On the positive side, we formulate the problem of computing sample-optimal scoring functions as a MIP. Its primary variables are the scores  $s_i$ . We describe key variables and constraints in the MIP in turn.<sup>10</sup>

Any scoring vector  $\vec{s} = (s_1, \dots, s_m)$  can be normalized without impacting the choice function, so we constrain  $\vec{s}$  as follows:

$$s_1 + \dots + s_m = 1, \quad s_i \geq s_{i+1} \quad \forall i \leq m-1, \quad \text{and} \quad s_m \geq 0. \quad (4)$$

Ties are again broken uniformly at random. Function  $f_{\vec{s}}$  selects an alternative for each  $\vec{\sigma}^i$ . To encode this, first abbreviate the score of  $a$  given  $\vec{\sigma}^i$  via the linear expression

$$\text{sc}(a, \vec{\sigma}^i) = \sum_{j=1}^m J_{aji} s_j \quad \forall a \in A, i \leq t, \quad (5)$$

where  $J_{aji}$  is the number of agents in  $\vec{\sigma}^i$  that rank  $a$  in position  $j$ . Note that  $J_{aji}$  is a constant and  $\text{sc}(a, \vec{\sigma}^i) \in [0, n]$  is continuous as a function of  $\vec{\sigma}^i$ . Let  $I_{abi}$ , for any alternatives  $a \neq b$  and  $i \leq t$ , be a binary variable indicating whether  $a$ 's score is *at least* that of  $b$  given  $\vec{\sigma}^i$ . We encode this as follows:

$$(n + \gamma)I_{abi} - \gamma \geq \text{sc}(a, \vec{\sigma}^i) - \text{sc}(b, \vec{\sigma}^i) \quad \forall i \leq t, a \neq b, \quad (6)$$

$$nI_{abi} - n \leq \text{sc}(a, \vec{\sigma}^i) - \text{sc}(b, \vec{\sigma}^i) \quad \forall i \leq t, a \neq b, \quad (7)$$

where  $\gamma$  is a (fixed) parameter that handles optimization-dependent floating point accuracy (corresponding to the level of discretization among scores). If the score difference is non-negative then constraint (6) forces  $I_{abi} = 1$  and (7) must be satisfied. If the difference is negative, then (7) forces  $I_{abi} = 0$  and (6) is satisfied. Let binary variable  $I_{ai}$  indicate if  $a$  is selected (possibly tied), given  $\vec{s}$  under  $\vec{\sigma}^i$ . We require:

$$m - 2 + I_{ai} \geq \sum_{b:b \neq a} I_{abi} \quad \text{and} \quad (m - 1)I_{ai} \leq \sum_{b:b \neq a} I_{abi} \quad \forall a, i \leq t. \quad (8)$$

Our objective is to choose  $\vec{s}$  to maximize the average social welfare over our samples; however, we must account for random tie-breaking, leading to the following objective:

$$\max_{\vec{s}, I} \sum_{i=1}^t \frac{\sum_a \text{sw}(a, \vec{u}^i) \cdot I_{ai}}{\sum_a I_{ai}}.$$

We can linearize the objective using indicator variables  $S_{ki}$ , for  $k \leq m$  and  $i \leq t$ , where  $S_{ki} = 1$  iff  $k = \sum_a I_{ai}$ , requiring that

$$\sum_{k=1}^m kS_{ki} = \sum_a I_{ai} \quad \text{and} \quad \sum_{k=1}^m S_{ki} = 1 \quad \forall i \leq t. \quad (9)$$

Our objective then becomes

$$\max_{\vec{s}, I, S} \sum_{i=1}^t \sum_{k=1}^m \frac{\sum_a \text{sw}(a, \vec{u}^i) \cdot I_{ai}}{k} \cdot S_{ki}.$$

Finally, let  $Z_{aki}$  indicate if  $I_{ai} \cdot S_{ki} = 1$ , which is encoded as

$$1 + Z_{aki} \geq I_{ai} + S_{ki} \quad \text{and} \quad 2Z_{aki} \leq I_{ai} + S_{ki} \quad \forall a \in A, k \leq m, i \leq t. \quad (10)$$

Pulling these together, our MIP is:

$$\begin{aligned} & \max_{\vec{s}, I, S, Z} \sum_{i=1}^t \sum_{k=1}^m \sum_a \frac{1}{k} \text{sw}(a, \vec{u}^i) \cdot Z_{aki} \\ & \text{subject to} \quad (4, 6, 7, 8, 9, 10), \end{aligned} \quad (11)$$

which has  $(2m^2 + m + 1)t$  variables and  $4m^2t + 2t + m + 1$  constraints.

<sup>10</sup> With suitable constraints on scores, the MIP can be used to compute optimal  $k$ -approval functions; however, direct evaluation of the small number of such restricted functions is feasible (if  $m$  is small).

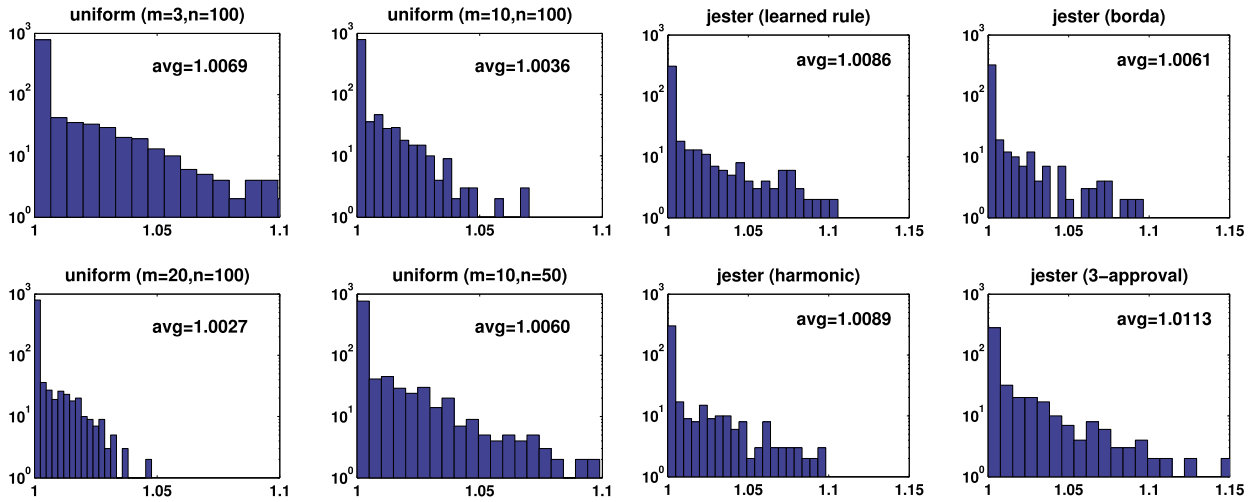


Fig. 2. Histograms of distortion ratios for *uniform* and *jester* experiments (average distortions are shown).

## 5.6. Experiments

We now empirically investigate the performance of both average-case optimal social choice functions and sample-optimal scoring functions by measuring their distortion. In the former case, we consider random utility profiles, while in the latter, we use a real data set with user-provided ratings of alternatives as a surrogate for utilities.

