A new binary programming formulation and social choice property for Kemeny rank aggregation

Yeawon Yoo
School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, yyoo12@asu.edu

Adolfo R. Escobedo
School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, adRes@asu.edu

Rank aggregation is widely used in group decision-making and many other applications where it is of interest to consolidate heterogeneous ordered lists. Oftentimes, these rankings may involve a large number of alternatives, contain ties, and/or be incomplete, all of which complicate the use of robust aggregation methods. In particular, these characteristics have limited the applicability of the aggregation framework based on the Kemeny-Snell distance, which satisfies key social choice properties that have been shown to engender improved decisions. This work introduces a binary programming formulation for the generalized Kemeny rank aggregation problem—whose ranking inputs may be complete and incomplete, with and without ties—and compare it to a modified version of a recently developed integer programming formulation for the generalized Kendall-tau distance. The new formulation has fewer variables and constraints, which leads to faster solution times. Moreover, we develop a new social choice property, the Non-strict Extended Condorcet Criterion, which can be regarded as a natural extension of the well-known Condorcet criterion and the Extended Condorcet criterion. Unlike its parent properties, the new property is adequate for handling complete rankings with ties. The property is leveraged to develop a structural decomposition algorithm, through which certain large instances of the NP-hard Kemeny rank aggregation problem can be solved exactly in a practical amount of time. To test the practical implications of the new formulation and social choice property, we work with instances constructed from a probabilistic distribution and benchmark instances from PrefLib, a library of preference data.

Key words: Group decision-making; Rank aggregation; Computational social choice; Combinatorial optimization.

1. Introduction

Group decision-making has been studied extensively since the shaping of democratic society. To give equal rights to each individual, rather than a selected few, many people have devoted their efforts to develop fair and consistent systems that aggregate the opinions of several individuals to make egalitarian social decisions. Eliciting and/or expressing the preferences over a set of alternatives or candidates as rankings (e.g., candidate i is in first place, candidate j is in second place, etc.) is popular across many decision-making contexts due in part to the scale-free characteristics of these evaluations and their efficient encapsulation of large numbers of pairwise comparisons. Therefore,

rank aggregation is a common and widely studied topic in group decision-making. A famous early result is that of Arrow et al. (1951) who studied the theoretical implications of the concept of a social welfare function (SWF), which maps individual rankings into a single ranking that should represent the best compromise among the given rankings. Therein, the author provided a set of fundamental conditions that a SWF should satisfy and demonstrated that they could not be satisfied simultaneously by any SWF. Despite this "impossibility" result, the study and use of rank aggregation methods have steadily increased across numerous fields because the need to consolidate such data for effective decision-making remains essential. For example, multi-criteria decision-making (MCDM) is closely related to rank aggregation; in a general MCDM problem, a set of actions or policies is evaluated according to multiple criteria. Different MCDM methods may yield different aggregate evaluations and, therefore, finding an overall aggregated ranking of alternatives is an ongoing point of concern (Mohammadi and Rezaei 2020).

Rank aggregation has been widely used across a number of practical group decision-making settings. For instance, Fields et al. (2013) consider a health care problem of improving nurse triage and prioritization in the emergency department of a hospital. When there are more patients waiting in the emergency room than available resources or staff, it is important to order the patients based on the severity of their condition. However, different nurses at times provide differing prioritization of patients, which can be represented by rankings, and thus it is necessary to resolve the conflicts among the multiple rankings. As additional examples, rank aggregation has been applied to evaluate research proposals (Cook et al. 2007), to judge student paper competitions (Hochbaum and Levin 2006), and to improve the annual draft preparation decision-making process in Major League Baseball (Streib et al. 2012).

