Single-peaked consistency and its complexity

Bruno Escoffier*, Jérôme Lang †, Meltem Öztürk‡

Abstract

A common way of dealing with the paradoxes of preference aggregation consists in restricting the domain of admissible preferences. The most well-known such restriction is *single-peakedness*. In this paper we focus on the problem of determining whether a given profile is single-peaked with respect to some axis, and on the computation of such an axis. This problem has already been considered in [2]; we give here a more efficient algorithm and address some related issues, such as the number of orders that may be compatible with a given profile, or the communication complexity of preference aggregation under the single-peakedness assumption.

Key words : single peakedness, preferences, complexity

1 Introduction

Aggregating preferences for finding a consensus between several agents is an important topic at the boarder between social choice and artificial intelligence. Given the preferences of a set of agents (or voters) over a set of alternatives (or candidates), preference aggregation aims at determining a collective preference relation representing as much as possible the individual preferences, whereas voting rules consists in finding a socially preferred candidate.

Among the paradoxes and impossibility theorems of preference agregation, the most famous may be the following three (in all three cases we assume that there are at least 3 alternatives):

*LAMSADE, Université Paris Dauphine, 75775 Paris `escoffier@lamsade.dauphine.fr`
†IRIT, Université Paul Sabatier, Toulouse, France lang@irit.fr `lang@irit.fr`
‡CRIL, Université Lille-Nord de France, Artois, 62307, France {`ozturk`}`@cril.fr`

- the *Condorcet paradox [3]*: a Condorcet cycle is a sequence of candidates $x_1, \ldots, x_k$ such that for all $i \leq k-1$, a majority of voters prefers x_i to x_{i+1}, and a majority of voters prefers x_k to x_1. Such cycles make it impossible to build a collective preference relation compatible with pairwise majority comparisons between candidates.

- *Arrow's theorem [1]*: any unanimous aggregation function for which the pairwise comparison between two alternatives is independent or irrelevant alternatives is dictatorial;

- *Gibbard and Satterthwaite's theorem* [7, 8]: any surjective and nondictatorial voting rule is manipulable.

A *profile* consists of a collection of preference relations over the candidates (one per voter). In the above results, any profile is admissible. However, in some contexts, voters' preferences may have a special structure restricting the domain of admissible profiles. The most well known such restriction is *single-peakedness*. It assumes that there is a natural linear axis, independent of the voters, on which alternatives are positioned: one may for instance think of a left-right axis as in politics, or a numerical axis (when the voters have to decide for instance about an amount of money to spend). A voter has a single-peaked preferences with respect to such an axis if, on each side of the "peak" (that is, the preferred candidate), his preference grows with the proximity to the peak. It is well-known that Condorcet cycles cannot occur when preferences are single-peaked; therefore, one escapes from the Condorcet paradox as well as Arrow's and Gibbard-Satterthwaite's theorem.

However, this way of escaping the paradoxes and impossibility theorems assumes that the axis on which the candidates are positioned is known in advance. In contexts where it is partially or fully unknown, one should identify it before any aggregation process is started. Therefore, we consider the problem of determining whether, given the preferences of some agents on a set of alternatives, these preferences are single-peaked with respect to some axis (which we refer to as *single-peaked consistency*), and if so, how one of the possible axes can be determined. This problem has been considered in [2] (as well as the problem of determining whether a profile is single peaked w.r.t. a *tree* [9], which is weaker than single peakedness w.r.t. an axis). They give an algorithm in $O(m.n^2)$ where n (resp. m) is the number of candidates (resp. voters), based on matrix representation. We give here a different algorithm, both more intuitive and efficient since it works in time $O(m.n)$. While the difference between $O(m.n)$ and $O(m.n^2)$ is pratically not very significant for standard political elections where n is typically small, this is no longer the case when the set of alternatives (or "candidates ") has a combinatorial structure, which is often the case in AI applications. A related problem is addressed by Conitzer [4]: without the prior knowledge of the axis, but knowing the preference relation of one agent (which gives some incomplete information about the axis), how can we elicit as efficiently as possible the preferences of a second agent?

