Computing Pareto Optimal Committees

Haris Aziz, Jérôme Lang, Jérôme Monnot

Computing Pareto Optimal Committees

Haris Aziz
NICTA and UNSW
Sydney, Australia
haris.aziz@data61.csiro.au

Jérôme Lang and Jérôme Monnot
LAMSADE, Université Paris-Dauphine
Paris, France
{lang, jerome.monnot}@lamsade.dauphine.fr

Abstract

Selecting a set of alternatives based on the preferences of agents is an important problem in committee selection and beyond. Among the various criteria put forth for desirability of a committee, Pareto optimality is a minimal and important requirement. As asking agents to specify their preferences over exponentially many subsets of alternatives is practically infeasible, we assume that each agent specifies a weak order on single alternatives, from which a preference relation over subsets is derived using some preference extension. We consider four prominent extensions (responsive, leximax, best, and worst). For each of them, we consider the corresponding Pareto optimality notion, and we study the complexity of computing and verifying Pareto optimal outcomes. We also consider strategic issues: for three of the set extensions, we present linear-time, Pareto optimal and strategyproof algorithms that work even for weak preferences.

1 Introduction

Pareto optimality is a central concept in economics and has been termed the *"single most important tool of normative economic analysis"* [Moulin, 2003]. An outcome is Pareto optimal if there does not exist another outcome that all agents like at least as much and at least one agent strictly prefers. Although Pareto optimality has been considered extensively in single-winner voting and other social choice settings such as fair division or hedonic games, it has received only little attention in *multiwinner voting*, in which the outcomes are *sets* of alternatives. Multiwinner voting applies to selecting a set of plans or a committee, hiring team members, movie recommendations, and more. For convenience, we use the terminology "committee" even if our results have an impact far beyond committee elections.

In *single-winner voting* setting, agents express preferences over alternatives and a single alternative is selected. Pareto optimality in this context is straightforward to define, achieve, and verify. In *multiwinner voting*, a well-known difficulty is that it is unrealistic to assume that agents will report preferences over all possible committees, since there are an exponential number of them. For this reason, most approaches assume that they only report a small part of their preferences, and that some *extension principle* is used to induce a preference over all possible subsets from this 'small input' over single alternatives [Barberà *et al.*, 2004]. Such preference extensions are also widely used in other social choice settings such as fair division or matching. The most two widely used choices of 'small inputs' in multiwinner voting are *rankings (linear orders) over alternatives* and *sets of approved alternatives*. In this paper we make a choice that generalizes both of them: agents report *weak orders over single alternatives*. Then we consider four prominent preference extension principles: the *responsive* extension, where a set of alternatives S is at least as preferred as a set of alternatives T if S is obtained from T by repeated replacements of an alternative by another alternative which is at least as preferred; the *optimistic*, or *'best'* (respectively *pessimistic*, or *'worst'*) extension, which orders subsets of alternatives according to their most (respectively, least) preferred element; and the *leximax* extension, a lexicographic refinement of the optimistic extension.

The responsive extension [Barberà *et al.*, 2004; Roth and Sotomayor, 1990] can be seen as the ordinal counterpart of additivity. The leximax extension has been considered in various papers [Bossert, 1995; Lang *et al.*, 2012; Klamler *et al.*, 2012]. The 'best' set extension has been considered in a number of approaches such as full proportional representation [Chamberlin and Courant, 1983; Monroe, 1995] and other committee voting settings [Elkind *et al.*, 2015]. The 'worst' set extension, also used by Klamler *et al.* [2012] and Skowron *et al.* [2015b], captures settings where the impact of a bad alternative in the selection overwhelms the benefits of good alternatives: for instance, when the decision about a crucial issue will be made by *one* of the members of the committee but the agent ignores which one and is risk-averse; or the case of a parent's preferences over a set of movies to be watched by a child. The 'best' and 'worst' set extensions have been used in coalition formation [Aziz and Savani, 2016; Cechlárová, 2008].

Although set extensions have been implicitly or explicitly considered in multiwinner voting, most of the computational work has dealt with *specific* voting rules (see the related work section). Instead, we concentrate on Pareto optimality, consider the computation and verification of Pareto optimal committees, as well as the existence of a polynomial-time and strategyproof algorithm that returns Pareto optimal outcomes.

Set Extension	Computation	Verification
Responsive (RS)	in P^{IC} (Th. 5)	coNP-C (Th. 2), W[2]-hard in P—dich. prefs and $tw \leq 2$ (Th. 4)
Leximax (LX)	in P^{IC} (Th. 6)	coNP-C (Th. 1), W[2]-hard
Best ($\mathcal{B}$)	NP-hard (Th. 8) in P for strict prefs	coNP-C, W[2]-hard (Th. 7)
Worst ($\mathcal{W}$)	in P^{IC} (Th. 10)	in P (Th.9)

Table 1: Complexity of computing and verifying Pareto optimal committees. P^{IC} (coined by Christos Papadimitriou in a seminar at Simons Institute in 2015) indicates a class of problems in which agents provide the input and the problems admit a strategyproof and polynomial-time algorithm.

Contributions We consider Pareto optimality with respect to the four aforementioned preference set extensions. We present various connections between the Pareto optimality notions. For each of the notions, we undertake a detailed study of complexity of computing and verifying Pareto optimal outcomes. Table 1 summarizes the complexity results.