Besides decision-making, rank aggregation has extensive applicability to other fields, such as bioinformatics (Lin 2010a,b, Marbach et al. 2012, Mandal and Mukhopadhyay 2017), information retrieval (Farah and Vanderpooten 2007, Yilmaz et al. 2008), and similarity search (Fagin et al. 2003, Ye et al. 2016, Gao and Xu 2019). In particular, these applications have demonstrated that robust rank aggregation methodologies can mitigate the influence of outliers and thwart manipulation and bias in practice (Dwork et al. 2001, Lin 2010a). From an operations research perspective, rank aggregation has been previously connected to the linear ordering problem (Martí and Reinelt 2011) and the theory of order polytopes (see Section 3.3) owing to its inherent combinatorial nature—i.e., a linear ordering is a permutation of the candidates (Fiorini and Fishburn 2004, Heiser 2004). Therefore, the rank aggregation problem is worth discussing in terms of both its potential impact on core methodological aspects and its practical benefits in a wide array of applications.

Distance-based and coefficient-based frameworks are the methodologies mainly used in robust rank aggregation due to their mathematically rigorous (i.e., axiomatic) foundations. The distance-based framework seeks a solution that minimizes the *cumulative disagreement* with the input

rankings, while the coefficient-based framework seeks a solution that maximizes the *cumulative agreement* with the input rankings. Accordingly, these methods are often referred to as *consensus* ranking aggregation methods.

Kemeny and Snell (1962) introduced a distance-based framework founded on a set of intuitive metric axioms; its associated SWF has been verified to possess many theoretical and practical benefits. Indeed, the consensus ranking framework based on the Kemeny-Snell distance has competitive advantages over other aggregation frameworks. Known widely as *Kemeny (rank) aggregation*, this framework is less vulnerable to manipulation than scoring methods and more robust to outliers (Feld and Grofman 1988, Favardin et al. 2002, Endriss et al. 2016). Assuming there are no cycles in the majority's pairwise preferences and the input rankings are complete, Kemeny aggregation is guaranteed to return the optimal ranking that reflects the majority's pairwise preferences. On the other hand, scoring methods are not guaranteed to do so, for instance, the optimal ranking solution may not place the *Condorcet winner* (see Section 5.1) in first place. The satisfaction of this and other key social choice properties are fundamental reasons that Kemeny aggregation is less vulnerable to outliers relative to other methods. As an example, Dwork et al. (2001) showed that Kemeny rank aggregation is effective at combating spam (i.e., outliers) in meta-search. The Borda count method (de Borda 1781) serves as a notable example of the vulnerability to outliers of scoring methods. This method, which assigns a score to each candidate based on the number of opponents it beats in an evaluation and calculates a final score for each candidate by summing the scores earned over all evaluations, is widely employed even though it can yield very inconsistent outcomes (Dummett 1998, Favardin et al. 2002), especially when the rankings are incomplete (Moreno-Centeno and Escobedo 2016). We remark that Kemeny aggregation has been employed in the area of decision analysis. For example, it has been used in MCDM to aggregate experts' opinions regarding which decision criteria should have higher priority (Krylovas et al. 2014) and to consolidate contradicting lists of the best alternatives obtained from the application of different decision criteria (Nurmi and Meskanen 2000). Furthermore, various extensions of the Kemeny-Snell distance have been developed and applied—e.g., (Dwork et al. 2001, Cook 2006, Moreno-Centeno and Escobedo 2016)—based on different assumptions about the judges' evaluations.

Correlation coefficients provide another framework to measure agreement between rankings, and they have been investigated primarily in statistics literature—e.g., (Ahlgren et al. 2003, Kendall 1938, Yilmaz et al. 2008). Kendall (1938) developed a coefficient-based framework, which is closely linked to the Kemeny-Snell distance. The original methodology called Kendall-τ is a non-parametric correlation coefficient that measures the agreement among *strict rankings* (i.e., rankings that do not allow ties); it was extended to handle *non-strict rankings* (i.e., rankings that allow ties) in Kendall (1948). Emond and Mason (2002) provided another version of Kendall-τ correlation coefficient

for non-strict rankings and demonstrated that their τ-extended coefficient, τ_x, returns the same optimal solutions as the Kemeny aggregation framework, when the inputs are also complete.