Single peaked consistency is important in at least two contexts. First, some domains tend to have a single-peaked structure, but for some reason we may not know the axis : In this case, from a few votes (for instance obtained from a sample of votes), we may learn this axis. Second, in some domains it is unclear whether it is reasonable to assume single-peakedness: then, checking the single-peaked consistency of the preference relations of a few voters gives a good hint as to whether single-peakedness is reasonable.[3]

In Section 2, we define single-peaked consistency and give a constructive algorithm that checks whether a profile is single-peaked consistent, and if so, returns a compatible axis. This algorithm works in time $O(n.m)$, where n is the number of agents and m the number of alternatives. In Section 4 we study a few combinatorial aspects of single-peaked preferences; in particular, we give a result on the number of axes that are compatible with a tuple of single-peaked preferences. In Section 5 we give a simple additional result on the communication complexity of preference aggregation of single-peaked preferences. Finally we point to interesting extensions of our work.

2 Single-peaked preferences

Let $V = \{1, \ldots, m\}$ be a finite set of voters and $X = \{x_1, \ldots, x_n\}$ a finite a set of candidates (or alternatives), with $n \geq 3$.

Definition 1 *A* preference relation $\succ$ *on* X *is a linear order on* X. *The* peak *of a preference relation* $\succ$ *is the candidate* $x^* = peak(\succ)$ *such that* $x^* \succ x$ *for all* $x \in X \setminus \{x^*\}$. *A* profile *is a* m*-uple* $P = \langle \succ_1, \ldots, \succ_m \rangle$ *of preference relations on* X.

Definition 2 *An* axis O *(noted by* $>$*) is a linear order on* X. *Given two candidates* $x_i, x_j \in X$, *a preference relation* $\succ$ *on* X *whose peak is* x^*, *and an axis* O, *we say that* x_i *and* x_j are on the same side of the peak of $\succ$ *iff one of the following 2 conditions is satisfied: (1)* $x_i > x^*$ *and* $x_j > x^*$*; (2)* $x^* > x_i$ *and* $x^* > x_j$.

A preference relation $\succ$ *is* single-peaked *with respect to an axis* O *if and only if for all* $x_i, x_j \in X$ *such that* x_i *and* x_j *are on the same side of the peak* x^* *of* $\succ$, *one has* $x_i \succ x_j$ *if and only if* x_i *is closer to the peak than* x_j, *that is, if* $x^* > x_i > x_j$ *or* $x_j > x_i > x^*$.

For simplicity, we sometimes note (as in Example 1) $x_1 x_2 \ldots x_n$ instead of $x_1 \succ x_2 \succ \cdots \succ x_n$ or of $x_1 > x_2 > \cdots > x_n$.

[3]This is for instance of particular interest when alternatives are evaluated on several criteria; here, the hidden axis may be some (*a priori* unknown) combination of the different criteria (projection from a multi-dimensional to a monodimentional representation).

Example 1 *Let* $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ *and* $O = (x_1 > x_2 > x_3 > x_4 > x_5 > x_6)$. *The preferences* $x_2x_3x_4x_1x_5x_6$*;* $x_4x_3x_2x_5x_6x_1$*; and* $x_6x_5x_4x_3x_2x_1$ *are single-peaked with respect to* O *but not* $x_4x_3x_5x_1x_6x_2$*. Indeed,* x_1 *and* x_2 *are on the same side of the peak (*x_4*) but* x_2 *is not preferred to* x_1 *while it is closer to the peak than* x_1*.*

An interesting question is the existence of a *common* axis to all voters, such that the preferences of these voters are single-peaked with respect to this common axis.

Definition 3 *A profile* $\langle \succ_1, \ldots, \succ_m \rangle$ *is* single-peaked with respect to O *iff for each voter* i*,* $\succ_i$ *is single-peaked with respect to* O*.*

Whether single-peakedness seems justified or not strongly depends on the nature of X. It is often deemed reasonable if the axis represents an objective left-right political axis such that voters' preferences are determined only from the position of the candidates on the axis, or else, if X is a set of numerical values or more generally a set equipped with a natural ordering.

Conitzer [4] considers the *elicitation* of single-peaked preferences. The elicitation process is all the more efficient as the amount of communication required by the process is low. This amount of communication can be measured in terms of the number of *elementary queries* of the form "between the candidates x and y, which one do you prefer?"

3 Single-peaked consistency

A very natural question is the following: given a p-voter profile, is it single-peaked with respect to some (unknown) axis? This is defined formally as follows:

Definition 4 (single-peaked consistency) *A preference profile* $P = \langle \succ_1, \ldots, \succ_m \rangle$ *on* X *is* single-peaked consistent *if there exists an axis* O *such that for all* i*,* $\succ_i$ *is single-peaked with respect to* O*.*

When P is single-peaked with respect to the axis O, we say that O is *compatible with* P. For every axis O, we denote by $SP(O)$ the set of preference relations on X that are single-peaked with respect to O. For instance, if $n = 3$ and $O = x_1 > x_2 > x_3$, then $SP(O) = \{x_1x_2x_3, x_2x_1x_3,\ x_2x_3x_1, x_3x_2x_1\}$.

