# Are There Any Nicely Structured Preference Profiles Nearby?

**Robert Bredereck**[*]
TU Berlin, Germany
robert.bredereck@tu-berlin.de

**Jiehua Chen**[†]
TU Berlin, Germany
jiehua.chen@tu-berlin.de

**Gerhard J. Woeginger**[‡]
TU Eindhoven, Netherlands
gwoegi@win.tue.nl

## Abstract

We investigate the problem of deciding whether a given preference profile is close to a nicely structured preference profile of a certain type, as for instance single-peaked, single-caved, single-crossing, value-restricted, best-restricted, worst-restricted, medium-restricted, or group-separable profiles. We measure this distance by the number of voters or alternatives that have to be deleted so as to reach a nicely structured profile. Our results classify all considered problem variants with respect to their computational complexity, and draw a clear line between computationally tractable (polynomial time solvable) and computationally intractable (NP-hard) questions.

## 1 Introduction

The area of Social Choice (and in particular the subarea of *Computational* Social Choice) is full of so-called *negative* results. On the one hand there are many axiomatic impossibility results, and on the other hand there are many computational intractability results. For instance, the famous impossibility result of Arrow [1950] states that there is no perfectly fair way (satisfying certain desirable axioms) of aggregating the preferences of a society of voters into a single preference ordering. As another example, Bartholdi *et al.* [1989] establish that it is computationally intractable (NP-hard) to determine whether some particular candidate wins an election under a voting scheme designed by Lewis Carroll. Most of these negative results hold for general preference profiles where *any* combination of preference orderings may occur.

One branch of Social Choice studies *restricted domains* of preference profiles, where only certain nicely structured combinations of preference orderings are admissible. The standard example for this approach are *single-peaked* preference profiles as introduced by Black [1948]: A preference profile is single-peaked if there exists a linear ordering of the alternatives such that any voter's preference along this ordering is either always strictly increasing, always strictly decreasing, or first strictly increasing and then strictly decreasing. Single-peakedness implies a number of interesting properties, such as non-manipulability (Moulin [1980]) and transitivity of the majority rule (Inada [1969]). Under single-peaked profiles, Arrow's impossibility result collapses. In a similar spirit (but in the algorithmic branch), Walsh [2007], Brandt *et al.* [2010], and Faliszewski *et al.* [2011b] show that many electoral bribery, control and manipulation problems that are NP-hard in the general case become tractable under single-peaked profiles. Besides the single-peaked domain, the literature contains many other *restricted domains* of nicely structured preference profiles (see Section 2 for precise mathematical definitions).

- Sen [1966] and Sen and Pattanaik [1970] introduced the domain of *value-restricted* preference profiles which satisfy the following: for any triple of alternatives, one alternative is not considered as the most preferred by any individual (best-restricted), or one is not considered as the least preferred by any individual (worst-restricted), or one is not considered as the intermediate alternative by any individual (medium-restricted).
- Inada [1964; 1969] considered the domain of *group-separable* preference profiles which satisfy the following: the alternatives can be split into two groups such that every voter prefers every alternative in the first group to those in the second group, or prefers every alternative in the second group to those in the first group. Every group-separable profile is also medium-restricted.
- *Single-caved* [Inada, 1964] preference profiles result from a single-peaked profiles by reversing the preferences of every voter. Sometimes single-caved profiles are also called single-dipped [Klaus *et al.*, 1997].
- *Single-crossing* preference profiles go back to a seminal paper of Roberts [1977] on income taxation. A preference profile is single-crossing if there exists a linear ordering of the voters such that for any pair of alternatives along this ordering, either all voters have the same opinion on the ordering of these two alternatives or there is a single spot where the voters switch from preferring one alternative to the other one.

Just like single-peakedness, each of these restrictions guarantees many nice properties, such as the transitivity of the sim-

[*]Supported by the DFG project PAWS (NI 369/10).
[†]Supported by the Studienstiftung des Deutschen Volkes.
[‡]Supported by DIAMANT (a mathematics cluster of the Netherlands Organization for Scientific Research NWO), and by the Alexander von Humboldt Foundation.

To appear in *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI '13)*, Beijing, China, August 2013. 

| Restriction | Voters deletion | Alternatives del. |
|---|---|---|
| Single-peaked | NP-c (*, Cor. 9) | P (*) |
| Single-caved | NP-c (*, Cor. 9) | P (*) |
| Single-crossing | P (Thm. 10) | NP-c (Thm. 13) |
| Group-separable | NP-c (Cor. 9) | NP-c (Cor. 12) |
| Worst-restricted | NP-c (Thm. 8) | NP-c (Thm. 11) |
| Medium-restricted | NP-c (Thm. 8) | NP-c (Thm. 11) |
| Best-restricted | NP-c (Thm. 8) | NP-c (Thm. 11) |
| Value-restricted | NP-c (Thm. 8) | NP-c (Thm. 11) |

Table 1: Summary of the results where NP-c means NP-complete and P means polynomial-time solvable. Entries marked by "*" are due to Erdélyi *et al.* [2013].

ple majority rule. Unfortunately, real-world elections are almost never single-peaked, value-restricted, group-separable, single-caved or single-crossing. Often there are maverick voters whose vote is determined by race, religion or gender. Mavericks destroy nice combinatorial structures in the preference profiles. In a very recent line of research, Faliszewski *et al.* [2011a] searched a cure against such mavericks, and arrived at *nearly* single-peaked preference profiles: A profile is nearly single-peaked, if it is very close to a single-peaked profile. Of course, there are many mathematical ways of measuring the distance of profiles. Perhaps, the most natural way is (i) by deletion of voters or (ii) by deletion of alternatives. Faliszewski *et al.* analyze various control and bribery problems under such nearly single-peaked profiles, and derive a number of somewhat unexpected results.

In a closely related paper (with a disjoint set of results), Erdélyi *et al.* [2013] study various notions of nearly single-peaked profiles. Besides deletion of voters and deletion of alternatives, they also study distance measures that are based on swapping alternatives in the preferences of some voters, or on introducing additional political axes. However, their investigations are limited to single-peaked profiles.