Our first experiment (*uniform*) investigates the uniform utility distribution described in Section 4. While Borda optimizes expected social welfare in this model (Corollary 4.5), it may not minimize distortion. We empirically measure its expected distortion by randomly generating  $t = 1000$  profiles from the uniform model for various values of  $m$  and  $n$ , and computing the distortion of Borda count (vis-à-vis the socially optimal alternative). Results are shown in the four leftmost histograms of Fig. 2. Each histogram shows the distortions of the 1000 utility profiles, for a fixed  $m$  and  $n$  (note the logarithmic scaling on the y-axis). Clearly, overall distortion is very small: average distortion is much less than 1.01 in each case, and never exceeds 1.1 for any random profile. We also see that average distortion decreases as either  $m$  or  $n$  increases.

Our second experiment uses the *jester* dataset [14], which consists of 4.1M user ratings of 100 different jokes by over 60,000 users. Ratings are on a continuous scale between  $[-10, 10]$ , which we rescale to the range  $[0, 20]$ . We define the set of alternatives to be the eight most-rated jokes, and draw agents from the set of 50,699 users who rated all eight. We create a sample of 100 “training profiles” from this data set, each with 100 voters, and use this sample to learn an approximately optimal scoring function.<sup>11</sup> The score vector that results is  $\vec{s}^* = (0.25, 0.15, 0.14, 0.13, 0.12, 0.11, 0.1, 0.0)$ . Note the significant dip from  $s_1$  to  $s_2$ , the gradual drop to  $s_7$ , then the significant drop to  $s_8$ , which is rather “un-Borda-like.” We divide the remaining users into 406 test profiles (each with 100 users), and evaluate the distortion of the learned function  $f_{\vec{s}^*}$  on each. For comparison, we also evaluate the Borda, harmonic and 3-approval functions on the same profiles. Results are shown in the four rightmost histograms of Fig. 2. We see clearly that distortion is almost negligible for the  $f_{\vec{s}^*}$ , Borda and harmonic functions, with average distortion less than 1.009 (and at worst roughly 1.1). By contrast, 3-approval is somewhat worse, with average distortion of 1.0113 (and in the worst case about 1.15). The sample-optimal function  $f_{\vec{s}^*}$  performs slightly worse than Borda, due to mild overfitting on the training profiles (note that the theoretical sample complexity for this problem is much greater than the 100 samples used). These results are of course limited, and merely suggestive; but they do indicate that scoring functions, either empirically optimized, or relying on stylized scoring vectors like Borda and harmonic score, can very closely approximate optimal social choice functions in practice.

## 6. Discussion

Our work offers three different but related perspectives on utilitarian social choice. Each model makes fundamentally different assumptions about the mechanism’s knowledge of the agents’ utility information. In the worst-case model, we study the distortion of randomized social choice functions assuming no information about the underlying utilities. In the average-case model, we derive the optimal social choice function with respect to distributions that are i.i.d. and neutral. Finally, in the learning-theoretic model, we develop a method for approximately optimizing (scoring-based) social choice functions under arbitrary utility distributions, establish sample complexity bounds and provide encouraging empirical results.

Our work raises a number of important questions and directions for future research. Access to sampled utility profiles, as assumed in our learning-theoretic model, may be difficult to obtain in practice. However, techniques from decision analysis

<sup>11</sup> CPLEX 12.2 on a modern workstation took 23.6 hrs. to solve the resulting MIP (accuracy gap of 1.52%).

and preference elicitation using lotteries, or more readily comprehensible queries involving simple comparisons, can be used to assess the utility functions of specific agents [9], while econometric techniques often use revealed preference or stated choice data to develop probabilistic models of utilities [20]. Applying these methods to the design of optimal (ranking-based) social choice functions is an important direction for future work.

One of our motivations is to reduce the cognitive and communication burden associated with utilities or valuations by allowing the agents to specify rankings. This burden can be further reduced by intelligent elicitation of *partial* ranking information [19,22]. Our utilitarian model offers a novel perspective on vote elicitation and raises the possibility of designing schemes that perform well with respect to utilitarian social welfare.

The utilitarian perspective also suggests new ways of assessing the potential manipulation of social choice functions. By assuming agents have utility functions, and probabilistic information about the utilities of their counterparts, one can quantify the gains of potential misreports in terms of expected utility, providing a Bayesian view of manipulation [23]. The design of scoring functions that make appropriate trade-offs between degree of optimality and degree of manipulability is another important problem to which our methods may be adapted.

## Acknowledgements

We thank Yury Makarychev, Hervé Moulin, and Tasos Sidiropoulos for helpful discussions, and the anonymous referees for valuable comments. Boutilier and Lu were supported by NSERC. Procaccia was supported in part by the NSF under grants IIS-1350598 and CCF-1215883, and by a Sloan Research Fellowship. Sheffet was supported in part by the NSF under grants CCF-1101215 and CCF-1116892 as well as by the MSR-CMU Center for Computational Thinking.

## Appendix A. Proof of Theorem 3.3

We begin by defining a randomized social choice function  $f$ . For a non-negative integer  $i$ , define the sets of alternatives  $B_i(\vec{\sigma})$  as follows. Set  $B_0(\vec{\sigma})$  consists of all alternatives in  $A$ . For  $i \geq 1$ , an alternative  $a \in B_{i-1}(\vec{\sigma})$  belongs to  $B_i(\vec{\sigma})$  if there exists a utility profile  $\vec{u}^a \in p^{-1}(\vec{\sigma})$  in which  $a$  has maximum social welfare and, furthermore,

$$\frac{|B_\ell(\vec{\sigma})|}{3\sqrt{m}} \text{sw}(a, \vec{u}^a) \geq \sum_{a' \in B_\ell(\vec{\sigma})} \text{sw}(a', \vec{u}^a) \quad (12)$$

for all  $\ell \in \{0, i-1\}$ .

Given the preference profile  $\vec{\sigma} \in (\mathcal{S}_m)^n$ , our randomized social choice function  $f$  first computes the sets  $B_0(\vec{\sigma}), B_1(\vec{\sigma}), \dots, B_k(\vec{\sigma})$  where  $k$  is the smallest integer such that either  $|B_k(\vec{\sigma})| < 3\sqrt{m}$  or  $B_{k+1}(\vec{\sigma}) = \emptyset$ . Deciding whether an alternative of  $B_{i-1}(\vec{\sigma})$  is included in  $B_i(\vec{\sigma})$  can be done in polynomial time by checking the feasibility of a straightforward linear program. Then, the function  $f$  picks an integer  $i \in \{0, \dots, k\}$  uniformly at random, and selects equiprobably among the alternatives in  $B_i(\vec{\sigma})$ .

In the remainder of the proof we show that for every  $\vec{\sigma} \in (\mathcal{S}_m)^n$ ,  $\text{dist}(f, \vec{\sigma}) = \mathcal{O}(\sqrt{m} \cdot \log^* m)$ . Specifically, we will establish that the distortion is less than  $3\sqrt{m}(\log^* m + 1)$  for  $m \geq 9$ .

Given a preference profile  $\vec{\sigma}$ , consider a utility profile  $\vec{u}$  consistent with  $\vec{\sigma}$ , and let  $a^*$  be an alternative with the highest social welfare. Let  $t$  be the largest integer in  $\{0, 1, \dots, k\}$  such that  $a^* \in B_t(\vec{\sigma})$ . First consider the special case where  $t = k$  and  $|B_k(\vec{\sigma})| < 3\sqrt{m}$ ;  $f$  picks alternative  $a^*$  with probability higher than  $\frac{1}{3(k+1)\sqrt{m}}$ .

