HAL open science

Measuring Nearly Single-Peakedness of an Electorate: Some New Insights

Bruno Escoffier, Olivier Spanjaard, Magdaléna Tydrichová

To cite this version:

Bruno Escoffier, Olivier Spanjaard, Magdaléna Tydrichová. Measuring Nearly Single-Peakedness of an Electorate: Some New Insights. Algorithmic Decision Theory - 7th International Conference, ADT, Nov 2021, Toulouse, France. pp.19-34, 10.1007/978-3-030-87756-9_2 . hal-03477649

HAL Id: hal-03477649

https://hal.science/hal-03477649v1

Submitted on 9 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Measuring Nearly Single-peakedness of an Electorate: Some New Insights

Bruno Escoffier[1,2][0000−0002−6477−8706], Olivier Spanjaard[✉1][0000−0002−9948−090X], Magdaléna Tydrichová[1][0000−0002−0329−0264]

1. Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
2. Institut Universitaire de France, Paris, France
{bruno.escoffier,olivier.spanjaard,magdalena.tydrichova}@lip6.fr

Abstract. After introducing a new distance measure of a preference profile to single-peakedness, directly derived from the very definition of single-peaked preferences by Black [4], we undertake a brief comparison with other popular distance measures to single-peakedness. We then tackle the computational aspects of the optimization problem raised by the proposed measure, namely we show that the problem is NP-hard and we propose an integer programming formulation. Finally, we carry out numerical tests on real and synthetic voting data. The obtained results show the interest of the proposed measure, but also shed new light on the advantages and drawbacks of some popular distance measures.

1 Introduction

The study of *structured preferences* in social choice [7, 21] starts from the observation that, although the opinions of individuals on candidates in an election are heterogeneous, the voters often agree on the way the candidates are related to each other, more precisely on the *ideological proximities* between them. Various preference structures can be considered to model these proximities, among which are single-peaked preferences [4] and its extensions (see e.g. [18]), as well as Euclidean models where the ideological positions of voters and candidates are viewed as points in an Euclidean space [8].

A preference structure is also called a *domain restriction* in social choice theory, because it restricts the domain of possible preferences for the voters by assuming a consistency of the preferences with the proximities between candidates. Domain restrictions often make it possible to overcome social choice paradoxes (such as the famous Arrow's impossibility theorem [1]), and impact the computational complexity of determining the winner of an election, as well as the complexity of manipulating an election. For more details regarding these computational aspects, we refer the reader to the survey by Elkind et al. [12].

Another issue addressed in the survey is that of recognizing structures in preferences, i.e., determining whether a set of preferences has a given structural property and, if yes, returning the corresponding structure (left-right axis,

graph or positions in the Euclidean space). The work we present here deals with this recognition problem. We focus on the most well-known domain restriction, that of single-peaked preferences. More precisely, we introduce a new distance measure to single-peakedness for an election.

As emphasized by Feld and Grofman [15], the assumption that preferences are perfectly single-peaked is indeed very strong if the alternatives are candidates in an election (the case of numerical alternatives, such as tax levels, is obviously different). We recall that preferences are said to be single-peaked if 1) all voters agree on a left-right axis on the alternatives, and 2) the preferences of all voters decrease along the axis when moving away from their most preferred alternative to the right or left. Single-peakedness in the strictest sense thus requires that no individual preference deviates (even slightly) from the single-peakedness condition. Given an axis A, the number of rankings consistent with A (i.e., such that condition 2 holds) is 2^{m-1}, over $m!$ possible rankings in total, where m is the number of alternatives. The proportion of consistent rankings within all possible rankings thus quickly becomes tiny when m increases ($2^{m-1}/m! \approx 0.01$ for $m = 7$), as well as the likelihood that no voter deviates from this subset of preferences. This observation is corroborated by the numerical tests carried out by Sui et al. [20] on 2002 Irish General Election data in Dublin West and Dublin North, where the best axes explain only 2.9% and 0.4% of voters' preferences.

Conitzer [7] distinguishes between two interpretations of nearly single-peakedness (see e.g. [13] for a systematic study of nearly single-peaked electorates): an interpretation where preferences are said nearly single-peaked if only a few voters' preferences deviates from a given axis A and the other voters' preferences are perfectly single-peaked w.r.t. A (the numerical tests reported above corresponds to this interpretation); another interpretation where one allows all voters' preferences to deviate to some extent from a given axis A. The distance measure we propose in this paper falls under the second interpretation, which has been less studied and tested than the first one.

Given an axis A on the candidates and a set $\mathcal{P}$ of preferences, the idea is to measure how far from single-peakedness w.r.t. A each individual preference is. Put another way, each preference in the electorate partially fits with the axis (according to a non-binary measure), and one sums up the degrees of fitness of preferences in $\mathcal{P}$ to obtain the "degree of single-peakedness" of $\mathcal{P}$ w.r.t. A. More precisely, one defines a *distance* to single-peakedness, i.e., the degree is 0 if $\mathcal{P}$ is single-peaked w.r.t. A. We are thus seeking a procedure that returns both a degree of single-peakedness of a profile and an axis that witnesses the obtained value. These outputs allow the analysis of a political landscape, by answering the questions: How close to single-peakedness is an electorate? How the voters perceive the ideological proximities between candidates?