**Results of this paper.** We investigate the problem of deciding the distance (by deletion of voters and by deletion of alternatives) of a given preference profile to a nicely structured one of a certain type (like being single-crossing, value-restricted, or group-separable). We focus on some of the most fundamental definitions of distance measures and on the most popular restricted domains. Our results draw a clear line between computationally tractable (polynomial-time solvable) and computationally intractable (NP-hard) questions as they classify all considered problem variants with respect to their computational complexity. See Table 1 for an overview.

This paper is organized as follows. Section 2 summarizes all the basic definitions and notations. Section 3 surveys our main results. We conclude in Section 4. Due to lack of space, we defer some proofs to a full version of the paper.

## 2 Preliminaries and Basic Notations

Let $a_1, \ldots, a_m$ be $m$ alternatives and let $v_1, \ldots, v_n$ be $n$ voters. A *preference profile* specifies the *preference orderings* of the voters, where voter $v_i$ ranks the alternatives according to a strict linear order $\succ_i$. For alternatives $a$ and $b$, $a \succ_i b$ means that voter $v_i$ strictly prefers $a$ to $b$. We omit the subscript $i$ if it is clear from the context whose preference ordering we are referring to.

Given two sets $A$ and $B$ of alternatives, we write $A \succ_i B$ to express that voter $v_i$ prefers set $A$ to set $B$, that is, for each alternative $a \in A$ and each alternative $b \in B$ it holds that $a \succ_i b$. In a similar way, we use $A \succ_i a$ to denote that voter $v_i$ prefers every alternative in $A$ to alternative $a$ and $a \succ_i A$ for the reverse situation. If we define the canonical ordering of the alternatives in $A$, then $\langle A \rangle$ denotes this canonical linear ordering. Further, $\langle A_1 \rangle \succ \langle A_2 \rangle$ denotes the preference ordering that is consistent with $\langle A_1 \rangle$ as well as $\langle A_2 \rangle$ and prefers all vertices in $A_1$ to all vertices in $A_2$.

Next, we review some concrete preference profiles of small size with special properties studied in the literature [Ballester and Haeringer, 2011; Bredereck *et al.*, 2012]. We call such profiles *configurations*.

**Value-restricted profiles.** The first three configurations describe profiles with three alternatives where each alternative is at the best, medium, or worst position in some voter's preference ordering.

**Definition 1** (Best-diverse configuration)**.** *A profile with three voters $v_1, v_2, v_3$ and three distinct alternatives $a, b, c$ is a* best-diverse configuration *if it satisfies the following:*

$$v_1 : a \succ_1 \{b, c\}; \quad v_2 : b \succ_2 \{a, c\}; \quad v_3 : c \succ_3 \{a, b\}.$$

**Definition 2** (Medium-diverse configuration)**.** *A profile with three voters $v_1, v_2, v_3$ and three distinct alternatives $a, b, c$ is a* medium-diverse configuration *if it satisfies the following:*

*$v_1$: $b \succ_1 a \succ_1 c$ or $c \succ_1 a \succ_1 b$;*
*$v_2$: $a \succ_2 b \succ_2 c$ or $c \succ_2 b \succ_2 a$;*
*$v_3$: $a \succ_3 c \succ_3 b$ or $b \succ_3 c \succ_3 a$.*

**Definition 3** (Worst-diverse configuration)**.** *A profile with three voters $v_1, v_2, v_3$ and three distinct alternatives $a, b, c$ is a* worst-diverse configuration *if it satisfies the following:*

$$v_1 : \{b, c\} \succ_1 a; \quad v_2 : \{a, c\} \succ_2 b; \quad v_3 : \{a, b\} \succ_3 c.$$

We use these three configurations to characterize several restricted domains: A profile is *best-restricted* (resp. *medium-restricted*, *worst-restricted*) *with respect to a triple $T$ of alternatives* if it contains no three voters that form a best-diverse configuration (resp. a medium-diverse configuration, a worst-diverse configuration) with respect to $T$. A *best-restricted* (resp. *medium-restricted*, *worst-restricted*) *profile* is best-restricted (resp. medium-restricted, worst-restricted) with respect to every possible triple of alternatives.

A profile is *value-restricted* [Sen, 1966] if for any triple $T$ of alternatives, it is best-restricted, medium-restricted, or worst-restricted with respect to $T$.

**Single-peaked profiles and single-caved profiles.** The single-peaked property requires the existence of a "natural" linear ordering of the alternatives: A profile is *single-peaked* [Black, 1948] if there is an ordering $\mathcal{L}$ of alternatives such that for each voter $v$, $\mathcal{L}$ can be split into two orderings $\mathcal{L}_1$ and $\mathcal{L}_2$, and $v$ prefers each alternative $a$ in $\mathcal{L}_1$ to any alternative to $a$'s left while he prefers each alternative $b$ in $\mathcal{L}_2$ to any alternative to $b$'s right. Single-peaked preferences are necessarily worst-restricted. In addition, we need the following configuration to fully characterize the single-peaked domain:

**Definition 4** ($\alpha$-configuration)**.** *A profile with two voters $v_1$ and $v_2$, and four distinct alternatives $a, b, c, d$ is an $\alpha$-configuration if it satisfies the following:*

$v_1$: $a \succ_1 b \succ_1 c$ *and* $d \succ_1 b$;
$v_2$: $c \succ_2 b \succ_2 a$ *and* $d \succ_2 b$.

The $\alpha$-configuration represents a situation where two voters have opposite opinions on the ordering of three alternatives $a, b$ and $c$ but agree that a fourth alternative $d$ is "better" than the one ranked in the middle. The profile is not single-peaked, as alternatives $b$ and $d$ must be put between alternatives $a$ and $c$, but then voter $v_1$ prevents us from putting $b$ next to $a$ and voter $v_2$ prevents us from putting $d$ next to $a$.

Ballester and Haeringer [2011] show that a profile is single-peaked if and only if it is worst-restricted and contains no $\alpha$-configurations. They also show that a profile is *single-caved* if and only if it is best-restricted and contains no $\bar{\alpha}$-configurations, where an $\bar{\alpha}$-configuration is a $\alpha$-configuration with both preference orderings being reversed.