In any other case, alternative  $a^*$  does not belong to set  $B_{t+1}(\vec{\sigma})$ . This means that  $\vec{u}$  (and every other utility profile in  $p^{-1}(\vec{\sigma})$  in which  $a^*$  has the highest social welfare) satisfies

$$\frac{1}{|B_\ell(\vec{\sigma})|} \sum_{a' \in B_\ell(\vec{\sigma})} \text{sw}(a', \vec{u}) > \frac{1}{3\sqrt{m}} \text{sw}(a^*, \vec{u})$$

for some  $\ell \in \{0, t\}$ . With probability  $\frac{1}{k+1}$ , the algorithm picks a random alternative from  $B_\ell(\vec{\sigma})$ . Therefore, its expected social welfare is at least

$$\frac{1}{(k+1)|B_\ell(\vec{\sigma})|} \sum_{a' \in B_\ell(\vec{\sigma})} \text{sw}(a', \vec{u}) > \frac{1}{3(k+1)\sqrt{m}} \text{sw}(a^*, \vec{u}).$$

The theorem will follow by proving that  $k \leq \log^* m$ . To do this, we assume that  $m \geq 9$ , and show that, for every integer  $i \geq 1$  such that  $|B_{i-1}(\vec{\sigma})| \geq 3\sqrt{m}$ , it holds that

$$\frac{2|B_i(\vec{\sigma})|}{3\sqrt{m}} \leq \log_2 \frac{2|B_{i-1}(\vec{\sigma})|}{3\sqrt{m}}. \quad (13)$$

This implies that  $|B_{\log^*(2\sqrt{m}/3)}(\vec{\sigma})| < 3\sqrt{m}$  and, hence,

$$k \leq \log^* (2\sqrt{m}/3) \leq \log^* m.$$

The remainder of the proof is therefore devoted to establishing Equation (13).

In the following, we will show that inequality (13) is true for  $i \geq 1$  assuming that  $|B_{i-1}(\vec{\sigma})| \geq 3\sqrt{m}$ . For each alternative  $a \in B_i(\vec{\sigma})$ , let  $\vec{u}^a$  be a utility profile satisfying inequality (12). For such an alternative, denote by  $N_1(a)$  the set of agents which rank less than  $\frac{|B_{i-1}(\vec{\sigma})|}{\sqrt{m}}$  alternatives of  $B_{i-1}(\vec{\sigma})$  above  $a$ . Let  $N_2(a) = N \setminus N_1(a)$ . Also, for  $s \in \{1, 2\}$ , let

$$U_s = \sum_{a \in B_i(\vec{\sigma})} \sum_{j \in N_s(a)} u_j^a(a).$$

First, observe that for each agent  $j \in N_2(a)$ , there are at least  $\frac{|B_{i-1}(\vec{\sigma})|}{\sqrt{m}}$  alternatives  $a'$  of  $B_{i-1}(\vec{\sigma})$  with  $u_j^a(a') \geq u_j^a(a)$ . Hence, we have

$$\begin{aligned} \sum_{a \in B_i(\vec{\sigma})} \sum_{a' \in B_{i-1}(\vec{\sigma})} \text{sw}(a', \vec{u}^a) &= \sum_{a \in B_i(\vec{\sigma})} \sum_{a' \in B_{i-1}(\vec{\sigma})} \sum_{j \in N} u_j^a(a') \\ &\geq \sum_{a \in B_i(\vec{\sigma})} \sum_{j \in N_2(a)} \sum_{a' \in B_{i-1}(\vec{\sigma})} u_j^a(a') \\ &\geq \frac{|B_{i-1}(\vec{\sigma})|}{\sqrt{m}} \sum_{a \in B_i(\vec{\sigma})} \sum_{j \in N_2(a)} u_j^a(a) = \frac{|B_{i-1}(\vec{\sigma})|}{\sqrt{m}} U_2. \end{aligned} \quad (14)$$

In addition, since each alternative in  $B_i(\vec{\sigma})$  satisfies inequality (12) for  $\ell = i - 1$ , we have

$$\sum_{a \in B_i(\vec{\sigma})} \sum_{a' \in B_{i-1}(\vec{\sigma})} \text{sw}(a', \vec{u}^a) \leq \sum_{a \in B_i(\vec{\sigma})} \frac{|B_{i-1}(\vec{\sigma})|}{3\sqrt{m}} \text{sw}(a, \vec{u}^a) = \frac{|B_{i-1}(\vec{\sigma})|}{3\sqrt{m}} (U_1 + U_2). \quad (15)$$

It follows from (14) and (15) that  $U_1 \geq 2U_2$ . Now, using inequality (12) for  $\ell = 0$ , we have

$$\begin{aligned} 3U_1 &\geq 2(U_1 + U_2) = 2 \sum_{a \in B_i(\vec{\sigma})} \sum_{j \in N} u_j^a(a) = 2 \sum_{a \in B_i(\vec{\sigma})} \text{sw}(a, \vec{u}^a) \\ &\geq \frac{6\sqrt{m}}{|B_0(\vec{\sigma})|} \sum_{a \in B_i(\vec{\sigma})} \sum_{a' \in B_0(\vec{\sigma})} \text{sw}(a', \vec{u}^a) = \frac{6n}{\sqrt{m}} |B_i(\vec{\sigma})|, \end{aligned} \quad (16)$$

where the last equality follows from Assumption 3.1 and the fact that  $|B_0(\vec{\sigma})| = m$ . Also, observe that  $U_1$  is upper-bounded by the total harmonic score of alternatives in the first  $\lceil |B_{i-1}(\vec{\sigma})|/\sqrt{m} \rceil$  positions in each agent's preference, i.e.,

$$3U_1 \leq 3nH_{\lceil |B_{i-1}(\vec{\sigma})|/\sqrt{m} \rceil} \leq 9n \log_2 \frac{2|B_{i-1}(\vec{\sigma})|}{3\sqrt{m}}, \quad (17)$$

where  $H_k$  is the  $k$ th harmonic number. The second inequality follows from the fact that  $|B_{i-1}(\vec{\sigma})| \geq 3\sqrt{m}$  and  $H_{\lceil t \rceil} \leq 3 \log_2 \frac{2t}{3}$  for every  $t \geq 3$ . By (16) and (17), we obtain inequality (13) as desired. The proof of Theorem 3.3 is therefore complete.  $\square$

## Appendix B. Proof of Theorem 5.10

We will use a reduction from Independent Set in 3-regular graphs and an explicit inapproximability result due to Berman and Karpinski [6].

**Theorem 6.1.** (See Berman and Karpinski [6].) *Given a 3-regular graph  $G$  with  $284T$  nodes and  $426T$  edges for some integer  $T \geq 2$  and any constant  $\epsilon \in (0, 1/2)$ , it is  $\mathcal{NP}$ -hard to distinguish between the following two cases:*

- $G$  has an independent set of size at least  $(140 - \epsilon)T$ .
- Any independent set of  $G$  has size at most  $(139 + \epsilon)T$ .