Related Work While recognizing perfectly single-peaked preferences is a polynomial time problem [3, 10], determining the distance to single-peakedness (according to various measures) is often NP-hard. Various notions of nearly single-peakedness are present in the literature. We briefly review here notions that do not relax the assumption of a one-dimensional axis on all the candidates,

which excludes other approaches that return, e.g., an axis on a subset of candidates [13], an axis on clone sets [9], or multiple axes [2, 20]. Most of them have been introduced and/or studied by Faliszewski et al. [14], Erdélyi et al. [13] and Elkind and Lackner [11]. Faliszewski et al. [14] studied *k-voter deletion single-peakedness*, also known as *partial single-peakedness* in economics [19]. One says that an electorate is k-voter deletion single-peaked consistent if all but k of the voters preferences ("maverick" voters) are consistent with a common axis on the candidates. The smallest number k such that there exists an axis w.r.t. which the electorate is k-voter deletion single-peaked can be viewed as a distance to single-peakedness. Erdélyi et al. [13] as well as Bredereck et al. [6] have proved that determining this distance is NP-hard. Elkind and Lackner [11] have proposed a polynomial time 2-approximation algorithm for this distance, and have established fixed-parameter tractability results (complexity $\mathcal{O}^*(1.28^k)$ if $k < n/2$, and $\mathcal{O}^*(2.08^k)$ if $k \geq n/2$, where n is the number of voters).

Erdélyi et al. [13] introduced *k-local candidate deletion single-peakedness*. They first defined single-peaked consistency of a partial preference (linear order on a subset of candidates) w.r.t. an axis A on all candidates: a partial preference is single-peaked w.r.t. A if it is single-peaked w.r.t. the axis obtained from A by removing the missing candidates. Then they say that an electorate is k-local candidate deletion single-peaked consistent if, by removing at most k candidates from each preference, one obtains a set of partial preferences that are single-peaked with respect to a common axis. As above, the smallest k for which the property holds can be viewed as a distance. Here again, the authors have proved that determining this distance is an NP-hard problem.

The class of distance measures that is the closest to our work is that of swap distances. Erdélyi et al. [13] introduced *k-global swaps single-peakedness*, where k is the number of swaps of consecutive candidates that need to be performed in the preferences to make the election single-peaked. Following Faliszewski et al. [14], they also considered a "local budget" for swaps, i.e., they allow up to k swaps *per vote*. They call *k-local swaps* this notion of nearly single-peakedness. For both notions, Erdélyi et al. [13] have proved that computing the smallest k enabling to make the election single-peaked is NP-hard. Finally, let us mention the notion of *PerceptionFlip$_k$ single-peakedness* [14]. An electorate is PerceptionFlip$_k$ single-peaked if there exists an axis A such that, for each voter, the axis A can be transformed into an axis A' by at most k swaps of consecutive candidates in A so that the voter's preference is single-peaked with respect to A'. Erdélyi et al. [13] have proved that k-local swaps single-peakedness and PerceptionFlip$_k$ single-peakedness are equivalent, in the sense that an electorate is k-local swaps single-peaked iff it is PerceptionFlip$_k$ single-peaked.

Our Contribution The originality of the distance measure we introduce in the paper is that it directly follows from the very definition of Black's single-peakedness condition. For a given axis on the candidates, it consists in counting the number of violations of the single-peakedness condition in the preferences. We give a formal definition in Section 2, as well as some insights on the differences between this measure, k-voter deletion single-peakedness and k-global swap single-

peakedness. We tackle computational complexity in Section 3: we prove that, as for most of the proposed measures in the literature, computing an axis at minimum distance to an electorate is NP-hard for our measure. We nevertheless propose an exact method to compute such an axis, that turns out to be efficient in practice. Then, in Section 4, we present the results of numerical tests on both real and synthetic election data, to evaluate the relevance of the returned axes on the candidates, providing also comparisons with the other notions of nearly single-peakedness. Some proofs are missing due to lack of space.

2 Definition and Comparison with Voter Deletion and Global Swap

2.1 Definition

We start by recalling some basic terminology of social choice theory. Given a set $\mathcal{C}=\{c_1, c_2, \ldots, c_m\}$ of candidates and a set $\mathcal{V}$ of n voters ($|V|=n$), each voter $v \in \mathcal{V}$ ranks all the candidates from the most to the least preferred one. This ranking is called the *preference relation* of v. The (multi)set $\mathcal{P}$ of preference relations of all the voters in $\mathcal{V}$ is called a *profile*. The couple $(\mathcal{C}, \mathcal{P})$ is called an *election*. The definition of a *single-peaked* profile states as follows:

Definition 1 (Single-peakedness). *Let an* axis *A be a total order $\lhd_A$ over a set $\mathcal{C}=\{c_1, \ldots c_m\}$ of candidates. Let $>_v$ denote the preference relation (total order) of a voter v over $\mathcal{C}$. Let c^* denote the most preferred candidate of v (also called the* peak *of v), i.e., $c^* >_v c$ for all $c \neq c^*$. The preference $>_v$ is* single-peaked with respect to *A if for any $c_i, c_j \in C$, if $c_j \lhd_A c_i \lhd_A c^*$ or $c^* \lhd_A c_i \lhd_A c_j$ then $c^* >_v c_i >_v c_j$ holds. A profile $\mathcal{P}$ is said to be single-peaked with respect to A if every vote is single-peaked with respect to A.*

Definition 2 (Betweenness relation). *The* betweenness relation induced by an axis *A is the relation R_A defined by:*

$$R_A = \{(c_i, c_j, c_k) \in \mathcal{C}^3 : c_i \lhd_A c_j \lhd_A c_k \text{ or } c_k \lhd_A c_j \lhd_A c_i\}.$$

Put another way, $(c_i, c_j, c_k) \in R_A$ means that c_j is between c_i and c_k on the axis A (note that c_i, c_j and c_k do not need to be consecutive on A). The notion of *A-forbidden triple* that we introduce now will make it possible to measure the consistency of a profile with an axis:

Definition 3 (A-forbidden triple). *Let c^* be the peak of a voter v. If $c^* >_v c_i >_v c_j$ and $(c^*, c_j, c_i) \in R_A$, then the triple $T=(c^*, c_i, c_j)$ is called A-forbidden in v.*

Counting the number of A-forbidden triples in a profile $\mathcal{P}$ (by summing over all $>_v \in \mathcal{P}$) amounts to counting the number of violations of the definition of single-peakedness w.r.t. A. Note that the number of A-forbidden triples per voter is upper bounded by $(m-1)(m-2)/2$ as c^* is unique.