**Group-separable profiles.** The *group-separable* property requires that the set $A$ of alternatives can be partitioned into two subsets $A_1$ and $A_2$ such that for each voter $v_i$: $A_1 \succ_i A_2$ or $A_2 \succ_i A_1$. Group-separable profiles are necessarily medium-restricted. In addition, we need the following configuration to fully characterize this property:

**Definition 5** ($\beta$-configuration)**.** *A profile with two voters $v_1$ and $v_2$ and four distinct alternatives $a, b, c, d$ is a $\beta$-configuration if it satisfies the following:*

$v_1$: $a \succ_1 b \succ_1 c \succ_1 d$; $\quad v_2$: $b \succ_2 d \succ_2 a \succ_2 c$.

The $\beta$-configuration represents a situation where the best and least preferred alternatives of the first voter are $a$ and $d$ which are different from the ones of the second voter: $b$ and $c$. Both voters agree that $b$ is better than $c$, but disagree whether $d$ is better than $a$. This profile is not group-separable: We cannot partition the four alternatives into a one-alternative set and a three-alternatives set as each alternative is ranked in the middle once, and we cannot partition them into two subsets of size two each since voter $v_1$ prevents us from putting alternatives $a$ and $c$ or alternatives $a$ and $d$ together and voter $v_2$ prevents us from putting alternatives $a$ and $b$ together.

A profile is group-separable if and only if it contains neither medium-diverse configurations nor $\beta$-configurations [Ballester and Haeringer, 2011].

**Single-crossing profiles.** The single-crossing property requires the existence of a "natural" linear ordering of the voters. A profile is *single-crossing* if there exists a linear ordering of the voters such that for any two alternatives along this ordering, there is a single spot where the voters switch from preferring one alternative to the other one. To characterize single-crossing preferences, we need the following two configurations.

**Definition 6** ($\gamma$-configuration)**.** *A profile with three voters $v_1$, $v_2$, $v_3$ and six (not necessarily distinct) alternatives $a, b, c, d, e, f$ is a $\gamma$-configuration if it satisfies the following:*

$v_1$: $b \succ_1 a$ *and* $c \succ_1 d$ *and* $e \succ_1 f$;
$v_2$: $a \succ_2 b$ *and* $d \succ_2 c$ *and* $e \succ_2 f$;
$v_3$: $a \succ_3 b$ *and* $c \succ_3 d$ *and* $f \succ_3 e$.

The $\gamma$-configuration represents a situation where each voter disagrees with the other two voters on the ordering of exactly two distinct alternatives. The profile is not single-crossing, as none of the three voters can be put between the other two: The pair $\{a, b\}$ prevents us from putting $v_1$ into the middle, the pair $\{c, d\}$ forbids voter $v_2$ in the middle, and the pair $\{e, f\}$ forbids $v_3$ in the middle.

**Definition 7** ($\delta$-configuration)**.** *A profile with four voters $v_1$, $v_2$, $v_3$, $v_4$ and four (not necessarily distinct) alternatives $a, b, c, d$ is a $\delta$-configuration if it satisfies the following:*

$v_1$: $a \succ_1 b$ *and* $c \succ_1 d$; $\quad v_2$: $a \succ_2 b$ *and* $d \succ_2 c$;
$v_3$: $b \succ_3 a$ *and* $c \succ_3 d$; $\quad v_4$: $b \succ_4 a$ *and* $d \succ_4 c$.

The $\delta$-configuration shows a different kind of voter behavior: Two voters disagree with the other two voters on the ordering of two alternatives, but also disagree between each other on the ordering of two further alternatives. As before, this profile is not single-crossing, as the pair $\{a, b\}$ forces us to place $v_1$ and $v_2$ next to each other, and to put $v_3$ and $v_4$ next to each other; the pair $\{c, d\}$ forces us to place $v_1$ and $v_3$ next to each other, and to put $v_2$ and $v_4$ next to each other. This means that no voter can be put into the first position.

A profile is single-crossing if and only if it contains neither $\gamma$-configurations nor $\delta$-configurations [Bredereck *et al.*, 2012].

**Two central problems.** As already discussed before, two natural ways of measuring the distance of profiles to some restricted domains is by deleting voters and by deleting alternatives. Hence, for $\Pi \in$ {worst-restricted, medium-restricted, best-restricted, value-restricted, single-peaked, single-caved, single-crossing, group-separable}, we study the following two types of modification problems: $\Pi$ MAVERICK DELETION and $\Pi$ ALTERNATIVE DELETION.

$\Pi$ MAVERICK DELETION
**Input:** A profile with $n$ voters and an integer $k \leq n$.
**Question:** Can we delete at most $k$ voters such that the resulting profile has the $\Pi$-property?

$\Pi$ ALTERNATIVE DELETION
**Input:** A profile with $m$ alternatives and an integer $k \leq m$.
**Question:** Can we delete at most $k$ alternatives such that the resulting profile has the $\Pi$-property?

## 3 Results

It is easy to see that both $\Pi$ MAVERICK DELETION and $\Pi$ ALTERNATIVE DELETION are in NP with $\Pi$ being one of the eight properties we consider: Given a preference profile, one can check in polynomial time whether it is $\Pi$, since the $\Pi$-property is characterized by a fixed number of forbidden substructures. Thus, in order to show the NP-completeness of $\Pi$ MAVERICK DELETION and $\Pi$ ALTERNATIVE DELETION, we only have to show their NP-hardness.

We will use the NP-complete VERTEX COVER (VC) problem [Garey and Johnson, 1979] to show many of our NP-hardness results: Given an undirected graph $G = (U, E)$ and a non-negative integer $k$, VC asks whether there is a *vertex cover* $U' \subseteq U$ of at most $k$ vertices, that is, each edge is incident to at least one vertex in $U'$.

**On deleting maverick voters.** We start our findings on intractability results with the four domain-restrictions which are characterized by configurations with three alternatives.

