

Beyond Pairwise Comparisons in Social Choice: A Setwise Kemeny Aggregation Problem

Hugo Gilbert

Gran Sasso Science Institute
67100 L'Aquila, Italy
hugo.gilbert@gssi.it

Tom Portoleau

LAAS-CNRS, IRIT-CNRS
Université de Toulouse
31400 Toulouse, France
tom.portoleau@laas.fr

Olivier Spanjaard

Sorbonne Université
CNRS, LIP6,
75005 Paris, France
olivier.spanjaard@lip6.fr

Abstract

In this paper, we advocate the use of setwise contests for aggregating a set of input rankings into an output ranking. We propose a generalization of the Kemeny rule where one minimizes the number of k -wise disagreements instead of pairwise disagreements (one counts 1 disagreement each time the top choice in a subset of alternatives of cardinality at most k differs between an input ranking and the output ranking). After an algorithmic study of this k -wise Kemeny aggregation problem, we introduce a k -wise counterpart of the majority graph. It reveals useful to divide the aggregation problem into several sub-problems. We conclude with numerical tests.

1 Introduction

Rank aggregation aims at producing a single ranking from a collection of rankings of a fixed set of alternatives. In social choice theory (e.g., Moulin 1991), where the alternatives are candidates to an election and each ranking represents the preferences of a voter, aggregation rules are called *Social Welfare Functions* (SWFs). Apart from social choice, rank aggregation has proved useful in many applications, including preference learning (Cheng and Hüllermeier 2009; Cléménçon, Korba, and Sibony 2018), collaborative filtering (Wang et al. 2014), genetic map creation (Jackson, Schnable, and Aluru 2008), similarity search in databases systems (Fagin, Kumar, and Sivakumar 2003) and design of web search engines (Altman and Tennenholtz 2008; Dwork et al. 2001). In the following, we use interchangeably the terms “input rankings” and “preferences”, “output ranking” and “consensus ranking”, as well as “alternatives” and “candidates”.

The well-known Arrow’s impossibility theorem states that there exists no aggregation rule satisfying a small set of desirable properties (Arrow 1950). In the absence of an “ideal” rule, various aggregation rules have been proposed and studied. Following Fishburn’s classification (1977), we can distinguish between the SWFs for which the output ranking can be computed from the *majority graph* alone, those for which the output ranking can be computed from the *weighted ma-*

majority graph alone, and all other SWFs¹. The majority graph is obtained from the input rankings by defining one vertex per alternative c and by adding an edge from c to c' if c is preferred to c' in a strict majority of input rankings. In the weighted majority graph, each edge is weighted by the majority margin. The many SWFs that rely on these graphs alone take therefore only pairwise comparisons into account to determine an output ranking. For a compendium of these SWFs, we refer to the book by Brandt et al. (2016).

The importance of this class of SWFs can be explained by their connection with the *Condorcet consistency* property, stating: if there is a Condorcet winner (i.e., an alternative with outgoing edges to every other ones in the majority graph), then it should be ranked first in the output ranking. Nevertheless, as shown by Baldiga and Green (2013), the lack of Condorcet consistency is not necessarily a bad thing, because this property may come into contradiction with the objective of maximizing voters’ agreement with the output ranking. The following example illustrates this point.

Example 1 (Baldiga and Green, 2013). *Consider an election with 100 voters and 3 candidates c_1, c_2, c_3 , where 49 voters have preferences $c_1 \succ c_2 \succ c_3$, 48 have preferences $c_3 \succ c_2 \succ c_1$ and 3 have preferences $c_2 \succ c_3 \succ c_1$. Candidate c_2 is the Condorcet winner, but is the top choice of only 3 voters. In contrast, candidate c_1 is in slight minority against c_2 and c_3 , but c_1 is the top choice of 49 voters. This massive gain in agreement may justify to put c_1 instead of c_2 in first position of the output ranking.*

Following Baldiga and Green (2013), we propose to handle this tension between the pairwise comparisons (leading to ranking c_2 first) and the plurality choice (leading to ranking c_1 first) by using SWFs that take into account not only pairwise comparisons but *setwise* contests. More precisely, given input rankings on a set C of candidates and $k \in \{2, \dots, |C|\}$, the idea is to consider the plurality score of each candidate c for each subset $S \subseteq C$ such that $2 \leq |S| \leq k$, where the plurality score of c for S is the number of voters for which c is the top choice in S . The results of setwise con-

¹Fishburn’s classification actually applies to social choice functions, which prescribe a subset of winning alternatives from a collection of rankings, but the extension to SWFs is straightforward.

Table 1: Results of setwise contests in Example 1.

set	c_1	c_2	c_3
$\{c_1, c_2\}$	49	51	–
$\{c_1, c_3\}$	49	–	51
$\{c_2, c_3\}$	–	52	48
$\{c_1, c_2, c_3\}$	49	3	48

tests for the preferences of Example 1 are given in Table 1 for $k=3$. Note that the three top rows obviously encode the same information as the weighted majority graph while the bottom row makes it possible to detect the tension between the pairwise comparisons and the plurality choice.

One can then define a new class of SWFs, those that rely on the results of setwise contests alone to determine an output ranking. The many works that have been carried out regarding voting rules based on the (weighted) majority graph can be revisited in this broader setting. This line of research has already been investigated by Lu and Boutilier (2010) and Baldiga and Green (2013). However, note that both of these works consider a setting where candidates may become unavailable after voters express their preferences. We do not make this assumption. We indeed believe that this new class of SWFs makes sense in the standard setting where the set of candidates is known and deterministic, as it amounts to generate an output ranking by examining the choices that are made by the voters on subsets of candidates of various sizes (while usually only pairwise choices are considered).

A natural SWF in this class consists in determining an output ranking that minimizes the number of disagreements with the results of setwise contests for sets of cardinality at most k . This is a k -wise generalization of the Kemeny rule, obtained as a special case for $k=2$. We recall that the Kemeny rule consists in producing a ranking that minimizes the number of *pairwise* disagreements (Kemeny 1959).

Example 2. *Let us come back to Example 1 and assume that we use the 3-wise Kemeny rule. Consider the output ranking $r = c_1 \succ c_2 \succ c_3$. For set $S = \{c_1, c_2\}$, the number of disagreements with the results of setwise contests is 51 because c_2 is the top choice in S for 51 voters (see Table 1) while it is c_1 for r . Similarly, the number of disagreements induced by $\{c_1, c_3\}$, $\{c_2, c_3\}$ and $\{c_1, c_2, c_3\}$ are respectively 51, 48 and $3+48$. The total number of disagreements is thus $51+51+48+3+48=201$. This is actually the minimum number of disagreements that can be achieved for these input rankings, which makes r the k -wise Kemeny ranking.*

The purpose of this paper is to study the k -wise Kemeny aggregation problem. Section 2 formally defines the problem and reports on related work. Section 3 is devoted to some axiomatic considerations of the corresponding voting rule, and to an algorithmic study of the problem. We then investigate a k -wise variant of the majority graph in Section 4. We prove that determining this graph is easy for $k=3$ but becomes NP-hard for $k>3$, and we show how to use it in a preprocessing step to speed up the computation of the output ranking. Numerical tests are presented in Section 5. Due to lack of space, missing proofs can be found in the long version of the paper (Gilbert, Portoleau, and Spanjaard 2019).