Example 1. Let us consider the following profile with 5 candidates $\{1, 2, \ldots, 5\}$ and 20 voters with the following preferences:

- $5 > 4 > 3 > 2 > 1$ for 1 voter (type I);
- $1 > 2 > 3 > 4 > 5$ for 10 voters (type II);
- $1 > 5 > 3 > 2 > 4$ for 9 voters (type III).

Let us consider the axis A: $1 \lhd_A 2 \lhd_A 3 \lhd_A 4 \lhd_A 5$. The voter of type I is single-peaked with respect to A, so there is no A-forbidden triple for her. Voters of type II are also single-peaked w.r.t. A, so the A-forbidden triples only occur with voters of type III. There, for each of them there are 4 A-forbidden triples: $(1, 3, 2)$ (as $1 > 3 > 2$ for the voters, but $1 \lhd_A 2 \lhd_A 3$), $(1, 5, 2)$, $(1, 5, 3)$ and $(1, 5, 4)$. Then the number of A-forbidden triples in the profile is 36.

More generally, we define a new notion of nearly single-peakedness, that also entails a distance measure to single-peaked profiles. Let $FT(\mathcal{P}, A)$ denote the number of A-forbidden triples in $\mathcal{P}$, and $FT(\mathcal{P}) = \min_A FT(\mathcal{P}, A)$.

Definition 4 (k-forbidden triples single-peakedness). *We say that a profile $\mathcal{P}$ is* k-forbidden triples single-peaked consistent *if $FT(\mathcal{P}) \leq k$.*

We will denote by $A_{FT}(\mathcal{P})$ an optimal axis, i.e., $A_{FT}(\mathcal{P}) \in \arg\min_A FT(\mathcal{P}, A)$.

Example 2 (Example 1, continued). Now, let us consider the following axis B: $4 \lhd_B 2 \lhd_B 1 \lhd_B 3 \lhd_B 5$. Voters of type II are single-peaked w.r.t. B. For voters of type III there is only one B-forbidden triple ($(1, 5, 3)$), and there are 4 forbidden triples for the unique voter of type I. In total, there are 13 B-forbidden triples, much smaller than 36 A-forbidden triples. Indeed, 19 voters are single-peaked or very close to being so w.r.t. B. Actually, B is an optimal axis, i.e., $FT(\mathcal{P}) = 13$ and $A_{FT}(\mathcal{P}) = B$.

2.2 Forbidden Triples, Voter Deletion and Global Swap

Let us highlight some differences between k-forbidden triples single-peakedness and other notions of nearly single-peakedness. We focus on the notions of k-voter deletion single-peakedness (we denote by VD the corresponding distance measure), k-global swaps single-peakedness (GS), and of course k-forbidden triples single-peakedness (FT). They all result in a single axis on all candidates (contrary to, e.g., multi-dimensional single-peakedness [20] and k-candidate deletion single-peakedness [13]). Formally, $VD(\mathcal{P}, A) = \sum_{>\in\mathcal{P}} \delta(>, A)$ where $\delta(>, A) = 0$ if $>$ is single-peaked w.r.t. A, otherwise 1, and $GS(\mathcal{P}, A) = \sum_{>\in\mathcal{P}} d_{swap}(>, A)$, where $d_{swap}(>, A)$ is the minimum number of swaps of consecutive candidates to make $>$ single-peaked w.r.t. A.

A major difference between VD, on the one hand, and FT and GS, on the other hand, lies on the fact that the latter ones are much smoother. Namely, they quantify how far a preference is from being single-peaked with respect to an axis, by a distance which lies from 0 (single-peaked) to $\Theta(m^2)$. On the opposite side, VD only looks if a preference is single-peaked or not, so the distance is 0 or 1. This may prevent VD to find interesting axis, with respect to which almost all the preferences are almost single-peaked. We illustrate this point on Example 1.

Example 3 (Example 1, continued). There is no axis compatible with voters of type II and III, so the axis $1 \lhd_A 2 \lhd_A 3 \lhd_A 4 \lhd_A 5$ is the (unique, up to reversal) axis compatible with voters of type I and II. Thus, it is optimal for VD. As pointed before, this axis is not satisfactory, as almost half of the voters have preferences very far from being single-peaked with respect to this axis. The axis B optimal for FT seems to better fit nearly single-peakedness, as among the 20 voters, 10 are single-peaked w.r.t. it and 9 are very close to being single-peaked.

GS and FT are intuitively closer to each other than VD, but still have some important differences. A first difference is computational, and will be dealt with in the next sections: while both of them are NP-hard, FT is much easier to compute in practice. FT and GS have also qualitative differences, and we illustrate this through a property for which they behave differently.

Let us consider the following *unpopularity property*. It states that beyond a certain level of unpopularity a candidate c can hardly be viewed as intermediate between others, and thus there should be an optimal axis where c is at an extremity. Let us say that a candidate is *unpopular* if she is never ranked in first position, and ranked last by at least $\lceil n/2 \rceil$ voters[1].