The main problem associated with this definition is to determine if a given profile is single-peaked consistent. We now present the main result of this article, *i.e.* the resolution of this problem. More precisely, we propose an algorithm working in time $O(mn)$ which, given a profile, outputs an axis compatible with this profile if it exists, and find a contradiction otherwise. The axis is built recursively, starting from the candidates ranked in last position by one or more voters. Indeed, we have the following easy lemma.

Lemma 1 *Let x be a candidate ranked in last position by a voter i. If the axis O is compatible with $\succ_i$, then x is either in the leftmost or in the rightmost position in O.*

Proof. If x is neither in the leftmost nor in the rightmost position, then there exist a candidate y on the left of x and a candidate z on the right of x (in O). But $y \succ_i x$ and $z \succ_i x$, contradiction with the fact that $\succ_i$ is single-peaked with respect to O. ∎

As a consequence of Lemma 1: in a single-peaked consistent profile, at most two candidates are ranked last by at least one voter.

Before giving the algorithm, we first explain in detail the first (and easiest) iteration. Let L be the set of all candidates ranked last by at least one voter. We consider the three (exhaustive) possible cases:

- $|L| \geq 3$: then P is not single-peaked consistent, due to Lemma 1.
- $L = \{x\}$: we place indifferently x either in the leftmost or in the rightmost position of the axis; this choice does not create any constraint in the remainder of the construction of the axis. Indeed, the problem is equivalent to first finding an axis compatible with the profiles restricted to the other candidates, and then adding x.
- $L = \{x_1, x_2\}$: we place x_1 and x_2 in the leftmost and the rightmost position of the axis. P is compatible with an order O if and only if it is compatible with the inverse of O; as a consequence, the choice (x_1 in leftmost or rightmost position) does not matter.

Then, the candidates of L being positioned, we iterate the process considering the restriction of the preference relations to the other candidates. Of course, this first iteration is simple because no other candidate is already positioned in the axis.

More generally, at each step of the algorithm, we have a set T of candidates already positioned at the extremal positions of the axis. Without loss of generality, let $T = \{x_1, x_2, \ldots, x_i, x_j, x_{j+1}, \ldots, x_n\}$ the candidates already positioned in the axis under construction: we have $x_1 > x_2 > \ldots > x_i$ in the leftmost positions of the axis O, and $x_j > x_{j+1} > \ldots > x_n$ in the rightmost positions. The other candidates in $\overline{T} = X \setminus T$ will be positioned between x_i and x_j positioned in the leftmost/rightmost position). Then, at this iteration:

- either we find a full compatible axis and P is single-peaked consistent;
- or we find a contradiction and P is not single-peaked consistent;
- or we position one or two new candidates to the right of i and/or to the left of j.

The soundness of the algorithm will follow from the recursive proof of the following hypothesis. At each iteration, the axis under construction verifies the two following properties:

- There exists a compatible axis for P if and only if there exists a compatible axis which extends the axis under construction.
- For any voter k, $x_1 \prec_k x_2 \prec_k \ldots \prec_k x_i$ and $x_j \succ_k x_{j+1} \succ_k \ldots \succ_k x_n$.

In particular, from the second item we deduce that the candidates in T, i and j excepted, are not the peak of any voter.

Let us now analyze the different possible configurations. Let L be the set of candidates ranked last by at least one voter (restricted to the candidates in $\overline{T}$). Based on Lemma 1, we have 3 possible cases:

1. $|L| \geq 3$: contradiction, 3 candidates must be either in position $i+1$ or $j-1$.
2. $L = \{x, y\}$: either x is in position $i+1$ and y in position $j-1$, or *vice versa*, or we will find a contradiction. Let us consider a voter k who ranked x last (among the candidates in $\overline{T}$):
   (a) $x \prec_k x^i$ and $x \prec_k x^j$: this is not possible since necessarily x^i or x^j is ranked worse than x by k (x_i or x_j was the candidate ranked last by k at the previous iteration).
   (b) $x^i \prec_k x$ and $x^j \prec_k x$: x being the last candidate in $\overline{T}$, and since $x^1 \prec_k x^2 \prec_k \ldots \prec_k x^i$ and $x^j \succ_k x^{j+1} \succ_k \ldots \succ_k x^n$, then any axis compatible with voter k on $\overline{T}$ will be compatible on all the candidates. Having positioned the first candidates does not create any constraint . Indeed, all the candidates in $\overline{T}$ are ranked better than all the candidates in T by voter k. As a consequence, for voter k, having x in position $i+1$ and y in position $j-1$ or *vice versa* does not matter.
   (c) $x^i \prec_k x \prec_k x^j \prec_k y$: x is necessarily in position $i+1$. Indeed, having x in position $j-1$ leads to a contradiction: x is positioned between y and x^j in the axis, but $x \prec_k y$ and $x \prec_k x^j$. Then, necessarily x is in position $i+1$ and y in position $j-1$. Symmetrically, if $x^j \prec_k x \prec_k x^i \prec_k y$, then x is necessarily in position $j-1$.
   (d) $x^i \prec_k x \prec_k y \prec_k x^j$ (or the symmetrical case) : x^j is necessarily the peak for the voter k (the candidate positioned immediately to the left is worse, and the candidate x^{j+1} (if any) positioned immediately to the right is also worse, by our recursive hypothesis), hence the candidates in $\overline{T}$ are necessarily positioned between positions i and j following the increasing order of voter k. We test if this axis is compatible with the preferences of other voters. If so, we have a compatible axis, otherwise we conclude that P is not single-peaked consistent.

We repeat step 2 for all voters. If case 2d occurs (for at least one voter), *then the algorithm ends* (either we found an axis, or a contradiction). Otherwise, either we find a contradiction (x have to be placed in two different positions), and the algorithm stops, or we position candidates x and y on the axis.

To conclude, note that if we are not in case 2d , the induction hypothesis $x^1 \prec_k x^2 \prec_k \ldots \prec_k x^i$ and $x^j \succ_k x^{j+1} \succ_k \ldots \succ_k x^n$ remains true after positioning x and y (otherwise, in case 2d the algorithm stops).

3. $L = \{x\}$, *i.e.* each voter ranked x last (in $\overline{T}$). Several cases may occur for voter k:

    (a) $x \prec_k x^i$ and $x \prec_k x^j$: as previously, this case is impossible.

    (b) $x^i \prec_k x$ and $x^j \prec_k x$: no constraint.

    (c) $x^i \prec_k x \prec_k x^j$ (or inverse): x is necessarily in position $i+1$.

Hence, if no contradiction is obtained and no compatible order is found, we position one or two new candidates.

Steps 2 and 3 are repeated until all the candidates are positioned or a contradiction occurs. The previous analysis enables us to state the following result:

Proposition 1 *Let P be a preference profile. The previous algorithm outputs an axis compatible with P if any, or finds a contradiction otherwise.*

Example 2 *Let $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ and consider two voters with the following preferences: $x_6 \prec_1 x_5 \prec_1 x_4 \prec_1 x_1 \prec_1 x_3 \prec_1 x_2$ and $x_1 \prec_2 x_6 \prec_2 x_5 \prec_2 x_2 \prec_2 x_3 \prec_2 x_4$*

- *Iteration 1: The set L of worst candidates is $L = \{x_1, x_6\}$. T being empty, we can choose the positions of x_1 and x_6, for instance respectively in the leftmost and rightmost positions. Partial axis: $x_1 > .... > x_6$.*

- *Iteration 2: $\overline{T} = \{x_2, x_3, x_4, x_5\}$ and $L = \{x_5\}$. For voter 1, $x_6 \prec_1 x_5 \prec_1 x_1$, hence necessarily x_5 is in fifth position in the axis. For voter 2, $x_1 \prec_2 x_5$ and $x_6 \prec_2 x_5$ hence for the voter 2 the positioning does not matter. Partial axis: $x_1 > ... > x_5 > x_6$.*

- *Iteration 3: $\overline{T} = \{x_2, x_3, x_4\}$ and $L = \{x_2, x_4\}$. For voter 1, $x_5 \prec_1 x_4 \prec_1 x_1 \prec_1 x_2$, hence necessarily x_4 is in fourth position, and therefore x_2 is in second position. For voter 2, $x_1 \prec_2 x_5 \prec_2 x_2 \prec_2 x_4$ hence for her the positioning does not matter. Partial axis: $x_1 > x_2 > . > x_4 > x_5 > x_6$*

- *Iteration 4:* $\overline{T} = \{x_3\}$. *We verify that with* x_3 *in third position, the partial axis* $x_2 > x_3 > x_4$ *is compatible with the two votes. Then, the axis* $x_1 > x_2 > x_3 > x_4 > x_5 > x_6$ *is compatible with the profile constituted by the preference relations of the 2 voters.*

Example 3 *Let us consider five candidates and two voters, with* $x_1 \prec_1 x_2 \prec_1 x_3 \prec_1 x_4 \prec_1 x_5$ *and* $x_4 \prec_2 x_3 \prec_2 x_2 \prec_2 x_1 \prec_2 x_5$

- *Iteration 1:* $L = \{x_1, x_4\}$*: we choose* $x_1 > ... > x_4$.