The proof of Theorem 6.1 in [6] can be thought of as a polynomial-time reduction that transforms an instance  $\phi$  of the generic  $\mathcal{NP}$ -hard problem of Satisfiability to an instance of Independent set on a  $284T$ -node 3-regular graph. The reduction is such that  $G$  has an independent set of size  $(140 - \epsilon)T$  if  $\phi$  is satisfiable while any independent set of  $G$  has size at most  $(139 + \epsilon)T$  otherwise. Our reduction will use the particular instances in the proof of [6]. Given such an instance with an integer parameter  $T \geq 2$  and any constant  $\epsilon \in (0, 1/2)$ , our reduction constructs in polynomial time a

|          | $N_1$          | $N_2$         | $N_3$          | $N_4$         | $N_5$      |
|----------|----------------|---------------|----------------|---------------|------------|
| Rank     | $\times(2T-2)$ | $\times 140T$ | $\times(2T-2)$ | $\times 140T$ | $\times 1$ |
| 1        | $a_1$          | $a_1$         | $b_1$          | $b_1$         | $b_1$      |
| 2        | $a_2$          | $a_2$         | $b_2$          | $b_2$         | $c_2$      |
| $\vdots$ | $\vdots$       | $\vdots$      | $\vdots$       | $\vdots$      | $\vdots$   |
| $v$      | $a_v$          | $a_v$         | $b_v$          | $a_1$         | $c_v$      |
| $v+1$    | $a_{v+1}$      | $b_1$         | $b_{v+1}$      | $b_v$         | $c_{v+1}$  |
| $v+2$    | $a_{v+2}$      | $a_{v+1}$     | $b_{v+2}$      | $b_{v+1}$     | $c_{v+2}$  |
| $\vdots$ | $\vdots$       | $\vdots$      | $\vdots$       | $\vdots$      | $\vdots$   |
| $ V +1$  | $b_1$          | $a_{ V }$     | $a_1$          | $b_{ V }$     | $a_1$      |
| $ V +2$  | $c_2$          | $c_2$         | $a_2$          | $a_2$         | $b_2$      |
| $ V +3$  | $c_3$          | $c_3$         | $a_3$          | $a_3$         | $b_3$      |
| $\vdots$ | $\vdots$       | $\vdots$      | $\vdots$       | $\vdots$      | $\vdots$   |
| $2 V $   | $c_{ V }$      | $c_{ V }$     | $a_{ V }$      | $a_{ V }$     | $b_{ V }$  |
| $2 V +1$ | $b_2$          | $b_2$         | $c_2$          | $c_2$         | $a_2$      |
| $\vdots$ | $\vdots$       | $\vdots$      | $\vdots$       | $\vdots$      | $\vdots$   |
| $3 V -1$ | $b_{ V }$      | $b_{ V }$     | $c_{ V }$      | $c_{ V }$     | $a_{ V }$  |

**Fig. 3.** The preference profile corresponding to node  $v$ . The corresponding utility profile  $\bar{u}^v$  is such that  $\text{sw}(a_1, \bar{u}^v) = 70T$ ,  $\text{sw}(b_1, \bar{u}^v) = T$ , and  $\text{sw}(a, \bar{u}^v) \leq T$  for alternative  $a$  different than  $a_1$  and  $b_1$ , so that it uniquely induces  $\bar{\sigma}^v$ .

family of  $710T$  utility profiles that uniquely defines a family  $\mathcal{R}$  of corresponding preference profiles for some integer  $T \geq 2$ , so that:

- If the input graph has an independent set of size at least  $(140 - \epsilon)T$ , then there is a scoring vector  $\bar{s}$  such that the total expected social welfare of the winning alternatives under  $f_{\bar{s}}$  over all profiles in  $\mathcal{R}$  is at least  $(39,764 - 69\epsilon)T^2$ .
- If any independent set of  $G$  has size at most  $(139 - \epsilon)T$ , then for every scoring vector  $\bar{s}$ , the total expected social welfare of the winning alternatives under  $f_{\bar{s}}$  over all profiles in  $\mathcal{R}$  is at most  $(39,695 + 69\epsilon)T^2$ .

In this way we will conclude that distinguishing between these two cases for the profiles returned by our reduction is an  $\mathcal{NP}$ -hard problem. In other words, the problem is hard to approximate within a factor better than  $39,764/39,695 \approx 1.00174$ .

**The reduction** Let  $G = (V, E)$  be a 3-regular graph with  $284T$  nodes (and  $426T$  edges). We assume that the nodes of  $V$  are identified by the integers  $1, 2, \dots, |V|$ . We construct  $|V|$  preference profiles corresponding to the nodes in  $V$  and  $|E|$  preference profiles corresponding to the edges of  $E$  (i.e.,  $710T$  profiles in total). Denote by  $\mathcal{R}$  the family of these preference profiles. All profiles have  $284T - 3$  agents and  $3|V| - 1 = 852T - 1$  alternatives. Among them, there are two special alternatives  $a_1$  and  $b_1$  as well as alternatives  $a_i$ ,  $b_i$ , and  $c_i$  for  $i = 2, \dots, |V|$ . The corresponding utility profiles are such that the total utility of each agent for all alternatives is exactly 1 and, furthermore, these utilities are all different so that the corresponding preference profile is uniquely defined.

Specifically, for each node  $v$ , the preference profile  $\bar{\sigma}^v$  is defined as follows. The agents are partitioned into 5 sets  $N_1, \dots, N_5$ :

- Set  $N_1$  consists of  $2T - 2$  agents that rank alternative  $a_i$  at position  $i$  for  $i = 1, \dots, |V|$ , alternative  $b_1$  at position  $|V| + 1$ , and alternatives  $c_i$  and  $b_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .
- Set  $N_2$  consists of  $140T$  agents that rank alternative  $a_i$  at position  $i$  for  $i = 1, \dots, v$ , alternative  $b_1$  at position  $v + 1$ , alternative  $a_i$  at position  $i + 1$  for  $i = v + 1, \dots, |V|$ , and alternatives  $c_i$  and  $b_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .
- Set  $N_3$  consists of  $2T - 2$  agents that rank alternative  $b_i$  at position  $i$  for  $i = 1, \dots, |V|$ , alternative  $a_i$  at position  $|V| + i$  for  $i = 1, \dots, |V|$ , and alternative  $c_i$  at position  $2|V| + i - 1$  for  $i = 2, \dots, |V|$ .
- Set  $N_4$  consists of  $140T$  agents that rank alternative  $a_1$  at position  $v$ , alternative  $b_i$  at position  $i$  for  $i = 1, \dots, v - 1$ , alternative  $b_i$  at position  $i + 1$  for  $i = v + 1, \dots, |V| + 1$ , and alternatives  $a_i$  and  $c_2$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .
- Set  $N_5$  consists of one agent that ranks alternative  $b_1$  first, alternative  $c_i$  at position  $i$  for  $i = 2, \dots, |V|$ , alternative  $a_1$  at position  $|V| + 1$ , and alternatives  $b_i$  and  $c_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .

See Fig. 3 for an example. The corresponding utility profile  $\bar{u}^v$  is such that  $\text{sw}(a_1, \bar{u}^v) = 70T$ ,  $\text{sw}(b_1, \bar{u}^v) = T$ , and  $\text{sw}(a, \bar{u}^v) \leq T$  for alternative  $a \in A \setminus \{a_1, b_1\}$ , so that it uniquely induces  $\bar{\sigma}^v$ . Observe that the  $142T - 2$  agents of  $N_1 \cup N_2$  always rank alternative  $a_1$  first and the  $2T - 1 > T$  agents of  $N_3 \cup N_5$  always rank alternative  $b_1$  first; hence the definition of  $\bar{u}^v$  is clearly feasible and the construction can be implemented in polynomial time.

|          | $N_1$           | $N_2$           | $N_3$      | $N_4$           | $N_5$           |
|----------|-----------------|-----------------|------------|-----------------|-----------------|
| Rank     | $\times(71T-1)$ | $\times(71T-1)$ | $\times 1$ | $\times(71T-1)$ | $\times(71T-1)$ |
| 1        | $a_1$           | $a_1$           | $a_1$      | $b_1$           | $b_1$           |
| 2        | $a_2$           | $a_2$           | $c_2$      | $b_2$           | $b_2$           |
| $\vdots$ | $\vdots$        | $\vdots$        | $\vdots$   | $\vdots$        | $\vdots$        |
| $v$      | $b_1$           | $a_v$           | $c_v$      | $b_v$           | $b_v$           |
| $v+1$    | $a_v$           | $a_{v+1}$       | $c_{v+1}$  | $a_1$           | $b_{v+1}$       |
| $v+2$    | $a_{v+1}$       | $a_{v+2}$       | $c_{v+2}$  | $b_{v+1}$       | $b_{v+2}$       |
| $\vdots$ | $\vdots$        | $\vdots$        | $\vdots$   | $\vdots$        | $\vdots$        |
| $w$      | $a_{w-1}$       | $b_1$           | $c_w$      | $b_{w-1}$       | $b_w$           |
| $w+1$    | $a_w$           | $a_w$           | $c_{w+1}$  | $b_w$           | $a_1$           |
| $w+2$    | $a_{w+1}$       | $a_{w+1}$       | $c_{w+2}$  | $b_{w+1}$       | $b_{w+1}$       |
| $\vdots$ | $\vdots$        | $\vdots$        | $\vdots$   | $\vdots$        | $\vdots$        |
| $ V +1$  | $a_{ V }$       | $a_{ V }$       | $b_1$      | $b_{ V }$       | $b_{ V }$       |
| $ V +2$  | $b_2$           | $b_2$           | $a_2$      | $c_2$           | $c_2$           |
| $ V +3$  | $b_3$           | $b_3$           | $a_3$      | $c_3$           | $c_3$           |
| $\vdots$ | $\vdots$        | $\vdots$        | $\vdots$   | $\vdots$        | $\vdots$        |
| $2 V $   | $b_{ V }$       | $b_{ V }$       | $b_{ V }$  | $c_{ V }$       | $c_{ V }$       |
| $2 V +1$ | $c_2$           | $c_2$           | $b_2$      | $a_2$           | $a_2$           |
| $\vdots$ | $\vdots$        | $\vdots$        | $\vdots$   | $\vdots$        | $\vdots$        |
| $3 V -1$ | $c_{ V }$       | $c_{ V }$       | $b_{ V }$  | $a_{ V }$       | $a_{ V }$       |

**Fig. 4.** The preference profile  $\vec{\sigma}^e$  corresponding to edge  $e = (v, w)$  with  $v < w$ . The corresponding utility profile  $\vec{u}^e$  is such that  $\text{sw}(a_1, \vec{u}^e) = 70T$ ,  $\text{sw}(b_1, \vec{u}^e) = T$ , and  $\text{sw}(a, \vec{u}^e) \leq T$  for alternative  $a$  different than  $a_1$  and  $b_1$ , so that it uniquely induces  $\vec{\sigma}^e$ .

For each edge  $e = (v, w)$  with  $v < w$ , the preference profile  $\vec{\sigma}^e$  is defined as follows. The agents are partitioned into 5 sets  $N_1, \dots, N_5$ :

- Set  $N_1$  consists of  $71T - 1$  agents that rank alternative  $b_1$  at position  $v$ , alternative  $a_i$  at position  $i$  for  $i = 1, \dots, v - 1$ , alternative  $a_i$  at position  $i + 1$  for  $i = v, \dots, |V|$ , and alternatives  $b_i$  and  $c_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .
- Set  $N_2$  consists of  $71T - 1$  agents that rank alternative  $a_i$  at position  $i$  for  $i = 1, \dots, w - 1$ , alternative  $b_1$  at position  $w$ , alternative  $a_i$  at position  $i + 1$  for  $i = w + 1, \dots, |V|$ , and alternatives  $b_i$  and  $c_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .
- Set  $N_3$  consists of one agent that ranks alternative  $a_1$  first, alternative  $c_i$  at position  $i$  for  $i = 2, \dots, |V|$ , alternative  $b_1$  at position  $|V| + 1$ , and alternatives  $a_i$  and  $b_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .
- Set  $N_4$  consists of  $71T - 1$  agents that rank alternative  $b_i$  at position  $i$  for  $i = 1, \dots, v$ , alternative  $a_1$  at position  $v + 1$ , alternative  $b_i$  at position  $i + 1$  for  $i = v + 1, \dots, |V|$ , and alternatives  $c_i$  and  $a_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .
- Set  $N_5$  consists of  $71T - 1$  agents that rank alternative  $b_i$  at position  $i$  for  $i = 1, \dots, w$ , alternative  $a_1$  at position  $w + 1$ , alternative  $b_i$  at position  $i + 1$  for  $i = w + 1, \dots, |V|$ , and alternatives  $c_i$  and  $a_i$  at positions  $|V| + i$  and  $2|V| + i - 1$ , respectively, for  $i = 2, \dots, |V|$ .

See Fig. 4 for an example. The corresponding utility profile  $\vec{u}^e$  is such that  $\text{sw}(a_1, \vec{u}^e) = 70T$ ,  $\text{sw}(b_1, \vec{u}^e) = T$ , and  $\text{sw}(a, \vec{u}^e) \leq T$  for  $a \in A \setminus \{a_1, b_1\}$ , so that it uniquely induces  $\vec{\sigma}^e$ . Observe that the  $71T$  agents of  $N_1 \cup N_3$  always rank alternative  $a_1$  first and the  $142T - 2$  agents of  $N_4 \cup N_5$  always rank alternative  $b_1$  first; hence the definition of  $\vec{u}^e$  is clearly feasible and the construction can be implemented in polynomial time.

*Proof of correctness* We proceed with some definitions. For a preference profile  $\vec{\sigma} \in \mathcal{R}$ , we use the notation  $\vec{u}(\vec{\sigma})$  to denote its corresponding utility profile. Also, define

$$\text{SW}(\vec{s}) = \sum_{\vec{\sigma} \in \mathcal{R}} \mathbb{E}[\text{sw}(f_{\vec{s}}(\vec{\sigma}), \vec{u}(\vec{\sigma}))]$$

to be the total expected social welfare of the winning alternative under score  $\vec{s}$ . We remark that in order to simplify the proof (and avoid reasoning about monotonicity of the score vector entries), we consider each scoring rule as computing a linear combination of  $k$ -approval scores for  $k = 1, \dots, 3|V| - 1$ . The non-negative entries of the score vector are the multipliers of the corresponding approval score in this linear combination. In particular, denote by  $\text{app}_k(a, \vec{\sigma})$  the number of agents that rank alternative  $a$  in some of the first  $k$  positions in the preference profile  $\vec{\sigma}$  (i.e., the  $k$ -approval score of alternative  $a$ ). Then, the score under the score vector  $\vec{s}$  is defined as  $\text{sc}_{\vec{s}}(a, \vec{\sigma}) = \sum_{k=1}^{3|V|-1} \text{app}_k(a, \vec{\sigma}) \cdot s_k$ . Clearly, this definition is equivalent to the standard one for scoring rules (by adjusting the score vector entries appropriately).



The next three claims provide information about the winning alternatives in each profile depending on the entries of the score vector  $\vec{s}$ .

