Approximating Common Voting Rules by Sequential Voting in Multi-Issue Domains

Lirong Xia **Vincent Conitzer**
Department of Computer Science
Duke University
Durham, NC 27708, USA
{lxia,conitzer}@cs.duke.edu

Jérôme Lang
LAMSADE
Université Paris-Dauphine
75775 Paris Cedex 16, France
jerome.lang@dauphine.fr

Abstract

When agents need to make decisions on multiple issues, one solution is to vote on the issues sequentially. In this paper, we investigate how well the winner under the sequential voting process approximates the winners under some common voting rules. Some common voting rules, including Borda, k-approval, Copeland, maximin, Bucklin, and Dodgson, admit natural *scoring functions* that measures the "quality" of the alternatives, thus can serve as a basis for approximation results. We focus on multi-issue domains where each issue is binary and the agents' preferences are $\mathcal{O}$-legal, separable, represented by LP-trees, or lexicographic. Our results show significant improvements in the approximation ratios when the preferences are represented by LP-trees, compared to the approximation ratios when the preferences are $\mathcal{O}$-legal. However, assuming that the preferences are separable (respectively, lexicographic) does not significantly improve the approximation ratios compared to the case where the preferences are $\mathcal{O}$-legal (respectively, are represented by LP-trees).

1 Introduction

In many situations, a set of agents (voters) has to decide collectively on the value of each one of a finite set of variables, or *issues*. Each of the issues can take its value from a given finite domain. Typical examples of such situations include *multiple referenda*, were local communities have to make decisions on possibly interrelated issues, or *committee elections*, where a set of voters has to choose a representative committee.

Arguably the simplest solution consists in voting separately on each variable in parallel. This solution is implemented in many real-life settings, presumably because of its simplicity. As [Brams, Kilgour, and Zwicker, 1998] and later [Lacy and Niou, 2000] show, this solution can lead to extremely undesirable outcomes; examples of this phenomenon are called *multiple election paradoxes*. The first type of multiple election paradox, analyzed in [Brams, Kilgour, and Zwicker, 1998], can be seen in the following example. Suppose we have 3 binary issues A, B, C whose domains are, respectively, $\{a, \bar{a}\}$, $\{b, \bar{b}\}$, $\{c, \bar{c}\}$, on which the voters vote in parallel. Let there be 3 voters, one voting $ab\bar{c}$, one $a\bar{b}c$, and one $\bar{a}bc$. Then the winning outcome is abc, *although* abc *did not receive a single vote*. A more severe example arises with 4 issues (Example 3 in [Brams, Kilgour, and Zwicker, 1998], where *the winning alternative is the unique alternative receiving the fewest plurality votes*. Therefore, *parallel voting can elect the plurality loser*, and since the example does not require any restriction of the voters' preferences below their top alternative, this property holds *even if all voters have separable preferences*. (A voter's preferences are *separable* if her preferences for each issue do not depend on the values of the other issues.)

However, the impact of this result is arguably limited, because it only focuses on plurality. We may wonder to what extent it extends to other voting rules. We will focus on those rules based on a *score*: given a rule r consisting in electing an alternative that maximizes a score function S_r, and given a profile consisting of separable preference relations on a multi-issue domain (composed of binary issues), what can we say about the score of the winner obtained by applying issue-wise majority? How does it compare to the score of the winner according to r? A natural way to pursue this question is to analyze the worst possible ratio between the score of the issue-wise majority winner and the score of the winner according to r. This will help us to identify those rules that issue-wise majority approximates best, which will help us to better understand the properties of issue-wise majority.

Now, there is no reason to consider only profiles of separable preferences. As shown in [Lang and Xia, 2009], provided that there exists an order $\mathcal{O}$ on the p binary issues, say $\mathbf{x}_1 > \cdots > \mathbf{x}_p$, such that for every $i \leq p$, every agent's preference for $\mathbf{x}_i$ does not depend on the values of $\{\mathbf{x}_{i+1}, \ldots, \mathbf{x}_p\}$ (in which case the profile is called *$\mathcal{O}$-legal*), then *sequential majority voting* can be defined in a natural way: elicit the voters' preferences for $\mathbf{x}_1$ and fix the value (0_1 or 1_1) according to the majority rule (possibly with a tie-breaking mechanism if we have an even number of voters); then, elicit the voters' preferences for $\mathbf{x}_2$ given the value collectively chosen for $\mathbf{x}_1$; etc. This results in a *sequential majority winner* (with respect to the order $\mathcal{O} = \mathbf{x}_1 > \cdots > \mathbf{x}_p$). Just as we argued above for issue-wise majority, we would like to know how well sequential majority approximates standard voting rules. This will give us insight into the fundamental properties of sequential majority voting and help to justify (or not) its use. Specifically, if we can prove that sequential majority voting approximates a given voting rule r well, then, given that applying sequential majority voting is computationally and communicationally cheap (while r will generally not be, given the prohibitive size of the domain), this is a good reason to use sequential majority voting instead of applying r directly.

Some common voting rules, include Borda, k-approval, Copeland, maximin, Bucklin, and Dodgson, admit natural *scoring functions* that can serve as a basis for approximation results. These scoring functions, besides being natural, also provide metrics for evaluating the quality of the alternatives. Our results are summarized in Table 1. We note that m is the number of alternatives, that is, $m = 2^p$, where p is the number of issues; n is the number of voters. We assume that n is odd (so that there can be no ties in the majority elections in sequential voting), and $n \geq 2\log m + 1$.

Rule	$\mathcal{O}$-legal	Separable	LP-trees	Lexico
Borda	$\Theta(\sqrt{m})$		$3/2 + o(1)$	
k-approval	∞ $(k < m - 2\sqrt{m})$		∞ $(k < m/4)$ $\Theta(n)$ $(m/4 \leq k < m/2)$ $\Theta(1)$ $(m/2 \leq k \leq m)$	
Bucklin	$\Theta(m)$		$2 + o(1)$	
Copeland	$\Theta(m/\log m)$		1	
Maximin	$2n/(n+1)$		1	
Dodgson	$\Omega(m)$		1	

Table 1: The approximation ratio obtained by the sequential winner, for several common rules with a natural scoring function.