**Theorem 8.** $\Pi$ MAVERICK DELETION *is NP-complete for every* $\Pi \in \{$*best-restricted, medium-restricted, worst-restricted, value-restricted*$\}$.

*Proof.* We reduce from VC to show the NP-hardness result. Let $(G, k)$ denote a VC-instance with vertex set $U = \{u_1, \ldots, u_r\}$ and edge set $E = \{e_1, \ldots, e_s\}$; without loss of generality $r \geq 4$. The set of alternatives consists of three *edge alternatives* $a_j, b_j$, and $c_j$ for each edge $e_j \in E$. The voter set one-to-one corresponds to vertex set $U$. In total, the number $m$ of alternatives is $3s$ and the number $n$ of voters is $r$. All voters prefer $\{a_j, b_j, c_j\}$ to $\{a_{j'}, b_{j'}, c_{j'}\}$ whenever $j < j'$. Moreover, voter $v_i$ has $a_j \succ b_j \succ c_j$ if $v_i \notin e_j$. Otherwise, let edge $e_j = \{v_i, v_{i'}\}$. Voter $v_i$ ranks $c_j \succ a_j \succ b_j$ if $i < i'$, and ranks $b_j \succ c_j \succ a_j$, otherwise. In this way, the two vertex voters in $e_j$ and any other voter $v_z$ not in $e_j$ form a worst-diverse configuration, a medium-diverse configuration as well as a best-diverse configuration regarding the three edge alternatives $a_i$,$b_i$, and $c_i$. The parameter $k$ is the same. The whole construction runs in polynomial time. Due to lack of space, its correctness proof is deferred to a full version of the paper. □

The profile constructed in the proof of Thm. 8 does not contain $\alpha$-configurations, $\bar{\alpha}$-configurations, or $\beta$-configurations. Hence, NP-hardness transfers to single-caved, group-separable, and single-peaked cases, respectively. Note that NP-hardness for SINGLE-PEAKED MAVERICK DELETION is already known by a different proof of Erdélyi *et al.* [2013]. However, their proof does not work for $\Pi$ MAVERICK DELETION with $\Pi \in \{$best-restricted, medium-restricted, worst-restricted, group-separable$\}$.

**Corollary 9.** $\Pi$ MAVERICK DELETION *is NP-complete for every* $\Pi \in \{$*single-caved, group-separable, single-peaked*$\}$.

In contrast to all NP-complete $\Pi$ MAVERICK DELETION problems above, SINGLE-CROSSING MAVERICK DELETION is tractable. The algorithm, which is similar to the single-crossing detection algorithm of Elkind *et al.* [2012], does not only solve the decision problem but the optimization problem asking for the smallest number $k$ of voters to delete in order to make the profile single-crossing.

**Theorem 10.** SINGLE-CROSSING MAVERICK DELETION *is solvable in* $O(n^3 \cdot m^2)$ *time, where* $n$ *denotes the number of voters and* $m$ *denotes the number of alternatives.*

*Proof.* In the following, we assume that the voters have pairwise distinct preference orderings. By using arc weights in the graphs to be constructed, the algorithm can be extended to also work for general preference profiles. Let $v_i, v_{i'}$, and $v_{i''}$ be three distinct voters. We say that voter $v_{i'}$ is $\succ_i$*-swap-transferable* to voter $v_{i''}$ if one can transform the preference ordering of $v_{i'}$ to the one of $v_{i''}$ by repetitive swapping of two alternatives $a_j$ and $a_{j'}$ with $a_j \succ_i a_{j'}$ and $a_j \succ_{i'} a_{j'}$.

Let $L = \langle v_1, \ldots, v_{i'}, \ldots, v_{i''}, \ldots \rangle$ be a linear ordering of voters with $v_{i'}$ and $v_{i''}$ being two distinct voters. Then, by definition of single-crossing profiles, $v_{i'}$ is $\succ_1$-swap-transferable to $v_{i''}$, but $v_{i''}$ is not $\succ_1$-swap-transferable to $v_{i'}$.

The idea is to guess (by testing all) the first voter in a single-crossing ordering and to compute a maximum set of possible successive voters. This idea is realized as follows.

For each voter $v_i$, build a directed graph $D_i$ with one vertex for each voter. Add an arc from vertex $x$ to vertex $y$ if $x$ is $\succ_i$-swap-transferable to $y$. By the definition of the swap-operation, the graph becomes acyclic. Now, a vertex ordering in a longest directed path among these graphs represents a single-crossing ordering of a subset of voters of maximum size. Thus, the minimum number $k$ of voters to delete to make the profile single-crossing is $n - \ell$ with $n$ being the number of voters and $\ell$ the length of a longest path.

Constructing $D_i$ takes $O(n^2 \cdot m^2)$ time: Check for each ordered pair $(x, y)$ of vertices whether the voter corresponding to $x$ is $\succ_i$-swap-transferable to $y$ by checking any two alternatives. Computing the longest path in each $D_i$ takes $O(n^2)$ time. Thus, checking all $n$ graphs takes $O(n^3 \cdot m^2)$ time. □

**On deleting alternatives.** Another way of obtaining nicely structured preference profiles is to delete alternatives. The corresponding problems are studied in the remainder of this section.

**Theorem 11.** $\Pi$ ALTERNATIVE DELETION *is NP-complete for every* $\Pi \in \{$*best-restricted, medium-restricted, worst-restricted, value-restricted*$\}$.

*Proof.* We show a polynomial-time many-one reduction from VC to $\Pi$ ALTERNATIVE DELETION with $\Pi \in \{$medium-restricted, value-restricted, worst-restricted$\}$. For BEST-RESTRICTED ALTERNATIVE DELETION one has to reverse all preference orderings in the forthcoming construction.