Property 1 (Unpopularity). Let $(\mathcal{C}, \mathcal{P})$ be an election. A distance d verifies the unpopularity property if for any unpopular candidate c there exists an axis A minimizing $d(\mathcal{P}, A)$ where c is at an extremity.

It is well-known that, if a profile is single-peaked, then such an unpopular candidate is indeed necessarily at an extremity of any compatible axis [10]. Interestingly, we show that dealing with nearly single-peakedness, among the considered measures, FT is the only one for which the unpopularity property holds.

Theorem 1 *FT satisfies the unpopularity property, while GS and VD do not.*

Proof. (sketch) We prove that the property holds for FT.

Let c be a candidate never ranked in first position, and ranked in last position by at least half of the voters. Let A be an arbitrary axis such that c is not one of its extremities. Let us denote by m_1 the number of candidates on the left of c in A, and by m_2 the number of candidates on the right of c. We define two axes A_l and A_r obtained from A by putting c respectively on the extreme left position for A_l, and on the extreme right position for A_r. We prove that at least one of the axis A_l, A_r is at least as good as A. To do so, for each voter v, we count down the difference of the number of forbidden triples with respect to A and with respect to A_l and A_r. It consists in counting for each of axes the number of triples involving the candidate c. In fact, as c is never ranked first and the restrictions of A, A_l and A_r on $\mathcal{C} \setminus \{c\}$ lead in the same axis, we observe that the

[1] As in any axiomatic approach, the specific situation considered here does not need to often happen in practice: it is a thought experiment in which one considers a hypothetical situation and examines whether the measure would behave well in such a case.

triple (c_i, c_j, c_k) (with c_i, c_j, c_k different from c) is forbidden with respect to A if and only if it is forbidden with respect to A_l (resp. A_r). Let v be an arbitrary voter and c^* the peak of v. Four configurations are possible:

(i) v ranks c in last position and c^* is on the left of c in A;
(ii) v ranks c in last position and c^* is on the right of c in A;
(iii) v does not rank c in last position and c^* is on the left of c in A;
(iv) v does not rank c in last position and c^* is on the right of c.

The table 1 expresses $FT(>_v, A_l)$ and $FT(>_v, A_r)$ in function of $FT(>_v, A)$ - if the exact value can not be given, the upper bound (representing the worst case) is given. For v of type (i), (c^*, c, c') is forbidden (w.r.t. A) if and only if c' is on the right of c in A (there are m_2 such positions). The candidate c is not involved in any forbidden triple with respect to an axis A_l or A_r, as it is placed on the extremity. The same reasoning applies to v of type (ii). For a v of type (iii), in the worst case c is not involved in any forbidden triple with respect to A, but moving it on the left (resp. right) extremity will create up to $m_1 - 1$ (resp. m_2) new forbidden triples. We reason the same way for type (iv).

type	$FT(>_v, A_l)$	$FT(>_v, A_r)$
(i)	$FT(>_v, A) - m_2$	$FT(>_v, A) - m_2$
(ii)	$FT(>_v, A) - m_1$	$FT(>_v, A) - m_1$
(iii)	$\leq FT(>_v, A) + m_1 - 1$	$\leq FT(>_v, A) + m_2$
(iv)	$\leq FT(>_v, A) + m_1$	$\leq FT(>_v, A) + m_2 - 1$

Table 1. Values of $FT(>_v, A_l)$ and $FT(>_v, A_r)$ in function of $FT(>_v, A)$, according to the type of v.

Assume that $m_1 \leq m_2$. We prove that A_l is always at least as good as A, i.e., $FT(\mathcal{P}, A_l) \leq FT(\mathcal{P}, A)$, which is written:

$$\sum_{>_v \in \mathcal{P}} FT(>_v, A_l) \leq \sum_{>_v \in \mathcal{P}} FT(>_v, A).$$

Thanks to Table 1 it is sufficient to prove that:

$$n^{(i)} m_2 + n^{(ii)} m_1 \geq (n^{(iii)} + n^{(iv)}) m_1 - n^{(iii)}$$

with n^t the number of voters of type t. By assumption, $n^{(i)} + n^{(ii)} \geq \frac{n}{2}$. As we assume that $m_1 \leq m_2$, the inequality holds all the time.

If $m_1 \geq m_2$, we prove in the same manner that A_r is always at least as good as A. □

We will present several other qualitative differences between the measures in the experimental section.

3 Computational Aspects

Given a preference profile $\mathcal{P}$, we first note that determining $FT(\mathcal{P}, A)$ for a given axis A can be handled in polynomial time $\mathcal{O}(nm^2)$, where $n = |\mathcal{P}|$ and $m = |\mathcal{C}|$,

by brute force enumeration of all triples. This complexity can be improved to $\mathcal{O}(nm\sqrt{\log m})$ (proof omitted).

Let us now focus on the complexity of determining if $FT(\mathcal{P}) \leq k$ for a given k, i.e., the following problem:

FT Single-peaked Consistency
Input: An election $(\mathcal{C}, \mathcal{P})$ and an integer k.
Output: Yes if $FT(\mathcal{P}) \leq k$, otherwise no.

We show that, similarly to other measures such as GS or VD, this is an NP-complete problem:

Theorem 2 FT Single-peaked Consistency *is NP-complete.*

We now provide an Integer Program (IP) formulation, that turns out to be efficient in practice in our experiments on this consistency problem. For each pair $\{c_i, c_j\}$ of candidates (with $i \neq j \in \{1, \ldots, m\}$), we introduce a binary variable x_{ij} describing their relative position on the sought axis A. More precisely, the constraints of type 1 and type 2 detailed below will ensure that: $x_{ij} = 1$ if $c_i \lhd_A c_j$, and 0 otherwise.