It can be seen from the table that for many common rules, sequential majority voting is not a good approximation when profiles are all $\mathcal{O}$-legal or even if they are all separable (in which case sequential voting coincides with parallel issue-wise majority). However, when profiles are lexicographic, or composed of LP-trees [Booth et al., 2010] with the same structure, we obtain much more positive results. (We will define these concepts shortly.) For most voting rules we study, there is a huge improvement in the approximation ratio. That is, compared to the case where the profile is separable, the quality of the winner under sequential voting is much closer to the quality of the winner under the common voting rule, measured by their respective scoring function. In particular, in these cases, there always exists a *Condorcet winner*, and the sequential majority rule always selects it. Therefore, the sequential majority rule coincides with every *Condorcet consistent* voting rule, e.g., Copeland, maximin, and Dodgson. As can be seen from the table, the ratio is also much improved for rules that are not Condorcet consistent. These positive results suggest that, among voting methods with a low cost in terms of computation and communication, sequential majority voting is a promising one—at least in settings where the voters' preferences are lexicographic, or, more generally, where they can be represented by LP-trees with the same structure.

The idea of approximating common voting rules that are based on scoring functions is not new to this paper. Approximately computing the Dodgson score has been studied in [Caragiannis et al., 2009; 2010]; approximately computing the Young score has been studied in [Caragiannis et al., 2010]; approximating some common voting rules by strategy-proof voting rules has been studied in [Procaccia, 2010]; and approximating Copeland by voting trees has been studied in [Fischer, Procaccia, and Samorodnitsky, forthcoming 2011]. We note that in all these papers, the set of alternatives has no combinatorial structure, and a voter is free to choose any linear order over the alternatives. In contrast, in our paper, we focus on multi-issue domains (so that the number of alternatives is already exponentially large), and we restrict the voters' preferences.

2 Preliminaries

2.1 Basics of voting

Let $\mathcal{X}$ be the set of *alternatives*, $|\mathcal{X}| = m$. A vote is a linear order over $\mathcal{X}$. The set of all linear orders over $\mathcal{X}$ is denoted by $L(\mathcal{X})$. For any $c \in \mathcal{X}$ and $V \in L(\mathcal{X})$, we let $rank_V(c)$ denote the position of c in V. An *n-profile* P is a collection of n votes for some $n \in \mathbb{N}$, that is, $P \in L(\mathcal{X})^n$. A *voting rule* r is a mapping that assigns to each n-profile a unique winning alternative. That is, $r : L(\mathcal{X})^n \to \mathcal{X}$. A *scoring function* S is a mapping $L(\mathcal{X})^n \times \mathcal{X} \to \mathbb{R}$. Often, a voting rule is defined to be the mapping that finds the alternative that maximizes/minimizes the score according to a particular scoring function.[1] Below are some common voting rules. In all these voting rules, we assume that n is odd.

• *Positional scoring rules*: Given a *scoring vector* $\vec{v} = (v(1), \dots, v(m))$, for any vote $V \in L(\mathcal{X})$ and any $c \in \mathcal{X}$, let $S_{\vec{v}}(V, c) = v(rank_V(c))$. For any profile $P = (V_1, \dots, V_n)$, let $S_{\vec{v}}(P, c) = \sum_{j=1}^{n} S_{\vec{v}}(V_j, c)$. The rule will select $c \in \mathcal{X}$ so that $S_{\vec{v}}(P, c)$ is maximized. Some examples of positional scoring rules are *Borda*, for which the scoring vector is $(m-1, m-2, \dots, 0)$ and the scoring function is denoted by S_{Borda}; k-approval (App_k, with $k \leq m$), for which $v(1) = \cdots = v(k) = 1$ and $v(k+1) = \cdots = v(m) = 0$, and the scoring function is denoted by S^k_{App}; and *plurality*, for which the scoring vector is $(1, 0, \dots, 0)$.

• *Bucklin*: An alternative c's Bucklin score $S_{\text{Bl}}(P, c)$ is the smallest number l such that more than half of the voters rank c in their top l positions. The winner is an alternative that has the lowest Bucklin score.

• *Copeland*: For any two alternatives c and d, we can simulate a *pairwise election* between them, by seeing how many votes rank c ahead of d, and how many rank d ahead of c; the winner of the pairwise election is the one ranked higher more often. Then, an alternative c's Copeland score $S_{\text{C}}(P, c)$ is the number of times it wins in pairwise elections. Since we assume an odd number of voters, there can be no pairwise ties. The winner is an alternative that has the highest Copeland score.

• *Maximin*: Let $N_P(c, d)$ denote the number of votes that rank c ahead of d. The S_{MM} score of an alternative c is defined to be $S_{\text{MM}}(P, c) = \max\{N_P(c', c) : c' \in \mathcal{X}, c' \neq c\}$. The winner is an alternative c that has the lowest S_{MM} score.[2]

[1] Technically, any voting rule can be defined as the maximizer/minimizer of some scoring function (for example, the score of an alternative can be 1 if it wins and 0 if it loses), but the rules studied in this paper have very natural scoring functions that will serve well as bases for approximation. It should also be noted that a scoring function can produce ties; this does not matter from the perspective of approximation.

[2] Usually, the score of c is defined to be the minimum number of times that c beats another alternative in their pairwise election, and

• *Dodgson*: Given a profile P, an alternative c is the *Condorcet winner* if it beats all other alternatives in pairwise elections. The Dodgson score of an alternative c is the minimum number of swaps of neighboring alternatives in the votes needed to make c a Condorcet winner. Let $S_{\text{D}}(P, c)$ denote the Dodgson score. The winner is an alternative c that has the lowest Dodgson score.

A voting rule r is *Condorcet consistent* if it always select the Condorcet winner, whenever one exists. For example, Copeland, maximin, and Dodgson are Condorcet consistent.