2 Preliminaries

Adopting the terminology of social choice theory, we consider an election with a set V of n voters and a set C of m candidates. Each voter v has a complete and transitive preference order r_v over candidates (also called ranking). The collection of these rankings defines a preference profile \mathcal{P} .

Notations and Definitions

Let us introduce some notations related to rankings. We denote by $\mathcal{R}(C)$ the set of $m!$ rankings over C . Given a ranking r and two candidates c and c' , we write $c \succ_r c'$ if c is in a higher position than c' in r . Given a ranking r and a candidate c , $\text{rk}(c, r)$ denotes the rank of c in r . For instance, $\text{rk}(c, r_v) = 1$ if c is the preferred candidate of voter v (the candidate ranked highest in r_v). Given a ranking r and a set $S \subseteq C$, we define r_S as the restriction of r to S and $t_r(S)$ as the top choice (i.e., preferred candidate) in S according to r . Similarly, given a preference profile \mathcal{P} and a set $S \subseteq C$, we define \mathcal{P}_S as the restriction of \mathcal{P} to S . Lastly, we denote by $\text{tail}_k(r)$ (resp. $\text{head}_k(r)$) the subranking compounded of the k least (resp. most) preferred candidates in r .

We are interested in SWFs which, given a preference profile \mathcal{P} , should return a consensus ranking which yields a suitable compromise between the preferences in \mathcal{P} . One of the most well-known SWFs is the *Kemeny rule*, which selects a ranking r with minimal Kendall tau distance to \mathcal{P} . Denoting by $\delta_{\text{KT}}(r, r')$ the Kendall tau distance between rankings r and r' , the distance $\delta_{\text{KT}}(r, \mathcal{P})$ between a ranking r and a profile \mathcal{P} reads as:

$$\delta_{\text{KT}}(r, \mathcal{P}) = \sum_{r' \in \mathcal{P}} \delta_{\text{KT}}(r, r')$$

$$\text{where } \delta_{\text{KT}}(r, r') = \sum_{\{c, c'\} \subseteq C} \mathbb{1}_{t_r(\{c, c'\}) \neq t_{r'}(\{c, c'\})}$$

Stated differently, δ_{KT} measures the distance between two rankings by the number of pairwise disagreements between them. The distance between a ranking and a preference profile is then obtained by summation.

However, the Kendall tau distance only takes into account pairwise comparisons, which may entail counterintuitive results as illustrated by Example 1. To address this issue, the Kendall tau distance can be generalized to take into consideration disagreements on sets of cardinal greater than two. Given a set $S \subseteq C$ and $t \leq m$, we denote by $\Delta^t(S)$ the set of subsets of S of cardinal lower than or equal to t , i.e., $\Delta^t(S) = \{S' \subseteq S \text{ s.t. } |S'| \leq t\}$. When S is not specified, it is assumed to be C , i.e., $\Delta^t = \Delta^t(C)$. For $k \geq 2$, the k -wise Kendall tau distance δ_{KT}^k between r and r' is defined by:

$$\delta_{\text{KT}}^k(r, r') = \sum_{S \in \Delta^k} \mathbb{1}_{t_r(S) \neq t_{r'}(S)}$$

In other words, δ_{KT}^k measures the distance between two rankings by the number of top-choice disagreements on sets of cardinal lower than or equal to k .

Note that δ_{KT}^k has all the properties of a distance: non-negativity, identity of indiscernibles, symmetry and triangle inequality. Secondly, as mentioned in the introduction, we

have $\delta_{KT}^2 = \delta_{KT}$. Thirdly and maybe most importantly, we point out that the distances induced by $\delta_{KT}^k(r, r')$ can be computed in $O(m^3)$ by using the following formula:

$$\begin{aligned} \delta_{KT}^k(r, r') &= \sum_{\{c, c'\} \subseteq C} \mathbb{1}_{c \succ_r c'} \mathbb{1}_{c' \succ_{r'} c} |\Delta^{k-2}(B_c(r) \cap B_{c'}(r'))| \\ &= \sum_{\{c, c'\} \subseteq C} \mathbb{1}_{c \succ_r c'} \mathbb{1}_{c' \succ_{r'} c} \sum_{i=0}^{k-2} \binom{|B_c(r) \cap B_{c'}(r')|}{i} \end{aligned} \quad (1)$$

where $B_c(r) = \{x \in C \text{ s.t. } c \succ_r x\}$ is the set of candidates that are ranked below c in r . Let us give some intuition for this formula. For any pair $\{c, c'\}$ of candidates such that $c \succ_r c'$ and $c' \succ_{r'} c$, we count the number of sets in Δ^k on which there is a disagreement because the top choice is c for r while it is c' for r' . Such sets are of the form $S \cup \{c, c'\}$, where $S \in \Delta^{k-2}(B_c(r) \cap B_{c'}(r'))$, otherwise c and c' would not be the top choices. Hence the formula.

The distance δ_{KT}^k induces a new SWF, the k -wise Kemeny rule, which, given a profile \mathcal{P} , returns a ranking r with minimal distance δ_{KT}^k to \mathcal{P} , where:

$$\delta_{KT}^k(r, \mathcal{P}) = \sum_{r' \in \mathcal{P}} \delta_{KT}^k(r, r')$$

Note that this coincides with the rule we used in the introduction, by commutativity of addition:

$$\sum_{r' \in \mathcal{P}} \sum_{S \in \Delta^k} \mathbb{1}_{t_r(S) \neq t_{r'}(S)} = \sum_{S \in \Delta^k} \sum_{r' \in \mathcal{P}} \mathbb{1}_{t_r(S) \neq t_{r'}(S)}$$

Determining a consensus ranking for this rule defines the k -wise Kemeny Aggregation Problem (k -KAP for short).

k-WISE KEMENY AGGREGATION PROBLEM

INSTANCE: A profile \mathcal{P} with n voters and m candidates.

SOLUTION: A ranking r of the m candidates.

MEASURE: $\delta_{KT}^k(r, \mathcal{P})$ to minimize.

Related Work

Several other variants of the Kemeny rule have been proposed in the literature, either to obtain generalizations able to deal with partial or weak orders (Dwork et al. 2001; Zwicker 2018), to penalize more some pairwise disagreements than others (Kumar and Vassilvitskii 2010), or to account for candidates that may become unavailable after voters express their preferences (Baldiga and Green 2013; Lu and Boutilier 2010).

Indeed, despite its popularity, the Kemeny rule has received several criticisms. One of them is that the Kendall tau distance counts equally the disagreements on every pair of candidates. This property is undesirable in many settings. For instance, with a web search engine, a disagreement on a pair of web pages with high positions in the considered rankings should have a higher cost than a disagreement on pairs of web pages with lower ones. This drawback motivated the introduction of weighted Kendall tau distances by Kumar and Vassilvitskii (2010). A more thorough comparison between our work and theirs can be found in the extended version of the paper (Gilbert, Portoleau, and Spanjaard 2019).

