## Provided for non-commercial research and education use. Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.
Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:
http://www.elsevier.com/copyright

# On the computation of median linear orders, of median complete preorders and of median weak orders 

Olivier Hudry<br>Télécom ParisTech 46, rue Barrault, 75634 Paris Cedex 13, France

## A R T I C L E I N F O

## Article history:

Received 10 May 2010
Received in revised form
18 May 2011
Accepted 10 June 2011
Available online 21 July 2011


#### Abstract

Given a finite set $X$ and a collection $\Pi$, called a profile, of binary relations defined on $X$ (which can be linear orders, complete preorders, any relations, and so on), a relation $R$ is said to be median if it minimizes the total number of disagreements with respect to $\Pi$. In the context of voting theory, $X$ can be considered as a set of candidates and the relations of $\Pi$ as the preferences of voters, while a median relation can be adopted as the collective preference with respect to the voters' opinions. If the relations of $\Pi$ are tournaments (which includes linear orders), then there always exists a median complete preorder (i.e. a median complete and transitive relation) which is in fact a linear order. Moreover, if there is no tie when aggregating the tournaments of $\Pi$, then all the median complete preorders are linear orders. We show the same when the median is assumed to be a weak order (a weak order being the asymmetric part of a complete preorder). We then deduce from this that the computation of a median complete preorder or of a median weak order of a profile $\Pi$ of $m$ linear orders is NP-hard for any even $m$ greater than or equal to 4 or for odd $m$ large enough with respect to $|X|$ (about $|X|^{2}$ ). We then sharpen these complexity results when coping with other kinds of profiles $\Pi$ for odd values of $m$. In particular, when the relations of $\Pi$ and the median relation are complete preorders, we obtain the same results for the NP-hardness of Kemeny's problem.


© 2012 Published by Elsevier B.V.

## 1. Introduction

Assume that we are given a finite set $X=\{1,2, \ldots, n\}$ of $n$ candidates and a collection (or multi-set) $\Pi=\left(R_{1}, R_{2}, \ldots, R_{m}\right)$, called a profile, of the preferences $R_{i}$ of $m$ voters $(1 \leq i \leq m)$ who want to rank the $n$ candidates. These preferences can be binary relations without noticeable properties, or can be linear orders, or complete preorders, and so on (see below for the definition of these ordered structures). Note that the relations involved in the profile may be the same: two different voters may share the same preference.

In order to aggregate these $m$ relations into a collective ranking, we may apply Condorcet's pairwise comparison method (Condorcet, 1785) (according to McLean et al. (2008), this method had been suggested a long time ago by Ramon Llull: "Ramon Llull (ca 1232-1316) (...) made contributions which have been believed to be centuries more recent. Llull promotes the method of pairwise comparison, and proposes the Copeland rule to select a winner"; for references upon the historical context, see also for instance (Black, 1958; McLean, 1995; McLean and Urken, 1995; Monjardet, 1991, 2008)). Such a pairwise comparison method consists in computing, for each pair of candidates $\{x, y\}$ (with $x \neq y$ ), the number $m_{x y}$ of voters who prefer $x$ to $y$ and the number $m_{y x}$

[^0]The aim of this paper is to study the complexity of the computation of median complete preorders and of median weak orders (for algorithmic considerations, see for instance Ailon (2010) or Charon and Hudry (2010) and references therein). With respect to a linear order, a complete preorder relaxes the antisymmetry property and thus allows ties between candidates (a complete preorder is a linear order if and only if it is antisymmetric). A weak order is the antisymmetric part of a complete preorder. It is known that the computation of a median linear order is NP-hard, even if all the preferences $R_{i}(1 \leq i \leq m)$ are linear orders (see Bartholdi III et al. (1989), Biedl et al. (2009), Dwork et al. (2001), Hemaspaandra et al. (2005), Hudry (1989, 2008), Wakabayashi (1986, 1998); see Charon and Hudry (2010) for more general references on the properties of median linear orders). Moreover, Wakabayashi showed in Wakabayashi (1986, 1998) that the computation of a median complete preorder of a profile of $m$ binary relations is NP-hard if $m$ is at least about $n^{4}$. This result has been improved and extended in Hudry $(1989,2008)$ in three directions. First, it has been extended to a profile of $m$ linear orders or to a profile of $m$ relations extending the structure of linear order (for instance, a profile of complete preorders, or of partial orders, or of interval orders, and so on); secondly, it has been extended to lower values of $m$ by showing that the problem remains NP-hard when $m$ is at least about $n^{2}$ for even values of $n$ (and still at least about $n^{4}$ when $n$ is odd); thirdly, it has been extended to medians which must be weak orders, under the same conditions as for median complete preorders.

## Summary of results and outline of the paper

In this paper, after giving some definitions and notation in Section 2, we show in Section 3 that, when $\Pi$ is a profile of linear orders, or more generally a profile of tournaments, there always exists a median complete preorder which is a linear order; moreover, if the majority relation $M^{*}$ contains no tie (this is systematically the case for instance when $m$ is odd), any median complete preorder of a profile of tournaments is a linear order; this answers a question set in Hudry (2008) (Remark 19). From this, we deduce in Section 4 that the computation of a median complete preorder or a median weak order of a profile of $m$ linear orders is NP-hard for any fixed even $m$ greater than or equal to 4 and for $m$ large enough (about $n^{2}$ ) when $m$ is odd. In Section 5 , we sharpen the results dealing with profiles of linear orders when we consider profiles of tournaments, and we do the same in Section 6 for other kinds of profiles, considering the main partially ordered structures used as models of preferences (see Bouyssou et al. (2006) and Caspard et al. (2007)). The results are summarized in the conclusion (Section 7).

## 2. Definitions and notation

Let $X=\{1,2, \ldots, n\}$ be a finite set with $n$ elements. A binary relation $R$ defined on $X$ is a subset of the Cartesian product $X \times X$. If $(x, y)$ belongs to $R$, then we write $x R y$; otherwise, we write $x \bar{R} y$. Given a binary relation $R$, we may define an asymmetric relation $R^{a}$ (called the asymmetric part of $R)$ by: $x R^{a} y \Leftrightarrow(x R y$ and $y \bar{R} x)$.

Basic properties that $R$ may fulfil are:

* reflexivity: $R$ is reflexive if, for any $x \in X$, we have $x R x$;
* irreflexivity: $R$ is irreflexive if, for any $x \in X$, we have $x \bar{R} x$;
* antisymmetry: $R$ is antisymmetric if, for any $(x, y) \in X^{2}$ with $x \neq y$, we have $x R y \Rightarrow y \bar{R} x$;
* completeness: $R$ is complete if, for any $(x, y) \in X^{2}$ with $x \neq y$, we have $x R y$ or $y R x$;
* transitivity: $R$ is transitive if, for any $(x, y, z) \in X^{3}$ with $x \neq y \neq$ $z \neq x$, we have the implication $(x R y$ and $y R z) \Rightarrow x R z$.

From the point of view of the theory of NP-completeness, reflexivity or irreflexivity do not matter when dealing with median relations: if reflexivity is required, it will be necessary to add the missing loops ( $x, x$ ); if irreflexivity is required, it will be necessary to delete the existing loops $(x, x)$ (see Hudry (2008) for more details). So we do not pay much attention to these two properties (or to the lack of these properties) in the sequel: the results for the complexity of the problems studied in Section 4 will remain the same if we require reflexivity or irreflexivity or neither of these two properties.

From these basic properties, we may define partially ordered structures (see Bouyssou et al. (2006) or Caspard et al. (2007)). As the names used to denote the usual structures are not always the same in the literature, we specify them below. For the reflexivity or irreflexivity, we choose the irreflexive version because it seems to be the most convenient for the proofs below (once again, it does not matter anyway):

* tournament: a tournament is a relation which is irreflexive, antisymmetric and complete; $\mathcal{T}$ will denote the set of tournaments;
* linear order: a linear order is a relation which is irreflexive, antisymmetric, complete and transitive; $\mathcal{L}$ will denote the set of linear orders;
* partial order: a partial order is a relation which is irreflexive, antisymmetric and transitive; $\mathcal{O}$ will denote the set of partial orders;
* interval order: an interval order is a partial order $R$ verifying the following property: $\forall(x, y, z, t) \in X^{4}$ with $x, y, z$ and $t$ pairwise distinct, $(x R y$ and $z R t) \Rightarrow(x R t$ or $z R y) ; \ell$ will denote the set of interval orders;
* semiorder: a semiorder is an interval order $R$ verifying the following property: $\forall(x, y, z, t) \in X^{4}$ with $x, y, z$ and $t$ pairwise distinct, $(x R y$ and $y R z) \Rightarrow(x R t$ or $t R z) ; \&$ will denote the set of semiorders;
* quasi-order: a quasi-order is a complete relation of which the asymmetric part is a semiorder; $Q$ will denote the set of quasiorders;
* preorder: a preorder is a relation which is irreflexive and transitive; $\mathcal{P}$ will denote the set of preorders;
* complete preorder: a complete preorder is a relation which is irreflexive, complete and transitive; $\mathcal{C}$ will denote the set of complete preorders;
* weak order: a weak order is the asymmetric part of any complete preorder; $\mathcal{W}$ will denote the set of weak orders.
Moreover, let $\mathcal{C o}$ (respectively $\mathcal{A}$ ) denote the set of complete (respectively antisymmetric) relations. Note the following inclusions: $\mathcal{L} \subset\{\subset \ell \subset \mathcal{O}=\mathcal{P} \cap \mathcal{A} ; \mathcal{L} \subset \mathcal{W} \subset \mathcal{A} ; \mathcal{L} \subset \mathcal{C}=$ $\mathcal{P} \cap \mathcal{C o} ; \mathcal{L} \subset \mathcal{T}=\mathcal{C} 0 \cap \mathcal{A} ; \mathcal{L} \subset Q \subset \mathcal{C o}$.