2.2 Multi-binary-issue domains

In this paper, the set of all alternatives $\mathcal{X}$ is a *multi-binary-issue domain*. That is, let $\mathcal{I} = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ ($p \geq 2$) be a set of *issues*, where each issue $\mathbf{x}_i$ takes a value in a binary *local domain* $D_i = \{0_i, 1_i\}$. The set of alternatives is $\mathcal{X} = D_1 \times \cdots \times D_p$, that is, an alternative is uniquely identified by its values on all issues. For any $Y \subseteq \mathcal{I}$ we denote $D_Y = \prod_{\mathbf{x}_i \in Y} D_i$.

A CP-net [Boutilier et al., 2004] $\mathcal{N}$ over $\mathcal{X}$ consists of two components: (a) a directed graph $G = (\mathcal{I}, E)$ and (b) a set of conditional linear preferences $\succ^i_{\vec{d}}$ over D_i, for any $i \leq p$ and any setting $\vec{d}$ of the parents of $\mathbf{x}_i$ in G (denoted by $\mathsf{Par}_G(\mathbf{x}_i)$). These conditional linear preferences $\succ^i_{\vec{d}}$ over D_i form the *conditional preference table* for issue $\mathbf{x}_i$, denoted by $\mathsf{CPT}(\mathbf{x}_i)$. When G is acyclic, $\mathcal{N}$ is said to be an *acyclic CP-net*.

The preference relation $\succ_{\mathcal{N}}$ induced by $\mathcal{N}$ is the transitive closure of $\{(a_i, \vec{d}, \vec{z}) \succ (b_i, \vec{d}, \vec{z}) \mid i \leq p; \vec{d} \in D_{\mathsf{Par}_G(\mathbf{x}_i)}; a_i, b_i \in D_i, a_i \succ^i_{\vec{d}} b_i; \vec{z} \in D_{-(\mathsf{Par}_G(\mathbf{x}_i) \cup \{\mathbf{x}_i\})}\}$. If $\succ_{\mathcal{N}}$ is asymmetric then $\mathcal{N}$ is *consistent*. It is known that if G is acyclic, then $\mathcal{N}$ is consistent [Boutilier et al., 2004].

We say that a CP-net $\mathcal{N}$ is *compatible* with (or, *follows*) $\mathcal{O}$, if $\mathbf{x}_i$ being a parent of $\mathbf{x}_j$ in the graph implies that $i < j$. That is, preferences over issues only depend on the values of earlier issues in $\mathcal{O}$. A CP-net is *separable* if there are no edges in its graph, which means that there are no preferential dependencies among issues. A linear order V over $\mathcal{X}$ *extends* a CP-net $\mathcal{N}$, denoted by $V \sim \mathcal{N}$, if it extends the partial order that $\mathcal{N}$ induces. If $\mathcal{N}$ is compatible with $\mathcal{O}$, then we say that V is *$\mathcal{O}$-legal*. V is *separable* if it extends a separable CP-net. To present our results, we will use notations that represent the projection of a vote/CP-net onto an issue $\mathbf{x}_i$ (that is, the voter's local preferences over $\mathbf{x}_i$) given the setting of all parents of $\mathbf{x}_i$, defined as follows. For any issue $\mathbf{x}_i$, any setting $\vec{d}$ of $\mathsf{Par}_G(\mathbf{x}_i)$, and any linear order V that extends $\mathcal{N}$, we let $V|_{\mathbf{x}_i:\vec{d}}$ and $\mathcal{N}|_{\mathbf{x}_i:\vec{d}}$ denote the the projection of V (or, equivalently, $\mathcal{N}$) to $\mathbf{x}_i$, given $\vec{d}$. That is, each of these notations evaluates to the linear order $\succ^i_{\vec{d}}$ in the CPT associated with $\mathbf{x}_i$.

The *$\mathcal{O}$-lexicographic extension* of an $\mathcal{O}$-legal CP-net $\mathcal{N}$ is a linear order V over $\mathcal{X}$ such that for any $1 \leq i \leq p$, any $\vec{d_i} \in D_1 \times \cdots \times D_{i-1}$, any $a_i, b_i \in D_i$, and any $\vec{y}, \vec{z} \in D_{i+1} \times \cdots \times D_p$, if $a_i \succ_{\mathcal{N}|_{\mathbf{x}_i:\vec{d_i}}} b_i$, then $(\vec{d_i}, a_i, \vec{y}) \succ_V (\vec{d_i}, b_i, \vec{z})$. Intuitively, in the lexicographic extension of $\mathcal{N}$, $\mathbf{x}_1$ is the most important issue, $\mathbf{x}_2$ is the next-most important issue, etc; a desirable change to an earlier issue always outweighs any changes to later issues. We note that the $\mathcal{O}$-lexicographic extension of any CP-net is unique w.r.t. the order $\mathcal{O}$. We say that $V \in L(\mathcal{X})$ is *$\mathcal{O}$-lexicographic* (or *lexicographic* for short, when there is no risk of confusion) if it is the $\mathcal{O}$-lexicographic extension of an $\mathcal{O}$-legal CP-net $\mathcal{N}$. For example, $0_1 0_2 \succ 1_1 0_2 \succ 0_1 1_2 \succ 1_1 1_2$ is separable (0_1 and 0_2 are always preferred) but not ($\mathbf{x}_1 > \mathbf{x}_2$)-lexicographic ($0_1 0_2 \succ 1_1 1_2$ but $1_1 0_2 \succ 0_1 1_2$). On the other hand, $0_1 0_2 \succ 0_1 1_2 \succ 1_1 1_2 \succ 1_1 0_2$ is ($\mathbf{x}_1 > \mathbf{x}_2$)-lexicographic but not separable.

A profile P is $\mathcal{O}$-legal/separable/lexicographic if each of its votes is $\mathcal{O}$-legal/separable/lexicographic. For any $\mathcal{O}$-legal profile P, $P|_{\mathbf{x}_i:\vec{d}}$ is the profile over D_i that is composed of the projections of all votes in P on $\mathbf{x}_i$, given $\vec{d}$. That is, suppose $P = (V_1, \ldots, V_n)$, and for any $1 \leq i \leq p$, V_i extends $\mathcal{N}_i$. Then, we have $P|_{\mathbf{x}_i:\vec{d}} = (V_1|_{\mathbf{x}_i:\vec{d}}, \ldots, V_n|_{\mathbf{x}_i:\vec{d}}) = (\mathcal{N}_1|_{\mathbf{x}_i:\vec{d}}, \ldots, \mathcal{N}_n|_{\mathbf{x}_i:\vec{d}})$.