Let us illustrate with the following example, that the k -wise Kendall tau distance is also well suited to penalize more the disagreements involving alternatives at the top of the input rankings.

Example 3. Consider rankings r_1, r_2, r_3 defined by $c_1 \succ_{r_1} c_2 \succ_{r_1} c_3$, $c_1 \succ_{r_2} c_3 \succ_{r_2} c_2$, and $c_2 \succ_{r_3} c_1 \succ_{r_3} c_3$. We have $\delta_{KT}(r_1, r_2) = \delta_{KT}(r_1, r_3) = 1$ while $\delta_{KT}^3(r_1, r_2) = 1 < 2 = \delta_{KT}^3(r_1, r_3)$ because r_1 and r_3 disagree on both subsets $\{c_1, c_2\}$ and $\{c_1, c_2, c_3\}$. Put another way, $\delta_{KT}^3(r_1, r_3) > \delta_{KT}^3(r_1, r_2)$ because r_1 and r_3 disagree on their top-ranked alternatives whereas r_1 and r_2 disagree on the alternatives ranked in the last places.

The two works closest to ours are related to another extension of the Kemeny rule. This extension considers a setting in which, besides the fact that voters have preferences over a set C , the election will in fact occur on a subset $S \subseteq C$ drawn according to a probability distribution (Baldiga and Green 2013; Lu and Boutilier 2010). The optimization problem considered is then to find a consensus ranking r which minimizes, in expectation, the number of voters' disagreements with the chosen candidate in S (a voter v disagrees if $t_{r_v}(S) \neq t_r(S)$). The differences between the work of Baldiga and Green (2013) and the one of Lu and Boutilier (2010) is then twofold. Firstly, while Baldiga and Green mostly focused on the axiomatic properties of this aggregation procedure, the work of Lu and Boutilier has more of an algorithmic flavor. Secondly, while Baldiga and Green mostly study a setting in which the probability $\mathbb{P}(S)$ of S is only dependent on its cardinality (i.e., $\mathbb{P}(S)$ is only a function of $|S|$), Lu and Boutilier study a setting that can be viewed as a special case of the former, where each candidate is absent of S independently of the others with a probability p (i.e., $\mathbb{P}(S) = p^{|C \setminus S|} (1-p)^{|S|}$). The Kemeny aggregation problem can be formulated in both settings, either by defining $\mathbb{P}(S) = 0$ for $|S| \geq 3$, or by defining a probability p that is "sufficiently high" w.r.t. the size of the instance (Lu and Boutilier 2010). Lu and Boutilier conjectured that the determination of a consensus ranking is NP-hard in their setting, designed an exact method based on mathematical programming, two approximation greedy algorithms and a PTAS.

Our model can be seen as a special case of the model of Baldiga and Green where the set S is drawn uniformly at random within the set of subsets of C of cardinal smaller than or equal to a given constant $k \geq 2$. While it cannot be casted in the specific setting studied by Lu and Boutilier, our model is closely related and may be used to obtain new insights on their work.

3 Aggregation with the k -wise Kemeny Rule

In this section, we investigate the axiomatic properties of the k -wise Kemeny rule, and then we turn to the algorithmic study of k -KAP.

Axiomatic Properties of the k -wise Kemeny Rule

Several properties of the k -wise Kemeny rule have already been studied by Baldiga and Green (2013), because their setting includes the k -wise Kemeny rule as a special case.

Among other things, they showed that the rule is not *Condorcet consistent*. That is to say, a Condorcet winner may not be ranked first in any consensus ranking even when one exists, as illustrated by Example 2.

The authors also show that the k -wise Kemeny rule is neutral, i.e., all candidates are treated equally, and that for $k \geq 3$ it is different from any positional method or any method that uses only the pairwise majority margins (among which is the standard Kemeny rule). We provide here some additional properties satisfied by the k -wise Kemeny rule:

- *Monotonicity*: up-ranking cannot harm a winner; down-ranking cannot enable a loser to win.
- *Unanimity*: if all voters rank c before c' , then c is ranked before c' in any consensus ranking.
- *Reinforcement*: let $\mathcal{R}_{\mathcal{P}}^*$ and $\mathcal{R}_{\mathcal{P}'}$ denote the sets of consensus rankings for preference profiles \mathcal{P} and \mathcal{P}' respectively. If $\mathcal{R}_{\mathcal{P}}^* \cap \mathcal{R}_{\mathcal{P}'}^* \neq \emptyset$ and \mathcal{P}'' is the profile obtained by concatenating \mathcal{P} and \mathcal{P}' , then $\mathcal{R}_{\mathcal{P}''}^* = \mathcal{R}_{\mathcal{P}}^* \cap \mathcal{R}_{\mathcal{P}'}^*$.

Besides, the k -wise Kemeny rule does not satisfy *Independence of irrelevant alternatives*, i.e., the relative positions of two candidates in a consensus ranking can depend on the presence of other candidates. Let us illustrate this point with the following example.

Example 4. *Considering the preference profile from Example 1, the only consensus ranking for δ_{KT}^3 is $c_1 \succ c_2 \succ c_3$. Yet, without c_3 the only consensus ranking would be $c_2 \succ c_1$.*

Lastly, note that there exists a noise model such that the k -wise Kemeny rule can be interpreted as a maximum likelihood estimator (Conitzer, Rognlie, and Xia 2009). In this view of voting, one assumes that there exists a “correct” ranking r , and each vote corresponds to a noisy perception of this correct ranking. Consider the conditional probability measure \mathbb{P} on $\mathcal{R}(C)$ defined by $\mathbb{P}(r'|r) \propto e^{-\delta_{\text{KT}}^k(r,r')}$. It is easy to convince oneself that the k -wise Kemeny rule returns a ranking r^* that maximizes $\mathbb{P}(\mathcal{P}|r^*) = \prod_{r' \in \mathcal{P}} \mathbb{P}(r'|r^*)$ and is thus a maximum likelihood estimate of r .

Computational Complexity of k -KAP

We now turn to the algorithmic study of k -KAP. After providing a hardness result, we will design an efficient Fixed Parameter Tractable (FPT) algorithm for parameter m .

While k -KAP is obviously NP-hard for $k = 2$ as it then corresponds to determining a consensus ranking w.r.t. the Kemeny rule, we strengthen this result by showing that it is also NP-hard for *any* constant value $k \geq 3$. The proof, uses a reduction from 2-KAP.

Theorem 1. *For any constant $k \geq 3$, k -KAP is NP-hard, even if the number of voters equals 4 or if the average range of candidates equals 2 (where the range of a candidate c is defined by $\max_{r \in \mathcal{P}} \text{rk}(c, r) - \min_{r \in \mathcal{P}} \text{rk}(c, r) + 1$ and the average is taken over all candidates).*

Despite this result, k -KAP is obviously FPT w.r.t. the number m of candidates, by simply trying the $m!$ rankings in $\mathcal{R}(C)$. We now design a dynamic programming procedure which significantly improves this time complexity.