In a complete preorder $C$, it is possible to partition $X$ into $k$ subsets $X_{1}, X_{2}, \ldots, X_{k}$ for an appropriate value of $k$ in such a way that two distinct elements $x$ and $y$ of a same subset are in symmetric relation $(x C y$ and $y C x)$ while in contrast two distinct elements $x \in X_{i}$ and $y \in X_{j}$ of different subsets $X_{i}$ and $X_{j}$ with $1 \leq i<j \leq k$ are in antisymmetric relation: $x C y$ and $y \bar{C} x$. We shall call such a partition the canonic partition of $C$ and we write $X_{1} \succ X_{2} \succ \cdots \succ X_{k}$ to recall that the elements $x$ of $X_{i}$ and $y$ of $X_{j}$ with $1 \leq i<j \leq k$ are in the relations $x C y$ and $y \bar{C} x$. If an element $x$ of $X$ belongs to the subset $X_{i}$, we say that $X_{i}$ is the class of $x$ (with respect to $C$ ). For a linear order, $k$ is equal to $n$ and the subsets $X_{1}, X_{2}, \ldots, X_{n}$ are singletons. Conversely, if $k$ is equal to $n$, then necessarily $C$ is a linear order. For a linear order $L$, we write also $L=x_{1} \succ x_{2} \succ \cdots \succ x_{n}$ to denote $L$. More generally, for a partially ordered relation, we write $x \succ y$ whenever $x$ is preferred to $y$ according to this relation. For any order $L=x_{1} \succ x_{2} \succ \cdots \succ x_{n}, \bar{L}$ will denote the order $x_{n} \succ x_{n-1} \succ \cdots \succ x_{1}$. The symbols $<, \leq,>$ and $\geq$ will be used for the usual inequalities between integers.

To define a median relation, we use the symmetric difference distance $\delta$ between two relations $R$ and $R^{\prime}$. This distance is defined by
$\delta\left(R, R^{\prime}\right)=\left|R \Delta R^{\prime}\right|$,
where $\Delta$ denotes the usual symmetric difference between sets. This distance, which has good axiomatic properties (see Barthélemy (1979) and Barthélemy and Monjardet (1981)), measures the number of disagreements between $R$ and $R^{\prime}$ :
$\delta\left(R, R^{\prime}\right)=\mid\left\{(x, y) \in X^{2}:\left[x R y\right.\right.$ and $\left.x \bar{R}^{\prime} y\right]$ or $\left[x \bar{R} y\right.$ and $\left.\left.x R^{\prime} y\right]\right\} \mid$.
Then we define a remoteness (Barthélemy and Monjardet, 1981) $\rho(\Pi, R)$ between the profile $\Pi=\left(R_{1}, R_{2}, \ldots, R_{m}\right)$ and a binary relation $R$ by
$\rho(\Pi, R)=\sum_{i=1}^{m} \delta\left(R_{i}, R\right)$.
So, the remoteness $\rho(\Pi, R)$ measures the total number of disagreements between $\Pi$ and $R$. Given a prescribed set $\mathcal{R}$ of relations (in the following, we will consider the cases $\mathcal{R}=\mathcal{L}, \mathcal{R}=$ $\mathcal{C}$ and $\mathcal{R}=\mathcal{W}$ ), an $\mathcal{R}$-median relation, or simply a median relation when there is no ambiguity, is a relation $R^{*}$ belonging to $\mathcal{R}$ and minimizing $\rho$ :
$\rho\left(\Pi, R^{*}\right)=\min _{R \in \mathcal{R}} \rho(\Pi, R)$.
We may state $\rho(\Pi, R)$ for any relation $R$ thanks to variables describing $R$. Let $r=\left(r_{x y}\right)_{(x, y) \in X^{2}}$ be the characteristic matrix associated with $R$, i.e. the matrix defined by: $r_{x y}=1$ if $x R y$ and $r_{x y}=0$ otherwise. Similarly, for $1 \leq i \leq m$, let $r_{x y}^{i}$ be equal to 1 if $x R_{i} y$ and to 0 otherwise; note the equalities $r_{x y}^{i}+r_{y x}^{i}=1$ for any distinct $x$ and $y$ when the relations $R_{i}$ are tournaments, and $r_{x x}^{i}=0$ for any $x$ when the relations $R_{i}$ are irreflexive. Because the quantities $r_{x y}^{i}$ and $r_{x y}$ are equal to 1 or 0 , we have

$$
\begin{aligned}
\delta\left(R_{i}, R\right) & =\sum_{(x, y) \in X^{2}}\left|r_{x y}^{i}-r_{x y}\right|=\sum_{(x, y) \in X^{2}}\left(r_{x y}^{i}-r_{x y}\right)^{2} \\
& =\sum_{(x, y) \in X^{2}} r_{x y}^{i}+\sum_{(x, y) \in X^{2}}\left(1-2 r_{x y}^{i}\right) \cdot r_{x y} .
\end{aligned}
$$

From this we obtain
$\rho(\Pi, R)=\sum_{i=1}^{m} \delta\left(R_{i}, R\right)=\lambda_{\Pi}+\sum_{(x, y) \in X^{2}} m_{x y}^{\Pi} \cdot r_{x y}$
where $\lambda_{\Pi}=\sum_{i=1}^{m} \sum_{(x, y) \in X^{2}} r_{x y}^{i}$ is a constant for any given profile $\Pi$ and with, for $(x, y) \in X^{2}, m_{x y}^{\Pi}=m-2 \sum_{i=1}^{m} r_{x y}^{i}$. When there is no ambiguity, we simply write $\lambda$ and $m_{x y}$ instead of $\lambda_{\Pi}$ and $m_{x y}^{\Pi}$. The quantity $m_{x y}$ may be interpreted as the difference between the total number $m$ of voters and twice the number of the voters who prefer $x$ to $y$. Moreover, when all the relations $R_{i}$ for $1 \leq i \leq m$ are tournaments, then note the relation $m_{y x}=-m_{x y}$ and the equality $\lambda=m n(n-1) / 2$ (both relations coming from the equality $r_{x y}^{i}+$ $r_{y x}^{i}=1$, true for any distinct $x$ and $y$, and from the equality $r_{x x}^{i}=0$, true for any $x$ ); then $m_{x y}$ can also be interpreted in this case as the difference between the number of voters who prefer $y$ to $x$ and the number of voters who prefer $x$ to $y$ (e.g., if all the voters prefer $x$ to $y, m_{y x}$ is equal to $m$ and $m_{x y}$ is equal to $-m$ ). In the following, the matrix $M_{\Pi}=\left(m_{x y}^{\Pi}\right)_{(x, y) \in X^{2}}$, or simply $M=\left(m_{x y}\right)_{(x, y) \in X^{2}}$ if there is no ambiguity, will be called the summarizing matrix of $\Pi$.

## 3. Links between median complete preorders or median weak orders and median linear orders for a profile of tournaments

In this section, all the relations $R_{i}(1 \leq i \leq m)$ belonging to the profile $\Pi$ are assumed to be tournaments (which includes the
particular case of linear orders). The next lemma shows that, when dealing with such a profile $\Pi$ of tournaments, the remoteness of any relation $R$ from $\Pi$ is the same if we consider the asymmetric part of $R$.