We can now define the sequential majority rule $\text{Seq}^{maj}_{\mathcal{O}}$. For any $\mathcal{O}$-legal profile P, $\text{Seq}^{maj}_{\mathcal{O}}(P) = (d_1, \ldots, d_p) \in \mathcal{X}$ is defined as follows: letting maj denote the majority rule, for every $i \leq p$, $d_i = maj(P|_{\mathbf{x}_i:d_1 \cdots d_{i-1}})$. That is, the winner is selected in p steps, one for each issue, in the following way: in step i, d_i is selected by applying the majority rule to the preferences of voters over D_i, conditioned on the values $d_1, \ldots, d_{i-1}$ that have already been determined for the issues that precede $\mathbf{x}_i$.

In this paper, we also study the case where the voters' preferences are represented by *Lexicographic Preference trees (LP-trees)* [Booth et al., 2010]. LP-trees are a generalization of lexicographic orders. An LP-tree is composed of two parts: (1) a tree T where each node t is labeled by an issue, denoted by $\mathsf{Iss}(t)$, such that each issue appears once and only once on each branch; each non-leaf node either has two outgoing edges, labeled by 0 and 1 respectively, or one outgoing edge, labeled by $\{0, 1\}$. (2) A *conditional preference table* $\mathsf{CPT}(t)$ for each node t, which is defined as follows. Let $\mathsf{Anc}(t)$ denote the set of issues labeling the ancestors of t. Let $\mathsf{Inst}(t)$ (resp., $\mathsf{NonInst}(t)$) denote the set of issues in $\mathsf{Anc}(t)$ that have two (resp., one) outgoing edge(s). There is a set $\mathsf{Par}(t) \subseteq \mathsf{NonInst}(t)$ such that $\mathsf{CPT}(t)$ is composed of the agent's local preferences over $D_{\mathsf{Iss}(t)}$ for all valuations of $\mathsf{Par}(t)$. That is, suppose $\mathsf{Iss}(t) = \mathbf{x}_i$, then for every valuation $\vec{u}$ of $\mathsf{Par}(t)$, there is an entry in the CPT which is either $\vec{u} : 0_i \succ 1_i$ or $\vec{u} : 1_i \succ 0_i$. Again, we can define the restriction of an LP-tree/profile of LP-trees to t given $\vec{u}$.

An LP-tree $\mathcal{T}$ represents a linear order $\succ_{\mathcal{T}}$ over $\mathcal{X}$ as follows. Let $\vec{d}$ and $\vec{e}$ be two different alternatives. We start at the root node t_{root} and trace down the tree according to the value of $\vec{d}$, until we find the first node t^* such that $\vec{d}$ and $\vec{e}$ differ on $\mathsf{Iss}(t^*)$. That is, w.l.o.g. letting $\mathsf{Iss}(t_{\text{root}}) = \mathbf{x}_1$, if $d_1 \neq e_1$, then we let $t = t_{\text{root}}$; otherwise, we follow the edge d_1 to examine the next node, etc. Once t^* is found, we let $U = \mathsf{Par}(t^*)$ and let d_U denote the sub-vector of $\vec{d}$ whose components correspond to the nodes in U. In $\mathsf{CPT}(t^*)$, if $d_U : d_{t^*} \succ e_{t^*}$, then

maximin selects the alternative that maximizes this score. However, if we use this definition, then the score of the maximin winner can be zero. To avoid this problem, in this paper, we define the score in a slightly different way to obtain a more informative bound on the approximation ratio. The rule itself, of course, is unchanged.

$\vec{d} \succ_{\mathcal{T}} \vec{e}$; otherwise, $\vec{e} \succ_{\mathcal{T}} \vec{d}$. We note that any lexicographic order is both $\mathcal{O}$-legal and can be represented by an LP-tree. However, LP-trees and $\mathcal{O}$-legal orders are not comparable in general. We use $\mathcal{T}$ and $\succ_{\mathcal{T}}$ interchangeably.

Example 1. *Suppose there are three issues. An LP-tree is illustrated in Figure 1. We have* $\mathsf{Iss}(t) = \mathbf{x}_2$*,* $\mathsf{Anc}(t) = \{\mathbf{x}_1, \mathbf{x}_2\}$*,* $\mathsf{Inst}(t) = \{\mathbf{x}_1\}$*,* $\mathsf{NonInst}(t) = \{\mathbf{x}_2\}$*, and* $\mathsf{Par}(t) = \{\mathbf{x}_2\}$*. The linear order represented by the LP-tree is* $000 \succ 001 \succ 010 \succ 011 \succ 111 \succ 101 \succ 100 \succ 110$*, where* 000 *is the abbreviation for* $0_1 0_2 0_3$*, etc.*

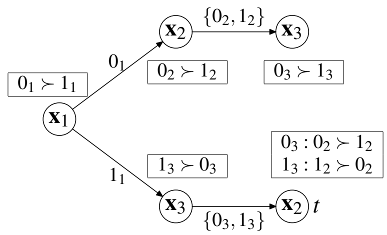

Figure 1: An LP-tree.

The sequential majority rule can easily be extended to aggregate preferences represented by LP-trees with the same structure. Let $P_T = (\mathcal{T}_1, \ldots, \mathcal{T}_n)$ denote a profile of LP-trees with the same structure T (meaning that the graph of each LP-tree is T). We note that across these LP-trees the parents as well as the CPTs of each node can be different. In this paper, we will always assume that all LP-trees in a profile have the same structure. The sequential majority rule Seq_T^{maj} selects the winner $\vec{d}$ in the following p steps. Let t denote the current node, starting at the root node t_1. In the first step, we use the majority rule to select the value of $\mathsf{Iss}(t_1)$, denoted by $d_{\mathsf{Iss}(t_1)}$, from the projection of P_T onto t_1; in the second step, we follow the path $d_{\mathsf{Iss}(t_1)}$ and reach a node $t = t_2$. Then, we use the majority rule to select $d_{\mathsf{Iss}(t_2)}$ from the projection of P_T onto t_2 given $d_{\mathsf{Iss}(t_1)}$; etc.