An important take-home message of the results is that testing Pareto optimality or obtaining Pareto improvements over status-quo committees is computationally hard even though computing *some* Pareto optimal committee is easy. For responsive and leximax extensions we give a *complete* characterization of the complexity of testing Pareto optimality when preferences are dichotomous and the size of top equivalence class is two: unless P = NP, Pareto optimality can be tested in polynomial time if and only if the size of the first equivalence classes is at most two. For the 'best' extension, we show that even computing a Pareto optimal outcome is NP-hard. Another interesting contrast with the responsive set extension is that even when preferences are dichotomous and the size of top equivalence class is two, testing Pareto optimality is coNP-complete. In contrast to the other extensions, for the 'worst' extension, both problems of computing and verifying Pareto optimal outcomes admit polynomial-time algorithms.

We also consider the requirement of strategyproofness on top of Pareto optimality. We show that there exist linear-time Pareto optimal and strategyproof algorithms for committee voting even for weak preferences for three of the four set extensions. The algorithms can be considered as careful adaptations of serial dictatorship for committee voting.

2 Related Work

A first related stream of work involves studying *specific committee elections rules* from a computational point of view (generally with little or no focus on Pareto optimality). Our focus on determining whether a committee is Pareto optimal or on finding a Pareto optimal committee, is in some sense orthogonal to the study of committee election rules. The simplest (and most widely used) rules for electing a committee, called *best-k* rules, compute a score for each alternative based on the ranks, and the alternatives with the best k scores are elected [Elkind *et al.*, 2014; Faliszewski *et al.*, 2016]. Scoring-based extension principles have also been used by Darmann [2013]. Note that the output of a best k-rule is obviously Pareto-optimal for the preferences induced by this scoring function, but not necessarily with respect to other set extensions.

Klamler *et al.* [2012] compute optimal committees under a weight constraint for a *single* agent (therefore optimality is equivalent to Pareto optimality), using several preference extensions including 'worst', 'best', and leximax.

The 'best' ($\mathcal{B}$) extension principle has been used in a number of papers on committee elections by *full proportional representation*, starting with [Chamberlin and Courant, 1983] and studied from a computational point of view in a long series of papers (*e.g.*, [Procaccia *et al.*, 2008; Lu and Boutilier, 2011; Betzler *et al.*, 2013; Skowron *et al.*, 2015a; Elkind and Ismaili, 2015]. These rules obviously output Pareto optimal committees *for* $\mathcal{B}$, but not necessarily for other extensions.

Some works are based on the *Hamming extension*. Each agent specifies his ideal committee and he prefers committees with less Hamming distance from the ideal committee. The Hamming distance notion can be used to define specific rules such as *minimax approval voting* [Brams *et al.*, 2007], which selects the committee minimizing the maximum Hamming distance for the agents. Although the output of minimax approval voting is not always Pareto-optimal for the Hamming extension, there are good Pareto-optimal approximations of it [Caragiannis *et al.*, 2010]. Note that for dichotomous preferences, the Hamming extension coincides with the responsive and the leximax extensions, therefore our computational results for responsive set extension for dichotomous preferences also hold for the Hamming and leximax extensions.

A second line of work concerns understanding the classes of rules that result in Pareto optimal outcomes. Most works along this line bear on a different type of committee elections, called *designated-seat voting*, where candidates must declare the seat they contest [Benoît and Kornhauser, 2010].[1] Results about the existence or non-existence of Pareto optimal rules have been presented [Özkal-Sanver and Sanver, 2006; Benoît and Kornhauser, 2010; Cuhadaroğlu and Lainé, 2012].

3 Setup

We consider a set of agents $N = \{1, \ldots, n\}$, a set of alternatives $A = \{a_1, \ldots, a_m\}$ and a preference profile $\succsim = (\succsim_1, \ldots, \succsim_n)$ such that each $\succsim_i$ is a complete and transitive relation over A. We write $a \succsim_i b$ to denote that agent i values a at least as much as b and use $\succ_i$ for the strict part of $\succsim_i$, i.e., $a \succ_i b$ iff $a \succsim_i b$ but not $b \succsim_i a$. Finally, $\sim_i$ denotes i's indifference relation, i.e., $a \sim_i b$ iff both $a \succsim_i b$ and $b \succsim_i a$.

The relation $\succsim_i$ results in equivalence classes $E_i^1, E_i^2, \ldots, E_i^{k_i}$ for some k_i such that $a \succ_i a'$ if $a \in E_i^l$ and $a' \in E_i^{l'}$ for some $l < l'$. We will use these equivalence classes to represent the preference relation of an agent as a preference list $i : E_i^1, E_i^2, \ldots, E_i^{k_i}$. For example, we will denote the preferences $a \sim_i b \succ_i c$ by the list $i : \{a, b\}, \{c\}$. An agent i's preferences are *strict* if the size of each equivalence class is 1. An agent i's preferences are

[1] If there are exactly two candidates per seat, then designated voting is equivalent to *multiple referenda*, where a decision has to be taken on each of a series of yes-no issues.

dichotomous if he partitions the alternatives into just two equivalence classes, i.e., $k_i = 2$. Let $Topwidth(\succsim)$ be the maximum size of the most preferred equivalence class, i.e., $Topwidth(\succsim) = \max_{i \leq n} |E_i^1|$. For any $S \subseteq A$, we will denote by $\max_{\succsim_i}(S)$ and $\min_{\succsim_i}(S)$ the alternatives in S that are maximally and minimally preferred by i respectively. Thus, if q and r are respectively the smallest and the largest indices such that $E_i^q \cap S \neq \emptyset$ and $E_i^r \cap S \neq \emptyset$, then $\max_{\succsim_i}(S) = E_i^q \cap S$ and $\min_{\succsim_i}(S) = E_i^r \cap S$. For $k \leq m$, let $S_k(A) = \{W \subseteq A : |W| = k\}$.