Lemma 1. Let $\Pi=\left(T_{1}, T_{2}, \ldots, T_{m}\right)$ be a profile of $m$ tournaments. Let $R$ be a binary relation and let $R^{a}$ be the asymmetric part of $R$. Then we have $\rho(\Pi, R)=\rho\left(\Pi, R^{a}\right)$.
Proof. We have (see above)
$\rho(\Pi, R)=\sum_{i=1}^{m} \delta\left(T_{i}, R\right)=\lambda+\sum_{(x, y) \in X^{2}} m_{x y} \cdot r_{x y}$
where $\left(r_{x y}\right)_{(x, y) \in X^{2}}$ is the characteristic matrix associated with $R$. We may split $\rho(\Pi, R)$ into two terms: one associated with the asymmetric part of $R$, i.e. $\rho\left(\Pi, R^{a}\right)$, and the other associated with the symmetric part of $R$. We obtain
$\rho(\Pi, R)=\lambda+\sum_{\substack{(x, y) \in X^{2} \\ \text { with } x R^{a_{y}}}} m_{x y} \cdot r_{x y}+\sum_{\substack{(x, y) \in X^{2} \\ \text { with } x R y \text { and } y R x}} m_{x y} \cdot r_{x y}$.
The quantity $\lambda+\sum_{\substack{(x, y)) \in x^{2} \\ \text { with } \times \mathrm{R}^{2} y}} m_{x y} \cdot r_{x y}$ is equal to $\rho\left(\Pi, R^{a}\right)$. On the other hand, let $x$ and $y$ be two elements of $X$ with $x R y$ and $y R x$. Then $r_{x y}$ is equal to 1 as well as $r_{x y}$. As we deal with a profile of tournaments, we have $m_{x y}=-m_{y x}$. So, $\sum_{\text {with }(x, y) \in x^{2}} m_{x y} \cdot r_{x y}$ is equal to 0 and we have $\rho(\Pi, R)=\rho\left(\Pi, R^{a}\right)$. $\quad \stackrel{\text { with x xy }}{\square}$

We may now state the following theorems (which generalize the result of Lemma 3 of Bartholdi III et al. (1989), dealing with the median complete preorders of a profile of linear orders), specifying some links between median complete preorders and median linear orders for profiles of linear orders, and more generally for profiles of tournaments.

Theorem 2. Let $\Pi$ be a profile of $m$ tournaments defined on $X$. Then there exists a median complete preorder which is a linear order.
Proof. Among the median complete preorders of $\Pi$, consider a median complete preorder $C$ of which the canonic partition $X_{1} \succ$ $X_{2} \succ \cdots \succ X_{k}$ contains a maximum number $k$ of classes. If $k$ is equal to $n, C$ is a linear order and we are done. So assume that $C$ is not a linear order: $k<n$. Necessarily, there exists an index $i$ with $\left|X_{i}\right| \geq 2$. Let $\alpha$ be any element of $X_{i}$ and consider the new complete preorder $C_{1}$ of which the canonic partition is $X_{1} \succ X_{2} \succ \cdots \succ$ $X_{i-1} \succ\{\alpha\} \succ X_{i}-\{\alpha\} \succ X_{i+1} \succ \cdots, \succ X_{k}$. In other words, we extract $\alpha$ from its current class $X_{i}$ with respect to $C$ and we create a new class for building $C_{1}$, reduced to $\alpha$, inserted just before what remains from $X_{i}$; note that what remains from $X_{i}$ is not empty and thus $C_{1}$ contains more classes than $C$.

With respect to the characteristic matrices $\left(c_{x y}\right)_{(x, y) \in X^{2}}$ of $C$ and $\left(c_{x y}^{1}\right)_{(x, y) \in X^{2}}$ of $C_{1}$, we have:

- for $x \in X_{i}-\{\alpha\}, c_{\alpha x}^{1}=c_{\alpha x}=1, c_{x \alpha}^{1}=0, c_{x \alpha}=1$,
- $c_{x y}^{1}=c_{x y}$ otherwise.

Then we obtain
$\rho\left(\Pi, C_{1}\right)=\lambda+\sum_{(x, y) \in X^{2}} m_{x y} \cdot c_{x y}^{1}=\rho(\Pi, C)-\sum_{x \in X_{i}-\{\alpha\}} m_{x \alpha}$.
Since the number of classes of $C_{1}$ is greater than that of $C$, then by definition of $C, C_{1}$ cannot be the median. Thus we have $\sum_{x \in X_{i}-\{\alpha\}} m_{x \alpha}<0$.

Now, consider the new complete preorder $C_{2}$ of which the canonic partition is $X_{1} \succ X_{2} \succ \cdots \succ X_{i-1} \succ X_{i}-\{\alpha\} \succ\{\alpha\} \succ$ $X_{i+1} \succ \cdots \succ X_{k}: \alpha$ is now extracted from its current class to be inserted, alone, after what remains from $X_{i}$. The characteristic matrix $\left(c_{x y}^{2}\right)_{(x, y) \in X^{2}}$ of $C_{2}$ is defined from the one of $C$ by:

- for $x \in X_{i}-\{\alpha\}, c_{x \alpha}^{2}=c_{x \alpha}=1, c_{\alpha x}^{2}=0, c_{\alpha x}=1$,
- $c_{x y}^{2}=c_{x y}$ otherwise.

Thus,
$\rho\left(\Pi, C_{2}\right)=\lambda+\sum_{(x, y) \in X^{2}} m_{x y} \cdot c_{x y}^{2}=\rho(\Pi, C)-\sum_{x \in X_{i}-\{\alpha\}} m_{\alpha x}$.
As we deal with a profile of tournaments, we have, for $x \in X_{i}-$ $\{\alpha\}, m_{x \alpha}=-m_{\alpha x}$. So
$\rho\left(\Pi, C_{2}\right)=\rho(\Pi, С)+\sum_{x \in X_{i}-\{\alpha\}} m_{x \alpha}<\rho(\Pi, С)$,
a contradiction with the fact that $C$ is a median complete preorder.
So the initial assumption was false, and $C$ is a linear order, which proves the statement of Theorem 2.

Theorem 3. Let $\Pi$ be a profile of $m$ tournaments defined on $X$ such that there is no tie: $\forall(x, y) \in X^{2}$ with $x \neq y, m_{x y} \neq 0$. Then all the median complete preorders are linear orders.
Proof. The proof of Theorem 2 shows that, for any median complete preorder $C$ of $\Pi$ which is not a linear order, the canonic partition $X_{1} \succ X_{2} \succ \cdots \succ X_{k}$ of $C$ (with $k<n$ ) is such that, if $X_{i}$ denotes a class with at least two elements, then for any element $\alpha$ of $X_{i}$, we have
$\sum_{\beta \in X_{i}-\{\alpha\}} m_{\alpha \beta}=\sum_{\beta \in X_{i}-\{\alpha\}} m_{\beta \alpha}=0$.
Assume that there exists a median complete preorder $C$ of $\Pi$ which is not a linear order and let $X_{1} \succ X_{2} \succ \cdots \succ X_{k}(k<n)$ still denote the canonic partition of $C$. Let $X_{i}$ be a class with at least two elements $x$ and $y$. Then consider the complete preorder $C_{1}$ of $\Pi$ of which the canonic partition is $X_{1} \succ X_{2} \succ \cdots \succ X_{i-1} \succ\{x\} \succ$ $X_{i}-\{x\} \succ X_{i+1}>\cdots>X_{k}$. As in the proof of Theorem 2, we have

$$
\begin{align*}
\rho\left(\Pi, C_{1}\right) & =\rho(\Pi, С)-\sum_{\beta \in X_{i}-\{x\}} m_{\beta x} \\
& =\rho(\Pi, С)-m_{x y}-\sum_{\beta \in X_{i}-\{x, y\}} m_{\beta x} . \tag{2}
\end{align*}
$$

By (1), we have $\sum_{\beta \in X_{i}-\{x\}} m_{\beta x}=0$ and thus $\rho\left(\Pi, C_{1}\right)=\rho(\Pi, C)$ : $C_{1}$ is also a median complete preorder of $\Pi$ with $X_{i}-\{x\}$ as one of its classes. By (2), we have also $\sum_{\beta \in X_{i}-\{x, y\}} m_{\beta x}=-m_{x y}$, while $m_{x y}$ is not equal to 0 by hypothesis. Thus $X_{i}-\{x, y\}$ is not empty and we may apply (1) to the class $X_{i}-\{x\}$ of $C_{1}$ with $\alpha=y$ : $\sum_{\beta \in X_{i}-\{x, y\}} m_{y \beta}=0$. On the other hand, the application of (1) to the class $X_{i}$ of $C$ with $\alpha=y$ gives $\sum_{\beta \in X_{i}-\{y\}} m_{y \beta}=0$. But $\sum_{\beta \in X_{i}-\{y\}} m_{y \beta}$ is also equal to $\sum_{\beta \in X_{i}-\{x, y\}} m_{y \beta}+m_{y x}$, involving the equality $m_{y x}=0$, a contradiction with the hypothesis of the statement of Theorem 3. The conclusion follows.

Theorem 3 applies in particular when we consider a profile of an odd number of tournaments (or linear orders). We then obtain Corollary 4 :

Corollary 4. Let $\Pi$ be a profile of $m$ tournaments defined on $X$ with $m$ odd. Then all the median complete preorders are linear orders.
Proof. Note that the parity of $m_{x y}$ is the same of that of $m$. If $m$ is odd, then the $m_{x y}$ 's cannot be equal to 0 and Theorem 3 yields the conclusion.