It is important to delve into three special characteristics of rankings addressed in this paper: high dimension, ties, and incompleteness. Large rank aggregation problems are prevalent in practice. For example, it is not uncommon for a federal funding agency to receive hundreds of submissions to a single program. Cushman et al. (2015) mentioned that the number of submitted proposals to the AAG program in the National Science Foundation Astronomical Sciences Division Astronomy was 731 in 2014. As a second example from portfolio decision analysis, Keisler (2004) and Schilling et al. (2007) consider problems with 500 options and 173 options, respectively; in the latter, it is mentioned that individual decision-makers had to assess their preferences for more than 50 alternatives. Furthermore, rank aggregation is applicable not only to preferential rankings obtained from human judges but also to a variety of ordinal data encountered in a wide array of non-human contexts such as bioinformatics, web-search engines, and recommendation systems. In bioinformatics, rank aggregation is used to integrate a long list of genes from genomic experiments and find putative genes for specific diseases, where a list of genes may consist of thousands of elements (Lin 2010b, Wald et al. 2012, Kolde et al. 2012). Rank aggregation can be also used in metasearch, where a user query is sent to multiple search engines and then the separate ranked lists from the search engines are aggregated into a representative collective list (Dwork et al. 2001, Desarkar et al. 2016). These and other examples underscore the need to consider large rank aggregation problems in practical decision analysis research.

Incomplete rankings are another common occurrence across various decision-making contexts. When the overall number of alternatives to evaluate is large, it may not be feasible or prudent for any single judge to provide a complete ranking of these alternatives. Indeed, according to Miller's law (Miller 1956), an average human can hold in short-term memory and judge properly only 7 ± 2 alternatives. In addition to this cognitive limitation, there are various other constraints (e.g., time) that would motivate the evaluation of a smaller subset of the alternatives (i.e., an incomplete ranking). Similarly, having the flexibility to tie some of the alternatives (i.e., a non-strict ranking) may help make an evaluation task more manageable. In practice, it is common for groups of alternatives to be perceived as being indistinguishable from one another and, therefore, it may not be possible for judges to order them strictly (Kendall 1945). Additionally, in a wide array of contexts, a set of evaluations may have very few distinct values and, hence, the corresponding rankings obtained from them may have many ties (Fagin et al. 2004).

The main contributions of this paper address these three special characteristics, which frequently occur in practice and can make rank aggregation especially difficult. First, this paper provides a binary programming formulation for non-strict complete and incomplete rankings; note that a

large majority of previous works consider only the case where the input rankings are strict and complete. The formulation is derived from an equivalent correlation coefficient interpretation of the problem, and its constraints are shown to be fundamentally connected to the weak order polytope (see Fiorini and Fishburn (2004)), specifically with its basic family of facet-defining inequalities. Second, we define a social choice property that allows certain instances that rank a large number of alternatives and may contain ties to be decomposed into smaller subproblems. In addition, we prove that Kemeny aggregation is consistent with the newly defined property when the rankings are non-strict and complete. The property provides several practical benefits for dealing with large instances of this problem. Specifically, for certain instances it ensures that some subsets of alternatives will always be preferred over other subsets—this without having to perform the aggregation. This has the benefit of allowing decision-makers to rule out many irrelevant alternatives from consideration. It can also help with the issue of multiple alternative optimal solutions inherent in Kemeny aggregation. In particular, decision-makers may validly conclude that a large number of the alternatives will never occupy the top positions of any optimal ranking with the use of this property. Third, the paper performs computational experiments on a wide range of instances to showcase the practical benefits of the binary programming formulation and social choice property. Altogether, our contributions enable the exact solution and analysis of certain types of instances with hundreds or thousands of alternatives in a practical amount of time.

The organization of the rest of the paper is as follows. Section 2 introduces distance metrics for group decision-making and describes the notation used throughout the paper. Section 3 derives a binary programming formulation for the generalized Kemeny rank aggregation problem. Section 4 summarizes the computational experiments performed to compare this new formulation to a modified version of a recently developed formulation for the generalized Kendall-τ distance. Section 5 introduces a social choice property that is applicable to non-strict complete rankings, an associated structural decomposition algorithm, and the mathematical proof establishing that the Kemeny optimal ranking aligns with this new property. Section 6 summarizes a set of computational experiments for demonstrating the various benefits of this new property. Lastly, Section 7 concludes the work and discusses some of the limitations of the proposed methodologies.