3 Approximating common rules when profiles are $\mathcal{O}$-legal/separable

In this section, we study how well the sequential majority rule $\text{Seq}_{\mathcal{O}}^{maj}$ approximates certain common voting rules when the profiles are $\mathcal{O}$-legal or even separable. We first give the general definition of the approximation ratio.

Definition 1. *Let r_1 and r_2 be two voting rules, and let S be a scoring function. We say that r_1 is a θ-approximation to r_2 w.r.t. S if*

$$\max_P \left\{ \max \left(\frac{S(P, r_1(P))}{S(P, r_2(P))}, \frac{S(P, r_2(P))}{S(P, r_1(P)} \right) \right\} = \theta$$

In general, P is taken over n-profiles. In this section, P is taken over all n-profiles that are $\mathcal{O}$-legal (or that are separable). We note that any separable profile is also $\mathcal{O}$-legal. Therefore, suppose r_1 is a θ_1-approximation (resp., θ_2-approximation) to r_2 w.r.t. S when the profiles are $\mathcal{O}$-legal (resp., separable), then $\theta_2 \leq \theta_1$. In other words, a lower bound on the approximation ratio for separable profiles is also a lower bound on the approximation ratio for $\mathcal{O}$-legal profiles; conversely, an upper bound on the approximation ratio for $\mathcal{O}$-legal profiles is also an upper bound on the approximation ratio for separable profiles.

Generally, we will be interested in approximating a rule r that maximizes (or minimizes) the scoring function S. The reason that we need to mention S separately in the definition, rather than just saying that we try to approximate r, is that r maximizes (or minimizes) many different scoring functions—for example, for any number $K \in \mathbb{R}$, r also maximizes (or minimizes) $S + K$. However, usually there is one such scoring function that is particularly natural. Throughout the paper, **we assume that $\mathbf{n}$ is sufficiently large, that is, $\mathbf{n \geq 2p + 1 = 2\log m + 1}$**. This avoids trivial versions of the question such as when there is only one voter (in which case both $\text{Seq}_{\mathcal{O}}^{maj}$ and any other voting rule in this paper select the top-ranked alternative of this voter). **We also assume that $\mathbf{n}$ is odd**, so that there are no ties in the rounds of sequential majority voting. We first give the following bounds, which are folklore results in social choice, and whose proof is straightforward.

Proposition 1. *For any profile P: (1) if $r_{\vec{v}}$ is the positional scoring rule associated with the vector $\vec{v} = (v(1), \ldots, v(m))$, then $S_{\vec{v}}(P, r_{\vec{v}}(P)) \geq \lceil \frac{n \sum_{i=1}^m v(i)}{m} \rceil$. In particular, (2) $S_{Borda}(P, Borda(P)) \geq \lceil (m-1)n/2 \rceil$ and (3) $S_{App}^k(P, App_k(P)) \geq \lceil kn/m \rceil$. We also have (4) $S_C(P, Copeland(P)) \geq \lceil (m-1)/2 \rceil$; (5) $S_{Bl}(P, Bucklin(P)) \leq \lceil (m+1)/2 \rceil$; (6) $S_{MM}(P, Maximin(P)) \leq n-1$; (7) $S_D(P, Dodgson(P)) \leq (m-1)(\lfloor n/2 \rfloor + 1)$.*

The following proposition (which follows from Theorem 3 in [Xia, Conitzer, and Lang, 2010]) states that the sequential winner can be ranked in an exponentially low position in every vote in a separable profile. We recall that in multi-binary-issue domains, $m = 2^p$.

Proposition 2 (Follows from Theorem 3 in [Xia, Conitzer, and Lang, 2010])**.** *There exists a separable n-profile P such that $Seq_{\mathcal{O}}^{maj}(P)$ is ranked within the bottom $2^{\lfloor p/2 \rfloor + 1} + 1$ positions in every vote in P.*

By Proposition 1 and Proposition 2, we immediately obtain the following proposition.

Proposition 3. *When profiles are separable, $Seq_{\mathcal{O}}^{maj}$ is an $\Omega(\sqrt{m})$-approximation to Borda w.r.t. S_{Borda}; for any $k < m - 2\sqrt{m}$, it is an ∞-approximation to App_k w.r.t. S_{App}^k; for any positional scoring rule $r_{\vec{v}}$, it is a $\Omega\left(\frac{\sum_{i=1}^m v(i)}{m \cdot v(m - 2\sqrt{m})}\right)$-approximation to $r_{\vec{v}}$. it is an $\Omega(\sqrt{m})$-approximation to Copeland w.r.t. S_C.*

Proof. Let P be the separable profile in Proposition 1. Let $c = \text{Seq}_{\mathcal{O}}^{maj}(P)$. For Borda, the lower bound follows from the observation that $S_{\text{Borda}}(P, c) \leq 2^{\lfloor p/2 \rfloor + 1} n \leq 2\sqrt{m}n = O(\sqrt{m}n)$ and $S_{\text{Borda}}(P, \text{Borda}(P)) = \Omega(mn)$. For k-approval, when $k < m - 2^{\lfloor p/2 \rfloor + 1}$, $S_{\text{App}}^k(P, c) = 0$. We note that $m - 2^{\lfloor p/2 \rfloor + 1} \geq m - 2\sqrt{m}$. For any positional scoring rule $r_{\vec{v}}$, the lower bound follows after the observation that $S_{\vec{v}}(P, c) \leq n\vec{v}(m - 2^{\lfloor p/2 \rfloor + 1}) \leq n\vec{v}(m - 2\sqrt{m})$. For Copeland, the total number of times that c is ranked higher than an alternative (across all votes) is no more than $2^{\lfloor p/2 \rfloor + 1} n$. Therefore, $S_{\text{C}}(P, c) \leq 2^{\lfloor p/2 \rfloor + 1} n/(n/2) \leq 4\sqrt{m}$. Hence, $S_{\text{C}}(P, \text{Copeland}(P))/S_{\text{C}}(P, c) = \Omega(\sqrt{m})$. □

The result for k-approval when $k < m - 2\sqrt{m}$ considerably strengthens the result obtained for plurality in [Brams, Kilgour, and Zwicker, 1998] (see Example 4), thus showing that the multiple election paradoxes go far beyond plurality voting.