Proposition 1. *If r^* is an optimal ranking for k -KAP, then $\delta_{\text{KT}}^k(r^*, \mathcal{P}) = d_{\text{KT}}^k(C)$, where, for any subset $S \subseteq C$, $d_{\text{KT}}^k(S)$ is defined by the recursive relation:*

$$d_{\text{KT}}^k(S) = \min_{c \in S} [d_{\text{KT}}^k(S \setminus \{c\}) + \sum_{r \in \mathcal{P}_S} \sum_{c' \succ_r c} \sum_{i=0}^{k-2} \binom{|S| - \text{rk}(c', r) - 1}{i}] \quad (2)$$

$$d_{\text{KT}}^k(\emptyset) = 0.$$

Proof. Given $S \subseteq C$ and $c \in S$, let us define $\mathcal{R}_c(S)$ as $\{r \in \mathcal{R}(S) \text{ s.t. } t_r(S) = c\}$. The set $\Delta^k(S)$ can be partitioned into $\Delta_c^k(S) = \{S' \subseteq \Delta^k(S) \text{ s.t. } c \in S'\}$ and $\Delta_{\bar{c}}^k(S) = \{S' \subseteq \Delta^k(S) \text{ s.t. } c \notin S'\} = \Delta^k(S \setminus \{c\})$. Given a preference profile \mathcal{P} over C and a ranking $\hat{r} \in \mathcal{R}_c(S)$, the summation defining $\delta_{\text{KT}}^k(\hat{r}, \mathcal{P}_S)$ breaks down as follows:

$$\delta_{\text{KT}}^k(\hat{r}, \mathcal{P}_S) = \sum_{r \in \mathcal{P}_S} \sum_{S' \in \Delta^k(S)} \mathbb{1}_{t_{\hat{r}}(S') \neq t_r(S')} = \delta_{\text{KT}}^k(\hat{r}_{S \setminus \{c\}}, \mathcal{P}_{S \setminus \{c\}}) + \sum_{r \in \mathcal{P}_S} \sum_{S' \in \Delta_c^k(S)} \mathbb{1}_{t_{\hat{r}}(S') \neq t_r(S')}. \quad (3)$$

Using the same reasoning as in Equation 1 on page 1984, the second summand in Equation 3 can be rewritten as follows:

$$\sum_{r \in \mathcal{P}_S} \sum_{c' \in S} \mathbb{1}_{c' \succ_r c} \sum_{i=0}^{k-2} \binom{|B_c(\hat{r}) \cap B_{c'}(r)|}{i}$$

because $t_{\hat{r}}(S') = c$ for all $S' \in \Delta_c^k(S)$. Note that $B_c(\hat{r}) = S \setminus \{c\}$ and $B_{c'}(r) = \{c'' \in S \text{ s.t. } c'' \succ_r c'\} \subseteq S$, thus $|B_c(\hat{r}) \cap B_{c'}(r)| = |S| - \text{rk}(c', r) - 1$. Hence, $\delta_{\text{KT}}^k(\hat{r}, \mathcal{P}_S)$ is equal to:

$$\delta_{\text{KT}}^k(\hat{r}_{S \setminus \{c\}}, \mathcal{P}_{S \setminus \{c\}}) + \sum_{r \in \mathcal{P}_S} \sum_{c' \succ_r c} \sum_{i=0}^{k-2} \binom{|S| - \text{rk}(c', r) - 1}{i} \quad (4)$$

Consider now a ranking $r^* \in \mathcal{R}(S)$ such that $\delta_{\text{KT}}^k(r^*, \mathcal{P}_S) = \min_{r \in \mathcal{R}(S)} \delta_{\text{KT}}^k(r, \mathcal{P}_S)$. We have:

$$\delta_{\text{KT}}^k(r^*, \mathcal{P}_S) = \min_{c \in S} \min_{\hat{r} \in \mathcal{R}_c(S)} \delta_{\text{KT}}^k(\hat{r}, \mathcal{P}_S) = \min_{c \in S} \left(\min_{\hat{r} \in \mathcal{R}(S \setminus \{c\})} \delta_{\text{KT}}^k(\hat{r}, \mathcal{P}_{S \setminus \{c\}}) + \sum_{r \in \mathcal{P}_S} \sum_{c' \succ_r c} \sum_{i=0}^{k-2} \binom{|S| - \text{rk}(c', r) - 1}{i} \right)$$

because the second summand in Equation 4 does not depend on \hat{r} (it only depends on c , which is the argument of the first min operator). If one denotes $\min_{r \in \mathcal{R}(S)} \delta_{\text{KT}}^k(r, \mathcal{P}_S)$ by $d_{\text{KT}}^k(S)$, one obtains Equation 2. This concludes the proof. \square

A candidate $c \in S$ that realizes the minimum in Equation 2 can be ranked in first position in an optimal ranking for \mathcal{P}_S . Once $d_{\text{KT}}^k(S)$ is computed for each $S \subseteq C$, a ranking r^* achieving the optimal value $d_{\text{KT}}^k(C)$ can thus be determined recursively starting from $S = C$. The complexity of the induced dynamic programming method is $O(2^m m^2 n)$ as there are 2^m subsets $S \subseteq C$ to consider and each value

$d_{\text{KT}}^k(S)$ is computed in $O(m^2n)$ by Equation 2. The min operation is indeed performed on m values and the sum $\sum_{c' \succ_r c} \sum_{i=0}^{k-2} \binom{|S| - \text{rk}_i(c', r) - 1}{i}$ is computed incrementally in $O(m)$, which entails an $O(mn)$ complexity for the second summand in Equation 2 (the n factor is due to the sum over all $r \in \mathcal{P}_S$). The computation of binomial coefficients $\binom{p}{i}$ for $i \in \{0, \dots, k-2\}$ and $p \in \{i, \dots, m-2\}$ is performed in $O(mk)$ in a preliminary step thanks to Pascal's formula.

4 The k -Wise Majority Digraph

We now propose and investigate a k -wise counterpart of the pairwise majority digraph, that will be used in a preprocessing procedure for k -KAP.

As stated in the introduction, the pairwise Kemeny rule is strongly related to the *pairwise majority digraph*. We denote by $\mathcal{G}_{\mathcal{P}}$ the pairwise majority digraph associated to profile \mathcal{P} . We recall that in this digraph, there is one vertex per candidate, and there is an arc from candidate c to candidate c' if a strict majority of voters prefers c to c' . In the weighted pairwise majority digraph, each arc (c, c') is weighted by $w_{\mathcal{P}}(c, c') := |\{r \in \mathcal{P} \text{ s.t. } c \succ_r c'\}| - |\{r \in \mathcal{P} \text{ s.t. } c' \succ_r c\}|$.