**Claim B.2.** Let  $\vec{s}$  be a score vector. In every profile of  $\mathcal{R}$ , one of the alternatives  $a_1$  and  $b_1$  has the highest score under  $\vec{s}$ . If an alternative  $a \notin \{a_1, b_1\}$  has the highest score under  $\vec{s}$  in some profile of  $\mathcal{R}$  then  $s_k = 0$  for  $k = 1, \dots, |V|$  and alternatives  $a_1$  and  $b_1$  have the highest score as well.

**Proof.** Observe that in each profile, the  $284T - 3$  agents have one of the alternatives  $a_1$  and  $b_1$  in their top position. Furthermore, all agents rank alternatives  $a_1$  and  $b_1$  in one of the first  $|V| + 1$  positions. Hence, the sum of the scores of these two alternatives is at least

$$(284T - 3) \sum_{k=1}^{|V|} s_k + 2(284T - 3) \sum_{k=|V|+1}^{3|V|-1} s_k.$$

For every other alternative  $a$ , the number of agents that rank it in some of the first  $|V|$  positions is at most  $142T - 2$ . Hence, the score of such an alternative is at most

$$(142T - 2) \sum_{k=1}^{|V|} s_k + (284T - 3) \sum_{k=|V|+1}^{3|V|-1} s_k.$$

Clearly, if  $s_k > 0$  for some  $k \in \{1, \dots, |V|\}$ , then at least one of the alternatives  $a_1$  and  $b_1$  has score strictly higher than  $a$ . Otherwise, the alternatives  $a_1$  and  $b_1$  still have the highest score (possibly together with some alternative  $a \notin \{a_1, b_1\}$ ).  $\square$

**Claim B.3.** Let  $v \in V$  be a node of  $G$ . Given a score vector  $\vec{s}$ , define

$$\text{Wit}_v(\vec{s}) = (140T - 1)s_v - \sum_{k \in V \setminus \{v\}} s_k.$$

Then

- Alternative  $a_1$  is the sole winner under  $\vec{s}$  in profile  $\vec{\sigma}^v$  if and only if  $\text{Wit}_v(\vec{s}) > 0$ .
- Alternative  $b_1$  is the sole winner under  $\vec{s}$  in profile  $\vec{\sigma}^v$  if and only if  $\text{Wit}_v(\vec{s}) < 0$ .
- Alternatives  $a_1$  and  $b_1$  are tied as highest-scoring alternatives under  $\vec{s}$  in profile  $\vec{\sigma}^v$  if and only if  $\text{Wit}_v(\vec{s}) = 0$ .

**Proof.** Consider a node  $v \in V$  and its corresponding preference profile  $\vec{\sigma}^v$ . By the definition of  $\vec{\sigma}^v$ , we have that the difference  $\text{app}_k(a_1, \vec{\sigma}^v) - \text{app}_k(b_1, \vec{\sigma}^v)$  of the  $k$ -approval scores of alternatives  $a_1$  and  $b_1$  is  $-1$  for  $k = 1, \dots, v - 1, v + 1, \dots, |V|$ ,  $140T - 1$  for  $k = v$ , and  $0$  for  $k = |V| + 1, \dots, 3|V| - 1$ . Therefore

$$\begin{aligned} \text{sc}_{\vec{s}}(a_1, \vec{\sigma}^v) - \text{sc}_{\vec{s}}(b_1, \vec{\sigma}^v) &= \sum_{k=1}^{3|V|-1} (\text{app}_k(a_1, \vec{\sigma}^v) - \text{app}_k(b_1, \vec{\sigma}^v))s_k \\ &= (140T - 1)s_v - \sum_{k \in V \setminus \{v\}} s_k \\ &= \text{Wit}_v(\vec{s}). \end{aligned}$$

Now the claim follows using [Claim B.2](#), because  $\text{Wit}_v(\vec{s}) \neq 0$  implies that  $s_k > 0$  for some  $k \in \{1, \dots, |V|\}$  and, hence, no alternative other than  $a_1$  and  $b_1$  can have maximum score.  $\square$

**Claim B.4.** Let  $e = (v, w) \in E$  be an edge of  $G$ . Given a score vector  $\vec{s}$ , define

$$\text{Wit}_e(\vec{s}) = (71T - 2)s_v + (71T - 2)s_w - \sum_{k \in V \setminus \{v, w\}} s_k.$$

Then

- Alternative  $a_1$  is the sole winner under  $\vec{s}$  in profile  $\vec{\sigma}^e$  if and only if  $\text{Wit}_e(\vec{s}) < 0$ .
- Alternative  $b_1$  is the sole winner under  $\vec{s}$  in profile  $\vec{\sigma}^e$  if and only if  $\text{Wit}_e(\vec{s}) > 0$ .
- Alternatives  $a_1$  and  $b_1$  are tied as highest-scoring alternatives under  $\vec{s}$  in profile  $\vec{\sigma}^e$  if and only if  $\text{Wit}_e(\vec{s}) = 0$ .

**Proof.** Consider an edge  $e = (v, w) \in E$  and its corresponding preference profile  $\vec{\sigma}^e$ . By the definition of  $\vec{\sigma}^e$ , the difference  $\text{app}_k(a_1, \vec{\sigma}^e) - \text{app}_k(b_1, \vec{\sigma}^e)$  of the  $k$ -approval scores of alternatives  $a_1$  and  $b_1$  is 1 for  $k = 1, \dots, v-1, v+1, \dots, w-1, w+1, \dots, |V|$ ,  $-71T+2$  for  $k = v$  or  $k = w$ , and 0 for  $k = |V|+1, \dots, 3|V|-1$ . We have

$$\begin{aligned} \text{sc}_{\vec{s}}(a_1, \vec{\sigma}^e) - \text{sc}_{\vec{s}}(b_1, \vec{\sigma}^e) &= \sum_{k=1}^{3|V|-1} (\text{app}_k(a_1, \vec{\sigma}^e) - \text{app}_k(b_1, \vec{\sigma}^e))s_k \\ &= -(71T-2)s_v - (71T-2)s_w + \sum_{k \in V \setminus \{v, w\}} s_k \\ &= -\text{Wit}_e(\vec{s}). \end{aligned}$$

As before, the claim now follows directly from [Claim B.2](#).  $\square$

The correctness of our reduction will follow by the next two lemmas.

**Lemma B.5.** For every independent set  $I$  of size  $K \in [71T, 140T)$  in  $G$ , there is a scoring protocol  $\vec{s}$  such that  $\text{SW}(\vec{s}) \geq 30,104T^2 + 69TK$ .

**Proof.** For every node  $v$  of  $G$ , we set  $s_v = 1/K$  if  $v \in I$ , and  $s_v = 0$  otherwise. We also set  $s_v = 0$  for every  $v$  in  $\{|V|+1, \dots, 3|V|-1\}$ . First, consider the preference profile  $\vec{\sigma}^v$  for some node  $v$  of  $G$ . We have

$$\text{Wit}_v(\vec{s}) = (140T-1)s_v - \sum_{k \in V \setminus \{v\}} s_k = 140Ts_v - \sum_{k \in V} s_k = 140Ts_v - 1.$$

Now, observe that the right-hand side in the above equality is strictly positive if  $v \in I$  (since  $s_k = 1/K$  and  $K < 140T$ ) and strictly negative if  $v \notin I$  (since  $s_v = 0$ ). Hence, by [Claims B.2 and B.3](#), we conclude that alternative  $a_1$  (respectively,  $b_1$ ) is the sole winner in profile  $\vec{\sigma}^v$  if  $v \in I$  (respectively, if  $v \notin I$ ).