4 Set Extensions and Pareto Optimality

Set Extensions *Set extensions* are used for reasoning about the preferences of an agent over *sets* of alternatives given their preferences over single alternatives. For fixed-size committee voting, the *responsive extension* is very natural and has been applied in various matching settings as well [Barberà *et al.*, 2004; Roth and Sotomayor, 1990]. For all $V, W \in S_k(A)$, we say that $W \succsim_i^{RS} V$ if and only if there is an injection f from V to W such that for each $a \in V$, agent i weakly prefers $f(a)$ to a, i.e. $f(a) \succsim_i a$.

We define the 'best' set extension and the 'worst' set extension which are denoted $\mathcal{B}$ and $\mathcal{W}$ respectively. For all $W, V \in S_k(A)$, $W \succsim_i^{\mathcal{B}} V$ if and only if $w \succsim_i v$ for $w \in \max_{\succsim_i}(W)$ and $v \in \max_{\succsim_i}(V)$. On the other side, $W \succsim_i^{\mathcal{W}} V$ if and only if $w \succsim_i v$ for $w \in \min \succsim_i(W)$ and $v \in \min_{\succsim_i}(V)$.

In the *leximax (LX)* extension, an agent prefers a committee that selects more alternatives from his most preferred equivalence class, in case of equality, the one with more alternatives for the second most preferred equivalence class, and so on. Formally, $W \succ_i^{LX} V$ iff for the smallest (if any) l with $|W \cap E_i^l| \neq |V \cap E_i^l|$ we have $|W \cap E_i^l| > |V \cap E_i^l|$.[2]

Efficiency based on Set Extensions With each set extension $\mathcal{E}$, we can define Pareto optimality with respect to $\mathcal{E}$. A committee $W \in S_k(A)$ is *Pareto optimal* with respect to $\mathcal{E}$, or simply *$\mathcal{E}$-efficient*, if there exists no committee $W' \in S_k(A)$ such that $W' \succsim_i^{\mathcal{E}} W$ for all $i \in N$ and $W' \succ_i^{\mathcal{E}} W$ for some $i \in N$. Note that for each of our set extensions, $\mathcal{E}$-efficiency coincides with standard Pareto optimality when $k = 1$. An outcome is a Pareto improvement over another if each agent weakly improves and at least one agent strictly improved.

In Figure 1, we illustrate the relations between the different efficiency notions. In one case, The inclusion relation follows from the fact that $\succsim_i^{LX}$ is a refinement of $\succsim_i^{RS}$. Most of the other observations can be proved by small examples consisting of two or three agents.

We also make the following general observation.

Lemma 1. *If there is a polynomial-time algorithm to compute a Pareto improvement over a committee, then there exists a polynomial-time algorithm to compute an $\mathcal{E}$-efficient committee under set extensions $\mathcal{E} \in \{RS, LX, \mathcal{W}, \mathcal{B}\}$.*

[2]One could define in a similar way a *leximin* refinement of $\succsim^{\mathcal{W}}$. For the sake of brevity we do not consider such a refinement here.

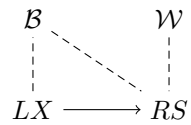


Figure 1: Relations between the four notions of efficiency. An arrow from $\mathcal{E}_1$ to $\mathcal{E}_2$ means that $\mathcal{E}_1$-efficiency implies $\mathcal{E}_1$-efficiency; a dashed line means there always exists a committee that is both $\mathcal{E}_1$- and $\mathcal{E}_2$-efficient; absence of arrow or line means that the sets of $\mathcal{E}_1$- and $\mathcal{E}_2$-efficient committees can be disjoint.

Proof. Here, we start from any committee and we recursively apply Pareto improvement until we reach a Pareto optimal committee. For the 'best' and 'worst' extensions, there can be at most mn Pareto improvements because for one agent there can be at most m improvements. Since an RS-improvement implies an LX-improvement, let us bound the number of LX-improvements. The maximum number of improvements is when we start from the set of worst alternatives and move to the set of best alternatives. The number of movements for the best alternative in the set is at most m, and similarly for other alternatives. Thus there can be at most m^2 improvements and in total there can be nm^2 Pareto improvements. □

We end this section by observing that, under any of the set extensions we consider, a set of Pareto optimal alternatives may be Pareto dominated. Consider the following example.

Example 1.

$1 : a, c, b, d$	$2 : a, d, b, c$
$3 : b, c, a, d$	$4 : b, d, a, c$

The set $\{c, d\}$ consists of Pareto optimal alternatives but is Pareto dominated by $\{a, b\}$ under any of our set extensions.

5 Responsive Set Extension

There is a trivial way to achieve Pareto optimality under the responsive set extension by taking any decreasing scoring vector consistent with the ordinal preferences, finding the total score of each alternative and returning the set of k alternatives with the maximum scores. For instance, on Example 1, the outcome of the rule that outputs the alternatives with the best k Borda scores is $\{a, b\}$.

Theorem 1. *A Pareto optimal committee under the responsive set extension committee can be computed in linear time.*

In many situations, one may already have a status-quo committee and one may want to find a Pareto improvement over it. This problem of testing Pareto optimality and finding a Pareto improvement under the responsive set extension turns out to be a much harder task. Note that if there exists a polynomial-time algorithm to compute a Pareto improvement, then it means that testing Pareto optimality is also polynomial-time solvable.