Example 5. Consider a profile \mathcal{P} with 10 voters and 6 candidates such that:

- 4 voters have preferences $c_1 \succ c_2 \succ c_4 \succ c_3 \succ c_5 \succ c_6$;
- 4 voters have preferences $c_1 \succ c_3 \succ c_2 \succ c_4 \succ c_5 \succ c_6$;
- 1 voter has preferences $c_6 \succ c_1 \succ c_2 \succ c_4 \succ c_3 \succ c_5$;
- 1 voter has preferences $c_6 \succ c_1 \succ c_4 \succ c_3 \succ c_2 \succ c_5$.

The pairwise majority digraph $\mathcal{G}_{\mathcal{P}}$ is on the left of Figure 1.

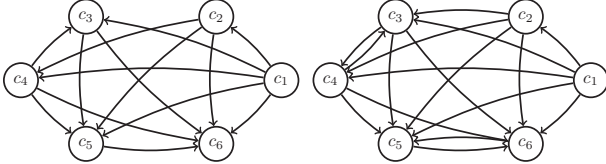


Figure 1: k -wise majority digraph in Example 5 for $k = 2$ (left) and $k = 3$ (right).

From $\mathcal{G}_{\mathcal{P}}$, we can define a set of *consistent* rankings:

Definition 1. Let \mathcal{G} be a digraph whose vertices correspond to the candidates in C . Let $B_1(\mathcal{G}), \dots, B_t(\mathcal{G})$ be the subsets of C corresponding to the Strongly Connected Components (SCCs) of \mathcal{G} , and $\mathcal{O}(\mathcal{G})$ denote the set of linear orders $\prec_{\mathcal{G}}$ on $\{1, \dots, t(\mathcal{G})\}$ such that if there exists an arc (c, c') from $c \in B_i(\mathcal{G})$ to $c' \in B_j(\mathcal{G})$ then $i <_{\mathcal{G}} j$. Given $\prec_{\mathcal{G}} \in \mathcal{O}(\mathcal{G})$, we say that a ranking r is consistent with $\prec_{\mathcal{G}}$ if the candidates in B_i are ranked before the ones of B_j when $i <_{\mathcal{G}} j$.

The following result states that, for any $\prec_{\mathcal{G}_{\mathcal{P}}} \in \mathcal{O}(\mathcal{G}_{\mathcal{P}})$, there exists a consensus ranking for δ_{KT} among the rankings consistent with $\prec_{\mathcal{G}_{\mathcal{P}}}$.

Theorem 2 (Theorem 16 in reference Charon and Hudry, 2010). Let \mathcal{P} be a profile over C and assume that the SCCs of $\mathcal{G}_{\mathcal{P}}$ are numbered according to a linear order $\prec_{\mathcal{G}_{\mathcal{P}}} \in \mathcal{O}(\mathcal{G}_{\mathcal{P}})$. Consider the ranking r^* , consistent with $\prec_{\mathcal{G}_{\mathcal{P}}}$, obtained by the concatenation of rankings $r_1^*, \dots, r_{t(\mathcal{G}_{\mathcal{P}})}^*$

where $\delta_{\text{KT}}(r_i^*, \mathcal{P}_{B_i(\mathcal{G}_{\mathcal{P}})}) = \min_{r \in \mathcal{R}(B_i(\mathcal{G}_{\mathcal{P}}))} \delta_{\text{KT}}(r, \mathcal{P}_{B_i(\mathcal{G}_{\mathcal{P}})})$. We have:

$$\delta_{\text{KT}}(r^*, \mathcal{P}) = \min_{r \in \mathcal{R}(C)} \delta_{\text{KT}}(r, \mathcal{P})$$

That is, r^* is a consensus ranking according to the Kemeny rule. Furthermore, if $\mathcal{O}(\mathcal{G}_{\mathcal{P}}) = \{\prec_{\mathcal{G}_{\mathcal{P}}}\}$ and $w_{\mathcal{P}}(c, c') > 0$ for all $c \in B_i(\mathcal{G}_{\mathcal{P}})$ and $c' \in B_j(\mathcal{G}_{\mathcal{P}})$ when $i <_{\mathcal{G}_{\mathcal{P}}} j$, then all consensus rankings are consistent with $\prec_{\mathcal{G}_{\mathcal{P}}}$.

This result does not hold anymore if one uses δ_{KT}^k (with $k \geq 3$) instead of δ_{KT} , as shown by the following example.

Example 6. Let us denote by \mathcal{P} the profile of Example 1. The pairwise majority digraph $\mathcal{G}_{\mathcal{P}}$ has three SCCs $B_1(\mathcal{G}_{\mathcal{P}}) = \{c_2\}$, $B_2(\mathcal{G}_{\mathcal{P}}) = \{c_3\}$ and $B_3(\mathcal{G}_{\mathcal{P}}) = \{c_1\}$. In this example, $\mathcal{O}(\mathcal{G}_{\mathcal{P}}) = \{\prec_{\mathcal{G}_{\mathcal{P}}}\}$ where $1 <_{\mathcal{G}_{\mathcal{P}}} 2 <_{\mathcal{G}_{\mathcal{P}}} 3$. The only ranking consistent with $\prec_{\mathcal{G}_{\mathcal{P}}}$ is $c_2 \succ c_3 \succ c_1$ while the only consensus ranking w.r.t. the 3-wise Kemeny rule is $c_1 \succ c_2 \succ c_3$.

In order to adapt Theorem 2 to the k -wise Kemeny rule, we now introduce the concept of *k -wise majority digraph*. Let $\Delta_{cc'}^k(S) = \{S' \in \Delta^k(S) \text{ s.t. } \{c, c'\} \subseteq S'\}$. If S is not specified, it is assumed to be C . Given a ranking r , we denote by $\Delta_r^k(S, c, c')$ the set $\{S' \in \Delta_{cc'}^k(S) \text{ s.t. } t_r(S') = c\}$. Given a profile \mathcal{P} , we denote by $\phi_{\mathcal{P}}^k(S, c, c')$ the value $\sum_{r \in \mathcal{P}} |\Delta_r^k(S, c, c')|$ and by $w_{\mathcal{P}}^k(S, c, c')$ the difference $\phi_{\mathcal{P}}^k(S, c, c') - \phi_{\mathcal{P}}^k(S, c', c)$. This definition implies that $w_{\mathcal{P}}^k(S, c', c) = -w_{\mathcal{P}}^k(S, c, c')$. The value $w_{\mathcal{P}}^k(S, c, c')$ is the net agreement loss that would be incurred by swapping c and c' in a feasible solution r of k -KAP where $\text{rk}(c', r) = \text{rk}(c, r) + 1$ and $S = B_{c'}(r) \cup \{c, c'\}$. If $\max_{S \in \Delta_{cc'}^m} w_{\mathcal{P}}^k(S, c, c') \geq 0$ (resp. $\min_{S \in \Delta_{cc'}^m} w_{\mathcal{P}}^k(S, c, c') > 0$) then, in a consensus ranking r for δ_{KT}^k where c and c' would be consecutive, it is possible (resp. necessary) that $c \succ_r c'$.