Thanks to Lemma 1, we can do the same for a profile of tournaments (or of course of linear orders) when we look for a median weak order:

## Theorem 5. Let $\Pi$ be a profile of $m$ tournaments defined on $X$.

1. There exists a median weak order which is a linear order.
2. If there is no tie, i.e. if we have: $\forall(x, y) \in X^{2}$ with $x \neq y, m_{x y} \neq 0$, then all the median weak orders are linear orders.
3. If $m$ is odd, all the median weak orders are linear orders.

Proof. Thanks to Lemma 1, the proofs of the three statements of Theorem 5 are similar to those of Theorem 2, Theorem 3 and Corollary 4 respectively.

## 4. Complexity results for profiles of linear orders

In this section, we pay attention to the complexity of the aggregation of linear orders into a median complete preorder or a median weak order (for the theory of NP-completeness, see for instance Barthélemy et al. (1996) or Garey and Johnson (1979)). We first improve the results of Hudry $(1989,2008)$ and Wakabayashi $(1986,1998)$ by specifying the minimum even value of $m$ such that this aggregation is an NP-hard problem.

To prove Theorem 8, we will use the following two results. The first one deals with the NP-completeness of the aggregation of an even number $m$ of linear orders with $m \geq 4$ into a linear order (see Dwork et al. (2001) and Biedl et al. (2009) where a small error in the proof of Dwork et al. (2001) is corrected; see also Charon and Hudry (2010)).

Theorem 6. For any even $m$ with $m \geq 4$, the following problem is $N P$-complete:
Name: aggregation of a profile of $m$ linear orders into a linear order (APLO-LO-even);
Instance: a profile $\Pi$ of $m$ linear orders defined on a finite set $X$; an integer $k$;
Question: does there exist a linear order $L$ with $\rho(\Pi, L) \leq k$ ?
Note that, for $m=2$, APLO-LO-even is polynomial (see Charon and Hudry (2010)).

The second result deals with Slater's problem (Slater, 1961) which is also NP-complete (see Alon (2006), Charbit et al. (2007), Conitzer (2006) and Hudry (2010)). This problem can be stated as follows (see Charon and Hudry (2010) for different possibilities for stating this problem and Hudry (2009) for a survey on the complexities of tournament solutions):

## Theorem 7. The following problem is NP-complete:

Name: Slater's problem (SP);
Instance: a tournament $T$ defined on a finite set $X$ with $n$ elements; an integer $h$;
Question: does there exist a linear order $L$ with $\delta(T, L) \leq h$ ?
We must be more precise for the sequel about the encoding of a tournament $T$ in any instance of SP. The tournament $T$ can be described by a matrix $\left(t_{x y}\right)_{(x, y) \in X^{2}}$ with $t_{x y}=1$ if we have $x T y$ and $t_{x y}=-1$ otherwise for $x \neq y$. Notice the relation $t_{x y}=-t_{y x}$ for any $x$ and $y$ with $x \neq y$. As we have $\sum_{t_{x y}=1} t_{x y}=n(n-1) / 2$ and $\sum_{t_{x y}=1} t_{x y}=\sum_{\substack{t_{x y=1}=1 \\ x>y}} t_{x y}+\sum_{\substack{t_{x y=1}=1 \\ x<y}} t_{x y}$, it is easy (i.e., polynomial) to find a numbering of the vertices of $T$ such that the number of entries $t_{x y}$ with $t_{x y}=1$ and $x<y$ is at most $n(n-1) / 4$. We adopt such a numbering to code $T$, what can be done in polynomial time.

We may now prove some complexity results. We begin by giving a result coping with the case when the median relation must be a complete preorder.

Theorem 8. The following problem is NP-complete:
Name: aggregation of a profile of linear orders into a complete preorder (APLO-CP);
Instance: a profile $\Pi$ of $m$ linear orders defined on a finite set $X$ with $n$ elements; an integer $k$;
Question: does there exist a complete preorder $C$ with $\rho(\Pi, C) \leq k$ ?
Moreover, this problem remains NP-hard for any fixed even value of $m$ with $m \geq 4$ or for odd value of $m$ large enough $(m \geq 1+n(n-$ 1) $/ 2$ is sufficient). It is polynomial for $m \in\{1,2\}$.

Proof. We prove Theorem 8 in several steps.

## 1. Belonging of APLO-CP to NP.

First, note that APLO-CP belongs to NP. Indeed, let $(\Pi, k)$ be any instance of APLO-CP admitting the answer "yes". To check that the answer is really "yes", imagine that we are given a structure $C$ which should be a complete preorder at a remoteness less than or equal to $k$ from $\Pi$. It is easy to check that $C$ is complete and transitive in $O\left(n^{3}\right)$ (it is even possible to do it with a lower complexity, but this does not matter here) and to compute $\rho(\Pi, C)$ in $O\left(m n^{2}\right)$, while the (binary) size of the instance is about $m n^{2}+$ $\log k$. Hence the polynomiality of the checking, and the belonging of APLO-CP to NP.
2. NP-completeness of APLO-CP when $m$ is even with $m \geq 4$.

Now, assume that $m$ is even with $m \geq 4$. Let $(\Pi, \bar{k})$ be any instance of APLO-LO-even with $m$ linear orders and consider it as an instance of APLO-CP. This transformation (the identity!) is obviously polynomial. Moreover, if $(\Pi, k)$ admits the answer "yes" for APLO-LO-even, then there exists a linear order $L$ with $\rho(\Pi, L) \leq k$. By considering $L$ as a complete preorder, which is always possible, we obtain that $(\Pi, k)$ admits the answer "yes" for APLO-CP. Conversely, if $(\Pi, k)$ admits the answer "yes" for APLOCP , there exists a complete preorder $C$ with $\rho(\Pi, C) \leq k$. From Theorem 2, we deduce that there exists also a linear order $L$ with $\rho(\Pi, L) \leq k$ : then $(\Pi, k)$ admits the answer "yes" for APLO-LO-even. So this transformation keeps the answer. Hence the NPcompleteness of APLO-CP for any even $m \geq 4$, because of that of APLO-LO-even for the same values of $m$.

## 3. Polynomiality of APLO-CP for $m=2$.

The polynomiality of APLO-CP for $m=2$ can be shown in a similar way, by considering any instance of APLO-CP as an instance of APLO-LO-even. The polynomiality of APLO-LO-even when there are only two linear orders in the profile allows concluding. Details are left to the reader.
4. NP-completeness of APLO-CP when $m$ is odd with $m=n(n-$ 1) $/ 2+1$ or $m=n(n-1) / 2+2$.

To prove that APLO-CP is NP-complete when $m$ has an odd value equal to $n(n-1) / 2+1$ or to $n(n-1) / 2+2$, we polynomially transform SP. Let $(T, h)$ be any instance of SP where $T$ is encoded by a matrix $\left(t_{x y}\right)_{(x, y) \in X^{2}}$ with, for $x \neq y, t_{x y}=1$ if we have $x T y$ and $t_{x y}=-1$ otherwise and such that the number of entries $t_{x y}$ with $t_{x y}=1$ and $x<y$ is at most $n(n-1) / 4$. Let $M_{T}=\left(m_{x y}^{T}\right)_{(x, y) \in X^{2}}$ be the summarizing matrix associated with $T$ : we have $m_{x y}^{T}=-t_{x y}$. We are going to polynomially build a profile $\Pi$ of $m$ linear orders such that the summarizing matrix $\left(m_{x y}^{I}\right)_{(x, y) \in X^{2}}$ satisfies the equality $m_{x y}^{\Pi}=m_{x y}^{T}$ for any $x$ and any $y$ with $x \neq y$ (note that McGarvey's result (Mcgarvey, 1953) showing that any unweighted tournament is the majority tournament of a profile of linear orders is of no help here, though it can be extended to weighted tournaments with even weights, since we consider weighted tournaments with odd weights; other constructions can be found for instance in Dwork et al. (2001) and Hudry (1989, 2008), but involving more linear orders; the construction in Bartholdi III et al. (1989) does not apply when $m$ is odd; the ones in Wakabayashi $(1986,1998)$ do not build profiles of linear orders but ones of generic binary relations). So, let $x$ and $y$ be two distinct vertices of $T$ with $t_{x y}=1$ and $x<y$. We build two linear orders $L_{x y}^{1}$ and $L_{x y}^{2}$ defined by
$L_{x y}^{1}=\alpha_{1} \succ \alpha_{2} \succ \ldots>\alpha_{n-2} \succ x \succ y \quad$ and
$L_{x y}^{2}=x \succ y \succ \alpha_{n-2} \succ \ldots \succ \alpha_{2} \succ \alpha_{1}$,
with $\left\{\alpha_{1}, \alpha_{2}, \ldots, \alpha_{n-2}\right\}=X-\{x, y\}$. Let $L_{0}$ be the linear order defined by $n \succ n-1 \succ \cdots \succ 2 \succ 1$. The profile $\Pi$ that we build to define the instance of APLO-CP contains $L_{0}$ and, for any pair $\{x, y\}$ with $t_{x y}=1$ and $x<y$, the orders $L_{x y}^{1}$ and $L_{x y}^{2}$. Note that $\Pi$
contains just now an odd number of linear orders which is at most $n(n-1) / 2+1$ since, in addition to $L_{0}$, there are at most $n(n-1) / 4$ pairs $\{x, y\}$ with $t_{x y}=1$ and $x<y$ and that each pair gives two linear orders. If $\Pi$ contains in fact less than $n(n-1) / 2+1$, then we add to $\Pi$ the pair $\left(L_{0}, \bar{L}_{0}\right)$ as many times as necessary to reach an odd number of linear orders equal to $n(n-1) / 2+1$ or to $n(n-$ $1) / 2+2$, depending on the parity of $n(n-1) / 2$. Moreover, we set $k=h+n(n-1)(m-1) / 2$. The instance of APLO-CP is then $(\Pi, k)$.