- *Iteration 2:* $\overline{T} = \{x_2, x_3, x_5\}$ *with* $L = \{x_2, x_3\}$*. voter 1:* $x_1 \prec_1 x_2 \prec_1 x_3 \prec_1 x_4$ *hence* x_4 *is necessarily the peak of the voter 1. The unique axis possible is consequently* $x_1 > x_2 > x_3 > x_5 > x_4$*; it is not compatible with the preference relation of the second voter. This profile is not single-peaked consistent.*

Example 4 *Let us consider five candidates and two voters, with* $x_1 \prec_1 x_2 \prec_1 x_3 \prec_1 x_4 \prec_1 x_5$ *and* $x_4 \prec_2 x_2 \prec_2 x_3 \prec_2 x_1 \prec_2 x_5$*. Iteration 1 is as Example 3. For iteration 2:* $\overline{T} = \{x_2, x_3, x_5\}$ *with* $L = \{x_2\}$*. For voter 1,* $x_1 \prec_1 x_2 \prec_1 x_4$ *hence* x_2 *must be immediately to the right of* x_1*. For voter 2,* $x_4 \prec_2 x_2 \prec_2 x_1$ *hence* x_2 *must be immediately to the left of* x_4*. Contradiction. This profile is not single-peaked consistent.*

Example 4 shows that a 2-voters profile may not be consistent.

Now, we analyse the running time of the algorithm. At each iteration, either we find a compatible order, or a contradiction, or we position at least one new element. Assuming that each preference relation is given in decreasing order, we find the set L of worst candidates in time $O(m)$. Then, for each voter we do $O(1)$ comparaisons. Step 2d can be possibly longer, since we test the compatibility of an axis with the preference relations of all voters. This step is done in time $O(nm)$ ($O(n)$ for each voter), but it occurs at most once during the algorithm. Then, as long as this step does not occur we have $T(n,m) \leq T(n-1,m) + O(m)$. This sums up to $T(n,m) = O(nm)$, and the possible execution of step 2d still leads to $T(n,m) = O(nm)$. Therefore :

Proposition 2 *The single-peaked consistency problem can be solved in time* $O(nm)$.

Proposition 2 improves the $O(m.n^2)$ algorithm given in [2] and is established by a completely different method. Interestingly the algorithm in [9] for cumputing a *tree* with respect to which the profile is single peaked has similarities with ours. However, not only it works in $O(m.n^2)$ but it is designed to find a *tree* and does not guarantee to output an axis where there exists one.

Of course, there may exist *several* axes compatible with a given profile (the number of such axes is the topic of the next section), and given a profile, one might be interested in finding *all* the axes compatible with it[4]. It is easy to see that the method we proposed can be adapted to find all axes compatible with a profile P; indeed, it is sufficient to keep in steps 2b and 3b all the different possibilities when several choices are possible. As we will see in the next section, there can be an exponential number of compatible axes, hence of course the running time cannot be polynomially bounded.

Example 5 *Let us consider 7 candidates and two voters, with:*

$$x_4 \prec_1 x_3 \prec_1 x_5 \prec_1 x_6 \prec_1 x_2 \prec_1 x_1 \prec_1 x_7$$
$$x_5 \prec_2 x_6 \prec_2 x_4 \prec_2 x_3 \prec_2 x_2 \prec_2 x_7 \prec_2 x_1$$

The modified algorithm gives the 8 compatible axes:

$$\begin{array}{ll} x_4x_3x_2x_1x_7x_6x_5 & x_5x_6x_1x_7x_2x_3x_4 \\ x_4x_3x_2x_7x_1x_6x_5 & x_5x_6x_7x_1x_2x_3x_4 \\ x_4x_3x_1x_7x_2x_6x_5 & x_5x_6x_2x_7x_1x_3x_4 \\ x_4x_3x_7x_1x_2x_6x_5 & x_5x_6x_2x_1x_7x_3x_4 \end{array}$$