The k -wise majority digraph associated to a profile \mathcal{P} over a set C of candidates is the weighted digraph $\mathcal{G}_{\mathcal{P}}^k = (\mathcal{V}, \mathcal{A})$, where $\mathcal{V} = C$ and $(c, c') \in \mathcal{A}$ iff:

$$\exists S \in \Delta_{cc'}^m \text{ s.t. } w_{\mathcal{P}}^k(S, c, c') > 0.$$

The weight $w_{\mathcal{P}}^k(c, c')$ of this arc is then given by:

$$w_{\mathcal{P}}^k(c, c') := \max_{S \in \Delta_{cc'}^m} w_{\mathcal{P}}^k(S, c, c').$$

Note that, if $k \geq 3$, we may obtain edges (c, c') and (c', c) both with strictly positive weights (which is impossible in the pairwise majority digraph). For instance, for the profile \mathcal{P} of Example 5, $w_{\mathcal{P}}^k(c_3, c_4) = w_{\mathcal{P}}^k(\{c_2, c_3, c_4\}, c_3, c_4) = 1$ and $w_{\mathcal{P}}^k(c_4, c_3) = w_{\mathcal{P}}^k(\{c_3, c_4, c_5\}, c_4, c_3) = 4$. The digraph $\mathcal{G}_{\mathcal{P}}^3$ is shown on the right of Figure 1. Besides, for any \mathcal{P} , $\mathcal{G}_{\mathcal{P}}^2$ is the pairwise majority digraph as $\Delta_{cc'}^2(S) = \{\{c, c'\}\}$ $\forall S \in \Delta_{cc'}^m$. Theorem 2 adapts as follows for an arbitrary k :

Theorem 3. Let \mathcal{P} be a profile over C and assume that the SCCs of $\mathcal{G}_{\mathcal{P}}^k$ are numbered according to a linear order $\prec_{\mathcal{G}_{\mathcal{P}}^k} \in \mathcal{O}(\mathcal{G}_{\mathcal{P}}^k)$. Among the rankings consistent with $\prec_{\mathcal{G}_{\mathcal{P}}^k}$, there exists a consensus ranking w.r.t. the k -wise Kemeny rule. Besides, if $\mathcal{O}(\mathcal{G}_{\mathcal{P}}^k) = \{\prec_{\mathcal{G}_{\mathcal{P}}^k}\}$ and $\min_{S \in \Delta_{cc'}^m} w_{\mathcal{P}}^k(S, c, c') > 0^2$ for all $c \in B_i(\mathcal{G}_{\mathcal{P}}^k)$ and $c' \in B_j(\mathcal{G}_{\mathcal{P}}^k)$ when $i <_{\mathcal{G}_{\mathcal{P}}^k} j$, then all consensus rankings are consistent with $\prec_{\mathcal{G}_{\mathcal{P}}^k}$.

²Or, equivalently, $\max_{S \in \Delta_{cc'}^m} w_{\mathcal{P}}^k(S, c', c) < 0$.

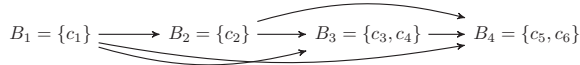


Figure 2: The meta-graph of SCCs of \mathcal{G}_P^3 in Example 5.

Example 7. The meta-graph of SCCs of \mathcal{G}_P^3 in Example 5 is represented in Figure 2. The above result implies that there exists a consensus ranking among $c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6$, $c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_6 \succ c_5$, $c_1 \succ c_2 \succ c_4 \succ c_3 \succ c_5 \succ c_6$ and $c_1 \succ c_2 \succ c_4 \succ c_3 \succ c_6 \succ c_5$.

To take advantage of Theorem 3, one could try 1) to index the SCCs of \mathcal{G}_P^k according to a linear order $\langle_{\mathcal{G}_P^k} \in \mathcal{O}(\mathcal{G}_P^k)$, and then 2) to work on each SCC separately, before concatenating the obtained rankings. However, for a consensus ranking consistent with $\langle_{\mathcal{G}_P^k}$, the relative positions of candidates in $B_i(\mathcal{G}_P^k)$ depend on the set of candidates in $B_{>i}(\mathcal{G}_P^k) := B_{i+1}(\mathcal{G}_P^k) \cup \dots \cup B_{t(\mathcal{G}_P^k)}(\mathcal{G}_P^k)$ (but not on their order). The influence of $B_{>i}(\mathcal{G}_P^k)$ can be captured in the dynamic programming procedure by applying a modified version of Equation 2 separately for each subset $B_{t(\mathcal{G}_P^k)}(\mathcal{G}_P^k)$ down to $B_1(\mathcal{G}_P^k)$. Formally, if r^* is optimal for k -KAP, then:

$$\delta_{\text{KT}}^k(r^*, \mathcal{P}) = \sum_{i=1}^{t(\mathcal{G}_P^k)} d_{\text{KT}}^k(B_i(\mathcal{G}_P^k))$$

where, for any subset $S \subseteq B_i(\mathcal{G}_P^k)$, $d_{\text{KT}}^k(S)$ is defined by $d_{\text{KT}}^k(\emptyset) = 0$ and ($B_{>i}$ stands for $B_{>i}(\mathcal{G}_P^k)$):

$$d_{\text{KT}}^k(S) = \min_{c \in S} [d_{\text{KT}}^k(S \setminus \{c\}) + \sum_{r \in \mathcal{P}_{S \cup B_{>i}}} \sum_{c' \succ_r c} \sum_{i=0}^{k-2} \binom{|S| + |B_{>i}| - \text{rk}(c', r) - 1}{i}]$$

It amounts to replacing S by $S \cup B_{>i}$ in the second summand of Equation 2 to take into account the existence of a consensus ranking where all the candidates of $B_{>i}$ are ranked after those of B_i . Let r_i^* be a ranking of $B_i(\mathcal{G}_P^k)$ such that $\delta_{\text{KT}}^k(r_{\geq i}^*, \mathcal{P}_{B_{\geq i}(\mathcal{G}_P^k)}) = d_{\text{KT}}^k(B_i(\mathcal{G}_P^k)) + \dots + d_{\text{KT}}^k(B_{t(\mathcal{G}_P^k)}(\mathcal{G}_P^k))$, where $r_{\geq i}^*$ is the ranking obtained by the concatenation of rankings $r_i^*, \dots, r_{t(\mathcal{G}_P^k)}^*$ in this order. The ranking $r_{\geq 1}^*$ of C is a consensus ranking w.r.t. the k -wise Kemeny rule. Given Theorem 3, the k -wise majority digraph thus seems promising to boost the computation of a consensus ranking. Unfortunately, the following negative result holds.

Theorem 4. Given a profile \mathcal{P} and two candidates c and c' , determining if $\max_{S \in \Delta_{cc'}^m} w_P^k(S, c, c') > 0$ is NP-hard for $k \geq 4$.