Theorem 2. *Checking whether a committee is Pareto optimal under the responsive set extension is coNP-complete, even for dichotomous preferences and $Topwidth(\succsim) \geq 3$, or for strict preferences.*

Proof. We only present the case where $Topwidth(\succsim) = 3$. The reduction is from the NP-complete problem VERTEX COVER [Garey and Johnson, 1979]. Given a simple graph $G = (V, E)$, the MINIMUM VERTEX COVER problem consists in finding a subset $C \subseteq V$ of minimum size such that every edge $e \in E$ is incident to some node of C. Its decision version VERTEX COVER consists, given a simple graph $G = (V, E)$ and an integer k, of deciding if there exists a vertex cover $C \subseteq V$ of G with $|C| \leq k$.

Let $\langle (V, E), k\rangle$ be an instance of VERTEX COVER, with $[x, y]$ being one arbitrary edge in E. We build the following instance of Pareto optimality under RS:

- $N = \cup_{e\in E} N_e \cup \{a\}$, where for each edge $e \in E$, N_e is a set of k agents, and a is a special agent.
- $A = V \cup D$, where $D = \{d_1, \ldots, d_k\}$.
- For each $e = [u, v] \in E$, the preferences of agent e^i, for $i = 1, \ldots, k$, and of agent a, are

$$\begin{array}{ll} e^i: & \{u, v, d_i\}, (D - d_i) \cup (V \setminus \{u, v\}) \\ a: & \{x, y\}, D \cup (V \setminus \{x, y\}) \end{array}$$

The reduction is clearly done within polynomial time and preferences are dichotomous. We can check easily that committee D (of size k) is not Pareto optimal under RS if and only if there exists a vertex cover of G of size at most k.

For strict preferences, in the previous reduction we replace $\{u, v, d_i\}, \ldots$ by $\{u\}, \{v\}, \{d_i\}, \ldots$ in the preferences of e^i. It is easy to see that the proof is similar. □

Using a similar reduction from the HITTING SET problem, we can also prove Theorem 3 that concerns a parametrized complexity intractability result [Downey and Fellows, 2013]. HITTING SET is defined as follows: given a ground set X of elements, and a collection $\mathcal{C} = \{C_1, \ldots, C_\ell\}$ of subsets of X, does there exist a $H \subset X$ such that $|H| \leq k$ and $H \cap C \neq \emptyset$ for all $C \in \mathcal{C}$?

Theorem 3. *Checking whether a committee is Pareto optimal under the responsive set extension is W[2]-complete under parameter k, even for dichotomous preferences.*

For dichotomous preferences we present a complete characterization of the complexity according to the $Topwidth(\succsim)$ parameter. If $Topwidth(\succsim) = 1$, then in any Pareto improvement over committee D, any alternative in D that is most preferred by some agent needs to be remain selected, and therefore the problem of checking RS-efficiency is easy. If $Topwidth(\succsim) \geq 3$, from Theorem 2, the problem is hard. Remains the case $Topwidth(\succsim) = 2$.

Theorem 4. *For dichotomous preferences, a Pareto improvement over a committee with respect to the responsive set extension can be computed in polynomial time when* $Topwidth(\succsim) \leq 2$.

Proof. Consider a preference profile $\succsim = (\succsim_1, \ldots, \succsim_n)$ where each $\succsim_i$ is dichotomous and verifies $Topwidth(\succsim) = 2$, and let $D \in S_k(A)$. For each $i \in N$, let (E_i^1, E_i^2) be the partition associated with $\succsim_i$.

First, if for all $i \in N$, $E_i^1 \subseteq D$, then D is obviously RS-efficient. Assume it is not the case, that is, (1) for some $i \in N$, $E_i^1 \setminus D \neq \emptyset$. Let

- $N' = \{i \in N : E_i^1 \cap D = E_i^1\}$, $W' = \cup_{i\in N'} E_i^1$ (by construction, $W' \subseteq D$), and $k' = |W'|$.
- $N'' = \{i \in N \setminus N' : E_i^1 \cap (D \setminus W') \neq \emptyset\}$ and $A'' = \cup_{i\in N''} E_i^1$.

Now, we build a graph $G = (V, E)$ with $V = \{v_1, \ldots, v_r\}$ isomorphic to A'', and $[v_p, v_q] \in E$ iff $E_i^1 = \{a_p, a_q\}$ for some $i \in N''$: each edge of G corresponds to the top two alternatives of some agent, provided one of them is in $D \setminus W'$. Let $\tau(G)$ be the size of an optimal vertex cover of G.

We first claim that there is a Pareto improvement over D if and only if one of follows two conditions is satisfied:

(i) $\tau(G) < k - k'$, or

(ii) $\tau(G) = k - k'$, and there is an optimal vertex cover of G containing either at least an element of E_i^1 for some $i \notin N' \cup N''$, or two elements of E_i^1 for some $i \in N''$.

We first show that (i) and (ii) are sufficient. If (i) holds then take a committee corresponding to a minimum vertex cover of G, add to it the k' alternatives of W', and add $(k-k') - \tau(G)$ alternatives, with at least one in $\cup_i (E_i^1 \setminus D)$; this is possible because of (1). If (ii) holds, then take a committee corresponding to a minimum vertex cover of G, and add to it the k' alternatives of W'. In both cases, the obtained committee contains E_i^1 for all $i \in N'$, contains at least one element of E_i^1 for all $i \in N'$, and contains either two elements of E_i^1 for some $i \in N''$, or an element of E_i^1 for some $i \notin N \cup N''$. Therefore it is a Pareto-improvement over D.