Additionally, for each voter $v \in \{1, \ldots, n\}$ and each pairwise preference $c_i >_v c_j$ with $\pi(v) \notin \{i, j\}$, where $\pi(v)$ is the index of the peak of v, we define a binary variable z_{vij} related to the triple $(c_{\pi(v)}, c_i, c_j)$. More precisely, the constraints of type 3 and type 4 detailed below will ensure that $z_{vij} = 1$ if $(c_{\pi(v)}, c_i, c_j)$ is A-forbidden in v, and 0 otherwise.

The sum of variables z_{vij} is the number of forbidden triples in the profile $\mathcal{P}$. The IP objective function is then $\min \sum_{(v,i,j) \in T} z_{vij}$, where $T = \{(v, i, j) : \pi(v) \notin \{i, j\}, i \neq j\}$.

We now detail the four types of constraints in the program:

1. For each pair $\{c_i, c_j\}$ of candidates, one and only one of the variables $\{x_{ij}, x_{ji}\}$ equals 1.
2. For each tuple (c_i, c_j, c_k), if $x_{ik} = 1$ and $x_{kj} = 1$ then $x_{ij} = 1$ (because $c_i \lhd_A c_k$ and $c_k \lhd_A c_j \Rightarrow c_i \lhd_A c_j$).
3. For each (v, i, j) such that $\pi(v) \notin \{i, j\}$ and $c_i >_v c_j$, if $c_{\pi(v)} \lhd_A c_j$ and $c_j \lhd_A c_i$ then $z_{vij} = 1$ ($(c_{\pi(v)}, c_i, c_j)$ is A-forbidden in v, on the right side of the peak).
4. For each (v, i, j) such that $\pi(v) \notin \{i, j\}$ and $c_i >_v c_j$, if $c_i \lhd_A c_j$ and $c_j \lhd_A c_{\pi(v)}$ then $z_{vij} = 1$ ($(c_{\pi(v)}, c_i, c_j)$ is A-forbidden in v, on the left side of the peak).

Altogether, we obtain the following IP, where $T = \{(v, i, j) : \pi(v) \notin \{i, j\}, i \neq j\}$:

$$\min \sum_{(v,i,j) \in T} z_{vij}$$

$$s.t. \begin{cases} x_{ij} + x_{ji} = 1 \quad \forall \{c_i, c_j\} \text{ with } i \neq j & (1) \\ x_{ij} \geq x_{ik} + x_{kj} - 1 \quad \forall (c_i, c_j, c_k) \text{ with } i \neq j \neq k & (2) \\ z_{vij} \geq x_{\pi(v)j} + x_{ji} - 1 \quad \forall (v, i, j) \in T \text{ with } c_i >_v c_j & (3) \\ z_{vij} \geq x_{j\pi(v)} + x_{ij} - 1 \quad \forall (v, i, j) \in T \text{ with } c_i >_v c_j & (4) \\ x_{ij} \in \{0, 1\} \ \forall i, j, \quad z_{vij} \in \{0, 1\} \ \forall (v, i, j) \in T \end{cases}$$

4 Experimental Study

We carried out numerical tests[2] on real and randomly generated preference profiles in order to compare experimentally the distance measures GS, VD, FT. For optimizing the VD distance, we used the `C++` code developed by Sui et al. [20], made available on the web[3]. For optimizing the FT distance, we used the Gurobi software to solve the IP formulation. Finally, for optimizing the GS distance, we used a brute force algorithm - to the best of our knowledge, no efficient algorithm is known for this problem so far. In particular, no efficient IP formulation is available in the literature for swap measures.

We study the quality of optimal axes on real data, compared to reference axes whose design is detailed below. To evaluate the quality of an axis, we use the following distance ρ between two axes A and A' defined on the same set of candidates: $\rho(A, A') = |R_A \cap R_{A'}|/|R_A|$ (note that $|R_A| = |R_{A'}|$). Put another way, we measure the proportion of the betweenness relation (see Definition 2) that is common to the optimal axis and the reference axis. We call this proportion, expressed in percentage in the sequel, *recognition rate*.

To go further and better understand the impact of the characteristics of the profiles on the numerical results, we also study the quality of optimal axes on profiles randomly generated according to diverse probability distributions for structured preferences.

4.1 Numerical Tests on Real Data

The real data sets were taken from the 2007 Glasgow city council election and a 2017 voting experiment during the French presidential election. The first data set is available on the PrefLib website[4], a library of preference data and links assembled by Mattei and Walsh [17]. The second data set comes from the website of the experiment called *Voter autrement*[5] [5].

The Glasgow election was separated in 21 wards (with one list of candidates per ward). The data from 20 of them were used in our tests. Each ward involved different candidates and voters, and elected 3 or 4 councillors using the Single Transferable Vote (STV) system. This implies that some political parties had several candidates for the same voting district. In order to fit the data with our setting, we restricted ourselves to the votes (ballots) consisting of complete rankings of the candidates. The number of candidates in the Glasgow data set ranges from 8 to 11, and the number of complete votes from 320 to 1003. In the *Voter autrement* data set, one file was usable for our purpose (file `stv111.csv`, here also reporting the results of an experiment about STV), with 11 candidates from as many distinct political parties and 4068 complete votes.

Regarding the computation times, an optimal axis for VD was computed in less than 3 seconds (sec.) on all data sets, and generally in around 30-40

[2] Tests performed on an Intel Core i7 (1.3 Ghz base, 3.9 Ghz turbo) with 8 GB RAM.
[3] http://www.cs.toronto.edu/ lex/code/asprgen.html
[4] https://www.preflib.org/data/index.php
[5] https://zenodo.org/record/1199545

sec. (resp. 82 sec.) for FT on a Glasgow ward (resp. on the French presidential election data). Of course, the brute force algorithm used for GS was much slower: determining an optimal axis with 6 candidates and 100 voters took about 20 sec., which became 2 minutes for 7 candidates, and 10 minutes for 8 candidates.