Hence, computing \mathcal{G}_P^k from \mathcal{P} is NP-hard for $k \geq 4$. In contrast, \mathcal{G}_P^3 can be computed in polynomial time. Indeed, given a set $S \subset C$ such that $\{c, c'\} \subseteq S$, adding to S an element $x \notin S$ increases $\phi_P^3(S, c, c')$ by one for each $r \in \mathcal{P}$ such that $c \succ_r c'$ and $c \succ_r x$. Let $\mathcal{P}_{c \succ_r c'} := \{r \in \mathcal{P} \text{ s.t. } c \succ_r c'\}$. A set S^* maximizing $w_P^3(S, c, c')$ is $S^* := \{c, c'\} \cup \{x \in C \text{ s.t. } |\mathcal{P}_{c \succ_r c'} \cap \mathcal{P}_{c \succ_r x}| > |\mathcal{P}_{c \succ_r c} \cap \mathcal{P}_{c' \succ_r x}|\}$.

Note that one can take advantage of the meta-graph of SCCs to trim the graph \mathcal{G}_P^k if one looks for a consensus ranking r^* consistent with a specific order $\langle_{\mathcal{G}_P^k} \in \mathcal{O}(\mathcal{G}_P^k)$.

It may indeed happen that, for an edge (c, c') , the weight $w_P^k(c, c') = w_P^k(S, c, c') > 0$ corresponds to a set S which contains candidates that will never be below c in r^* . Conversely, the set S may omit candidates that are necessarily below c in r^* . These constraints can be induced by either unanimity dominance relations or by $\langle_{\mathcal{G}_P^k}$. The following example illustrates this idea.

Example 8. Let us refine the digraph \mathcal{G}_P^3 previously obtained for the profile \mathcal{P} of Example 5. The SCCs are $B_1 = \{c_1\}$, $B_2 = \{c_2\}$, $B_3 = \{c_3, c_4\}$ and $B_4 = \{c_5, c_6\}$ and $\mathcal{O}(\mathcal{G}_P^k) = \{\langle_{\mathcal{G}_P^k}\}$, where $1 \langle_{\mathcal{G}_P^k} 2 \langle_{\mathcal{G}_P^k} 3 \langle_{\mathcal{G}_P^k} 4$. A set maximizing $w_P^3(S, c_3, c_4)$ is $S = \{c_2, c_3, c_4\}$. This set contains c_2 while it is necessarily above c_3 in a consistent ranking. Conversely, candidates c_5 and c_6 are omitted while they are necessarily below c_3 . By taking into account these constraints, we obtain that a set maximizing $w_P^3(S, c_3, c_4)$ is $S = \{c_3, c_4, c_5, c_6\}$, for which $w_P^3(S, c_3, c_4) = -4$. Hence, we can remove the arc (c_3, c_4) from \mathcal{G}_P^3 . Similarly, it is possible to show that the arc (c_6, c_5) can be removed from \mathcal{G}_P^3 . Thanks to these refinement steps, we can conclude that a consensus ranking is $r^* = c_1 \succ c_2 \succ c_4 \succ c_3 \succ c_5 \succ c_6$.

5 Numerical Tests

Our numerical tests³ have three objectives: we evaluate the computational performance of the dynamic programming approach of Section 3, we evaluate the impact of parameter k on the set of consensus rankings, and we assess the efficiency of the preprocessing technique of Section 4.

Generation of preference profiles. The preference profiles are generated according to the Mallows model (Mallows 1957), using the Python package PrefLib-Tools (Mattei and Walsh 2013). This model takes two parameters as input: a reference ranking σ (the mode of the distribution) and a dispersion parameter $\phi \in (0, 1)$. Given these inputs, the probability of generating a ranking r is proportional to $\phi^{\delta_{\text{KT}}(r, \sigma)}$. The more ϕ tends towards 0 (resp. 1), the more the preference rankings become correlated and resemble σ (resp. become equally probable, i.e., we are close to the *impartial culture assumption*). This model enables us to measure in a simple way how the level of correlation in the input rankings impacts our results. In all tests, the number n of voters is set to 50 and the ranking σ is set arbitrarily as the k -wise Kemeny rule is neutral. For each triple (m, k, ϕ) considered, the results are averaged over 50 preference profiles.

Practicability of the dynamic programming approach. We first evaluate our dynamic programming approach on instances with different values for m and k . Note that the computational performance measured here is not impacted by the level of correlation in the input rankings as it does

³Implementation in C++. All times are CPU seconds on an Intel Core I7-8700 3.20 GHz processor with 16GB of RAM.

Table 2: Average, max and min wall-clock times in seconds of the dynamic programming approach of Section 3 for varying values of m and k (Rows 3 to 5). Average number of consensus rankings for increasing values of m and k (Row 6).

m	6			10			14			18		
k	2	3	6	2	5	10	2	7	14	2	9	18
Average time	<0.01	<0.01	<0.01	0.07	0.08	0.08	2.52	2.54	2.61	70.93	72.26	74.95
Max time	<0.01	<0.01	<0.01	0.8	0.08	0.09	2.64	2.60	2.64	71.57	73.57	75.38
Min time	<0.01	<0.01	<0.01	0.7	0.07	0.08	2.49	2.49	2.57	70.27	71.91	74.33
$ \mathcal{R}^* _{\text{avg}}$	3.00	1.20	1.05	3.84	1.24	1.10	5.36	2.36	1.16	19.7	4.12	1.47

not change the number of states in dynamic programming nor the computation time to determine the optimal value in each state. Hence, we only consider instances generated under the impartial culture assumption, i.e., with $\phi \approx 1$. Table 2 (Rows 3-5) displays the average, max and min running times obtained for some representative (m, k) values. As expected, the running times increase exponentially with m . Conversely, parameter k seems to have a moderate impact on the running times. The dynamic programming approach enables us to solve k -KAP in a time of up to 3 sec. (resp. 76 sec.) for $m \leq 14$ (resp. $m \leq 18$).

Influence of k on the set of consensus rankings. Second, we study the impact of k on the set of optimal solutions to k -KAP. Indeed, one criticism for the Kemeny rule is that there exists instances for which the set of consensus rankings is compounded of many solutions which are quite different from one another. Thus, we investigate if increasing k helps in mitigating this issue. For this purpose, we consider the same instances as before and compute the average number of consensus rankings denoted by $|\mathcal{R}^*|_{\text{avg}}$. The results are displayed in the sixth row of Table 2. Interestingly, this measure decreases quickly with k . For instance, when $m = 18$, $|\mathcal{R}^*|_{\text{avg}}$ is divided by 5 when k increases from 2 to 9 and is below 2 when $k = m$. The intuition is that δ_{KT}^k becomes more fine-grained as k increases.