4 On the number of axes compatible with a profile

In Section 3, we proposed an algorithm for computing an axis compatible with a given profile, but such an axis is not necessarily unique. It is now worth to give bounds on the number of axes compatible with a given profile, as well as the prior probability that a profile is single-peaked consistent. As mentioned earlier, this set of compatible axes may be of some interest when new voters give their preferences. Obviously, the more compatible axes we have, the more likely this new profile is single-peaked consistent. On the other hand, the existence of several compatible axes may be considered as a drawback, for instance if our goal is to *learn* some structural information about the candidates. In this section, we focus on the minimum and maximum numbers of axes that are compatible with a set of k *distinct* votes for n candidates. Let $q(k, n)$ and $Q(k, n)$ be these respective numbers.

To begin with, remark that P is compatible with O then P is compatible with the inverse of O (denoted by O^{-1}). Moreover, of course, the more voters (or candidates), the

[4]This may be useful for instance if a new voter appears. In this case, it is very easy to find for instance if this new profile is single-peaked consistent.

less the number of compatible axes. Hence, q and Q are even and non increasing with k and n.

First, let us deal with the case of a single axis.

Lemma 2 $|SP(O)| = 2^{n-1}$

Proof. Let $O = x_1 > x_2 > \ldots > x_n$ and $\succ \in SP(O)$. $\succ$ is fully determined by (a) its peak x_i and (b) the positions of $x_1, \ldots, x_{i-1}$ in the remaining $n-1$ positions. Indeed, we know that $x_j \succ x_k$ for $x_k < x_j < x^*$ and for $x^* < x_j < x_k$, hence (a) and (b) suffice to describe $\succ$. There are $\binom{n-1}{i-1}$ possible positionings for $x_1, \ldots, x_{i-1}$, therefore, $\binom{n-1}{i-1}$ preference relations in $SP(O)$ whose peak is x_i. To get the cardinality of $SP(O)$ we have to sum up over i. ∎

By symmetry considerations, we obtain that there exist 2^{n-1} axes compatibles with a given preference relation. Hence, $q(1,n) = Q(1,n) = 2^{n-1}$. We also know (cf. Example 4 without x_5) that $q(2,4) = 0$, therefore, for every $k \geq 2$ and $n \geq 4$ we have $q(k,n) = 0$. The only missing case is $q(2,3)$, which can be easily shown to be equal to 2.

The case of $Q(k,n)$ is more interesting. We already know that $Q(1,n) = 2^{n-1}$, and, by Lemma 2, $Q(k,n) = 0$ for $k > 2^{n-1}$. We now show that the maximum number of compatible axes is globally inversely proportional to the number of distinct votes. More precisely, $Q(k,n) = 2^n/k$ when $k = 2^j$ $1 \leq j \leq n-1$ (Proposition 3). This gives bounds on $Q(k,n)$ for the other values of k. We first show this result for $k = 2^{n-1}$ (Lemma 3), and then some relations between the values of $Q(k,n)$ when n and/or k change (lemmas 4 and 5).

Lemma 3 $Q(2^{n-1}, n) = 2$

Proof (sketch). Let $O = x_1 > x_2 > \cdots > x_n$. Let us focus on the set of axes compatible with the 2^{n-1} preference relations (see Lemma 2) in $SP(O)$. Let x_i, x_j with $x_i > x_j$. The relation R: $x_j \succ x_{j+1} \succ \ldots x_n \succ x_{j-1} \succ \ldots \succ x_i \succ \ldots \succ x_1$ is compatible with O. Any axis O' such that $x_j >_{O'} x_i >_{O'} x_n$ is not compatible with R. Therefore, O is the only axis compatible with $SP(O)$ whose rightmost element is x_n. By symmetry, O^{-1} is the only one whose rightmost element is x_1. The result follows from Lemma 1. ∎

Lemma 4 *For all* $k, n \geq 1$, $Q(k, n+1) \geq 2Q(k,n)$

Proof. Consider a profile P of k preference relations on n candidates that are compatible with $Q(k,n)$ axes. We extend these k relations to $n+1$ candidates by positioning the new candidate x_{n+1} last in all relations. For each of the $Q(k,n)$ axes compatible with the initial k relations, we can add x_{n+1} either as the leftmost element or rightmost element. Therefore we obtain $2Q(k,n)$ distinct axes, compatible with k distinct preference relations. Thus, $Q(k,n+1) \geq 2Q(k,n)$. ∎