This transformation is polynomial with respect to the size of $(T, h)$, which is about $n^{2}+\log h$. Indeed, each linear order of $\Pi$ can be encoded by its adjacency matrix with $n^{2}$ bits. There are less than $n^{2}$ such linear orders. The size of $\Pi$ is thus less than $n^{4}$. Similarly, $k$ requires a number of bits which is about $\log (h+n(n-1)(m-1) / 2)$, i.e. about $\log \left(h+n^{4} / 4\right)$. Then the size of the whole instance $(\Pi, k)$ can be bounded by a polynomial in $n^{2}+\log h$.

Moreover, this transformation is such that, for any two distinct elements $x$ and $y$ of $X$, the quantities $m_{x y}^{\Pi}$ used to summarize the profile $\Pi$ are equal to $m_{x y}^{T}$. Indeed, consider the number of orders of $\Pi$ for which $x$ is preferred to $y$ with $x<y$. To compute this number, assume first that we have $t_{x y}=1$ (and hence $m_{x y}^{T}=-1$ ). Then $\Pi$ contains the two orders $L_{x y}^{1}$ and $L_{x y}^{2}$, which both prefer $x$ to $y$. The other orders of $\Pi$, except one occurrence of $L_{0}$, can be gathered by pairs such that one order prefers $x$ to $y$ while the other prefers $y$ to $x$. There are $(m-3) / 2$ such pairs. Moreover, $L_{0}$ prefers $y$ to $x$ since we have $x<y$. Thus, there are exactly $2+(m-3) / 2=(m+1) / 2$ voters who prefer $x$ to $y$ and $1+(m-3) / 2=(m-1) / 2$ voters who prefer $y$ to $x$. Consequently, $m_{x y}^{\Pi}$ is equal to -1 , i.e. also equal to $m_{x y}^{T}$. Assume now that we have $t_{x y}=-1$ (and hence $m_{x y}^{T}=1$ ) still with $x<y$. Then the two orders $L_{x y}^{1}$ and $L_{x y}^{2}$ do not belong to $\Pi$ (nor the orders $L_{y x}^{1}$ and $L_{y x}^{2}$ ). So we may gather all the orders of $\Pi$, except one occurrence of $L_{0}$, by pairs such that one order prefers $x$ to $y$ while the other prefers $y$ to $x$, and there are $(m-1) / 2$ such pairs. Moreover, $L_{0}$ still prefers $y$ to $x$. So there are exactly $(m-1) / 2$ voters who prefer $x$ to $y$ and $(m+1) / 2$ voters who prefer $y$ to $x$. This involves that $m_{x y}^{\Pi}$ is equal to 1, i.e. also equal to $m_{x y}^{T}$. In conclusion, whatever the value of $t_{x y}, m_{x y}^{\Pi}$ is equal to $m_{x y}^{T}$ for $x$ and $y$ with $x \neq y$, and thus for any distinct $x$ and $y$, because of the relations $m_{y x}^{T}=$ $-m_{x y}^{T}$ and $m_{y x}^{\Pi}=-m_{x y}^{\Pi}$ applicable for any $x$ and $y$ with $x \neq y$.

We may now prove easily that $(\Pi, k)$ admits the same answer as $(T, h)$. The computations done at the end of Section 2 specify the expression for the remoteness for any complete preorder $C$ of which the characteristic matrix is $\left(c_{x y}\right)_{(x, y) \in X^{2}}$ :

$$
\begin{aligned}
\rho(\Pi, C) & =\lambda_{\Pi}+\sum_{(x, y) \in X^{2}} m_{x y}^{\Pi} \cdot c_{x y} \\
& =\operatorname{mn}(n-1) / 2+\sum_{(x, y) \in X^{2}} m_{x y}^{\Pi} \cdot c_{x y}
\end{aligned}
$$

Because of the relation $m_{x y}^{\Pi}=m_{x y}^{T}$, we have also
$\rho(\Pi, С)=m n(n-1) / 2+\sum_{(x, y) \in X^{2}} m_{x y}^{T} \cdot c_{x y}$.
Assume that $(T, h)$ admits the answer "yes": there exists a linear order $L$ with $\delta(T, L) \leq h$. Let $\left(\omega_{x y}\right)_{(x, y) \in X^{2}}$ be the characteristic matrix of $L$ and $\left(r_{x y}\right)_{(x, y) \in X^{2}}$ be that of $T$. Then we have

$$
\begin{aligned}
\delta(T, L) & =\sum_{(x, y) \in X^{2}} r_{x y}+\sum_{(x, y) \in X^{2}}\left(1-2 r_{x y}\right) \cdot \omega_{x y} \\
& =n(n-1) / 2+\sum_{(x, y) \in X^{2}} m_{x y}^{T} \cdot \omega_{x y}
\end{aligned}
$$

As a linear order is a complete preorder, consider $L$ as a complete preorder. So the inequality $\delta(T, L) \leq h$ involves $\rho(\Pi, L) \leq$
$h+n(n-1)(m-1) / 2=k$. Consequently, $(\Pi, k)$ admits the answer "yes".

Conversely, assume that ( $\Pi, k$ ) admits the answer "yes". This means that there exists a complete preorder $C$ with $\rho(\Pi, C) \leq k$. By Theorem 2, as we deal with a profile of linear orders, there exists a linear order $L$ with $\rho(\Pi, L) \leq k$. With the same notation as above, we get

$$
\begin{aligned}
\delta(T, L) & =n(n-1) / 2+\sum_{(x, y) \in X^{2}} m_{x y}^{T} \cdot \omega_{x y} \\
& =\rho(\Pi, L)-n(n-1)(m-1) / 2
\end{aligned}
$$

from which we deduce
$\delta(T, L) \leq k-n(n-1)(m-1) / 2=h$.
So ( $T, h$ ) admits the answer "yes".
This completes step 4 of the proof.
5. NP-completeness of APLO-CP when $m$ is odd with $m>m=$ $n(n-1) / 2+2$

For larger values of $m$, we proceed by induction. Let $(\Pi, k)$ be an instance of APLO-CP such that $\Pi$ contains $m-2$ linear orders. Then we build a new profile $\Pi^{\prime}$ with $m$ linear orders by adding the pair $(L, \bar{L})$ to $\Pi$, where $L$ denotes any linear order. It is easy to check the following equality, for any complete preorder $C: \rho\left(\Pi^{\prime}, C\right)=$ $\rho(\Pi, C)+n(n-1)$. Then, for any integer $k$, the instance $(\Pi, k)$ admits the answer "yes" if and only if the instance ( $\Pi^{\prime}, k+n(n-1)$ ) admits the answer "yes". As adding two linear orders to a profile can be done in polynomial time, this transformation is polynomial and keeps the answer. Hence the complexity of APLO-CP when we consider a profile $\Pi^{\prime}$ with $m$ linear orders from the one of the same problem when considering profiles with $m-2$ linear orders, and then the complexity of APLO-CP when we consider a profile with an odd number $m$ of linear orders when $m$ is greater than or equal to $n(n-1) / 2+1$.
6. Polynomiality of APLO-CP for $m=1$

The polynomiality of APLO-CP for $m=1$ comes from the fact that any instance ( $\Pi, k$ ) of APLO-CP for $m=1$, with $\Pi=(L)$ and $L$ a linear order, admits the answer "yes". Indeed, as $L$ can be considered as a complete preorder, $L$ is a median complete preorder of $\Pi$ and the minimum remoteness is equal to 0 .