Now, consider the preference profile  $\vec{\sigma}^e$  for some edge  $e = (v, w)$  of  $G$ . We have

$$\begin{aligned} \text{Wit}_e(\vec{s}) &= (71T-2)s_v + (71T-2)s_w - \sum_{k \in V \setminus \{v, w\}} s_k \\ &= (71T-1)(s_v + s_w) - \sum_{k \in V} s_k \\ &= (71T-1)(s_v + s_w) - 1 \\ &\leq \frac{71T-1}{K} - 1 \\ &< 0. \end{aligned}$$

The first inequality holds since at most one of the adjacent nodes  $v$  and  $w$  belongs in the independent set  $I$  and the second one since  $K \geq 71T$ . By [Claims B.2 and B.4](#), we conclude that alternative  $a_1$  is the sole winner under  $\vec{s}$  in profile  $\vec{\sigma}^e$ , for every edge  $e$  of  $E$ .

Summing the expected social welfare of the winner under  $\vec{s}$  over all profiles, we have

$$\begin{aligned} \text{SW}(\vec{s}) &= \sum_{v \in I} \text{sw}(a_1, \vec{u}^v) + \sum_{v \in V \setminus I} \text{sw}(b_1, \vec{u}^v) + \sum_{e \in E} \text{sw}(a_1, \vec{u}^e) \\ &= 70TK + T(|V|-K) + 70T|E| \\ &= 30,104T^2 + 69TK \end{aligned}$$

as desired.  $\square$

**Lemma B.6.** Let  $\vec{s}$  be a score vector such that  $\text{SW}(\vec{s}) \geq 30,104T^2 + 69TK$ . Then  $G$  has an independent set of size  $K$ .

**Proof.** First, we will prove that  $s_k > 0$  for some  $k \in \{1, \dots, |V|\}$ . Indeed, if this not the case, by [Claims B.3 and B.4](#),  $a_1$  and  $b_1$  are both highest-scoring alternatives (possibly, together with some other alternative). Then, the expected social welfare of the winner under  $\vec{s}$  in any profile  $\vec{\sigma} \in \mathcal{R}$  is at most  $\frac{1}{2}(\text{sw}(a_1, \vec{u}(\vec{\sigma})) + \text{sw}(b_1, \vec{u}(\vec{\sigma}))) = 71T/2$ . In total, we have that  $\text{SW}(\vec{s}) \leq (|E| + |V|)71T/2 = 25,205T^2$ , contradicting the assumption of the lemma.

Furthermore, we can assume that  $\text{Wit}_v(\vec{s}) \neq 0$  for every node  $v \in V$  and  $\text{Wit}_e(\vec{s}) \neq 0$  for every edge  $e \in E$ . If this is not the case, we can transform the score vector in order to get another one that satisfies these conditions as well as the assumption of the lemma. This can be done as follows. Let  $\eta$  be the minimum non-zero value of the quantities  $|\text{Wit}_v(\vec{s})|$  and  $|\text{Wit}_e(\vec{s})|$

over all nodes  $v \in V$  and edges  $e \in E$ . Now, let  $v \in V$  be a node with  $s_v > 0$  (by the argument above, such a node certainly exists). Define two score vectors  $\vec{s}'$  and  $\vec{s}''$  as follows: set  $s'_v = s_v - \frac{\eta}{200T}$  and  $s'_v = s_v + \frac{\eta}{200T}$ ; all other scores in  $\vec{s}'$  and  $\vec{s}''$  are the same as those of  $\vec{s}$ . Observe that for every non-zero  $\text{Wit}_v(\vec{s})$ ,  $\text{Wit}_v(\vec{s}')$  and  $\text{Wit}_v(\vec{s}'')$  are non-zero and have the same sign as  $\text{Wit}_v(\vec{s})$ . Similarly, for every non-zero  $\text{Wit}_e(\vec{s})$ ,  $\text{Wit}_e(\vec{s}')$  and  $\text{Wit}_e(\vec{s}'')$  are non-zero and have the same sign as  $\text{Wit}_e(\vec{s})$ . This means that each such profile  $\vec{\sigma}^v$  or  $\vec{\sigma}^e$  has the same sole winner under scores  $\vec{s}$ ,  $\vec{s}'$ , and  $\vec{s}''$ . For every node  $v$  such that  $\text{Wit}_v(\vec{s}) = 0$ ,  $\text{Wit}_v(\vec{s}')\text{Wit}_v(\vec{s}'') < 0$ , i.e., they are non-zero and have different signs. Similarly, for every edge  $e$  such that  $\text{Wit}_e(\vec{s}) = 0$ ,  $\text{Wit}_e(\vec{s}')\text{Wit}_e(\vec{s}'') < 0$ . Hence, alternatives  $a_1$  and  $b_1$  are the only winners in profiles  $\vec{\sigma}^v$  and  $\vec{\sigma}^e$  under  $\vec{s}'$  and  $\vec{s}''$  while they both have the same highest score under  $\vec{s}$ . Clearly,  $\text{SW}(\vec{s}') + \text{SW}(\vec{s}'') \geq 2\text{SW}(\vec{s})$  which implies that one of the score vectors  $\vec{s}'$  or  $\vec{s}''$  has the desired properties.

So, in the following, we assume that  $\text{Wit}_v(\vec{s}) \neq 0$  and  $\text{Wit}_e(\vec{s}) \neq 0$  for every node  $v \in V$  and every edge  $e \in E$ . Define  $H = \{e \in E : \text{Wit}_e(\vec{s}) < 0\}$  and  $I = \{v \in V : \text{Wit}_v(\vec{s}) > 0\}$ . Hence, the sole winner under  $\vec{s}$  in profile  $\vec{\sigma}^v$  is alternative  $a_1$  if  $e \in H$  and alternative  $b_1$  otherwise. Similarly, the sole winner in profile  $\vec{\sigma}^v$  is alternative  $a_1$  if  $v \in I$  and alternative  $b_1$  otherwise. It follows that

$$\begin{aligned} \text{SW}(\vec{s}) &= \sum_{v \in I} \text{sw}(a_1, \vec{u}^v) + \sum_{v \in V \setminus I} \text{sw}(b_1, \vec{u}^v) + \sum_{e \in H} \text{sw}(a_1, \vec{u}^e) + \sum_{e \in E \setminus H} \text{sw}(b_1, \vec{u}^e) \\ &= 70T|I| + T(|V| - |I|) + 70T|H| + T(|E| - |H|) \\ &= T|V| + 69T|I| + T|E| + 69T|H| \end{aligned}$$

Also, observe that the assumption of the lemma implies that

$$\begin{aligned} \text{SW}(\vec{s}) &\geq 30,104T^2 + 69TK \\ &= 70T|E| + T|V| + 69TK. \end{aligned}$$

By the above two equations, we obtain that  $|I| \geq K + |E| - |H|$ . Now, consider the subgraph  $G'$  of  $G$  induced by the nodes of  $I$  and let  $E(G')$  be its set of edges. Since for each edge  $e = (v, w)$  of  $E(G')$ , nodes  $v$  and  $w$  belong to  $I$ , it holds that

$$\begin{aligned} \text{Wit}_e(\vec{s}) &= (71T - 2)s_v + (71T - 2)s_w - \sum_{k \in V \setminus \{v, w\}} s_k \\ &= \frac{1}{2} (\text{Wit}_v(\vec{s}) + \text{Wit}_w(\vec{s})) + (T - 1)(s_v + s_w) \\ &> 0, \end{aligned}$$

i.e.,  $E(G') \subseteq |E \setminus H|$ . Since  $|I| - |E(G')|$  is a lower bound on the number of connected components in  $G'$ , the graph  $G'$  (and, consequently, graph  $G$ ) has an independent set of size  $|I| - |E(G')| \geq |I| - |E \setminus H| \geq K$ .  $\square$