Impact of the 3-wise majority graph. Lastly, we study the impact of the preprocessing method proposed in Section 4 for $k = 3$. This preprocessing uses the k -wise majority digraph to divide k -KAP into several subproblems which can be solved separately by dynamic programming. Hopefully, when voters' preferences are correlated (i.e., for "small" ϕ values), these subproblems become smaller and more numerous, making the preprocessing more efficient. The results are shown in Table 3, where the results obtained without preprocessing are also given in the last column. The obtained running times are highly dependent on ϕ . For instance, with $m = 18$, the average running time for solving 3-KAP is above 1 minute if $\phi = 0.95$ while it is below 1 second if $\phi \leq 0.85$. This gap is necessarily related to the preprocessing step, since ϕ has no impact on the running time of the dynamic programming approach. To explain this significant speed-up, we display in Table 4 the average size of the largest SCC of the 3-wise majority digraph at the end of the preprocessing step. Unsurprisingly, this average size turns out to be correlated with ϕ : when $\phi \leq 0.5$, the size of the largest SCC is almost always 1. Hence, the preprocess-

ing step is likely to yield directly a consensus ranking. In contrast, when $\phi = 0.95$, the average size of the largest SCC is close to m , thus the impact of the preprocessing is low.

Table 3: Average, max and min wall-clock times (in seconds) for the 3-wise Kemeny rule with preprocessing.

m	ϕ	0.5	0.8	0.85	0.9	0.95	w/o preproc.
6	Avg time	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	Max time	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	Min time	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
10	Avg time	0.03	0.03	0.04	0.07	0.10	0.07
	Max time	0.03	0.03	0.10	0.15	0.17	0.08
	Min time	0.03	0.03	0.03	0.03	0.03	0.07
14	Avg time	0.09	0.09	0.11	0.94	2.21	2.52
	Max time	0.09	0.13	0.25	3.14	3.26	2.59
	Min time	0.08	0.08	0.09	0.10	0.26	2.49
18	Avg time	0.20	0.20	0.55	14.87	61.72	71.17
	Max time	0.31	0.21	8.46	79.87	80.11	71.61
	Min time	0.19	0.19	0.19	0.22	6.02	71.02

Table 4: Average size of the largest SCC after preprocessing.

$m \setminus \phi$	0.47	0.81	0.85	0.88	0.95
6	<1.1	1.84	1.88	2.72	3.28
10	<1.1	1.64	3.28	5.32	8.20
14	<1.1	2.68	3.84	9.12	12.91
18	<1.1	2.84	4.27	9.80	17.44

6 Conclusion

In this paper, we advocate using the results of *setwise* contests between candidates to design social welfare functions that are less myopic than those only based on pairwise comparisons. In this direction, we have studied a k -wise generalization of the Kemeny rule, and established that determining a consensus ranking is NP-hard for any $k \geq 3$. After proposing a dynamic programming procedure, we have investigated a k -wise variant of the majority graph, from which we developed a preprocessing step. Computing this graph is a polynomial time problem for $k = 3$ but becomes NP-hard for $k \geq 4$. The numerical tests show the practicality of the approach for up to 18 candidates. A natural research direction is to investigate the complexity of determining a consensus ranking for δ_{KT}^k when $k = m$, because our hardness result only holds for *fixed* values of k . Another avenue to explore is to propose alternative definitions of k -wise majority graphs that are easier to compute for $k > 3$. Finally, other social welfare functions based on the results of setwise contests are worth investigating in our opinion, both from the axiomatic and the computational points of view.

7 Acknowledgments

This work has been partially supported by the Italian MIUR PRIN 2017 Project ALGADIMAR “Algorithms, Games, and Digital Markets”.

References

- Altman, A., and Tennenholtz, M. 2008. Axiomatic foundations for ranking systems. *Journal of Artificial Intelligence Research* 31:473–495.
- Arrow, K. J. 1950. A difficulty in the concept of social welfare. *Journal of political economy* 58(4):328–346.
- Baldiga, K. A., and Green, J. R. 2013. Assent-maximizing social choice. *Social Choice and Welfare* 40(2):439–460.
- Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D. 2016. *Handbook of computational social choice*. Cambridge University Press.
- Charon, I., and Hudry, O. 2010. An updated survey on the linear ordering problem for weighted or unweighted tournaments. *Annals of Operations Research* 175(1):107–158.
- Cheng, W., and Hüllermeier, E. 2009. A new instance-based label ranking approach using the mallows model. In *Proceedings of the 6th International Symposium on Neural Networks, ISNN 2009, Wuhan, China, May 26-29, 2009*, 707–716.
- Cléménçon, S.; Korba, A.; and Sibony, E. 2018. Ranking median regression: Learning to order through local consensus. In *Proceedings of the 29th international conference on Algorithmic Learning Theory, ALT 2018, Lanzarote, Canary Islands, Spain, April 7-9, 2018*, 212–245.
- Conitzer, V.; Rognlie, M.; and Xia, L. 2009. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence, IJCAI 2009, Pasadena, California, USA, July 11-17, 2009*, 109–115.
- Dwork, C.; Kumar, R.; Naor, M.; and Sivakumar, D. 2001. Rank aggregation methods for the web. In *Proceedings of the 10th international conference on World Wide Web, WWW 2001, Hong Kong, China, May 1-5, 2001*, 613–622.
- Fagin, R.; Kumar, R.; and Sivakumar, D. 2003. Efficient similarity search and classification via rank aggregation. In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data, San Diego, California, USA, June 9-12, 2003*, 301–312.
- Fishburn, P. C. 1977. Condorcet social choice functions. *SIAM Journal on applied Mathematics* 33(3):469–489.
- Gilbert, H.; Portoleau, T.; and Spanjaard, O. 2019. Beyond pairwise comparisons in social choice: A setwise kemeny aggregation problem. *CoRR* abs/1911.06226.
- Jackson, B. N.; Schnable, P. S.; and Aluru, S. 2008. Consensus genetic maps as median orders from inconsistent sources. *IEEE/ACM Transactions on computational biology and bioinformatics* 5(2):161–171.
- Kemeny, J. G. 1959. Mathematics without numbers. *Daedalus* 88(4):577–591.
- Kumar, R., and Vassilvitskii, S. 2010. Generalized distances between rankings. In *Proceedings of the 19th international conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010*, 571–580.
- Lu, T., and Boutilier, C. 2010. The unavailable candidate model: a decision-theoretic view of social choice. In *Proceedings of the 11th ACM conference on Electronic Commerce, EC 2010, Cambridge, Massachusetts, USA, June 7-11, 2010*, 263–274.
- Mallows, C. L. 1957. Non-null ranking models. I. *Biometrika* 44(1/2):114–130.
- Mattei, N., and Walsh, T. 2013. Preflib: A library of preference data [HTTP://PREFLIB.ORG](http://preflib.org). In *Proceedings of the 3rd international conference on Algorithmic Decision Theory, ADT 2013, Bruxelles, Belgium, November 13-15, 2013*, 259–270.
- Moulin, H. 1991. *Axioms of cooperative decision making*. Number 15 in Econometric Society Monographs. Cambridge University Press.
- Wang, S.; Sun, J.; Gao, B. J.; and Ma, J. 2014. Vsrnk: A novel framework for ranking-based collaborative filtering. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5(3):51.
- Zwicker, W. S. 2018. Cycles and intractability in a large class of aggregation rules. *Journal of Artificial Intelligence Research* 61:407–431.