Lemma 5 *For all $n \geq 2$ and all k :*

$$Q(k, n+1) \leq max\{Q(\lceil k/2 \rceil, n), 2Q(k, n)\}.$$

Proposition 3 *For all $n \geq 2$, all $j \in [1, n-1]$: $Q(2^j, n) = 2^{n-j}$*

Proof (sketch). Let j between 1 and $n-1$. By Lemma 3, $Q(2^j, j+1) = 2$. Thanks to Lemma 4, we get $Q(2^j, n) \geq 2^{n-j}$. Using Lemma 5, we can show that it is in fact an equality. ∎

In particular, we get that for each k between 2 and 2^{n-1}, $2^{n-1}/k < Q(k, n) < 2^{n+1}/k$ (or, if we want tighter bounds: $2^{n-\lfloor \log_2(k) \rfloor - 1} < Q(k, n) \leq 2^{n-\lfloor \log_2(k) \rfloor}$).

Lemma 2 enables us to give an approximation of the probability that a randomly generated k-voter, n-candidate profile is single-peaked consistent. Suppose P is drawn randomly with a uniform probability: for each voter i, the probability that a given preference relation R is the preference relation of voter i is $\frac{1}{n!}$, the preference relations of two different voters being independent, therefore each possible profile has a probability of $\left(\frac{1}{n!}\right)^k$. From Lemma 2 we get that given an axis O and a preference relation R, the probability that $R \in SP(O)$ is $\frac{2^{n-1}}{n!}$. Now, the probability that a k-voter profile is compatible with a fixed axis O is $\left(\frac{2^{n-1}}{n!}\right)^k = \frac{2^{k(n-1)}}{n!^k}$. This implies that the probability that a k–voter profile on n candidates is single-peaked consistent is smaller than $n!\frac{2^{k(n-1)}}{n!^k} = \frac{2^{k(n-1)}}{n!^{k-1}}$. (The exact probability is of course lower than that, but gets asymptotically close to this upper bound, when the number of voters grows.) Therefore, the probability of single-peaked consistency decreases exponentially with both with the number of voters and the number of candidates[5]. Finally, note that the probability of single-peaked consistency is lower than the probability of non-occurrence of the Condorcet paradox. which has received much more attention (see e.g. [6]).

5 Communication complexity of the aggregation of single-peaked preferences

We end this paper by a short additional result on the communication complexity of the aggregation of single-peaked preferences. As said in Section 1, the restriction to single-peaked profiles allows for escaping usual impossibility theorems, which means that there exist natural and satisfactory voting rules and aggregation functions under single-peakedness.

[5]Of course, the above computation relies on the assumption that the preference relations of the voters are independent, which is arguably not very realistic. Positive correlations between preference relations allow the probability of single-peaked consistency to decrease less fast.

First, it is well-known that, if the number of voters is odd (which we will now assume for the sake of simplicity), then the *median of the peaks* is the Condorcet winner and the *pairwise majority aggregation* of a profile P, defined by $x \succ_P^* y$ if and only if $|\{k \mid x \succ_k y\}| > \frac{m}{2}$ for all $x, y \in X$, is a linear order.

We are now interested in the *communication complexity* of the median voting rule and pairwise majority aggregation for single-peaked profiles. The deterministic communication complexity of a function is the minimal quantity of information (measured in number of bits) used by the a protocol that computes it. One can find a study on the communication complexity of several voting rules (without the single-peakedness restriction) in [5].

In this Section, we assume that the axis O is given (and is common knowledge to all voters).

Obviously, the deterministic communication complexity of the median of peaks for single-peaked profiles is at most $m.\lceil\log n\rceil$, since the median of peaks can simply be computed by asking voters to name their peak, which needs $\lceil\log n\rceil$ bits per voter. The lower bound is less obvious. It can be obtained by taking the same fooling set as in the proof of Theorem 3 in [5], and taking an axis whose median is a. This leads to the following result:

Proposition 4 *The deterministic communication complexity of the median of peaks is* $O(m.\log n)$ *and* $\Omega(m.\log n)$[6].

The (deterministic) communication complexity of pairwise majority aggregation is a little less obvious but still very simple:

Proposition 5 *The deterministic communication complexity of pairwise majority aggregation for single-peaked profiles is at most* $2m.\lceil\log n\rceil + 2m(n-2)$.

