Paradoxes of Multiple Elections: An Approximation Approach

Vincent Conitzer
Department of Computer Science
Duke University
Durham, NC 27708, USA
conitzer@cs.duke.edu

Lirong Xia
School of Engineering and Applied Sciences
Harvard University
Cambridge, MA 02138, USA
lxia@seas.harvard.edu

Abstract

When agents need to make decisions on multiple issues, applying common voting rules becomes computationally hard due to the exponentially large number of alternatives. One computationally efficient 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 that admit natural *scoring functions* that 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. We show some generalized *paradoxes of multiple elections*: Sequential voting does not approximate many common voting rules well even when the preferences are $\mathcal{O}$-legal or separable. However, these paradoxes are much alleviated or even completely avoided when the preferences are lexicographic or represented by LP-trees. Our results thus draw a border for conditions under which sequential voting rules, which have extremely low computational and communicational cost, are good approximations of some common voting rules w.r.t. their corresponding scoring functions.

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 focuses on plurality. We may wonder to what extent it extends to other voting rules. We will focus on those 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 answer these questions 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. These ratios will help us to identify voting rules that issue-wise majority approximates best, and thus 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 or 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.

Our Contributions

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. 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 and n is the number of voters.

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

Table 1: The approximation ratio obtained by the sequential winner for several common rules with a natural scoring function. We investigate two score functions for maximin.

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. 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.

In addition to these technical contributions, we also feel that the approximation approach taken in this paper has the following conceptual contributions. On the social choice side, we view traditional paradoxes of multiple elections as inapproximability results, which provides us a principled way to examine whether voting in multi-issue domains exhibits more general paradoxes. On the computer science side, we show that voting in multi-issue domains is a rich setting where the idea of approximation can be adopted to explore the tradeoff between computational complexity of the voting process and the quality of the winner.

Previous Work

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); approximating Copeland by voting trees has been studied in (Fischer, Procaccia, and Samorodnitsky forthcoming 2011); and scoring functions have been used as the basis for computing alternatives that minimize the maximum regret when there is some uncertainty in the profile (Lu and Boutilier 2011). We note that in all of 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 directly applying most common voting rules becomes computationally intractable), and we restrict the voters' preferences.

Discussions

One important question to ask is: Does this idea of approximation based on specific scoring functions make sense? Technically, any voting rule can be defined as the maximizer or minimizer of some scoring function (for example, the score of an alternative can be 1 if it wins and 0 if it loses). We feel that the approximation approach in multi-issue domains is justified in two ways. First, as we have mentioned earlier, in the social choice literature, paradoxes of multiple elections can be interpreted as a special type of inapproximability results w.r.t. plurality scores. It is natural to study the approximability and inapproximability w.r.t. some other well-defined metrics, e.g., Borda score, Bucklin score, Copeland score, etc. Second, as it has been argued in previous work, voting rules studied in this paper admit very natural scoring functions that measure the quality of an alternative. Hence, these scoring functions serve well as bases for approximation. For example, the Borda score of an alternative (which we will define shortly) can be seen as the social welfare in terms of a specific type of utility functions. Moreover, in combinatorial domains, the number of alternatives is already exponentially large, so that it is not clear how to efficiently compute the winners for these common voting rules.[1] It should also be noted that a scoring function can produce ties, but this does not matter from the perspective of approximation.

[1]Brams, Kilgour, and Zwicker (1997) argued that, when the number of issues is not large, applying common voting rules seems to be better than applying sequential majority.

Preliminaries

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 or minimizes the score according to a particular scoring function. Below are some common voting rules. In all these voting rules, we assume that n is odd.

• *Positional scoring rules*: Given a *scoring vector* $\vec{v} = (v(1), \ldots, v(m))$ composed of m integers, 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 $P = (V_1, \ldots, V_n)$, 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 $(m-1, m-2, \ldots, 0)$ 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 (=1-approval)*, for which the scoring vector is $(1, 0, \ldots, 0)$.

• *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. Alternatively, we can define the score of c to be the minimum number of times that c beats another alternative in their pairwise election, and maximin selects the alternative that maximizes this score. That is, let $S_{\text{MMA}}(P, c) = \min\{N_P(c, c') : c' \in \mathcal{X}, c' \neq c\}$, and the winner is an alternative c that has the highest S_{MMA}. The maximin rule itself, of course, is unchanged. As we will see later, different metrics give us different approximation ratios.