For each election (at the level of a ward or a country), we built a reference left-right axis on the candidates. To do so, we used Wikipedia as external source. The free encyclopedia provides indeed a political position (of course debatable) for each political party (e.g., left wing, centre, centre right, etc.). We assumed that the political position of an affiliated candidate corresponds to that of the belonging party, and we built an axis over the affiliated candidates based on these positions. We excluded the non-affiliated candidates from the data sets as we were not able to define a political position for them. Actually, the "Wikipedia axis" is not unique since several parties can be labeled by the same political position, or some parties can have several candidates in an election. For instance, a Wikipedia axis reads $((1,3),2,(4,5))$, where the numbers are the indices of candidates, and candidates $\{1,3\}$ as well as $\{4,5\}$ have indistinguishable political positions. This leads to a set of $2 \cdot 2 = 4$ *compatible* axes : $1 \lhd 3 \lhd 2 \lhd 4 \lhd 5$, $3 \lhd 1 \lhd 2 \lhd 4 \lhd 5$, $1 \lhd 3 \lhd 2 \lhd 5 \lhd 4$, and $3 \lhd 1 \lhd 2 \lhd 5 \lhd 4$.

Note that indistinguishable political positions do *not* mean here that the candidates share the same position on the political spectrum, but that we have a partial knowledge of the exact axis. The sets of candidates with indistinguishable political positions (as $\{1,3\}$ and $\{4,5\}$ above) are called *blocks* in the following. Given a distance measure d (in $\{VD, FT, GS\}$) and a profile $\mathcal{P}$, the recognition rate is formulated in the following manner to take into account blocks:

$\min\{\rho(A_d(\mathcal{P}), A') : A'$ compatible with the Wikipedia axis$\}$

where $A_d(\mathcal{P})$ is an optimal axis according to d.

Apart from the recognition rate, we also distinguish three classes of results for the optimal axis w.r.t. a distance:

- T (True): The optimal axis is compatible with the Wikipedia axis, e.g. $3 \lhd 1 \lhd 2 \lhd 4 \lhd 5$ for $((1,3),2,(4,5))$.
- EE (Exchanged Extremities): The optimal axis can be made compatible with the Wikipedia axis by swapping the far left and far right blocks, e.g. an optimal axis $5 \lhd 2 \lhd 6 \lhd 4 \lhd 3 \lhd 1$ for the Wikipedia axis $((1,3),2,6,4,5)$.
- F (False): The optimal axis is called false otherwise.

We distinguish class EE because the experiments revealed a difficulty in recognizing the two extreme blocks. To get an intuition of what is going on, consider a profile where the two "extreme" candidates are ranked in the two last positions by a large number of voters, in an arbitrary order, and the voters who rank one of them in first position do not want to rank anyone else. The data do not provide then much information to distinguish who is left wing and who is right wing.

The results obtained are summarized in Table 2. Note that only the results for the VD and FT measures are given in the table, because the brute force algorithm used for the GS measure was not able to compute an axis in a reasonable amount of time for more than 8 candidates. Regarding the two profiles with

8 candidates in the Glasgow data set, the results obtained with the GS measure are of class EE, while with FT one result is of class T and the other of class EE (the result is of class F in both cases with VD).

d	T	EE	F	Rate	T	EE	F	Rate
Glasgow city council					French election			
VD	2	1	17	57.25%	0	0	1	58.8%
FT	5	5	10	67%	0	0	1	74.6%

Table 2. Results on real election data. Rate means recognition rate.

Table 2 indicates how many times each class occurs for the VD and FT measures (over 20 preference profiles for the Glasgow city council election and 1 for the French presidential election), as well as the average recognition rate. The results tend to show that the recognition ability of the FT measure is better than that of VD. When the FT measure is used, an axis perfectly compatible with Wikipedia is recognized in nearly 24% of cases; it reaches 48% if one adds the cases when the extremities are swapped.

Let us detail now in a more down-to-earth manner the results obtained on the voting data from the French election. The Wikipedia axis is $W = ((1,2,3),4,5,(6,7),(8,9),10)$, with one non-affiliated candidate excluded from the voting data (for readability, the candidates are here numbered in function of their position in W). The axis A_{VD} minimizing VD is $7 \lhd 8 \lhd 5 \lhd 4 \lhd 3 \lhd 2 \lhd 1 \lhd 6 \lhd 9 \lhd 10$, the axis A_{FT} minimizing the FT measure is $8 \lhd 1 \lhd 2 \lhd 3 \lhd 4 \lhd 5 \lhd 6 \lhd 9 \lhd 7 \lhd 10$; both axes are not compatible with W, but A_{FT} is much better than A_{VD} in terms of recognition rate. While it can be objected that this result may follow from the fact that FT explicitly relates to triples while VD does not (although only triples involving the peak are used in FT, not the whole betweenness relation), note that, by swapping candidates 7 and 9 and moving candidate 8 in A_{FT}, an axis compatible with W is obtained, while many more fixes are needed in A_{VD}.

As for the recognition rate, it is slightly higher for the French election. It may reflect the fact that a national election is commonly more structured by the left-right political spectrum than a local election.