The same kind of proof applies to show that the aggregation of an odd number $m$ of linear orders into a median linear order is NPhard when $m$ is greater than or equal to $n(n-1) / 2+1$. We give below the sketch of another proof, based on Theorem 2.
Theorem 9. The following problem is NP-complete:
Name: aggregation of a profile of linear orders into a linear order (APLO-LO-odd);
Instance: a profile $\Pi$ of $m$ linear orders defined on a finite set $X$ with $n$ elements with $m$ odd and $m \geq 1+n(n-1) / 2$; an integer $k$;
Question: does there exist a linear order $L$ with $\rho(\Pi, L) \leq k$ ?
Proof. The proof is based on the NP-completeness of APLO-CP and on Theorem 2. Indeed, consider any instance of $\operatorname{APLO}-\mathrm{CP}(\Pi, k)$ with an odd number $m$ of linear orders, with $m \geq n(n-1) / 2+1$, as an instance of APLO-LO-odd. Then Theorem 2 shows that there exists a linear order $L$ with $\rho(\Pi, L) \leq k$ if and only if there exists a complete preorder $C$ with $\rho(\Pi, C) \leq k$. Details are left to the reader.

From the NP-completeness of APLO-CP, we may also deduce the NP-completeness of the similar problem when we require the median relation to be a weak order instead of a complete preorder. This will be a consequence of Lemma 1 :
Theorem 10. The following problem is NP-complete:
Name: aggregation of a profile of linear orders into a weak order (APLO-WO);
Instance: a profile $\Pi$ of $m$ linear orders defined on a finite set $X$ with $n$ elements; an integer $k$;

Question: does there exist a weak order $W$ with $\rho(\Pi, W) \leq k$ ?
Moreover, this problem remains NP-hard for any fixed even value of $m$ with $m \geq 4$ or for odd value of $m$ large enough ( $m \geq 1+n(n-1) / 2$ is sufficient). It is polynomial for $m \in\{1,2\}$.
Proof. This is a consequence of Lemma 1 and of the complexity of APLO-CP (Theorem 8). Indeed, if we consider any instance ( $\Pi, k$ ) of APLO-CP as an instance of APLO-WO (thus the transformation is the identity), Lemma 1 shows that ( $\Pi, k$ ) considered as an instance of APLO-CP admits the answer "yes" if and only if ( $\Pi, k$ ) considered as an instance of APLO-WO admits the answer "yes". Details (including the belonging of APLO-WO to NP) are left to the reader.

## 5. Extensions of Slater's problem: complexity results for profiles of tournaments

Stated as above, Slater's problem (Slater, 1961) corresponds to the case where the profile $\Pi$ contains only one tournament $T$ and we look for a linear order minimizing $\rho(\Pi, O)$, i.e. $\delta(T, O)$, over the set $\mathcal{L}$ of linear orders. This problem is NP-hard (see Theorem 7). We can wonder what happens if we relax some constraints in the required structure for the median relation. For instance, what is the complexity if, instead of a linear order, we look for a complete preorder or a weak order? Theorems 11 and 12 applied to $m$ equal to 1 show that these extensions of Slater's problem remain NPhard. More generally, this remains true for any set $\mathcal{R}$ including $\mathcal{T}$ (for instance the set $\mathcal{C}$ o of complete relations or the set $\mathcal{A}$ of antisymmetric relations).

Theorem 11. Let $m$ be any integer greater than or equal to 1 . Let $\mathcal{R}$ be any set of relations with $\mathcal{T} \subseteq \mathscr{R}$. The following problem is NPcomplete:
Name: aggregation of an $\mathcal{R}$-profile into a complete preorder (APT-CP); Instance: a profile $\Pi$ of $m$ relations belonging to $\mathcal{R}$; an integer $k$;
Question: does there exist a complete preorder $C$ with $\rho(\Pi, C) \leq k$ ?
Proof. As for the other aggregation problems considered here, the belonging of APT-CP to NP is easy and is left to the reader.

We transform Slater's problem SP into APT-CP as follows. Let ( $T, h$ ) be any instance of SP. We define an instance ( $\Pi, k$ ) of APTCP by duplicating $m$ times the tournament $T: \Pi=(T, T, \ldots, T)$, and by setting $k=m h$. Since $m$ is fixed, this transformation is polynomial. Let us check that it keeps the answer.

For this, assume that ( $T, h$ ) admits the answer "yes": there exists a linear order $L$ with $\delta(T, L) \leq h$. By considering $L$ as a complete preorder and because of the linearity of the remoteness, we get $\rho(\Pi, L)=m \cdot \delta(T, L) \leq m h=k$. Thus $(\Pi, k)$ admits the answer "yes".

Conversely, assume that ( $\Pi, k$ ) admits the answer "yes": there exists a complete preorder $C$ with $\rho(\Pi, C) \leq k$. Then, as $\Pi$ contains only tournaments, by Theorem 2, there exists a linear order $L$ with $\rho(\Pi, L) \leq k$. As we still have $\rho(\Pi, L)=m \cdot \delta(T, L)$, we get $\delta(T, L) \leq h$ and, thus, $(T, h)$ admits the answer "yes".

The transformation is polynomial and keeps the answer while APT-CP belongs to NP and SP is NP-complete. Hence the NPcompleteness of APT-CP.

A very similar proof, left to the reader, allows proving Theorem 12:

Theorem 12. Let $m$ be any integer greater than or equal to 1 . Let $\mathcal{R}$ be any set of relations with $\mathcal{T} \subseteq \mathcal{R}$. The following problem is NPcomplete:
Name: aggregation of an $\mathcal{R}$-profile into a weak order (APT-WO);
Instance: a profile $\Pi$ of $m$ relations belonging to $\mathcal{R}$; an integer $k$; Question: does there exist a weak order $W$ with $\rho(\Pi, W) \leq k$ ?

It is still easier to prove the following result, still obtained from Slater's problem SP by duplicating the tournament of any instance of SP:

Theorem 13. Let $m$ be any integer greater than or equal to 1 . Let $\mathcal{R}$ be any set of relations with $\mathcal{T} \subseteq \mathcal{R}$. The following problem is NPcomplete:
Name: aggregation of an $\mathcal{R}$-profile into a linear order (APT-LO); Instance: a profile $\Pi$ of $m$ relations belonging to $\mathcal{R}$; an integer $k$; Question: does there exist a linear order $L$ with $\rho(\Pi, L) \leq k$ ?

## 6. Complexity results for the computation of median relations of other kinds of profiles

We consider now the complexity of the aggregation of other kinds of relations. In this section, $\mathscr{R}$ will denote a set of relations of a given type (for instance, $\mathcal{R}$ could be the set $\mathcal{C}$ of complete preorders). The relations belonging to the profiles considered will all belong to $\mathcal{R}$; we shall say that we deal with an $\mathcal{R}$-profile. The subsections below are characterized by some subsets contained by $\mathcal{R}$. In all of these subsections, $\mathcal{R}$ will include at least $\mathcal{L}$. Thus, the results of Section 4 (the NP-completeness of the aggregation of $m$ linear orders when $m$ is even and greater than or equal to 4) can be obviously extended to $\mathcal{R}$-profiles when $\mathcal{R}$ contains $\mathcal{L}$ (it suffices to consider the identity for the polynomial transformation). The aim of Section 6 is to obtain, for fixed odd values of $m$, results similar to those of Section 4 applicable for fixed even values of $m$.

### 6.1. Median linear orders for $\mathcal{R}$ containing $\mathcal{C}, \mathcal{P}$ or $\mathcal{Q}$

In this subsection, we consider the case for which the profiles contain relations belonging to a set $\mathcal{R}$ including $\mathcal{L}$ and containing the "full" relation $F_{0}$ defined by: $\forall(x, y) \in X^{2}$ with $x \neq y, x F_{0} y$ ( $F_{0}$ is a complete preorder and a quasi-order); note that the characteristic matrix $\left(f_{x y}^{0}\right)_{(x, y) \in X^{2}}$ associated with $F_{0}$ is the matrix of which all the entries are equal to 1 except for $x$ equal to $y$, in which case the entry is equal to 0 . This is the case for instance for $\mathcal{R}=\mathcal{C}, \mathcal{R}=\mathcal{P}$ or $\mathcal{R}=\mathcal{Q}$, but also for the set of all the relations defined on $X$.

Theorem 14. Let $m$ be an integer greater than or equal to 4. Let $\mathcal{R}$ be any set of relations with $\mathcal{L} \subseteq \mathcal{R}$ and $F_{0} \in \mathcal{R}$. The following problem is NP-complete:
Name: aggregation of an $\mathcal{R}$-profile of $m$ relations into a linear order (APCP-LO);
Instance: a profile $\Pi$ of $m$ relations belonging to $\mathcal{R}$; an integer $k$;
Question: does there exist a linear order $L$ with $\rho(\Pi, L) \leq k$ ?
Proof. The belonging of APCP-LO to NP is easy and is left to the reader.