• *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 selects the Condorcet winner whenever one exists. For example, Copeland, maximin, and Dodgson are Condorcet consistent.

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 $D_Y = \prod_{\mathbf{x}_i \in Y} D_i$.

CP-nets are a popular language that is used to model preferences over multi-issue domains (Boutilier et al. 2004). A CP-net $\mathcal{N}$ over $\mathcal{X}$ consists of the following two parts:

(a) a directed graph $G = (\mathcal{I}, E)$, and

(b) for each issue $\mathbf{x}_i$, a *conditional preference table* $\mathsf{CPT}(\mathbf{x}_i)$, which consists of conditional linear preferences $\succ^i_{\vec{d}}$ over D_i, for every setting $\vec{d}$ of the parents of $\mathbf{x}_i$ in G (denoted by $\mathsf{Par}_G(\mathbf{x}_i)$).

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*) an ordering $\mathcal{O} = \mathbf{x}_1 > \mathbf{x}_2 > \cdots > \mathbf{x}_p$, if $\mathbf{x}_i$ being a parent of $\mathbf{x}_j$ in the graph of $\mathcal{N}$ 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, and so forth; 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 (respectively, separable or lexicographic) if each of its votes is $\mathcal{O}$-legal (respectively, separable or 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 $\text{Seq}_{\mathcal{O}}^{maj}$. For any $\mathcal{O}$-legal profile P, $\text{Seq}_{\mathcal{O}}^{maj}(P) = (d_1, \ldots, d_p) \in \mathcal{X}$ is defined as follows. Let 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 path from the root to a leaf; 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 $\mathsf{Inst}(t)$ (respectively, $\mathsf{NonInst}(t)$) denote the set of issues in $\mathsf{Anc}(t)$ that have two (respectively, 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, and so on. 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 $\vec{d} \succ_{\mathcal{T}} \vec{e}$; 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* $0_1 0_2 0_3$.

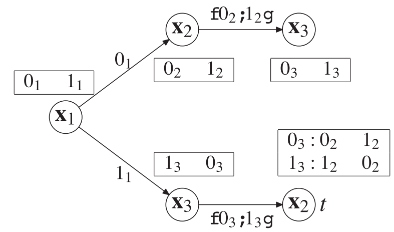

Figure 1: An LP-tree.

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 (but the labels on the nodes and branches must be the same). 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. We note that if a profile P composed of LP-trees with the same structure T is also $\mathcal{O}$-lexicographic, then $\text{Seq}_T^{maj}(P) = \text{Seq}_{\mathcal{O}}^{maj}(P)$.

Approximating Common Rules when Profiles are $\mathcal{O}$-legal or 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*

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

By definition, if r_2 selects an alternative whose S score is maximized, then $\theta = \max_P \left\{ \frac{S(P,r_2(P))}{S(P,r_1(P))} \right\}$; if r_2 selects an alternative whose S score is minimized, then $\theta = \max_P \left\{ \frac{S(P,r_1(P))}{S(P,r_2(P))} \right\}$.

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 (respectively, θ_2-approximation) to r_2 w.r.t. S when the profiles are $\mathcal{O}$-legal (respectively, 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, we have the following.*

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, $S_{Borda}(P, Borda(P)) \geq \lceil (m-1)n/2 \rceil$ and $S^k_{App}(P, App_k(P)) \geq \lceil kn/m \rceil$.*
2. *$S_C(P, Copeland(P)) \geq \lceil (m-1)/2 \rceil$.*
3. *$S_{Bl}(P, Bucklin(P)) \leq \lceil (m+1)/2 \rceil$.*
4. *$S_{MM}(P, Maximin(P)) \leq n - 1$.*
5. *$S_{MMA}(P, Maximin(P)) \geq 1$.*
6. *$S_D(P, Dodgson(P)) \leq (m-1)(\lfloor n/2 \rfloor + 1)$.*

The following proposition (which follows from Theorem 4 in (Xia, Conitzer, and Lang 2011)) 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 4 in (Xia, Conitzer, and Lang 2011)). *There exists a separable n-profile P such that $\text{Seq}_{\mathcal{O}}^{maj}(P)$ 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, $\text{Seq}_{\mathcal{O}}^{maj}$ 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^k_{App}; for any positional scoring rule $r_{\vec{v}}$, it is an $\Omega\left(\frac{\sum_{i=1}^m v(i)}{m \cdot v(m - 2\sqrt{m})}\right)$-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 $k < m - 2^{\lfloor p/2 \rfloor+1}$, $S^k_{\text{App}}(P, c) = 0$. 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 multiple election paradoxes go far beyond plurality voting.