Theorem 1. *When profiles are $\mathcal{O}$-legal (or profiles are separable), $Seq_{\mathcal{O}}^{maj}$ is a $\Theta(\sqrt{m})$-approximation to Borda w.r.t. S_{Borda}.*

Proof. We only need to prove the lower bound for separable profiles and the upper bound for $\mathcal{O}$-legal profiles. The lower bound has already been proved in Proposition 3. We next prove the upper bound for $\mathcal{O}$-legal profiles. Let $P = (V_1, \ldots, V_n)$ be an $\mathcal{O}$-legal profile. Without loss of generality, $\text{Seq}_{\mathcal{O}}^{maj}(P) = \vec{1} = (1_1, \ldots, 1_p)$. For any $j \leq n$, let $\mathcal{I}_j \subseteq \mathcal{I}$ denote the set of issues $\mathbf{x}_i$ such that $1_1 \cdots 1_{i-1} : 1_i \succ_{V_j} 0_i$. We have the following claim.

Claim 1. *For any $j \leq n$, there are at least $2^{|\mathcal{I}_j|} - 1$ alternatives ranked lower than $\vec{1}$ in V_j.*

Proof. For any $\vec{d} = (d_1, \ldots, d_p) \in \mathcal{X}$ such that $\vec{d}$ takes value 1 on all issues outside $\mathcal{I}_j$ (and $\vec{d}$ takes value 0 on at least one issue in $\mathcal{I}_j$), we next prove that $\vec{1} \succ_{V_j} \vec{d}$. Let $\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_L}$ be the issues for which $\vec{d}$ takes value 0 (with $L \geq 1$, $\{\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_L}\} \subseteq \mathcal{I}_j$, and $i_1 < i_2 < \cdots < i_L$). We recall that for any $\mathbf{x}_i \in \mathcal{I}_j$, $1_1 \cdots 1_{i-1} : 1_i \succ_{V_j} 0_i$. Therefore, for any $l \leq L$, we have the following preference relationship. $(1_1, \ldots, 1_{i_l-1}, 1_{i_l}, d_{i_l+1}, \ldots, d_p) \succ_{V_j} (1_1, \ldots, 1_{i_l-1}, 0_{i_l}, d_{i_l+1}, \ldots, d_p) = (1_1, \ldots, 1_{i_l-1}, d_{i_l}, d_{i_l+1}, \ldots, d_p)$. We obtain the following preference relationship by chaining the above preference relationships.

$$
\begin{aligned}
&(1_1, \ldots, 1_p)\\
\succ_{V_j} &(1_1, \ldots, 1_{i_L-1}, 0_{i_L}, 1_{i_L+1}, \ldots, 1_p)\\
&= (1_1, \ldots, 1_{i_L-1}, d_{i_L}, d_{i_L+1}, \ldots, d_p)\\
\succ_{V_j} &(1_1, \ldots, 1_{i_{L-1}-1}, 0_{i_{L-1}}, 1_{i_{L-1}+1}, \ldots, 1_{i_L-1}, d_{i_L}, \ldots, d_p)\\
&= (1_1, \ldots, 1_{i_{L-1}-1}, d_{i_{L-1}}, d_{i_{L-1}+1}, \ldots, d_p)\\
&\vdots\\
\succ_{V_j} &(1_1, \ldots, 1_{i_1-1}, 0_{i_1}, 1_{i_1+1}, \ldots, 1_{i_2-1}, d_{i_2}, \ldots, d_p)\\
&= (d_1, \ldots, d_p) = \vec{d}
\end{aligned}
$$

The claim follows from the fact that the number of such alternatives $\vec{d}$ is $2^{|\mathcal{I}_j|} - 1$. □

Because $\vec{1}$ is the sequential winner, $\sum_{j=1}^{n} |\mathcal{I}_j| \geq p(n+1)/2$. Note that $f(x) = 2^x$ is convex. We have the following calculation.

$$
\begin{aligned}
S_{\text{Borda}}(P, \vec{1}) \geq &\sum_{j=1}^{n} 2^{|\mathcal{I}_j|}\\
\geq &n 2^{\sum_{j=1}^{n} |\mathcal{I}_j|/n} \qquad \text{(Jensen's inequality)}\\
\geq &n 2^{(n+1)p/(2n)} > n 2^{p/2} = n\sqrt{m}
\end{aligned}
$$

We note that $S_{\text{Borda}}(P, \text{Borda}(P)) \leq n(m-1)$. Therefore, $1 \leq S_{\text{Borda}}(P, \text{Borda}(P))/S_{\text{Borda}}(P, \vec{1}) \leq \sqrt{m}$, which proves the upper bound. □

Due to the space constraint, the proofs for most theorems below are omitted.

Theorem 2. *When profiles are $\mathcal{O}$-legal (or profiles are separable), $Seq_{\mathcal{O}}^{maj}$ is a $\Theta(m)$-approximation to Bucklin w.r.t. S_{Bl}.*

Theorem 3. *When profiles are $\mathcal{O}$-legal (or profiles are separable), $Seq_{\mathcal{O}}^{maj}$ is a $\Theta(m/\log m)$-approximation to Copeland w.r.t. S_C.*

Theorem 4. *When profiles are $\mathcal{O}$-legal (or profiles are separable), $Seq_{\mathcal{O}}^{maj}$ is a $[2n/(n+1)]$-approximation to maximin w.r.t. S_{MM}.*

Theorem 5. *When profiles are $\mathcal{O}$-legal (or profiles are separable), $Seq_{\mathcal{O}}^{maj}$ is an $\Omega(m)$-approximation to Dodgson w.r.t. S_D.*

4 Approximating common rules when profiles are composed of LP-trees/lexicographic orders

In this section, we study the approximation ratio when the profiles are composed of LP-trees with the same structure, or are lexicographic. Suppose r_1 is a θ_1-approximation (resp., θ_2-approximation) to r_2 w.r.t. S when the profiles are composed of LP-trees with the same structure (resp., lexicographic orders), then $\theta_2 \leq \theta_1$, because any lexicographic order can be represented by an LP-tree. We note that if a profile P composed of LP-trees with the same structure T is also $\mathcal{O}$-lexicographic, then $\text{Seq}_T^{maj}(P) = \text{Seq}_{\mathcal{O}}^{maj}(P)$.