We transform the NP-complete problem APLO-LO-even with four orders into APCP-LO with $m$ relations belonging to $\mathcal{R}$. Let ( $\Pi=\left(L_{1}, L_{2}, L_{3}, L_{4}\right), h$ ) be any instance of APLO-LO-even with four linear orders. We build a new profile $\Pi^{\prime}$ by adding $m-4$ times $F_{0}$ to $\Pi$. Then we consider the instance ( $\Pi^{\prime}, k$ ) of APCP-LO with $k=h+(m-4) n(n-1) / 2$. Hence $\Pi^{\prime}$ contains $m$ relations belonging to $\mathscr{R}$. Since $m$ is fixed, the transformation is polynomial. Let us show that it keeps the answer.

Let $L$ be any linear order with $\left(\omega_{x y}\right)_{(x, y) \in X^{2}}$ as its characteristic matrix. Since the difference between $\Pi$ and $\Pi^{\prime}$ relies in the fact that $\Pi^{\prime}$ contains $m-4$ times the full relation $F_{0}$, it is easy to show the following equality:
$\rho\left(\Pi^{\prime}, L\right)=\rho(\Pi, L)+(m-4) n(n-1) / 2$.
Indeed, we may partition $\rho\left(\Pi^{\prime}, L\right)$ into the sum of the contributions of $L_{1}, L_{2}, L_{3}$ and $L_{4}$, which provides $\rho(\Pi, L)$, and of the
contributions of the $m-4$ relations $F_{0}$ (the amount of each copy of $F_{0}$ is equal to $n(n-1) / 2$ ). Because $k$ is equal to $h+(m-4) n(n-1) / 2$, we have $\rho\left(\Pi^{\prime}, L\right) \leq k$ if and only if we have $\rho(\Pi, L) \leq h$ and the answer is kept by the transformation.

The NP-completeness of APLO-LO-even with four orders involves that of APCP-LO for any $m$ greater than or equal to 4 .

### 6.2. Median linear orders for $\mathcal{R}$ containing $\ell, \ell, \mathcal{O}$ or $\mathcal{W}$

We consider now the case for which the profiles contain relations belonging to a set $\mathcal{R}$ including $\mathcal{L}$ and containing the "empty" relation $E_{0}$ defined by: $\forall(x, y) \in X^{2}, x \bar{E}_{0} y$; note that the characteristic matrix associated with $E_{0}$ is the matrix equal to 0 . This is the case for instance for $\mathcal{R}=\ell, \mathcal{R}=\ell, \mathcal{R}=\mathcal{O}$ or $\mathscr{R}=\mathcal{W}$, but also for the set of acyclic relations or, still, for the set of all the relations defined on $X$.

Theorem 15. Let $m$ be an integer greater than or equal to 4 . Let $\mathcal{R}$ be any set of relations with $\mathcal{L} \subseteq \mathscr{R}$ and $E_{0} \in \mathcal{R}$. The following problem is NP-complete:
Name: aggregation of an $\mathcal{R}$-profile into a linear order (APSO/WO-LO); Instance: a profile $\Pi$ of $m$ relations belonging to $\mathcal{R}$; an integer $k$; Question: does there exist a linear order $L$ with $\rho(\Pi, L) \leq k$ ?

Proof. The proof is quite similar to that of Theorem 14 and we do not detail all its steps. The main difference is in the relation which is added to the profile $\Pi$ of the transformed instance of APLO-LOeven. Instead of adding the full relation $F_{0}$ to $\Pi$, we add the empty relation $E_{0}$ to obtain the profile $\Pi^{\prime}$ appearing in the instance of APSO/WO-LO.

Thus, starting from any instance ( $\Pi=\left(L_{1}, L_{2}, L_{3}, L_{4}\right), h$ ) of APLO-LO-even with four linear orders, we build a new profile $\Pi^{\prime}$ by adding $m-4$ times $E_{0}$ to $\Pi$ and we set $k=h+(m-4) n(n-1) / 2$. As in the proof of Theorem 14, it is easy to show the following equality:
$\rho\left(\Pi^{\prime}, L\right)=\rho(\Pi, L)+(m-4) n(n-1) / 2$.
Indeed, we may still partition $\rho\left(\Pi^{\prime}, L\right)$ into the sum of the contributions of $L_{1}, L_{2}, L_{3}$ and $L_{4}$, which provides $\rho(\Pi, L)$, and of the contributions of the $m-4$ relations $E_{0}$ (the amount of each copy of $E_{0}$ is equal to $\left.n(n-1) / 2\right)$. Hence $\rho\left(\Pi^{\prime}, L\right) \leq k$ if and only if $\rho(\Pi, L) \leq h$.

### 6.3. Median complete preorders or weak orders for $\mathfrak{R}$ containing $\mathcal{L}$

In the following theorems, $\mathcal{R}$ denotes any set of relations defined on a same set $X$ containing all the linear orders defined on $X$; e.g., $\mathcal{R}$ can be the set $\mathcal{L}$ of linear orders on $X$, the set $\mathcal{C}$ of complete preorders on $X$, the set $\mathcal{T}$ of tournaments defined on $X$, and so on, including mixtures of such sets (e.g., $\mathcal{R}$ can be the set of complete preorders and tournaments on $X$ ). Observe that almost all the usual ordered structures include linear orders as special cases.

Theorem 16. Let $\mathcal{R}$ denote any set of relations with $\mathcal{L} \subseteq \mathcal{R}$. The following problem is NP-complete:
Name: aggregation of an $\mathcal{R}$-profile into a complete preorder (APR-CP); Instance: a profile $\Pi$ of $m$ relations defined on a finite set $X$ with $n$ elements and belonging to $\mathcal{R}$; an integer $k$;
Question: does there exist a complete preorder $C$ with $\rho(\Pi, C) \leq k$ ? Moreover, this problem remains NP-hard for any fixed even value of $m$ with $m \geq 4$ and for large enough odd values of $m(m \geq 1+n(n-1) / 2$ is sufficient).
Proof. This is a consequence of Theorem 8. Indeed, any instance of APLO-CP can be considered as an instance of APR-CP, since $\mathcal{R}$ is

| Set $\mathbb{R}$ for the profile | $m$ odd | $m$ even |
| :---: | :---: | :---: |
| antisymmetric relations $(\mathcal{A})$ | 1 (Th. 13) | $2(\mathrm{Th.13)}$ |
| any binary relations | 1 (Wakabayashi, 1986) | 2 (Wakabayashi, 1986) |
| complete preorders $(C)$ | 5 (Th. 14) | 4 (Th. 14) |
| complete relations $(C o)$ | 1 (Th. 13) | 2 (Th. 13) |
| interval orders $(I)$ | 5 (Th. 15) | 4 (Th. 15) |
| linear orders $(\mathcal{L})$ | $1+n(n-1) / 2$ or $2+n(n-1) / 2$ (Th. 9) | 4 (Dwork et al., 2001) |
| partial orders $(O)$ | 5 (Th. 15) | 4 (Th. 15) |
| preorders $(\mathcal{P})$ | 5 (Th. 14) | 4 (Th. 14) |
| quasi-orders $(Q)$ | 5 (Th. 14) | 4 (Th. 14) |
| semiorders $(S)$ | 5 (Th. 15) | 4 (Th. 15) |
| tournaments $(\mathcal{T})$ | 1 (Alon, 2006, Charbit et al., 2007, | 2 (Th. 13) |
|  | Conitzer, 2006) |  |
| weak orders $(\mathcal{W})$ | 5 (Th. 15) | 4 (Th. 15) |

Fig. 1. Minimum known values of $m$ for which the computation of a median linear order is NP-hard.

| Set $R$ for the profile | $m$ odd | $m$ even |
| :---: | :---: | :---: |
| antisymmetric relations ( $\mathcal{A}$ ) | 1 (Th. 11, 12) | 2 (Th. 11, 12) |
| any binary relations | 1 (Th. 11, 12) | 2 (Th. 11, 12) |
| complete preorders ( $C$ ) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 16, 18) | 4 (Th. 16, 18) |
| complete relations (Co) | 1 (Th. 11, 12) | 2 (Th. 11, 12) |
| interval orders (I) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 16,18$)$ | 4 (Th. 16, 18) |
| linear orders ( $\llcorner$ ) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 8,10$)$ | 4 (Th. 8, 10) |
| partial orders ( $O$ ) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 16, 18) | 4 (Th. 16, 18) |
| preorders ( $\mathcal{P}$ ) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 16, 18) | 4 (Th. 16, 18) |
| quasi-orders (Q) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 16, 18) | 4 (Th. 16, 18) |
| semiorders (S) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 16, 18) | 4 (Th. 16, 18) |
| tournaments ( $\mathcal{T}$ ) | 1 (Th. 11, 12) | 2 (Th. 11, 12) |
| weak orders (W) | $1+n(n-1) / 2$ or $2+n(n-1) / 2($ Th. 16,18$)$ | 4 (Th. 16, 18) |

Fig. 2. Minimum known values of $m$ for which the computation of a median complete preorder or a median weak order is NP-hard.
assumed to include $\mathcal{L}$. This transformation, i.e. the identity, is of course polynomial and keeps the answer. As it is easy to prove that APR-CP belongs to NP, the NP-completeness of APLO-CP involves that of APR-CP.