Theorem 1. *When profiles are $\mathcal{O}$-legal (or profiles are separable), $\text{Seq}_{\mathcal{O}}^{maj}$ 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.

$$S_{\text{Borda}}(P,\vec{1}) \geq \sum_{j=1}^{n} 2^{|\mathcal{I}_j|} \geq n2^{\sum_{j=1}^{n}|\mathcal{I}_j|/n} \quad \text{(Jensen's inequality)} \geq n2^{(n+1)p/(2n)} > n2^{p/2} = n\sqrt{m}$$

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. □

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}.*

Proof. The upper bound is trivial. We now prove the lower bound for separable profiles. For each $j \leq (n-1)/2$, let $\mathcal{N}_j^*$ be the separable CP-net where for every $i \leq p$, $1_i \succ_{\mathcal{N}_j^*|_{\mathbf{x}_i}} 0_i$. For each $i \leq p$, let $\mathcal{N}_{(n-1)/2+i}^*$ be the separable CP-net whose projection on $\mathbf{x}_i$ is $1_i \succ 0_i$, and for every $i' \neq i$, its projection on $\mathbf{x}_{i'}$ is $0_{i'} \succ 1_{i'}$. For each j such that $(n-1)/2+p+1 \leq j \leq n$, let $\mathcal{N}_j^*$ be the separable CP-net where for every $i \leq p$, its projection on $\mathbf{x}_i$ is $0_i \succ 1_i$.

Let $V_1^*, \ldots, V_n^*$ be extensions of $\mathcal{N}_1^*, \ldots, \mathcal{N}_n^*$, respectively, such that for every $i \leq p$, $\vec{0}$ (respectively, $\vec{1}$) is ranked in the second (respectively, second from the bottom) position in $V_{(n-1)/2+i}^*$. Let $P^* = (V_1^*, \ldots, V_n^*)$. It follows that $\text{Seq}_{\mathcal{O}}^{maj}(P^*) = \vec{1}$ and for any $j \geq (n-1)/2+p+1$, $\vec{0}$ (respectively, $\vec{1}$) is ranked in the top (respectively, bottom) position in V_j^*. We note that for any $i \leq p$, the top-ranked (respectively, bottom-ranked) alternative in $V_{(n-1)/2+i}^*$ differs from $\vec{0}$ (respectively, $\vec{1}$) only on issue $\mathbf{x}_i$. It follows that $S_{\text{Bl}}(P^*,\vec{0}) = 2$ and $S_{\text{Bl}}(P^*,\vec{1}) = m-1$. Hence, the lower bound for separable profiles is $(m-1)/2 = \Omega(m)$. This proves the theorem. □

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.*

Proof. We will use the following lemma, whose proof is straightforward and is thus omitted.

Lemma 1. *Let $\mathcal{N}$ be an acyclic CP-net. For any alternative $\vec{d}$, let $D_{\mathcal{N}}(\vec{d}) = \{\vec{e} : \vec{d} \succ_{\mathcal{N}} \vec{e}\}$. There exists an extension V of $\mathcal{N}$ where $\vec{d}$ is ranked in the $(|D_{\mathcal{N}}(\vec{d})|+1)$th position from the bottom.*