Now, we show that (i) and (ii) are necessary. Let $W \in S_k(A)$ be a Pareto improvement of D containing a maximum number of alternatives from D. We have the following two properties: $W' \subseteq W$ and $W \setminus W'$ is a vertex cover of G. $W' \subseteq W$ holds, since otherwise there would be an $i \in N'$ such that $W' \succsim_i^{RS} W$ does not hold. For similar reasons, $C' = (W \setminus W') \cap A''$ is a vertex cover of G. If $|(W \setminus W') \cap A''| < \tau(G)$, then by adding to it any set of $D \setminus C'$ of size $k - k' - \tau(G)$ we obtain a set of size k which constitutes a Pareto improvement of D because now, $E_i^1 \subseteq W$ for some $i \in N''$. If $|(W \setminus W') \cap A''| = \tau(G)$, then $(W \setminus W') \cap A'' = W \setminus W'$ and necessarily either $E_i^1 \cap C \neq \emptyset$ for some $i \notin (N' \cup N'')$ or $E_i^1 \subseteq C$ for some $i \in N''$.

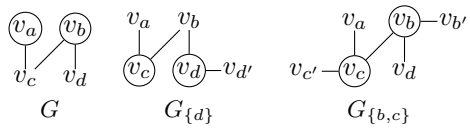
It remains to be shown that (i) and (ii) can be checked in polynomial time. (i) can be done in polynomial-time because G is bipartite: indeed, by construction, G is two-colorable with color sets $A'' \cap D$ and $A'' \setminus D$, and by König's theorem, for bipartite graphs, the problem of finding the minimum vertex cover is equivalent to computing a maximum matching, hence solvable in polynomial time. As for (ii), if $\tau(G) = k - k'$, we have to check whether for some optimal vertex cover C of G, either (ii.1) $E_i^1 \cap C \neq \emptyset$ holds for some $i \notin (N' \cup N'')$, or (ii.2) $E_i^1 \subseteq C$ for some $i \in N''$. In order to check (ii.1), for each $i \notin (N' \cup N'')$ such that there exists $x \in E_i^1 \cap A''$, we transform G into a new bipartite graph $G_{\{x\}}$ where we add a new vertex x' and an edge $[x, x']$. In order to check (ii.2), for each $i \notin N''$, let $E_i^1 = \{x, y\}$; we transform G into a new bipartite graph $G_{\{x,y\}}$ where we add two new vertices x' and y', and two edges $[x, x']$ and $[y, y']$. Finally, we test if $\tau(G) = \tau(G_{\{x\}})$ or if $\tau(G) = \tau(G_{\{x,y\}})$ for one

of these graphs, because all optimal vertex covers of $G_{\{x\}}$ (respectively $G_{\{x,y\}}$) must contain x (respectively $\{x, y\}$). $\square$

Example 2. *We illustrate the algorithm in the proof of Theorem 4. Let $k = 2$ and consider the dichotomous profile, where we specify only the top equivalence class of each agent:*

$$1 : \{a, c\} \qquad 2 : \{b, c\} \qquad 3 : \{b, d\}$$
$$4 : \{d, e\} \qquad 5 : \{e, f\}$$

Let $D = \{a, b\}$. We have $N' = W' = \emptyset$, $k' = 0$, $D \setminus W' = \{a, b\}$, $N'' = \{1, 2, 3\}$, and $A'' = \{a, b, c, d\}$. We construct the graph $G = (V, E)$: $V = \{v_a, v_b, v_c, v_d\}$ and $E = \{\{v_a, v_c\}, \{v_b, v_c\}, \{v_b, v_d\}\}$. We have $\tau(G) = 2 = k - k'$. Now we consider the four graphs $G_{\{d\}}$, resulting from the addition to G of a new vertex $v_{d'}$ and edge $[v_d, v_{d'}]$, and $G_{\{a,c\}}$, $G_{\{b,c\}}$ and $G_{\{b,d\}}$: $G_{\{a,c\}}$ results from the addition to G of two new vertices $v_{a'}, v_{c'}$ and edges $[v_a, v_{a'}]$ and $[v_c, v_{c'}]$, etc. Two of these graphs have an optimal cover of size 2: $G_{\{d\}}$, with optimal cover $\{v_c, v_d\}$, and $G_{\{b,c\}}$, with optimal cover $\{v_b, v_c\}$. Therefore, $\{c, d\}$ and $\{b, c\}$ are RS-Pareto-improvements over $\{a, b\}$, and $\{a, b\}$ is not RS-efficient.



Note that finding an algorithm that computes a Pareto improvement over a committee can be used to decide whether a given a committee D of size k, is Pareto optimal under the responsive set extension.

Pareto optimality and Strategyproofness We now try to achieve both RS-efficiency and strategyproofness simultaneously. A mechanism is strategyproof if no agent can get a more preferred outcome by misreporting his/her preference.

A naive way of achieving RS-efficiency and Pareto optimality is to enumerate the list of possible winning sets and implement serial dictatorship over the possible outcomes as is done in voting [Aziz *et al.*, 2013b]. However, the number of possible outcomes is exponential and responsive preferences result in a partial order over the possible winning sets and not a complete and transitive order. This problem is solved by Algorithm 1 which can be viewed as a computationally efficient serial dictatorship.