By choosing $\mathcal{R}$ equal to the set $C$ of complete preorders, we obtain the complexity of Kemeny's problem:

Corollary 17. The following problem is NP-complete:
Name: aggregation of a © -profile into a complete preorder (Kemeny's problem);
Instance: a profile $\Pi$ of $m$ complete preorders defined on a finite set $X$ with $n$ elements; an integer $k$;
Question: does there exist a complete preorder $C$ with $\rho(\Pi, C) \leq k$ ? Moreover, this problem remains NP-hard for any fixed even value of $m$ with $m \geq 4$ and for large enough odd values of $m(m \geq n(n-1) / 2+1$ is sufficient).

A similar result holds for the median which would be a weak order:

Theorem 18. Let $\mathcal{R}$ denote any set of relations with $L \subseteq \mathcal{R}$. The following problem is NP-complete:
Name: aggregation of an $\mathcal{R}$-profile into a weak order (APR-WO);
Instance: a profile $\Pi$ of $m$ relations defined on a finite set $X$ with $n$ elements and belonging to $\mathcal{R}$; an integer $k$;
Question: does there exist a weak order $W$ with $\rho(\Pi, W) \leq k$ ? Moreover, this problem remains NP-hard for any fixed even value of $m$ with $m \geq 4$ and for large enough odd values of $m(m \geq n(n-1) / 2+1$ is sufficient).

Proof. This is a consequence of Theorem 10. Indeed, any instance of APLO-WO can be considered as an instance of APR-WO. As for Theorem 16, this transformation is polynomial and keeps the answer. As it is easy to prove that APR-WO belongs to NP, the NPcompleteness of APLO-WO involves that of APR-WO.

## 7. Conclusion

We may summarize the results obtained in this paper.
First, it is shown (Section 3) that any profile of tournaments admits a median complete order and a median weak order which are linear orders. Moreover, if the number $m$ of voters is odd, then all the median complete orders and all the median weak orders are linear orders. This is not the case for any kind of median. For instance, an example on six candidates is given in Hudry (2008), for which no median partial order is a linear order.

The complexity results are summarized in Figs. 1 and 2. The first column specifies for both figures the set $\mathcal{R}$ to which the relations of the profile are assumed to belong. The second column of Fig. 1 gives the minimum known odd value of $m$ for which the computation of a median linear order is NP-hard, while the third column does the same for even values of $m$. Fig. 2 specifies the same results when computing a median complete preorder or a median weak order. The parentheses specify where the result can be found (with respect to the previous theorems or to the references below).

The results remain applicable if we require that the relations considered (the relations of the profiles or the median relations) are reflexive or irreflexive. They also remain applicable if we consider, for the relations of the profile, supersets of the sets involved in Figs. 1 and 2 or by mixing these sets (then we may adopt for $m$ the lowest of the values involved).

## Acknowledgments

I would like to thank Vincent Conitzer for the question he set me during the workshop "Computational Foundations of Social Choice" (Dagstuhl, Germany, March 2010). His question about the links between median linear orders and median complete preorders reactivated my interest for this issue, which already appears in Hudry (2008) (Remark 19). I would like to thank also the two referees for their valuable comments.

The research was supported by the ANR project "Computational Social Choice" no. ANR-09-BLAN-0305.

## References

Ailon, N., 2010. Aggregation of partial rankings, p-ratings and top-m lists. Algorithmica 57, 284-300.
Alon, N., 2006. Ranking tournaments. SIAM Journal on Discrete Mathematics 20 (1), 137-142.
Arrow, K.J., 1951. Social Choice and Individual Values. Wiley, New York, see rather the revised edition of 1963.
Barthélemy, J.-P., 1979. Caractérisations axiomatiques de la distance de la différence symétrique entre des relations binaires. Mathématiques et Sciences Humaines 67, 85-113.
Barthélemy, J.-P., Monjardet, B., 1981. The median procedure in cluster analysis and social choice theory. Mathematical Social Sciences 1, 235-267.
Barthélemy, J.-P., Cohen, G., Lobstein, A., 1996. Algorithmic Complexity and Communication Problems. UCL Press, London.
Bartholdi III, J.J., Tovey, C.A., Trick, M.A., 1989. Voting schemes for which it can be difficult to tell who won the election. Social Choice and Welfare 6, 157-165.
Biedl, T., Brandenburg, F.J., Deng, X., 2009. On the complexity of crossings in permutations. Discrete Mathematics 309 (7), 1813-1823.
Black, D., 1958. The Theory of Committees and Elections. Cambridge University Press, Cambridge, Massachusetts.
Bouyssou, D., Marchant, T., Pirlot, M., Tsoukias, A., Vincke, P., 2006. Evaluation and Decision Models with Multiple Criteria. Springer.
Caspard, N., Monjardet, B., Leclerc, B., 2007. Ensembles Ordonnés Finis: Concepts, Résultats et Usages. Springer.
Charbit, P., Thomassé, S., Yeo, A., 2007. The minimum feedback arc set problem is NP-hard for tournaments. Combinatorics, Probability and Computing 16 (1), 1-4.
Charon, I., Hudry, O., 2010. An updated survey on the linear ordering problem for weighted or unweighted tournaments. Annals of Operations Research 175, 107-158.
Condorcet, M.J.A.N., Caritat (marquis de), 1785. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Imprimerie royale, Paris.

Conitzer, V., 2006. Computing Slater rankings using similarities among candidates, in: Proceedings of the 21st National Conference on Artificial Intelligence AAAI06, Boston, MA, USA, pp. 613-619.
Dwork, C., Kumar, R., Naor, M., Sivakumar, D., 2001. Rank aggregation methods for the Web, in: Proceedings of the 10th International Conference on World Wide Web WWW10, pp. 613-622. See also, by the same authors, Rank aggregation revisited, manuscript.
Garey, M.R., Johnson, D.S., 1979. Computers and Intractability, A Guide to the Theory of NP-completeness. Freeman, New York.
Guilbaud, G.Th., 1952. Les théories de l'intérêt général et le problème logique de l'agrégation. Économie appliquée 5 (4), 501-584; Éléments de la théorie des jeux. Dunod, Paris, 1968; English translation: Theories of the general interest and the logical problem of aggregation. Electronic Journal for History of Probability and Statistics 4, 2008.
Hemaspaandra, E., Spakowski, H., Vogel, J., 2005. The complexity of Kemeny elections. Theoretical Computer Science 349, 382-391.
Hudry, O., 1989. Recherche d'ordres médians: complexité, algorithmique et problèmes combinatoires, Ph.D. Thesis, Télécom ParisTech, Paris.
Hudry, O., 2008. NP-hardness results on the aggregation of linear orders into median orders. Annals of Operations Research 163 (1), 63-88.
Hudry, O., 2009. A survey on the complexity of tournament solutions. Mathematical Social Sciences 57, 292-303.
Hudry, O., 2010. On the complexity of Slater's problems. European Journal of Operational Research 203, 216-221.
Kemeny, J.G., 1959. Mathematics without numbers. Daedalus 88, 577-591.
McGarvey, D., 1953. A theorem on the construction of voting paradoxes. Econometrica 21, 608-610.
McLean, I., 1995. The first golden age of social choice, 1784-1803. In: Barnett, W.A., Moulin, H., Salles, M., Schofield, N.J. (Eds.), Social Choice, Welfare, and Ethics: Proceedings of the Eighth International Symposium in Economic Theory and Econometrics. Cambridge University Press, Cambridge, pp. 13-33.
McLean, I., Lorrey, H., Colomer, J.M., 2008. Social choice in medieval Europe. Electronic Journal for History of Probability and Statistics 4 (1).
McLean, I., Urken, A., 1995. Classics of Social Choice. University of Michigan Press, Ann Arbor, Michigan.
Monjardet, B., 1991. Éléments pour une histoire de la médiane métrique, moyenne, milieu, centre. Histoire et usages, Coll. histoire des sciences et techniques $\mathrm{n}^{\circ} 5$ Editions, de l'EHESS, Paris.
Monjardet, B., 2008. "Mathématique sociale" and mathematics. A case study: Condorcet's effect and medians. Electronic Journal for History of Probability and Statistics 4 (1).
Slater, P., 1961. Inconsistencies in a schedule of paired comparisons. Biometrika 48, 303-312.
Wakabayashi, Y., 1986. Aggregation of binary relations: algorithmic and polyhedral investigations, Ph.D. Thesis, Augsbourg.
Wakabayashi, Y., 1998. The complexity of computing medians of relations. Resenhas 3 (3), 323-349.
Young, H.P., 1988. Condorcet theory of voting. American Political Science Review 82, 1231-1244.


[^0]:    E-mail addresses: hudry@enst.fr, hudry@infres.enst.fr.