We first prove the lower bound for separable profiles. We construct n separable CP-nets $\mathcal{N}_1^*, \ldots, \mathcal{N}_n^*$ as follows. The idea is to distribute $1 \succ 0$ (for all issues) as even as possible in these CP-nets. For each $1 \leq j \leq p$, we let $I_j = \{j, j+1, \ldots, j+\lfloor p/2 \rfloor\}$ and let $\mathcal{N}_j^*$ be the CP-net where for every $i \in I_j$, $\mathcal{N}_j^*|_{\mathbf{x}_i} = [1_i \succ 0_i]$. Here for any $l \leq p$, we have $\mathbf{x}_{l+p} = \mathbf{x}_l$, $1_{l+p} = 1_l$, and $0_{l+p} = 0_l$. For any $\mathcal{N}_j^*|_{\mathbf{x}_i}$ that is not defined in the previous step, we let $\mathcal{N}_j^*|_{\mathbf{x}_i} = [0_i \succ 1_i]$. Let $k_1 = \cdots = k_p = \lfloor n/2 \rfloor - \lfloor p/2 \rfloor$. For each j such that $p+1 \leq j \leq n$, we let $I_j = \{i_1, \ldots, i_{\lfloor p/2 \rfloor+1}\}$ be the set of indices of the highest k's. Then, for every $i \in I_j$ we let $k_i \leftarrow k_i - 1$, and if $k_i \geq 0$ then we let $\mathcal{N}_j^*|_{\mathbf{x}_i} = [1_i \succ 0_i]$; for any other $\mathcal{N}_j^*|_{\mathbf{x}_i}$ that is not defined in the previous step, we let $\mathcal{N}_j^*|_{\mathbf{x}_i} = [0_i \succ 1_i]$. Because $n \geq 2p+1$, $n(\lfloor p/2 \rfloor + 1) \geq p(\lfloor n/2 \rfloor + 1)$, which means that after $n-p$ steps, for all $i \leq p$, $k_i \leq 0$. For each $i \leq p$, $1_i \succ 0_i$ in exactly $\lfloor p/2 \rfloor + 1$ CP-nets. Let $V_1^*, \ldots, V_n^*$ be extensions of $\mathcal{N}_1^*, \ldots, \mathcal{N}_n^*$, respectively, such that for every $j \leq n$, $\vec{1} = (1_1, \ldots, 1_p)$ is ranked as low as possible in V_j^*. Let $P^* = (V_1^*, \ldots, V_n^*)$. It follows that $\text{Seq}_{\mathcal{O}}^{maj}(P^*) = \vec{1}$.

For any $j \leq n$, let I_j' denote the set of indices i such that $1_i \succ_{\mathcal{N}_j|_{\mathbf{x}_i}} 0_i$. We have $I_j' \subseteq I_j$. It follows that for any $j \leq n$, $|D_{\mathcal{N}_j}(\vec{1})| = 2^{|I_j'|} - 1$, and $\vec{e} \in D_{\mathcal{N}_j}(\vec{1})$ if and only if $\vec{e} \neq \vec{1}$, and $\vec{e}$ takes 0 only on the issues whose indices are in I_j'. We have the following claim.

Claim 2. *Let $\vec{e}$ be an alternative that takes 0 on at least two issues. $\vec{1} \succ \vec{e}$ in more than $n/2$ votes in P^*.*

Proof. We only prove the case where $\vec{e} = (0_1, 0_2, 1_3, \ldots, 1_p)$. The other cases can be proved similarly. By Lemma 1, $\vec{1} \succ \vec{e}$ in V_j if and only $\{1,2\} \subseteq I_j'$. We note that $|j \leq n : 2 \in I_j'| = (n+1)/2$ (we recall that n is odd), $2 \in I_2'$ and $1 \notin I_2'$. Therefore, $\vec{1} \succ \vec{e}$ in no more than $n/2$ votes in P. □

By Claim 2, $\vec{1}$ may only beats an alternative that takes 0 on only one issue. The number of such alternatives is p, which means that $S_{\text{C}}(P) \leq p = \log m$. By Proposition 1, $S_{\text{C}}(P, \text{Copeland}(P)) = \Omega(m)$. Hence, we have the lower bound $\Omega(m/\log m)$ for separable (therefore also $\mathcal{O}$-legal) profiles.

Next, we prove the upper bound for $\mathcal{O}$-legal profiles. For any $i \leq p$, let $\vec{e}_i$ denote the alternative that takes 0_i on $\mathbf{x}_i$, and takes 1 on all other issues. Let $P = (V_1, \ldots, V_n)$ be an $\mathcal{O}$-legal profile and w.l.o.g. $\text{Seq}_{\mathcal{O}}^{maj}(P) = \vec{1}$. For any $j \leq n$ and any $i \leq p$ such that the projection of V_j on $\mathbf{x}_i$ given $1_1 \cdots 1_{i-1}$ is $1_i \succ 0_i$, we have $\vec{1} \succ \vec{e}_i$. Because $\vec{1}$ is the sequential winner, for any $i \leq p$ the majority of voters prefer 1_i to 0_i given $1_1 \cdots 1_{i_1}$, which means that $\vec{1}$ beats $\vec{e}_i$ in their pairwise elections. Therefore, $S_{\text{C}}(P,\vec{1}) \geq p = \log m$. Together with the trivial upper bound $S_{\text{C}}(P, \text{Copeland}(P)) \leq m$, we obtain the upper bound on the approximation ratio for $\mathcal{O}$-legal (therefore also seperable) profiles. □