We have the following positive result, strengthening Proposition 3 in [Lang and Xia, 2009].

Theorem 6. *For any profile P of LP-trees with the same structure T, $Seq_T^{maj}(P)$ is the Condorcet winner for P.*

Therefore, for any Condorcet consistent voting rule r (including Copeland, maximin, and Dodgson), the sequential majority winner is the same as the winner under r, which means that when profiles are composed of LP-trees (with the same structure), the sequential majority rule is a 1-approximation to any Condorcet consistent voting rule.

Theorem 7. *When profiles are composed of LP-trees with the same structure T (resp., profiles are $\mathcal{O}$-lexicographic), Seq_T^{maj} (resp., $Seq_{\mathcal{O}}^{maj}$) is a $(3/2 + o(1))$-approximation to Borda w.r.t. S_{Borda}.*

Proof. We first prove the lower bound for lexicographic profiles. We define n CP-nets $\mathcal{N}_1^*, \ldots, \mathcal{N}_n^*$ as follows. For any $j \leq (n+1)/2$, let $1_1 \succ 0_1$ in $\mathcal{N}_j^*$; for any j such that $(n+1)/2 \leq j \leq n$ and any i such that $2 \leq i \leq p$, let $1_1 \cdots 1_{i-1} : 1_i \succ 0_i$ in $\mathcal{N}_j^*$; and for all $j \leq n$, let all the other local preferences in $\mathcal{N}_j^*$ not defined above be $0 \succ 1$. Let $V_1^*, \ldots, V_n^*$ be the lexicographic extensions of $\mathcal{N}_1^*, \ldots, \mathcal{N}_n^*$, respectively. Let $P^* = (V_1^*, \ldots, V_n^*)$. We have that $\text{Seq}_{\mathcal{O}}^{maj}(P^*) = \vec{1}$, $S_{\text{Borda}}(P^*, \vec{1}) = 2^{p-1}(n-1)/2 + (2^p - 1) + (2^{p-1} - 1)(n-1)/2 = 2^{p-1}(n+1) - (n+1)/2 = (m-1)(n+1)/2 = mn/2 + o(mn)$. We have $S_{\text{Borda}}(P^*, \vec{0}) = (2^{p-1} - 1)(n+1)/2 + (2^p - 1)(n-1)/2 = 3mn/4 + o(mn)$. Therefore, the approximation ratio is at least $\frac{3mn/4+o(mn)}{mn/2+o(mn)} = 3/2 + o(1)$.

Next, we prove the upper bound for profiles composed of LP-trees with the same structure. Let $P = (\mathcal{T}_1, \ldots, \mathcal{T}_n)$ be composed of n LP-trees with the same structure T. (We recall that for any LP-tree $\mathcal{T}$, we also use $\mathcal{T}$ to represent the linear order to which it corresponds.) W.l.o.g. $\text{Seq}_T^{maj}(P) = \vec{1}$, and in T, $\mathbf{x}_1$ is the issue labeling the root, and its outgoing edge labeled 1 or $\{0, 1\}$ goes to a node labeled by $\mathbf{x}_2$, whose outgoing edge labeled 1 or $\{0, 1\}$ goes to a node labeled by $\mathbf{x}_3$, etc. For any LP-tree $\mathcal{T}$ whose structure is T, we define $I(\mathcal{T})$ to be a set that is composed of all $i \leq p$ such that in the CPT of the node labeled by $\mathbf{x}_i$ along the branch $\mathbf{x}_1 \xrightarrow{1} \mathbf{x}_2 \xrightarrow{1} \cdots \xrightarrow{1} \mathbf{x}_p$, we have $1_i \succ 0_i$ given $1_1 \cdots 1_{i-1}$. We have the following two lemmas, whose proofs are straightforward and are therefore omitted.

Lemma 1. *For any LP-tree $\mathcal{T}$ with structure T, $S_{Borda}(\mathcal{T}, \vec{1}) = \sum_{i \in I(\mathcal{T})} 2^{p-i}$.*

Lemma 2. *Let $\vec{d} \neq \vec{1}$. For any $i \leq p$ such that $d_1 = 1_1, \ldots, d_i = 1_i$, $|S_{Borda}(\mathcal{T}, \vec{d}) - S_{Borda}(\mathcal{T}, \vec{1})| < 2^{p-i}$.*

Claim 2. *For any profile P composed of LP-trees with the same structure T, $S_{Borda}(P, Seq_T^{maj}(P)) \geq (2^p - 1)(n+1)/2$.*

Proof. W.l.o.g. $\text{Seq}_T^{maj}(P) = \vec{1}$. By Lemma 1, $S_{\text{Borda}}(P, \vec{1}) = \sum_{j=1}^{n} \sum_{i \in I(\mathcal{T}_j)} 2^{p-i}$. Because $\vec{1}$ is the sequential winner, for any $i \leq p$, i is in an $I(\mathcal{T}_j)$ at least $(n+1)/2$ times. Hence, we have that $S_{\text{Borda}}(P, \vec{1}) \geq \sum_{i \leq p} 2^{p-i}(n+1)/2 = (2^p - 1)(n+1)/2$. □

Now, we prove the upper bound by induction on p. We note that even though in this paper we assume that $p \geq 2$, in this proof, the base case is $p = 1$. It is easy to check that when $p = 1$, $\text{Seq}_T^{maj}(P) = maj(P) = \text{Borda}(P)$, which means that the approximation ratio is 1. Hence, the upper bound holds for $p = 1$.