The proof uses a protocol very similar to the one used in [4] for the elicitation of single-peaked preferences of a voter. We start by determining the median of peaks, which needs $m.\lceil\log n\rceil$ bits (see above). Then we communicate the result to each voter (which requires again $m.\lceil\log n\rceil$ bits). After this, the voters are asked $m-2$ successive pairwise comparisons, according to the following protocol, presented informally on an example: suppose the median of peaks is x_3 (the axis being $x_1 < x_2 < x_3 < x_4 < \ldots$). We set $rank(x_3) = 1$, and we ask to each voter her preference between x_2 and x_4. If there is a majority for x_2, then x_2 is the second "socially preferred candidate" and we set $rank(x_2) = 2$. Then, we ask to each voter her preference between x_1 and x_4, and so on. Each of these steps requires the central authority (CE) to send to each voter the information enabling her to know the two candidates she has to compare. For this, CE does not have to send the identity of the two candidates (which would require $2\lceil\log n\rceil$ bits)

[6]Actually, the same bounds would hold for the *nondeterministic* communication complexity – see [5].

but only one bit, indicating whether the winner of the previous step is the "right" candidate, or the "left" one (for instance, after the voters have been asked their preferences between x_2 and x_4, if there is a majority for x_4 then CE sends the information "right" to the voters, who now know the next comparison is between x_2 and x_5). Each voter sends the answer to CE, which requires one bit per voter. Hence each iteration requires $2m$ bits. There are exactly $n-2$ iterations, hence the protocol requires the communication of $m.\lceil \log n \rceil + 2m(n-2)$ bits. Finally, we see easily that $x \succ_P^* y$ if and only if $rank(x) < rank(y)$, hence the protocol computes $\succ_P^*$.

6 Discussion

In this article we have studied some combinatorial and algorithmic aspects of reasoning with single-peaked preferences. The main contribution is an algorithm that outputs an axis compatible with a profile (when there is one) in time $O(mn)$. We have identified the minimal and maximal number of axes that are simultaneously compatible with a profile (which, as a byproduct, gives an approximation of the probability of single-peaked consistency of a randomly generated profile). As a side result we have given some simple results on the communication complexity of the aggregation of single-peaked preferences.

This work deserves some further research in several directions. In particular, as said in Section 4, the probability that a profile single-peakes decreases dramatically with the number of voters and the number of candidates. However, in many practical cases, even if not *stricto sensu* single-peaked, the profile can be *close* (with respect to some metric) to being so. For instance, in a nation-wide political election, given the very high number of voters, the profile is surely not single-peaked. However, in this case, it may be the case that the profile is approximately single-peaked. To make this more precise, we need to define formal notions of "approximate single-peakedness", which are meant to measure how far a profile is from being single-peaked. Several definitions seem natural, such as (1) the minimum number of voters whose deletion gives a single-peaked profile, (2) the minimum number of candidates whose deletion gives a single-peaked profile, or (3) the minimum number of axes such that each preference relation of the profile is single-peaked with at least one axis. Computing these measures of single-peakedness lead to very interesting computational problems, for which our algorithm of Section 3 can be the starting point. For instance, for (1) and (2), we can design a branch-and-bound algorithm that generalizes our algorithm. As for (3), we can modify our algorithm to produce a set of axes which *covers* the whole profile (*i.e.* such that each preference relation of the profile is compatible with at least one axis).

Acknowledgement : The autors are grateful to the Pro ject ANR-05-BLAN-0384 for its financial support.

References

[1] K.J. Arrow. *Social choice and individual values*. J. Wiley, New York, 1951. 2nd edition, 1963.

[2] J. Bartholdi and M. Trick. Stable matching with preferences derived from a psychological model. *Operations Research Letters*, 5(4):165–169, 1986.

[3] Marquis de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie Royale, Paris, 1785.

[4] V. Conitzer. Eliciting single-peaked preferences using comparison queries. In *Proceedings of AAMAS-07*, pages 408–415, 2007.

[5] V. Conitzer and T. Sandholm. Communication complexity of common voting rules. In *Proceedings of (EC-05*, pages 78–87, 2005.

[6] W. Gehrlein. Condorcet's paradox and the likelihood of its occurrence: different perspectives on balanced preferences. *Theory and Decision*, 52(2):171–199, 2002.

[7] A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41:587–601, 1973.

[8] M.A. Satterthwaite. Strategy proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.

[9] M. Trick. Recognizing single-peaked preferences on a tree. *Mathematical Social Sciences*, 17(1):329–334, 1989.