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}; $Seq_{\mathcal{O}}^{maj}$ is an ∞-approximation to maximin w.r.t. S_{MMA}.*

Proof. We first prove the lower bound for separable profiles. Let $\mathcal{N}_1^*, \ldots, \mathcal{N}^*$ denote the CP-nets that are defined in the proof for Theorem 3. Let $V_1^*, \ldots, V_n^*$ denote extensions of $\mathcal{N}_1^*, \ldots, \mathcal{N}^*$, respectively, such that in each V_j^*, $\vec{1}$ ranked as low as possible and $\vec{0}$ is ranked as high as possible (these two objectives are consistent with each other, because there is no alternative $\vec{d}$ such that $\vec{1} \succ_{\mathcal{N}_j^*} \succ \vec{d} \succ_{\mathcal{N}_j^*} \vec{0}$.) Similarly to the proof for Theorem 3, we can prove that $S_{\text{MM}}(P^*,\vec{0}) =$

$(n+1)/2$ (via $(1_1, 0_2, \ldots, 0_p)$). We also note that for any $j \leq n$, $\vec{0} \succ_{V_j} \vec{1}$. Therefore, $S_{\text{MM}}(P^*, \vec{1}) = n$. This gives us the lower bound $2n/(n+1)$ for S_{MM}. Because in each vote $\vec{0}$ is ranked above $\vec{1}$, $S_{\text{MMA}}(\vec{1}) = 0$, and the maximin winner has a non-zero MMA score, which means that the approximation ratio for maximin w.r.t. S_{MMA} is ∞.

Next, we prove the upper bound for $\mathcal{O}$-legal profiles. Let P be an $\mathcal{O}$-legal profile. W.l.o.g. $\text{Seq}_{\mathcal{O}}^{maj}(P) = \vec{1}$. For any alternative $\vec{d} \neq \vec{1}$, let i^* be the smallest number such that $d_{i^*} = 0_{i^*}$. Let $\vec{e}$ be the alternative that differs from $\vec{d}$ only on $\mathbf{x}_{i^*}$. Because $\vec{1}$ is the sequential winner, in at least $(n+1)/2$ votes we have $1_1 \cdots 1_{i^*-1} : 1_{i^*} \succ 0_{i^*}$. Therefore, in each of these votes we must have that $\vec{e} \succ \vec{d}$, which means that $S_{\text{MM}}(P, \vec{d}) \geq (n+1)/2$. Because $S_{\text{MM}}(P, \vec{1}) \leq n$, the approximation ratio is at most $2n/(n+1)$. This proves the theorem. □

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

Proof. The lower bound can be proved using the same separable profile P^* defined in the proof for Theorem 3. □

Approximating Common Rules when Profiles are Composed of LP-Trees or 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. We will see that for each voting rule and its scoring function studied in this paper, the approximation ratio is significantly improved, which means that sequential voting rules provides much better approximations of common voting rules in such cases.

Suppose r_1 is a θ_1-approximation (respectively, θ_2-approximation) to r_2 w.r.t. S when the profiles are composed of LP-trees with the same structure (respectively, lexicographic orders), then $\theta_2 \leq \theta_1$, because any lexicographic order can be represented by an LP-tree.