4.2 Numerical Tests on Synthetic Data

We also generated synthetic election data to deepen the analysis of the recogition abilities of the VD, FT and GS measures. The aim is to model situations where the preferences are noisy but there is a strong underlying structure. Given an axis A, each preference relation is generated in two steps:

1. A candidate c^* is drawn uniformly at random in $\mathcal{C}$ and an auxiliary preference relation $>_0$ of peak c^* single-peaked w.r.t. A is generated uniformly at random.
2. A preference relation $>$ is drawn from the Mallows model centered around $>_0$.

We recall that the Mallows model defines a probability distribution on rankings. A central ranking $>_0$ has the highest probability, and the probability of other

rankings $>$ decreases in a Gaussian manner with the Kendall tau distance from $>_0$. The central ranking $>_0$ is often interpreted as a "ground truth", and rankings $>$ as noisy views of $>_0$. Formally, given a dispersion parameter $\theta \geq 0$, the probability $P(>)$ of a ranking $>$ is proportional to $e^{-\theta d(>,>_0)}$, where $d(.,.)$ is the Kendall tau distance. If $\theta = 0$, the uniform distribution is obtained. The greater the value of θ, the higher the probabilities of the rankings around $>_0$. It is known that using the Mallows model with parameters $>_0$ and θ is equivalent to generating a binary relation R where, for each pair c_i, c_j of candidates, if $c_i >_0 c_j$, then $c_i R c_j$ with probability $p = e^{\theta}/(1 + e^{\theta})$; if the obtained binary relation R is transitive then stop and return the corresponding ranking, otherwise repeat the process until R is transitive. For the sake of interpretability, in the tables, we give the value of p instead of θ.

We used the PerMallows R package[6] for generating rankings according to the Mallows model. For a fast generation of the profiles, the number of voters is set to 100 and the number m of candidates varies from 7 to 9. The above probability p takes its values in $\{0.7, 0.75, 0.8, 0.85, 0.9\}$. For each couple (m, p) of parameter values, 100 instances were generated and one counted the number of instances for which the axis is perfectly recognized.

As the GS brute force algorithm is not usable in practice for more than 7 candidates, we give only the results for VD and FT. However, we generated instances with 5 to 7 candidates and observed very similar results with GS and FT.

The results are reported in Table 3. It appears that the VD measure is the one for which axis A is the most often recognized. This result was quite unexpected because it is well-known, as mentioned in the introduction, that an optimal axis for VD explains only a few percentage of voter preferences in real election data - and this is the case in these election data. Nevertheless, the good behaviour of VD can be simply explained by the manner in which the preferences are generated here: the probability that a voter preference is perfectly compatible with A is low but is the highest among all the preferences, thus the law of large numbers plays in favor of VD, and this with all the more intensity as probability p is high.

To refine the analysis, we also studied the recognition rates for FT and VD, since it is a smoother criterion than the previous one. The results are reported in Table 4. The differences are then much narrower, which means that, for the instances where A is not perfectly recognized, the optimal axis for FT is very similar to A.

Computation Times for Optimizing the FT Distance We also carried out some numerical tests with a greater number of candidates and/or voters to evaluate how the running time scales for initializing and solving the IP of Section 3. In preliminary tests, the running times did not appear to vary significantly with the value of θ used for generating the profiles, thus we set $\theta = 0$, which corresponds to a uniform distribution on the rankings. The running times obtained seem to be

[6] https://cran.r-project.org/package=PerMallows

much more sensitive to the number of candidates than to the number of voters: on the one hand, for 100 voters, solving the IP took about 3-4 (resp. 10) minutes for 15 (resp. 20) candidates; on the other hand, for 11 candidates, it took about 50 seconds (resp. 2 minutes, 5 minutes) for 1000 (resp. 5000, 10000) voters.

	7 candidates		8 candidates		9 candidates	
p \ d	VD	FT	VD	FT	VD	FT
0.7	39%	9%	26%	5%	12%	3%
0.75	85%	46%	74%	29%	58%	16%
0.8	100%	93%	98%	81%	91%	67%
0.85	100%	98%	100%	98%	100%	95%
0.9	100%	100%	100%	100%	100%	100%

Table 3. Percentage of profiles where the axis is perfectly recognized, w.r.t. distance measure d and probability p.

	7 candidates		8 candidates		9 candidates	
p \ d	VD	FT	VD	FT	VD	FT
0.7	83.5%	75%	77.5%	72%	72%	69.1%
0.75	96%	89.1%	93.7%	86.7%	58%	83.2%
0.8	100%	99.1%	99.5%	97.3%	91%	96.2%
0.85	100%	99.8%	100%	99.9%	100%	99.7%
0.9	100%	100%	100%	100%	100%	100%

Table 4. Recognition rates w.r.t. measure d and probability p.

Robustness to Similar Candidates Another case which can make single-peaked preferences noisy is the presence of similar candidates. We say that candidates c and c' are *similar* if some voters perceive c as the left neighbour of c' on the left-right spectrum while others perceive the opposite. More generally, a subset of candidates are similar if they are consecutive on the left-right spectrum and the perception of their order changes with the voters. Such a subset of candidates is called *block* below. We studied here the robustness to the presence of similar candidates of each of the considered measures.

Let us call *weak axis* an axis where several candidates are similar, and describe such an axis with the same notation used for Wikipedia axes. We considered weak axes where the blocks contained approximately the same number of candidates - that means there were no political position shared by (considerably) more candidates than the others. In practice, we worked with the axes ((1,2),(3,4),(5,6)) for $m = 6$ candidates, ((1,2),(3,4),(5,6),(7,8)) for $m = 8$, and ((1,2),(3,4,5),(6,7,8),(9,10)) for $m = 10$. For each weak axis, number of voters and number of candidates, we generated 1000 profiles[7] and computed an optimal axis according to FT and VD. For each measure, we counted the number

[7] For a weak axis A, each preference relation in the profile is generated in two steps: (1) an axis A' compatible with A is generated uniformly at random; (2) a candidate c^*

of times the returned axis was compatible with the weak axis. Regarding the FT measure, the optimal axis was compatible with the weak axis *in all tests performed*, independently of the number of voters or candidates. In contrast, the VD measure is much less robust to the presence of similar candidates: the percentages of profiles for which the optimal axis for VD was compatible with the weak axis are given in Table 5.