Let $(G, k)$ denote a VC-instance with vertex set $U = \{u_1, \ldots, u_r\}$ and edge set $E = \{e_1, \ldots, e_s\}$. The set of alternatives consists of all vertices in $U$ and of $k + 1$ new dummy alternatives. Let $D$ denote the set of these new dummy alternatives. We arbitrarily fix a canonical ordering of $D$ and set $\langle U \rangle = \langle u_1, \ldots, u_r \rangle$. The number $m$ of constructed alternatives is $r + k + 1$. We introduce a voter $v_0$ with the special preference ordering $\langle D \rangle \succ \langle U \rangle$. Furthermore, for each edge $e_i = \{u_j, u_{j'}\}$ with $j < j'$, we introduce two edge voters $v_{2i-1}$ and $v_{2i}$ with preference orderings $u_j \succ u_{j'} \succ \langle D \rangle \succ \langle U \setminus e_i \rangle$ and $u_{j'} \succ \langle D \rangle \succ \langle U \setminus \{u_{j'}\} \rangle$, respectively. Together with voter $v_0$, these two voters $v_{2i-1}$ and $v_{2i}$ form a worst-diverse configuration and a medium-diverse configuration with respect to the two vertex alternatives $u_j$, $u_{j'}$ and any dummy alternative. In total, the number $n$ of constructed voters is $2s+1$. The parameter $k$ remains the same. This completes the construction.

Our reduction runs in polynomial time. It remains to show its correctness. In particular, we show that $(G, k)$ has a vertex cover of size at most $k$ if and only if the constructed profile can be made worst-restricted (resp. medium-restricted), and hence, value-restricted by deleting at most $k$ alternatives.

For the "only if" part, suppose that $U' \subseteq U$ with $|U'| \leq k$ is a vertex cover. First, we show that after deleting the vertex alternatives corresponding to $U'$, the resulting profile is worst-restricted and, hence, value-restricted. Suppose for the sake of contradiction that the resulting profile still contains a worst-diverse configuration $\sigma$. Since all voters have the same ranking over $D$, $\sigma$ contains at most one dummy alternative. But if $\sigma$ contains one dummy alternative $d \in D$, then there is a voter with $u \succ u' \succ d$, $u, u' \in U$ which means that edge $\{u, u'\}$ is not covered by $U'$. Hence, $\sigma$ contains no dummy alternative. This means that $\sigma$ contains three vertex alternatives $u_j$, $u_{j'}$, and $u_{j''}$ with $j < j' < j''$ and by the definition of the worst-diverse configuration, $\sigma$ concerns three voters with preferences $\{u_j, u_{j'}\} \succ u_{j''}$, $\{u_j, u_{j''}\} \succ u_{j'}$, and $\{u_{j'}, u_{j''}\} \succ u_j$, respectively. However, the last preference implies that $\{u_{j'}, u_{j''}\}$ is an edge which is not covered by $U'$—a contradiction.

Second, we show that after deleting the vertex alternatives corresponding to $U'$ the resulting profile is medium-restricted. Suppose for the sake of contradiction that the resulting profile still contains a medium-diverse configuration $\sigma'$. Since all voters have the same ranking over $D$ and no voter ranks $d \succ u \succ d'$ with $d, d' \in D$ and $u \in U$, $\sigma'$ can contain at most one dummy alternative. Now, if $\sigma'$ involves one dummy alternative $d \in D$ and two vertex alternatives $u_j, u_{j'} \in U$ with $j < j'$, then the voter ranking $u_{j'}$ between $u_j$ and $d$ must have $u_j \succ u_{j'} \succ d$. But this means that $\{u_j, u_{j'}\}$ is an uncovered edge—a contradiction. Hence, assume that $\sigma'$ contains no dummy alternative. This means that $\sigma'$ must involve three vertex alternatives $u_j, u_{j'}, u_{j''}$ with $j < j' < j''$. However, there is no voter with $u_j \succ u_{j''} \succ u_{j'}$ or $u_{j'} \succ u_{j''} \succ u_j$ in the resulting profile—a contradiction.

For the "if" part, suppose that the constructed profile is a yes-instance for the worst-restricted or the medium-restricted case. Let $U' \subseteq U$ be the set of deleted vertex alternatives with $|U'| \le k$. Then $U'$ is also a vertex cover of $G$. Assume towards a contradiction that $e_i = \{u_j, u_{j'}\}$ $(j < j')$ is an uncovered edge. Since $|D| > k$, at least one dummy alternative $d$ is not deleted. Then, $v_0$ and $v_{2i}, v_{2i-1}$ form a worst-diverse configuration as well as a medium-diverse configuration regarding $u_j, u_{j'}, d$—a contradiction. □

The profile which results from deleting alternatives from the profile constructed in the proof of Thm. 11 is not only medium-restricted, but it even contains no $\beta$-configurations. By the definition of group-separability, this means that the NP-completeness result of MEDIUM-RESTRICTED ALTERNATIVE DELETION also holds for the group-separable case.

**Corollary 12.** GROUP-SEPARABLE ALTERNATIVE DELETION *is NP-complete.*

While making a profile single-crossing by deleting as few maverick voters as possible is in P, the decision variant of this problem becomes NP-hard if one instead deletes alternatives.

**Theorem 13.** SINGLE-CROSSING ALTERNATIVE DELETION *is NP-complete.*

*Proof.* For the NP-hardness result we reduce from the NP-complete satisfiability problem MAXIMUM 2-SATISFIABILITY (MAX2SAT) [Garey and Johnson, 1979]. Given a set $U$ of Boolean variables, a collection $C$ of size-two clauses over $U$ and a positive integer $k'$, MAX2SAT asks whether there is a truth assignment for $U$ such that at least $k'$ clauses in $C$ are satisfied.