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, $\text{Seq}_T^{maj}(P)$ is the Condorcet winner for P.*

Proof. W.l.o.g. $\text{Seq}_T^{maj}(P) = \vec{1}$ and the root is labeled $\mathbf{x}_1$, which has an outgoing edge labeled 1 or $\{0,1\}$ to a node labeled $\mathbf{x}_2$, which has an outgoing edge labeled 1 or $\{0,1\}$ to a node labeled $\mathbf{x}_3$, etc. Let $\vec{d} \neq \vec{1}$. Let i^* be the smallest number such that $d_{i^*} = 0_{i^*}$. Let t denote the node reached from the root by following the branch labeled with 1 or $\{0,1\}$ a total of $i^* - 1$ times. It follows that $\mathsf{Iss}(t) = \mathbf{x}_{i^*}$. Then, because $\vec{1}$ is the sequential majority winner, the majority of voters prefer 1_{i^*} to 0_{i^*} at t, given that all issues in $\mathsf{Anc}(t)$ take value 1. Therefore, these votes also prefer $\vec{1}$ to $\vec{d}$, which means that $\vec{1}$ beats $\vec{d}$ in their pairwise election. Hence, $\vec{1}$ is the Condorcet winner. □

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 (respectively, profiles are $\mathcal{O}$-lexicographic), Seq_T^{maj} (respectively, $\text{Seq}_{\mathcal{O}}^{maj}$) 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 2. *For any LP-tree $\mathcal{T}$ with structure T, $S_{\text{Borda}}(\mathcal{T}, \vec{1}) = \sum_{i \in I(\mathcal{T})} 2^{p-i}$.*

Lemma 3. *Let $\vec{d} \neq \vec{1}$. For any $i \leq p$ such that $d_1 = 1_1, \ldots, d_i = 1_i$, $|S_{\text{Borda}}(\mathcal{T}, \vec{d}) - S_{\text{Borda}}(\mathcal{T}, \vec{1})| < 2^{p-i}$.*

Claim 3. *For any profile P composed of LP-trees with the same structure T, $S_{\text{Borda}}(P, \text{Seq}_T^{maj}(P)) \geq (2^p - 1)(n+1)/2$.*

Proof. W.l.o.g. $\text{Seq}_T^{maj}(P) = \vec{1}$. By Lemma 2, $S_{\text{Borda}}(P, \vec{1}) = \sum_{j=1}^{n} \sum_{i \in I(\mathcal{T}_j)} 2^{p-i}$. 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.

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 3, $S_{\text{Borda}}(P, \vec{d})/S_{\text{Borda}}(P, \vec{1}) < 3/2 + o(1)$.

Case 2: $d_1 = 1_1$. For any $j \le n$, 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 $P' = (\mathcal{T}'_1, \dots, \mathcal{T}'_n)$ is a profile of LP-trees with the same structure defined over the multi-issue domain $D_2 \times \cdots \times D_p$. Let $K = |j \le n : 1_1 \succ_{\mathcal{T}_j} 0_1|$. Let $\vec{d'}$ denote the alternative in $D_2 \times \cdots \times D_p$ such that $\vec{d} = (1_1, \vec{d'})$. By Lemma 2 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 $\frac{S_{\text{Borda}}(P, \vec{d})}{S_{\text{Borda}}(P, \vec{1})} = \frac{2^{p-1}K + S_{\text{Borda}}(P', \vec{d'})}{2^{p-1}K + S_{\text{Borda}}(P', \vec{1})} < \frac{S_{\text{Borda}}(P', \vec{d'})}{S_{\text{Borda}}(P', \vec{1})} \le \frac{3}{2} + o(1)$. 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*

$$\theta(n) = \begin{cases} \infty & \text{when } k < m/4 \\ \Theta(n) & \text{when } m/4 \le k < m/2 \\ \Theta(1) & \text{when } m/2 \le k \le m \end{cases}$$

If profiles are composed of LP-trees with the same structure T (respectively, profiles are $\mathcal{O}$-lexicographic), then Seq_T^{maj} (respectively, $Seq_{\mathcal{O}}^{maj}$) is a $\theta(n)$-approximation to k-approval w.r.t. S_{App}^k.