Theorem 5. *There exists a linear-time and strategyproof algorithm that returns a committee that is Pareto optimal under the responsive set extension.*

Proof. Consider Algorithm 1. We show that at each stage i', agent $\pi(i')$, implicitly refines the set of feasible committees to the maximal set of most preferred outcomes from the set by providing additional constraints. This is true for the base case $i' = 1$. Now assume it holds from 1 to i'. Note that L contains all those alternatives that are strictly less preferred by agents in $\{\pi(1), \ldots, \pi(i')\}$ than the ones they respectively *fixed*. Moreover, each agent in $\{1, \ldots, \pi(i')\}$ is indifferent between the alternatives in L. As for $\pi(i' + 1)$, he fixes the

Input: $(N, A, \succsim, k,$ permutation π of $N)$
Output: $W \in S_k(A)$.

L (last set to be refined) $\longleftarrow A$
r (number of alternatives yet to be fixed) $\longleftarrow k$; $W \longleftarrow \emptyset$
i' (index of the permutation π) $\longleftarrow 1$
while $r \neq 0$ or $i' \neq n$ **do**
  Agent $i = \pi(i')$ selects first t equivalence classes such that $|\bigcup_{j=1}^{t} E_i^j \cap L| \geq r$ and $|\bigcup_{j=1}^{t-1} E_i^j \cap L| < r$.
  $W \longleftarrow W \cup (\bigcup_{j=1}^{t-1} E_i^j \cap L)$ (we say agent i *fixes* the alternatives in $\bigcup_{j=1}^{t-1} E_i^j$);
  $r \longleftarrow |\bigcup_{j=1}^{t} E_i^j \cap L| - |\bigcup_{j=1}^{t-1} E_i^j \cap L|$
  $L \longleftarrow E_i^t$; $r_{i'} \longleftarrow r$
  Increment i' by one
if $r > 0$ **then**
  pick any r alternatives from L and add them to W
return W

Algorithm 1: Committee Voting Serial Dictatorship

best $|\bigcup_{j=1}^{t-1} E_{\pi(i'+1)}^j \cap L|$ alternatives in L where t is the value such that $|(\bigcup_{j=1}^{t} E_{\pi(i'+1)}^j) \cap L| \geq r_{i'}$ and $|\bigcup_{j=1}^{t-1} E_{\pi(i'+1)}^j \cap L| < r_{i'}$. For $E_{\pi(i'+1)}^t$, the agent only requires that $r_{i'+1} = |(\bigcup_{j=1}^{t-1} E_{\pi(i'+1)}^j) \cap L| - |(\bigcup_{j=1}^{t-1} E_{\pi(i'+1)}^j) \cap L|$ alternatives are selected from his equivalence class $E_{\pi(i'+1)}^t$ which is ensured by the definition of the algorithm. It follows from the argument that the returned set is Pareto optimal under the responsive set extension. For strategyproofness, when an agent $\pi(i')$ turn comes, it only has a choice over *fixing* the alternatives in L and requiring $r_{i'}$ alternatives from his equivalence class $E_{\pi(i')}^t$. In this case the algorithm already chooses one of the best possible committees for the agent. $\square$

Note that for $k = 1$, the algorithm is equivalent to serial dictatorship as formalized by Aziz *et al.* [2013a]. Note that a committee that is Pareto optimal under the responsive set extension may not be a result of serial dictatorship. This holds even for $k = 1$ and the basic voting setting.

The problem with the serial dictatorship algorithm formalized is that it overly favours the agent that is the first in the permutation. One way to limit his power is to let him choose only $\lceil k/n \rceil$ alternatives. We note that this attempt at having a fairer extension of serial dictatorship comes at an expense because strategyproofness is compromised. Consider the profile in which 1 has preferences a, b, c and 2 has preferences a, c, b. For $k = 2$, and permutation 12, the outcome is $\{a, c\}$. But if agent 1 reports b, a, c, then the outcome is $\{a, b\}$.

6 Leximax Set Extension

We point out that for dichotomous preferences, the responsive set extension coincides with the leximax set extension. Hence we get a corollary of our results for responsive preferences:

Corollary 1. *Checking whether a committee is LX-efficient is coNP-complete, even for dichotomous preferences and $Topwidth(\succsim) \geq 3$.*

Note that Algorithm 1 returns a LX-efficient committee.

Theorem 6. *There exists a linear-time and strategyproof algorithm that returns a LX-efficient committee.*

7 'Best' Set Extension

Next, we consider Pareto optimality with respect to $\mathcal{B}$, which has been used for defining many rules (see Section 2).

Theorem 7. *Unless P=NP, there is no polynomial-time algorithm to compute a Pareto improvement over a committee with respect to $\mathcal{B}$, even for dichotomous preferences and $Topwidth(\succsim) = 2$.*

Proof. We show that if it is not the case, then we can solve polynomially the vertex cover decision problem. Consider an instance of VERTEX COVER given by a simple graph $G = (V, E)$ with $V = \{v_1, \ldots, v_q\}$ and $E = \{e_1, \ldots, e_r\}$, and an integer k. Assume the existence of a polynomial-time algorithm `Algo` that computes a Pareto improvement over a committee with respect to $\mathcal{B}$ when $Topwidth(\succsim) \leq 2$: given a profile $\succsim$ and a set of k alternatives W, $\texttt{Algo}(\succsim, W)$ returns, in time polynomial in $|\succsim|$, `Yes` if W is Pareto optimal with respect to $\mathcal{B}$, and otherwise returns a k-set U of alternatives which Pareto dominates W. We will now prove by applying at most n times `Algo` with different inputs that we can decide in polynomial if G has a vertex cover $C \subseteq V$ of G with $|C| \leq k$. We construct the following profile P:

- The set of agents is $N = \{1, \ldots, q+r-k\}$, where agent $i \leq r$ corresponds to edge $e_i \in E$.
- The set of $2q - k$ alternatives is $A = V \cup D$ where $D = \{d_1, \ldots, d_{q-k}\}$.
- Let $e_i = [u, v] \in E$ be an edge of G; the preferences of agent i for $i = 1, \ldots, r$ are:

$$i: \ \{u, v\}, D \cup (V \setminus \{u, v\}).$$

  The preferences of the last set of $q - k$ agents $\{r + 1, \ldots, q + r - k\}$ are given by: for $i = 1, \ldots, q - k$,

$$r + i: \ d_i, V \cup (D \setminus \{d_i\}).$$

The reduction is clearly done within polynomial time and the set of preferences given by $\succsim$ are dichotomous.