Let $(U, C, k')$ be a MAX2SAT-instance with variable set $U = \{x_1, \ldots, x_r\}$ and clause set $C = \{c_1, \ldots, c_s\}$. There are two sets $O$ and $\overline{O}$ of $2(rs + r + s) + 1$ dummy alternatives each. For each variable $x_i \in U$, there are two sets $X_i$ and $\overline{X_i}$ of $s + 1$ *variable alternatives* each. We say that $X_i$ corresponds to $x_i$ and that $\overline{X_i}$ corresponds to $\overline{x_i}$. The canonical orderings $\langle O \rangle$, $\langle \overline{O} \rangle$, $\langle X_i \rangle$ and $\langle \overline{X_i} \rangle$, $i \in \{1, \ldots, r\}$, are arbitrary but fixed. Let $X$ be the union $\bigcup_{i=1}^{r} X_i \cup \overline{X_i}$ of all variable alternatives. The canonical ordering $\langle X \rangle$ is $\langle X_1 \rangle \succ \langle \overline{X_1} \rangle \succ \ldots \succ \langle X_r \rangle \succ \langle \overline{X_r} \rangle$. For each clause $c_j \in C$, there are two *clause alternatives* $a_j$ and $b_j$. Let $A$ denote the set of all clause alternatives. The canonical ordering $\langle A \rangle$ is $a_1 \succ b_1 \succ \ldots \succ a_s \succ b_s$. The total number $m$ of alternatives is $6(rs + r + s) + 2$.

The rough idea is that deleting all alternatives in $X_i$ corresponds to setting $x_i$ to true, and deleting all alternatives in $\overline{X_i}$ corresponds to setting $x_i$ to false. Furthermore, deleting $b_j$ or $a_j$ corresponds to not-satisfied clause $c_j$.

To this end, let the parameter $k$ be $r(s + 1) + (s - k')$. There are two sets $V$ and $W$ of voters with $|V| = 2r$ and $|W| = 4s$. Voter set $V$ consists of two voters $v_{2i-1}$ and $v_{2i}$ for each variable $x_i$. Their preference orderings are

$$\langle \boldsymbol{O} \rangle \succ \langle \overline{\boldsymbol{O}} \rangle \succ \langle X_1 \rangle \succ \langle \overline{X_1} \rangle \ldots \langle \overline{\boldsymbol{X_i}} \rangle \succ \langle \boldsymbol{X_i} \rangle \ldots \langle X_r \rangle \succ \langle \overline{X_r} \rangle \succ \langle A \rangle \text{ and}$$
$$\langle \overline{\boldsymbol{O}} \rangle \succ \langle \boldsymbol{O} \rangle \succ \langle X_1 \rangle \succ \langle \overline{X_1} \rangle \ldots \langle \overline{\boldsymbol{X_i}} \rangle \succ \langle \boldsymbol{X_i} \rangle \ldots \langle X_r \rangle \succ \langle \overline{X_r} \rangle \succ \langle A \rangle.$$

These two voters together with any other two voters $v_l$ and $v_{l'} \in V \setminus \{v_{2i-1}, v_{2i}\}$ with odd $l$ and even $l'$ form a $\delta$-configuration regarding $o \in O, \overline{o} \in \overline{O}, x \in X_i, \overline{x} \in \overline{X_i}$:

$$v_{2i-1}\colon o \succ \overline{o} \text{ and } \overline{x} \succ x; \quad v_{2i}: \overline{o} \succ o \text{ and } \overline{x} \succ x;$$
$$v_l \quad : o \succ \overline{o} \text{ and } x \succ \overline{x}; \quad v_{l'}: \overline{o} \succ o \text{ and } x \succ \overline{x}.$$

Voter set $W$ consists of four voters $w_{4j-3}, w_{4j-2}, w_{4j-1}$, and $w_{4j}$ for each clause $c_j$. These four voters have the same preference ordering $\langle \overline{O} \rangle \succ \langle O \rangle \succ \langle A_1 \rangle \succ \langle X \rangle \succ \langle A_2 \rangle$ over set $O \cup \overline{O} \cup A_1 \cup A_2 \cup X$, where $A_1 = \{a_{j'}, b_{j'} \mid j' < j\}$ and $A_2 = \{a_{j'}, b_{j'} \mid j' > j\}$. The positions of $a_j$ and $b_j$ are placed as follows: Let $\widehat{X}_1^j$ denote the set of variable alternatives corresponding to the literal in $c_j$ with lower index and $\widehat{X}_2^j$ denote the set of variable alternatives corresponding to the literal in $c_j$ with higher index. Voters $w_{4j-3}$ and $w_{4j-2}$ rank clause alternatives $a_j$ right below the last alternative in $\langle \widehat{X}_1^j \rangle$ while voters $w_{4j-1}$ and $w_{4j}$ rank it right above the first alternative in $\langle \widehat{X}_1^j \rangle$. As for alternative $b_j$, voters $w_{4j-3}$ and $w_{4j-1}$ rank $b_j$ right above the first variable alternative in $\langle \widehat{X}_2^j \rangle$ while voters $w_{4j-2}$ and $w_{4j}$ rank it right below the last variable alternative in $\langle \widehat{X}_2^j \rangle$. Thus, these four voters form a $\delta$-configuration regarding $a_j, b_j, x \in \widehat{X}_1^j$, and $y \in \widehat{X}_2^j$:

$$w_{4j-3}\colon x \succ a_j \text{ and } b_j \succ y; \quad w_{4j-2}: x \succ a_j \text{ and } y \succ b_j;$$
$$w_{4j-1}\colon a_j \succ x \text{ and } b_j \succ y; \quad w_{4j}: a_j \succ x \text{ and } y \succ b_j.$$

The reduction clearly runs in polynomial time. It remains to show that $(U, C, k')$ is a yes-instance for MAX2SAT if and only if the constructed profile together with $k$ is a yes-instance for SINGLE-CROSSING ALTERNATIVE DELETION.

For the "only if" part, suppose that there is a truth assignment $U \to \{\text{true}, \text{false}\}^r$ of the variables such that at least $k'$

clauses are satisfied. We delete all variable alternatives in $X_i$ if $x_i$ is assigned to true, and delete all variable alternatives in $\overline{X_i}$, otherwise. Furthermore, we delete the clause alternative $b_j$ if $c_j$ is not satisfied by the assignment. Let $X'$ be the set of remaining variable alternatives, and $A'$ the set of all remaining clause alternatives. Then the number of deleted alternatives is $|X|+|A|-(|X'|+|A'|) \leq r(s+1)+(s-k') = k$.