Proof. Case 1: $k < m/4$. We prove that there exists an $\mathcal{O}$-lexicographic profile $P^* = (V_1^*, \dots, V_n^*)$ such that $\text{Seq}_{\mathcal{O}}^{maj}(P^*) = \vec{1}$ and for all $j \le n$, $\vec{1}$ is not ranked within the top $2^{p-2} - 1$ positions in V_j. For any $j \le p$, let $1_1 \succ 0_1$ and $1_1 \cdots 1_{j-1} : 1_j \succ 0_j$ in $\mathcal{N}_j$; for any j such that $p + 1 \le j \le (n+1)/2$, let $1_1 \succ 0_1$ in $\mathcal{N}_j^*$; for any j such that $(n+1)/2 + 1 \le j \le n$, for every $i \le p$, let $1_1 \cdots 1_{i-1} : 1_i \succ 0_i$ in $\mathcal{N}_j^*$. All the other local preferences in $\mathcal{N}_1^*, \dots, \mathcal{N}_n^*$ that are not mentioned above are defined to be $0 \succ 1$. For any $j \le n$, let V_j^* be the lexicographic extension of $\mathcal{N}_j^*$ w.r.t. $\mathcal{O}$. It is easy to check that $\text{Seq}_{\mathcal{O}}^{maj}(P^*) = \vec{1}$ and $\vec{1}$ is not ranked within the top $2^{p-2} - 1$ positions in all votes.

Case 2: $m/4 \le k < m/2$. The lower bound is proved by the same profile P^* defined in Case 1. We next prove the upper bound for all profiles composed of LP-trees with the same structure T. Let $P' = (\mathcal{T}_1, \dots, \mathcal{T}_n)$ be such a profile and $\text{Seq}_T^{maj}(P) = \vec{1}$. Again, we assume that there is a branch $\mathbf{x}_1 \xrightarrow{1} \mathbf{x}_2 \xrightarrow{1} \cdots \xrightarrow{1} \mathbf{x}_p$. There exists $j \le n$ where $1_1 \succ 0_1$ in $\mathcal{T}_j$ and $1_1 : 1_2 \succ 0_2$ in the CPT of the child of the root following the edge labeled by 1 or $\{0, 1\}$. By Lemma 2 (and also note that the Borda score c in a vote V is the number of alternatives that is ranked below c in V), $\vec{1}$ is ranked within the top 2^{p-2} positions, which means that $S_{\text{App}}^k(P, \vec{1}) \ge 1$. This proves the upper bound.

Case 3: $m/2 \le k \le m$. The lower bound is trivial. We note that for any LP-tree $\mathcal{T}$, if the root is labeled $\mathbf{x}_1$ and $1_1 \succ 0_1$ in T, then $\vec{1}$ is ranked within the top $m/2$ positions. Hence, for any profile P composed of LP-trees with the same structure T, if $\text{Seq}_T^{maj}(P) = \vec{1}$, then $S_{\text{App}}^k(P, \vec{1}) \ge (n+1)/2$. This gives the upper bound. □

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

Proof. The lower bound can be proved by the same lexicographic profile P^* that is used to prove the lower bound in the proof of Theorem 7.

We next prove the upper bound for all profiles composed of LP-trees with the same structure. Let $P = (\mathcal{T}_1, \dots, \mathcal{T}_n)$ be a profile composed of n LP-trees with the same structure T. Suppose $\text{Seq}_T^{maj}(P) = \vec{1} \ne \vec{d} = \text{Bucklin}(P)$. W.l.o.g. there is a path $\mathbf{x}_1 \xrightarrow{1} \mathbf{x}_2 \xrightarrow{1} \cdots \xrightarrow{1} \mathbf{x}_p$ in T. For any $i \le p$, let t_i denote the node in T that is reached from the root via the path $\mathbf{x}_1 \xrightarrow{1} \mathbf{x}_2 \xrightarrow{1} \cdots \xrightarrow{1} \mathbf{x}_i$. That is, t_i is labeled by $\mathbf{x}_i$. For any $i \le p$, we say that an LP-tree $\mathcal{T}$ whose structure is T has depth i, if for every $i' \le i$, the CPT of $t_{i'}$ contains $1_{i'} \succ 0_{i'}$, given that all issues in $\mathsf{Par}(t_{i'})$ take 1. For any $i \le p$, let $K_l = \{j \le n : \mathcal{T}_j \text{ has depth } i\}$. Let l be the largest number such that $|K_l| \ge (n+1)/2$. Without loss of generality, let $K_l = \{1, \dots, q\}$ (where $q \ge (n+1)/2$). We have $S_{\text{Bl}}(P, \vec{1}) \le 2^{p-l}$.

If there exists $i \le l$ such that $d_i = 0_i$, then for any $j \le q$, $\vec{d}$ is ranked lower than the 2^{p-i}th position in $\mathcal{T}_j$, which means that $S_{\text{Bl}}(P, \vec{d}) > 2^{p-l} \ge S_{\text{Bl}}(P, \vec{1})$. This contradicts the assumption that $\vec{d} = \text{Bucklin}(P)$. Therefore, $d_1 = 1_1, \dots, d_l = 1_l$. We prove the upper bound in the following two cases.