Consider the following inductive procedure: $W_0 = V$ and for $i \geq 1$, $W_i = \texttt{Algo}\ (\succsim, W_{i-1})$ if W_{i-1} is not Pareto optimal with respect to $\mathcal{B}$, otherwise we return W_{i-1}. Let $W = W_{n-k}$ be the solution output after $n - k$ calls. Because `Algo` is polynomial, the whole procedure is polynomial.

We claim that G has a vertex cover of size k iff $D \subseteq W$. We will first prove by induction that at each step i, $W_i \setminus D$ is a vertex cover of G. For the initial step, it is valid because V is a vertex cover of G. Assume that it is true for $i < n - k$ and let us prove that $W_{i+1} \setminus D$ is a vertex cover of G. If it is not the case, some edge $e_j = [u, v] \in E$ is not covered. By assumption, e_j is covered by $W_i \setminus D$. This implies $W_i \succ_j^{\mathcal{B}} W_{i+1}$, which is a contradiction. Hence, $W_{i+1} \succsim_j^{\mathcal{B}} W_i$. From this hypothesis, we deduce $D \setminus W_{i+1} \succsim_j^{\mathcal{B}} D \setminus W_i$ for $j = n + 1, \ldots, 2n - k$ with a strict preference for some agent. Equivalently, $D \setminus W_i \subset D \setminus W_{i+1}$. In conclusion, after $n - k$ recursive calls, $|W \setminus D| \leq k$ if and only if $D \subseteq W$. □

Theorem 8. *Computing a $\mathcal{B}$-efficient committee is NP-hard, even for dichotomous preferences.*

Proof. We give a reduction from HITTING SET. Let $N = \{1, \ldots, \ell\}$, $A = X$ and for each $i \in N$, i's dichotomous preferences are $i: C_i, (X \setminus C_i)$. If there exists a polynomial-time algorithm to compute a $\mathcal{B}$-efficient committee, it will return a committee in which each agent gets a most preferred alternative if such a committee exists. But such a committee corresponds to a hitting set of size k. □

8 'Worst' Set Extension

In contrast to all the other set extensions considered in the paper, Pareto optimality with respect to the 'worst' set extension can be checked in polynomial time.

Theorem 9. *There exists a polynomial-time algorithm that checks whether there a committee is $\mathcal{W}$-efficient and computes a Pareto improvement over it if possible.*

Proof. Let $W \in S_k(A)$. For each $i \in N$, let $E_i^{t_i}$ be the least preferred equivalence class such that $E_i^{t_i} \cap W \neq \emptyset$. We want to check whether there is a k-set D of alternatives in which at least some agent $i \in N$ gets a strictly better outcome and all the other agents get at least as preferred an outcome. We check this as follows. For $i \in N$, let $B_i = A \setminus (\bigcup_{\ell=t_i}^{k_i} E_i^{\ell}) \cup \bigcup_{j \in N \setminus \{i\}} \bigcup_{\ell=t_j+1}^{k_j} E_j^{\ell})$. We check whether $|B_i| \geq k$ or not. If $|B_i| \geq k$, we know that there exists a subset of B_i, that is strictly more preferred by $i \in N$ and at least as preferred by each agent. If $|B_i| < k$, then this means that a Pareto improvement with i strictly improving is only possible if the size of the winning set is less than k which is not feasible. □

We now consider strategyproofness together with $\mathcal{W}$-efficiency. We first note that Algorithm 1 may not return a $\mathcal{W}$-efficient outcome. However, we construct a suitable strategyproof and $\mathcal{W}$-efficient by formalising an appropriate serial dictatorship algorithm for the worst set extension.

Theorem 10. *There exists a linear-time and strategyproof algorithm that returns a $\mathcal{W}$-efficient committee.*

Proof. Consider the agents in a permutation π. The set of alternatives A' is initialized to A. We reduce the set A' while ensuring that it of size at least k. The next agent i in the permutation comes and deletes the maximum number of least preferred equivalence classes from his preferences and the corresponding alternatives in A' while ensuring that $|A'| \geq k$. Each successive agent in the permutation gets a most preferred outcome while ensuring that agents before him in the permutation get at least as preferred an outcome as before. Thus the algorithm is strategyproof and Pareto optimal with respect to the 'worst' set extension. □

9 Conclusions

We considered Pareto optimality in multi-winner voting with respect to a number of prominent set extensions. We presented results on the relations between the notions as well as complexity of computing and verifying Pareto optimal outcomes. Further directions of future work include considering Pareto optimality with respect to other set extensions [Brandt and Brill, 2011].

Acknowledgments

Data61 (formerly known as NICTA) is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program. The authors thank Felix Brandt for useful pointers and comments. Jérôme Lang and Jérôme Monnot thank the ANR project CoCoRICo-CoDec.