The properties of our reduction follow from the last two lemmas. If  $G$  has an independent set of size  $(140 - \epsilon)T$ , using Lemma B.5, we obtain that there is a score vector  $\vec{s}$  such that  $\text{SW}(\vec{s}) \geq (39764 - 69\epsilon)T^2$ . On the other hand, if any independent set in  $G$  has size at most  $(139 + \epsilon)T$ , then  $\text{SW}(\vec{s}) \leq (39,695 + 69\epsilon)T^2$ . Assuming otherwise, we obtain a contradiction using Lemma B.6.  $\square$

## References

- [1] N. Alon, Tools from higher algebra, in: R.L. Graham, M. Grötschel, L. Lovász (Eds.), *Handbook of Combinatorics*, vol. 2, MIT Press, 1996, pp. 287–324.
- [2] M. Anthony, P.L. Bartlett, *Neural Network Learning: Theoretical Foundations*, Cambridge University Press, 1999.
- [3] J. Apesteguía, M.A. Ballester, R. Ferrer, On the justice of decision rules, *Rev. Econ. Stud.* 78 (2011) 1–16.
- [4] Y. Azrieli, S. Kim, Pareto efficiency and weighted majority rules, *Int. Econ. Rev.* 55 (2014) 1067–1088.
- [5] J. Bartholdi, C.A. Tovey, M.A. Trick, Voting schemes for which it can be difficult to tell who won the election, *Soc. Choice Welf.* 6 (1989) 157–165.
- [6] P. Berman, M. Karpinski, On some tighter inapproximability results, in: *Proceedings of the 26th International Colloquium on Automata, Languages and Programming, ICALP*, 1999, pp. 200–209.
- [7] R.F. Bordley, A pragmatic method for evaluating election schemes through simulation, *Am. Polit. Sci. Rev.* 77 (1983) 123–141.
- [8] S.J. Brams, P.C. Fishburn, *Approval Voting*, 2nd edition, Springer, 2007.
- [9] D. Brazunas, C. Boutilier, Assessing regret-based preference elicitation with the UTPREF recommendation system, in: *Proceedings of the 11th ACM Conference on Economics and Computation, EC*, 2010, pp. 219–228.
- [10] I. Caragiannis, A.D. Procaccia, Voting almost maximizes social welfare despite limited communication, *Artif. Intell.* 175 (9–10) (2011) 1655–1671.
- [11] V. Conitzer, M. Rognlie, L. Xia, Preference functions that score rankings and maximum likelihood estimation, in: *Proceedings of the 21st International Joint Conference on Artificial Intelligence, IJCAI*, 2009, pp. 109–115.
- [12] E. Elkind, P. Faliszewski, A. Slinko, On distance rationalizability of some voting rules, in: *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge, TARK*, 2009, pp. 108–117.
- [13] P. Faliszewski, A.D. Procaccia, AI's war on manipulation: are we winning?, *AI Mag.* 31 (4) (2010) 53–64.
- [14] K. Goldberg, T. Roeder, D. Gupta, C. Perkins, Eigentaste: a constant time collaborative filtering algorithm, *Inf. Retr.* 4 (2) (2001) 133–151.
- [15] J. Harsanyi, Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *J. Polit. Econ.* 63 (4) (1955) 309–321.
- [16] E. Hemaspaandra, L.A. Hemaspaandra, Dichotomy for voting systems, *J. Comput. Syst. Sci.* 73 (1) (2007) 73–83.
- [17] E. Hemaspaandra, L.A. Hemaspaandra, J. Rothe, Exact analysis of Dodgson elections: Lewis Carroll's 1876 voting system is complete for parallel access to NP, *J. ACM* 44 (6) (1997) 806–825.

- [18] C. Hillinger, The case for utilitarian voting, Discussion papers in econ. No. 653, U. Munich, 2005.
- [19] M. Kalech, S. Kraus, G.A. Kaminka, C.V. Goldman, Practical voting rules with partial information, *Auton. Agents Multi-Agent Syst.* 22 (1) (2011) 151–182.
- [20] J.J. Louviere, D.A. Hensher, J.D. Swait, *Stated Choice Methods: Analysis and Application*, Cambridge University Press, 2000.
- [21] T. Lu, C. Boutilier, The unavailable candidate model: a decision-theoretic view of social choice, in: *Proceedings of the 11th ACM Conference on Economics and Computation, EC, 2010*, pp. 263–274.
- [22] T. Lu, C. Boutilier, Robust approximation and incremental elicitation in voting protocols, in: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence, IJCAI, 2011*, pp. 287–293.
- [23] D. Majumdar, A. Sen, Bayesian incentive compatible voting rules, *Econometrica* 72 (2) (2004) 523–540.
- [24] S. Merrill, A comparison of efficiency of multicandidate electoral systems, *Am. J. Polit. Sci.* 28 (1984) 23–48.
- [25] N. Nisan, Introduction to mechanism design (for computer scientists), in: N. Nisan, T. Roughgarden, É. Tardos, V. Vazirani (Eds.), *Algorithmic Game Theory*, Cambridge University Press, 2007, chapter 9.
- [26] M. Pivato, Condorcet meets Bentham, Thema working paper No. 2014-17, 2014.
- [27] A.D. Procaccia, J.S. Rosenschein, The distortion of cardinal preferences in voting, in: *Proceedings of the 10th International Workshop on Cooperative Information Agents, CIA*, in: *Lecture Notes in Computer Science (LNCS)*, vol. 4149, Springer, 2006, pp. 317–331.
- [28] A.D. Procaccia, A. Zohar, Y. Peleg, J.S. Rosenschein, The learnability of voting rules, *Artif. Intell.* 173 (12–13) (2009) 1133–1149.
- [29] M. Regenwetter, B. Grofman, A.A.J. Marley, I. Tsetlin, *Behavioral Social Choice: Probabilistic Models, Statistical Inference, and Applications*, Cambridge University Press, 2006.
- [30] A. Slinko, How large should a coalition be to manipulate an election?, *Math. Soc. Sci.* 47 (3) (2004) 289–293.
- [31] I. Tsetlin, M. Regenwetter, B. Grofman, The impartial culture maximizes the probability of majority cycles, *Soc. Choice Welf.* 21 (2003) 387–398.
- [32] J. von Neumann, O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1944.
- [33] E.H. Warren, Lower bounds for approximation by nonlinear manifolds, *Trans. Am. Math. Soc.* 133 (1) (1968) 167–178.
- [34] R.J. Weber, Reproducing voting systems, Cowles Foundation discussion paper No. 498, 1978.
- [35] L. Xia, V. Conitzer, Generalized scoring rules and the frequency of coalitional manipulability, in: *Proceedings of the 9th ACM Conference on Economics and Computation, EC, 2008*, pp. 109–118.
- [36] H.P. Young, Social choice scoring functions, *SIAM J. Appl. Math.* 28 (4) (1975) 824–838.
- [37] H.P. Young, Condorcet's theory of voting, *Am. Polit. Sci. Rev.* 82 (4) (1988) 1231–1244.