Case 1: $d_{l+1} = 0_{l+1}$. Let $K_l^* \subseteq \{1, \dots, n\}$ be the set of numbers j such that in $\mathcal{T}_j$, $\mathsf{CPT}(t_{l+1})$ has an entry $1_{\mathsf{Par}(t_{l+1})} : 1_{l+1} \succ 0_{l+1}$. Because $\vec{1}$ is the sequential winner, $|K_l^*| \ge (n+1)/2$. We note that in every $j \in K_l^*$, $\vec{d}$ is ranked below the 2^{p-l-1}th position. Consequently, $S_{\text{Bl}}(P, \vec{d}) \ge 2^{p-l-1}$. We recall that $S_{\text{Bl}}(P, \vec{1}) \le 2^{p-l}$. It follows that the approximation ratio is at most $2^{p-l}/2^{p-l-1} = 2 + o(1)$.

Case 2: $d_{l+1} = 1_{l+1}$. For any $j \ge q$, both $\vec{1}$ and $\vec{d}$ are ranked below the 2^{p-l}th position in $\mathcal{T}_j$. For any $j \le q$, $j \notin K_l^*$, and any alternative $\vec{e}$ with $e_1 = 1_1, \dots, e_l = 1_l, e_{l+1} = 0_{l+1}$, we have that $\vec{e}$ is ranked above $\vec{d}$ in $\mathcal{T}_j$. The number of such alternatives $\vec{e}$ is 2^{p-l-1}. Therefore, $\vec{d}$ is ranked within top 2^{p-l-1} positions in no more than $|K_l \cap K_l^*|$ time. By the maximality of l, we have $|K_l \cap K_l^*| \le (n-1)/p$, which means that $S_{\text{Bl}}(P, \vec{d}) \ge 2^{p-l-1}$. Again, because $S_{\text{Bl}}(P, \vec{1}) \le 2^{p-l}$, the approximation ratio is at most $2^{p-l}/2^{p-l-1} = 2 + o(1)$.

Therefore, $2 + o(1)$ is an upper bound on the approximation ratio for profiles composed of LP-trees. □

Summary and Future Work

In this paper, we show how well sequential voting in multi-binary-domains approximates some common voting rules w.r.t. their respective scores. Our results can be interpreted as generalized paradoxes of multiple elections. We showed that when the profiles are $\mathcal{O}$-legal or separable, such paradoxes are quite strong. However, these paradoxes are much alleviated or even completely avoided when the preferences are lexicographic or represented by LP-trees.

Future research may focus on designing computationally tractable voting rules for multi-issue domains. For example, we can ask the following intriguing questions. 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?

Acknowledgments

We thank Jérôme Lang for fruitful discussions and comments. Vincent Conitzer acknowledges NSF IIS-0812113, IIS-0953756, and CCF-1101659, as well as an Alfred P. Sloan fellowship, for support. Lirong Xia is supported by NSF under Grant #1136996 to the Computing Research Association for the CIFellows Project. We thank all anonymous AAAI-11 and KR-12 reviewers of this paper for helpful comments.

References

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

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. *Journal of Artificial Intelligence Research* 21:135–191.

Brams, S. J.; Kilgour, D. M.; and Zwicker, W. S. 1997. Voting on referenda: the separability problem and possible solutions. *Electoral Studies* 16(3):359–377.

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

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 *Proc. of SODA*, 1058–1067.

Caragiannis, I.; Kaklamanis, C.; Karanikolas, N.; and Procaccia, A. D. 2010. Socially desirable approximations for Dodgson's voting rule. In *Proc. of EC*, 253–262.

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

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

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

Lu, T., and Boutilier, C. 2011. Robust approximation and incremental elicitation in voting protocols. In *Proc. of IJCAI*, 287–293.

Procaccia, A. D. 2010. Can approximation circumvent Gibbard-Satterthwaite? In *Proc. of AAAI*, 836–841.

Xia, L.; Conitzer, V.; and Lang, J. 2011. Strategic sequential voting in multi-issue domains and multiple-election paradoxes. In *Proc. of EC*, 179–188.