For each $j \in \{1,\ldots,s\}$, we define $z_j$ by $\langle z_j \rangle = \langle w_{4j-2}, w_{4j}, w_{4j-3}, w_{4j-1} \rangle$ if the literal in clause $c_j$ with lower index is satisfied; otherwise, $\langle z_j \rangle = \langle w_{4j-3}, w_{4j-2}, w_{4j-1}, w_{4j} \rangle$. The resulting profile is single-crossing with respect to the voter ordering $L := \langle v_1, v_3, \ldots, v_{2r-1}, v_2, v_4, \ldots, v_{2r}, z_1, z_2, \ldots, z_s \rangle$. To show this, we define the concept of "separation": If $\langle \mathcal{L} \rangle$ is a linear ordering of voters, then we say that a pair $\{a, b\}$ of distinct alternatives *separates ordering $\langle \mathcal{L} \rangle$ (into two orderings $\langle \mathcal{L}_1 \rangle$ and $\langle \mathcal{L}_2 \rangle$)* if $\langle \mathcal{L} \rangle = \langle \mathcal{L}_1, \mathcal{L}_2 \rangle$ and no voter in $\langle \mathcal{L}_1 \rangle$ agrees with any voter in $\langle \mathcal{L}_2 \rangle$ on the ordering of $a$ and $b$. Obviously, $\mathcal{L}$ is single-crossing if it can be separated by every possible pair of alternatives.

Suppose for the sake of contradiction that $L$ is not a single-crossing ordering which means that $L$ cannot be separated by a pair $\{a, a'\} \subset O \cup \overline{O} \cup X' \cup A'$ of alternatives. Note that all voters along $L$ up to and including voter $v_{2r-1}$ rank $\langle O \rangle \succ \langle \overline{O} \rangle \succ \langle X \rangle$ while all voters from $v_2$ onwards rank $\langle \overline{O} \rangle \succ \langle O \rangle \succ \langle X \rangle$. Hence, $a$ and $a'$ can neither both be in $O \cup \overline{O}$, nor both be in $X'$. Furthermore, $a$ and $a'$ cannot both be in $A'$, as all voters have the same ranking $\langle A \rangle$. Since all voters rank $(O \cup \overline{O}) \succ (X \cup A)$, $a$ and $a'$ are not in $O \cup \overline{O}$. This means, without loss of generality, $a \in X'$ and $a' \in A'$.

Assume that alternative $a' \in \{a_j, b_j\}$. Then, alternative $a$ cannot be in $X' \setminus (\widehat{X}_1^j \cup \widehat{X}_2^j)$ since pair $\{a', a''\}$ with $a'' \in X' \setminus (\widehat{X}_1^j \cup \widehat{X}_2^j)$ separates ordering $L$ into two orderings $L_1$ and $L_2$ where either $L_1 = \langle v_1, v_3, \ldots, v_{2r-1}, v_2, v_4, \ldots, v_{2r}, z_1, z_2, \ldots, z_j \rangle$ and $L_2 = \langle z_{j+1}, z_{j+2}, \ldots, z_s \rangle$, or $L_1 = \langle v_1, v_3, \ldots, v_{2r-1}, v_2, v_4, \ldots, v_{2r}, z_1, z_2, \ldots, z_{j-1} \rangle$ and $L_2 = \langle z_j, z_{j+1}, \ldots, z_s \rangle$. Thus, $a \in \widehat{X}_1^j \cup \widehat{X}_2^j$. If the literal in $c_j$ with lower index is satisfied, then all variable alternatives in $\widehat{X}_1^j$ are deleted and $z_j = \langle w_{4j-2}, w_{4j}, w_{4j-3}, w_{4j-1} \rangle$. Hence, $a \in \widehat{X}_2^j$. All voters along $L$ up to and including $w_{4j}$ prefer $a$ to $a'$, and all voters onwards $w_{4j-3}$ prefer $a'$ to $a$. Hence, $L$ is separated by $\{a, a'\}$. Otherwise, either $b_j$ or all alternatives in $\widehat{X}_2^j$ are deleted, thus, $z_j = \langle w_{4j-3}, w_{4j-2}, w_{4j-1}, w_{4j} \rangle$. If $b_j$ is deleted, then all voters along $L$ up to and including $w_{4j-2}$ prefer each $a \in \widehat{X}_1^j \cup \widehat{X}_2^j$ to $a' = a_j$, and all voters onwards $w_{4j-3}$ prefer $a' = a_j$ to each $a \in \widehat{X}_1^j \cup \widehat{X}_2^j$. Otherwise, all alternatives in $\widehat{X}_2^j$ are deleted. So, all voters along $L$ up to and including $w_{4j-2}$ prefer each $a \in \widehat{X}_1^j$ to each $a' \in \{a_j, b_j\}$, and all voters onwards $w_{4j-3}$ prefer each $a' \in \{a_j, b_j\}$ to each $a \in \widehat{X}_1^j \cup \widehat{X}_2^j$. Hence, $L$ is separated by $\{a, a'\}$. In summary, $L$ can always be separated by $\{a, a'\}$—a contradiction to the assumption that $L$ is not a single-crossing ordering.

For the "if" part, suppose that deleting a set $K$ of at most $k$ alternatives makes the remaining profile single-crossing. This means that by deleting $K$ one eliminates all $\delta$-configurations. Note that deleting $K \setminus (O \cup \overline{O})$ also results in a single-crossing profile: Assume towards a contradiction that there is a $\delta$-configuration involving a set $D$ of alternatives with $D \cap (K \setminus (O \cup \overline{O})) = \emptyset$. Clearly, $D$ contains exactly two alternatives in $O$ and in $\overline{O}$ each since otherwise $\sigma$ does not form a $\delta$-configuration or $K$ is not a solution. Since $|O| = |\overline{O}| > k$, there are two alternatives $o^* \in O$ and $\overline{o}^* \in \overline{O}$ which are not in $K$. If we replace the alternatives in $D$ that are from $O \cup \overline{O}$ with $o^*$ and $\overline{o}^*$, then we get a $\delta$-configuration for the profile which remains after deleting $K$—a contradiction. Hence, in the following we assume without loss of generality that none of the dummy alternatives is deleted.