Suppose the upper bound holds for $p - 1$. We next prove that it also holds for p. Let $\vec{d}$ be an arbitrary alternative. We prove the upper bound in the following cases.

Case 1: $d_1 = 0_1$. Because $\vec{d}$ is ranked within the bottom 2^{p-1} positions for at least $(n+1)/2$ times (in those LP-trees where $1_1 \succ 0_1$), we have $S_{\text{Borda}}(P, \vec{d}) < 2^{p-1}(n+1)/2 + 2^p(n-1)/2$. Therefore, by Claim 2, $S_{\text{Borda}}(P, \vec{d})/S_{\text{Borda}}(P, \vec{1}) < 3/2 + o(1)$.

Case 2: $d_1 = 1_1$. For any $j \leq n$, we let $\mathcal{T}'_j$ denote the sub-LP-tree of $\mathcal{T}_j$ whose root is the child of $\mathbf{x}_1$ following the edge labeled by 1 or $\{0, 1\}$. It follows that $P' = (\mathcal{T}'_1, \ldots, \mathcal{T}'_n)$ is a profile of LP-trees with the same structure defined over the multi-issue domain $D_2 \times \cdots \times D_p$. Let $K = |j \leq n : 1_1 \succ_{\mathcal{T}_j} 0_1|$. Let $\vec{d}'$ denote the alternative in $D_2 \times \cdots \times D_p$ such that $\vec{d} = (1_1, \vec{d}')$. By Lemma 1 we have $S_{\text{Borda}}(P, \vec{1}) = 2^{p-1}K + S_{\text{Borda}}(P', \vec{1})$ and $S_{\text{Borda}}(P, \vec{d}) = 2^{p-1}K + S_{\text{Borda}}(P', \vec{d}')$. Therefore, if $S_{\text{Borda}}(P, \vec{d}) > S_{\text{Borda}}(P, \vec{1})$, then $\frac{S_{\text{Borda}}(P, \vec{d})}{S_{\text{Borda}}(P, \vec{1})} = \frac{2^{p-1}K + S_{\text{Borda}}(P', \vec{d}')}{2^{p-1}K + S_{\text{Borda}}(P', \vec{1})} < \frac{S_{\text{Borda}}(P', \vec{d}')}{S_{\text{Borda}}(P', \vec{1})} \leq \frac{3}{2} + o(1)$. The last inequality follows from the induction hypothesis.

Therefore, the upper bound holds for all $p \in \mathbb{N}$. This completes the proof. □

Theorem 8. *Let* $\theta(n) = \begin{cases} \infty & \text{when } k < m/4 \\ \Theta(n) & \text{when } m/4 \leq k < m/2 \\ \Theta(1) & \text{when } m/2 \leq k \leq m \end{cases}$. *When profiles are composed of LP-trees with the same structure T (resp., profiles are $\mathcal{O}$-lexicographic), Seq_T^{maj} (resp., $Seq_{\mathcal{O}}^{maj}$) is a $\theta(n)$-approximation to k-approval w.r.t. S_{App}^k.*

Theorem 9. *When profiles are composed of LP-trees with the same structure T (resp., profiles are $\mathcal{O}$-lexicographic), Seq_T^{maj} (resp., $Seq_{\mathcal{O}}^{maj}$) is a $(2 + o(1))$-approximation to Bucklin w.r.t. S_{Bl}.*

5 Future work

There are many intriguing questions for future research. Can we find other restrictions on voter preferences such that sequential majority is a good approximation to common voting rules? Are there any voting rules with low overhead (in terms of computation and communication) that are good approximations to common voting rules when profiles are $\mathcal{O}$-legal? Can we generalize to multi-issue domains composed of non-binary issues? Can we design efficient algorithms that compute the winners for common voting rules directly when profiles are separable/$\mathcal{O}$-legal/lexicographic/represented by LP-trees?

References

[Booth et al., 2010] Booth, R.; Chevaleyre, Y.; Lang, J.; Mengin, J.; and Sombattheera, C. 2010. Learning conditionally lexicographic preference relations. In *ECAI*, 269–274.

[Boutilier et al., 2004] Boutilier, C.; Brafman, R.; Domshlak, C.; Hoos, H.; and Poole, D. 2004. CP-nets: A tool for representing and reasoning with conditional ceteris paribus statements. *JAIR* 21:135–191.

[Brams, Kilgour, and Zwicker, 1998] Brams, S.; Kilgour, D.; and Zwicker, W. 1998. The paradox of multiple elections. *Social Choice and Welfare* 15(2):211–236.

[Caragiannis et al., 2009] Caragiannis, I.; Covey, J. A.; Feldman, M.; Homan, C. M.; Kaklamanis, C.; Karanikolas, N.; Procaccia, A. D.; and Rosenschein, J. S. 2009. On the approximability of dodgson and young elections. In *SODA*, 1058–1067.

[Caragiannis et al., 2010] Caragiannis, I.; Kaklamanis, C.; Karanikolas, N.; and Procaccia, A. D. 2010. Socially Desirable Approximations for Dodgson's Voting Rule. In *EC*, 253–262.

[Fischer, Procaccia, and Samorodnitsky, forthcoming 2011] Fischer, F.; Procaccia, A. D.; and Samorodnitsky, A. forthcoming, 2011. A new perspective on implementation by voting trees. In *Random Structures and Algorithms*.

[Lacy and Niou, 2000] Lacy, D., and Niou, E. M. 2000. A problem with referendums. *Journal of Theoretical Politics* 12(1):5–31.

[Lang and Xia, 2009] Lang, J., and Xia, L. 2009. Sequential composition of voting rules in multi-issue domains. *Math. Social Sciences* 57(3):304–324.

[Procaccia, 2010] Procaccia, A. D. 2010. Can approximation circumvent gibbard-satterthwaite? In *AAAI*, 836–841.

[Xia, Conitzer, and Lang, 2010] Xia, L.; Conitzer, V.; and Lang, J. 2010. Strategic sequential voting in multi-issue domains and multiple-election paradoxes. In *LOFT*.