#cand. \ #voters	100	200	500	1000
6	47%	25%	23%	10%
8	17%	8%	0%	0%
10	8%	1%	0%	0%

Table 5. Percentages of profiles for which the optimal axis for VD was compatible with the weak axis.

5 Conclusion

We have proposed a new distance measure to single-peakedness, based on counting the number of violations of Black's definition. After a brief comparison with other existing measures, we have shown that determining an optimal axis for this measure is NP-hard. We have then presented an IP formulation, and carried out numerical tests on real and synthetic data. They show that the proposed measure compares favorably to other popular measures. In particular, the IP formulation is operational while no efficient procedure is known for minimizing the number of swaps in the preferences to make them single-peaked; it is more robust to noise in preferences than minimizing the number of votes to delete.

For future work, from a computational viewpoint, one may wonder whether problems that are NP-hard in general but polynomial time on single-peaked profiles remain tractable for nearly (w.r.t. FT) single-peaked electorates. Besides, a local version of the new measure (where, instead of the sum, we minimize the *maximum* over the voters of the number of forbidden triples), together with a comparison to local versions of VD and GS, might be investigated. Also, it would be interesting to undertake the same type of approach, based on the very definition of a restricted domain, for defining distance measures to other domains, such as the single-crossing domain [16], or single-peaked preferences *on a graph* [18].

References

1. Arrow, K.: Social Choice and Individual Values. John Wiley & Sons, Chapman & Hall (1951)

is drawn uniformly at random in $\mathcal{C}$, then a preference relation is generated uniformly at random among preferences of peak c^* single-peaked w.r.t. A'.

2. Barberà, S., Gul, F., Stacchetti, E.: Generalized median voter schemes and committees. Journal of Economic Theory 61(2), 262–289 (1993)
3. Bartholdi III, J., Trick, M.A.: Stable matching with preferences derived from a psychological model. Operations Research Letters 5(4), 165–169 (1986)
4. Black, D.: On the rationale of group decision-making. The Journal of Political Economy 56(1), 23–34 (1948)
5. Bouveret, S., Blanch, R., Baujard, A., Durand, F., Igersheim, H., Lang, J., Laruelle, A., Laslier, J.F., Lebon, I., Merlin, V.: Voter autrement 2017 - online experiment (Jul 2018), https://doi.org/10.5281/zenodo.1199545
6. Bredereck, R., Chen, J., Woeginger, G.J.: Are there any nicely structured preference profiles nearby? Mathematical Social Sciences 79, 61–73 (2016)
7. Conitzer, V.: Eliciting single-peaked preferences using comparison queries. Journal of Artificial Intelligence Research 35, 161–191 (2009)
8. Coombs, C.H.: Psychological scaling without a unit of measurement. Psychological review 57(3), 145 (1950)
9. Cornaz, D., Galand, L., Spanjaard, O.: Bounded Single-Peaked Width and Proportional Representation. In: ECAI. pp. 270–275. IOS Press (2012)
10. Doignon, J., Falmagne, J.: A polynomial time algorithm for unidimensional unfolding representations. J. Algorithms 16(2), 218–233 (1994)
11. Elkind, E., Lackner, M.: On detecting nearly structured preference profiles. In: Proceedings of the AAAI Conference on Artificial Intelligence (2014)
12. Elkind, E., Lackner, M., Peters, D.: Structured preferences. In: Trends in computational social choice, chap. 10, pp. 187–207. AI Access (2017)
13. Erdélyi, G., Lackner, M., Pfandler, A.: Computational aspects of nearly single-peaked electorates. In: Proc. of 27th AAAI. pp. 283–289. AAAI Press (2013), an extended version appeared in *J. of AI Research* in 2017
14. Faliszewski, P., Hemaspaandra, E., Hemaspaandra, L.A.: The complexity of manipulative attacks in nearly single-peaked electorates. In: TARK. pp. 228–237. ACM (2011), an extended version appeared in *Artif. Intel.* in 2014
15. Feld, S.L., Grofman, B.: Research note partial single-peakedness: An extension and clarification. Public Choice 51(1), 71–80 (1986)
16. Jaeckle, F., Peters, D., Elkind, E.: On recognising nearly single-crossing preferences. In: Proceedings of 32th AAAI (2018)
17. Mattei, N., Walsh, T.: Preflib: A library of preference data http://preflib.org. In: Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT 2013). pp. 259–270. Lecture Notes in Artificial Intelligence, Springer, Brussels, Belgium (2013)
18. Nehring, K., Puppe, C.: The structure of strategy-proof social choice—part I: General characterization and possibility results on median spaces. Journal of Economic Theory 135(1), 269–305 (2007)
19. Niemi, R.G.: Majority decision-making with partial unidimensionality. American Political Science Review 63(2), 488–497 (1969)
20. Sui, X., Francois-Nienaber, A., Boutilier, C.: Multi-dimensional single-peaked consistency and its approximations. In: Proceedings of the Twenty-Third IJCAI. pp. 375–382. AAAI Press, Beijing, China (2013)
21. Walsh, T.: Uncertainty in preference elicitation and aggregation. In: AAAI. vol. 7, pp. 3–8 (2007)