For each $x_i \in U$, all variable alternatives in either $X_i$ or $\overline{X_i}$ must be deleted to destroy all $\delta$-configurations involving alternatives in $O \cup \overline{O} \cup X_i \cup \overline{X_i}$. Let $X'$ be the set of all deleted variable alternatives and let $A'$ be the set of all deleted clause alternatives. Then, $|X'| \geq r(s+1)$ and $|A'| \leq k - r(s+1) = s - k'$. We show that by setting variable $x_i \in U$ to true if $X_i \subseteq X'$, and false otherwise, all clauses $c_j$ with $\{a_j, b_j\} \cap A' = \emptyset$ are satisfied. Suppose for the sake of contradiction that clause $c_j$ with $\{a_j, b_j\} \cap A' = \emptyset$ is not satisfied. This means both $\widehat{X}_1^j$ as well as $\widehat{X}_2^j$ are not completely in $X'$. But then, voters $w_{4i-3}, w_{4i-2}, w_{4i-1}$, and $w_{4i}$ form a $\delta$-configuration regarding $a_j, b_j, x, x'$ with $x \in \widehat{X}_1^j \setminus X'$ and $x' \in \widehat{X}_2^j \setminus X'$ —a contradiction. □

# 4 Conclusion

In terms of computational complexity theory, little is known about preference profiles which are "close" to being nicely structured. We showed that making a profile single-crossing by deleting as few voters as possible can be solved in polynomial time. In contrast, making a profile nicely structured by deleting at most $k$ voters or at most $k$ alternatives is NP-hard for all other considered cases. However, we mention in passing that all these problems become tractable when $k$ is small: All considered properties are characterized by a fixed number of forbidden substructures. Thus, by branching over all possible voters (resp. alternatives) of each forbidden substructure in the profile one obtains a fixed-parameter algorithm [Downey and Fellows, 1999; Flum and Grohe, 2006; Niedermeier, 2006] that is efficient for small distances. One line of future research is to investigate more sophisticated and more efficient (fixed-parameter) algorithms to compute the distance of a profile to a nicely-structured one.

A second line of research which has already been started by Erdélyi *et al.* [2013] for single-peaked profiles is to study further distance measures.

A third line of research is to investigate whether and in which way properties of nicely structured preference profiles transfer to profiles that are only close to being nicely structured. This has been started by Faliszewski *et al.* [2011a] for some notion of nearly single-peakedness which is different from, but related to ours. They investigate cases where the computational tractability of attacks on single-peaked profiles transfers to nearly single-peaked preferences, and cases where the vulnerability disappears even if the preference profile is extremely close to being single-peaked.

## References


[Arrow, 1950] Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58:328–346, 1950.

[Ballester and Haeringer, 2011] Miguel Ángel Ballester and Guillaume Haeringer. A characterization of the single-peaked domain. *Social Choice and Welfare*, 36(2):305–322, 2011.

[Bartholdi *et al.*, 1989] John J. Bartholdi, Craig A. Tovey, and Michael A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6:157–165, 1989.

[Black, 1948] Duncan Black. On the rationale of group decision making. *Journal of Political Economy*, 56:23–34, 1948.

[Brandt *et al.*, 2010] Felix Brandt, Markus Brill, Edith Hemaspaandra, and Lane A. Hemaspaandra. Bypassing combinatorial protections: Polynomial-time algorithms for single-peaked electorates. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence*, pages 715–722, 2010.

[Bredereck *et al.*, 2012] Robert Bredereck, Jiehua Chen, and Gerhard J. Woeginger. A characterization of the single-crossing domain. *Social Choice and Welfare*, 2012. Online available.

[Downey and Fellows, 1999] Rodney G. Downey and Michael R. Fellows. *Parameterized Complexity*. Springer, 1999.

[Elkind *et al.*, 2012] Edith Elkind, Piotr Faliszewski, and Arkadii M. Slinko. Clone structures in voters' preferences. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 496–513. ACM, 2012.

[Erdélyi *et al.*, 2013] Gábor Erdélyi, Martin Lackner, and Andreas Pfandler. Computational aspects of nearly single-peaked electorates. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, 2013. Accepted for publication.

[Faliszewski *et al.*, 2011a] P. Faliszewski, E. Hemaspaandra, and L.A. Hemaspaandra. The complexity of manipulative attacks in nearly single-peaked electorates. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 228–237, 2011.

[Faliszewski *et al.*, 2011b] Piotr Faliszewski, Edith Hemaspaandra, Lane A. Hemaspaandra, and Jörg Rothe. The shield that never was: Societies with single-peaked preferences are more open to manipulation and control. *Information and Computation*, 209(2):89–107, 2011.

[Flum and Grohe, 2006] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Springer, 2006.

[Garey and Johnson, 1979] Michael R. Garey and David S. Johnson. *Computers and Intractability—A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.

[Inada, 1964] Ken-ichi Inada. A note on the simple majority decision rule. *Econometrica*, 32:525–531, 1964.

[Inada, 1969] Ken-ichi Inada. The simple majority rule. *Econometrica*, 37:490–506, 1969.

[Klaus *et al.*, 1997] Bettina Klaus, Hans Peters, and Ton Storcken. Strategy-proof division of a private good when preferences are single-dipped. *Economics Letters*, 55:339–346, 1997.

[Moulin, 1980] Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35:437–455, 1980.

[Niedermeier, 2006] Rolf Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford University Press, February 2006.

[Roberts, 1977] Kevin W.S. Roberts. Voting over income tax schedules. *Journal of Public Economics*, 8:329–340, 1977.

[Sen and Pattanaik, 1970] Amartya K. Sen and Prasanta K. Pattanaik. Necessary and sufficient conditions for rational choice under majority decision. *Journal of Economic Theory*, 1:178–202, 1970.

[Sen, 1966] Amartya K. Sen. A possibility theorem on majority decisions. *Econometrica*, 34:491–499, 1966.

[Walsh, 2007] Toby Walsh. Uncertainty in preference elicitation and aggregation. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence*, pages 3–8, 2007.