References

[Aziz and Savani, 2016] H. Aziz and R. Savani. Hedonic games. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 15. Cambridge University Press, 2016. Forthcoming.

[Aziz *et al.*, 2013a] H. Aziz, F. Brandt, and M. Brill. The computational complexity of random serial dictatorship. *Economics Letters*, 121(3):341–345, 2013.

[Aziz *et al.*, 2013b] H. Aziz, F. Brandt, and M. Brill. On the tradeoff between economic efficiency and strategyproofness in randomized social choice. In *Proc. of 12th AAMAS Conference*, pages 455–462. IFAAMAS, 2013.

[Barberà *et al.*, 2004] S. Barberà, W. Bossert, and P. K. Pattanaik. Ranking sets of objects. In S. Barberà, P. J. Hammond, and C. Seidl, editors, *Handbook of Utility Theory*, volume II, chapter 17, pages 893–977. Kluwer Academic Publishers, 2004.

[Benoît and Kornhauser, 2010] J.-P. Benoît and L. Kornhauser. Only a dictatorship is efficient. *Games and Economic Behavior*, 70(2):261–270, 2010.

[Betzler *et al.*, 2013] N. Betzler, A. Slinko, and J. Uhlmann. On the computation of fully proportional representation. *JAIR*, 47:475–519, 2013.

[Bossert, 1995] W. Bossert. Preference extension rules for ranking sets of alternatives with a fixed cardinality. *Theory and Decision*, 39:301—317, 1995.

[Brams *et al.*, 2007] S. Brams, D. Kilgour, and R. Sanver. A minimax procedure for electing committees. *Public Choice*, 3-4(132):401–420, 2007.

[Brandt and Brill, 2011] F. Brandt and M. Brill. Necessary and sufficient conditions for the strategyproofness of irresolute social choice functions. In *Proc. of 13th TARK Conference*, pages 136–142. ACM Press, 2011.

[Caragiannis *et al.*, 2010] I. Caragiannis, D. Kalaitzis, and E. Markakis. Approximation algorithms and mechanism design for minimax approval voting. In *Proc. of 24th AAAI Conference*, pages 737–742, 2010.

[Cechlárová, 2008] K. Cechlárová. Stable partition problem. In *Encyclopedia of Algorithms*, pages 885–888. Springer, 2008.

[Chamberlin and Courant, 1983] J. R. Chamberlin and P. N. Courant. Representative deliberations and representative decisions: proportional representation and the Borda rule. *American Political Science Review*, 77(3):718–733, 1983.

[Cuhadaroğlu and Lainé, 2012] T. Cuhadaroğlu and J. Lainé. Pareto efficiency in multiple referendum. *Theory and Decision*, 72(4):525–536, 2012.

[Darmann, 2013] A. Darmann. How hard is it to tell which is a condorcet committee? *Mathematical Social Sciences*, 66(3):282–292, 2013.

[Downey and Fellows, 2013] R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.

[Elkind and Ismaili, 2015] E. Elkind and A. Ismaili. OWA-Based Extensions of the Chamberlin-Courant Rule. In *Proc. of 4th ADT*, pages 486–502. Springer-Verlag, 2015.

[Elkind *et al.*, 2014] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. Properties of multiwinner voting rules. In *Proc. of 13th AAMAS Conference*, pages 53–60, 2014.

[Elkind *et al.*, 2015] E. Elkind, J. Lang, and A. Saffidine. Condorcet winning sets. *Social Choice and Welfare*, 44(3):493–517, 2015.

[Faliszewski *et al.*, 2016] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Multiwinner analogues of the plurality rule: Axiomatic and algorithmic views. In *Proc. of 28th AAAI Conference*, 2016.

[Garey and Johnson, 1979] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[Klamler *et al.*, 2012] C. Klamler, U. Pferschy, and S. Ruzika. Committee selection under weight constraints. *Mathematical Social Sciences*, 64(1):48–56, 2012.

[Lang *et al.*, 2012] J. Lang, J. Mengin, and L. Xia. Aggregating conditionally lexicographic preferences on multi-issue domains. In *Principles and Practice of Constraint Programming*, pages 973–987, 2012.

[Lu and Boutilier, 2011] T. Lu and C. Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proc. of 22nd IJCAI*, pages 280–286. AAAI Press, 2011.

[Monroe, 1995] B. L. Monroe. Fully proportional representation. *The American Political Science Review*, 89(4):925–940, 1995.

[Moulin, 2003] H. Moulin. *Fair Division and Collective Welfare*. The MIT Press, 2003.

[Özkal-Sanver and Sanver, 2006] İ. Özkal-Sanver and R. Sanver. Ensuring Pareto-optimality by referendum voting. *Social Choice and Welfare*, 27:211–219, 2006.

[Procaccia *et al.*, 2008] A. D. Procaccia, J. S. Rosenschein, and A. Zohar. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30:353–362, 2008.

[Roth and Sotomayor, 1990] A. E. Roth and M. A. O. Sotomayor. *Two-Sided Matching: A Study in Game Theoretic Modelling and Analysis*. Cambridge University Press, 1990.

[Skowron *et al.*, 2015a] P. Skowron, P. Faliszewski, and A. Slinko. Achieving fully proportional representation: Approximability results. *Artif. Intell.*, 222:67–103, 2015.

[Skowron *et al.*, 2015b] P. K. Skowron, P. Faliszewski, and J. Lang. Finding a collective set of items: From proportional multirepresentation to group recommendation. In *Proc. of 23rd IJCAI*, pages 2131–2137. AAAI Press, 2015.