2. Distance metrics for decision-making

Distance functions are commonly used in decision-making to measure the disagreement between two subjective evaluations over a set of alternatives. This section introduces distance metrics used for rank aggregation; beforehand, Table 1 lists symbols and notations used throughout this paper.

V	A set of alternatives (i.e., $V=\{v_1,v_2,v_3,...,v_n\}$), where v_i denotes an alternative i and n is the number of alternatives
V^k	The k-th subset of alternatives, where $V^k \subseteq V$, $k \geq 1$
$\mathcal{P}(V)$	A family of all possible partitions of V (e.g., if $\{V^1,V^2\} \in \mathcal{P}(V)$, then $V^1 \cup V^2 = V$ and $V^1 \cap V^2 = \emptyset$)
L	A set of judges
A	A set of input rankings (ordinal-valued evaluations)
$\boldsymbol{a}^\ell$	The ranking from judge ℓ ($\ell=1,2,3,..,\lvert L\rvert$), where $\boldsymbol{a}^\ell \in A$
a_i^ℓ	Rank position of v_i in the evaluation from judge ℓ, where $\ell=1,2,3,..,\lvert L\rvert$
$v_i \succ v_j$	v_i is preferred over v_j (i.e., $a_i < a_j$ for some ranking $\boldsymbol{a}$)
$v_i \approx v_j$	v_i is tied with v_j (i.e., $a_i = a_j$ for some ranking $\boldsymbol{a}$)
$v_i \succeq v_j$	v_i is preferred over or tied with v_j (i.e., $a_i \leq a_j$ for some ranking $\boldsymbol{a}$)
p_{ij}	The number of judges who prefer v_i over v_j (i.e., $\lvert\{\boldsymbol{a}^\ell \in A : a_i^\ell < a_j^\ell\}\rvert$)
t_{ij}	The number of judges who tie v_i and v_j (i.e., $\lvert\{\boldsymbol{a}^\ell \in A : a_i^\ell = a_j^\ell\}\rvert$)
$v_i \overset{m}{\succ} v_j$	A majority of judges prefers, rather than disprefers, v_i over v_j (i.e., $p_{ij} > p_{ji}$)
$v_i \overset{m}{\approx} v_j$	No majority of judges prefers or disprefers v_i over v_j (i.e., $p_{ij} = p_{ji}$)
$v_i \overset{M}{\succ} v_j$	A decisive majority of judges prefers v_i over v_j (i.e., $p_{ij} > p_{ji} + t_{ij}$)
$v_i \overset{M}{\approx} v_j$	No decisive majority of judges prefers v_i over v_j, or vice versa (i.e., $t_{ij} \geq \lvert p_{ij} - p_{ji}\rvert$)

Table 1 Symbols and Notations

Kemeny and Snell (1962) introduced an axiomatic distance for non-strict complete rank aggregation, which is denoted as d_{KS} and is defined as follows:

$$d_{KS}(\boldsymbol{a},\boldsymbol{b}) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\text{sign}(a_i - a_j) - \text{sign}(b_i - b_j)\right|. \tag{1}$$

In essence, distance d_{KS} counts the pairwise inversions between ranking $\boldsymbol{a}$ and $\boldsymbol{b}$. Note that $\text{sign}(x)$ returns 1 if $x > 0$, 0 if $x = 0$, and -1 if $x < 0$. As discussed in Section 1, d_{KS} possesses certain unique advantages over other rank aggregation methods. Specifically, d_{KS} uniquely satisfies a set of intuitive axioms: non-negativity, commutativity (symmetry), triangular inequality, anonymity, scaling (Kemeny and Snell 1962, Cook and Seiford 1978, Cook and Kress 1985). Additionally, the aggregated outcome obtained with the Kemeny-Snell distance uniquely satisfies a set of desirable social choice properties conjointly: neutrality, consistency, and the Condorcet property (Young and Levenglick 1978). It has been shown that scoring methods (e.g., the Borda count method, plurality rule, average methods) and other distance-based methods cannot satisfy all of these properties conjointly (Goodin and List 2006, Brandt et al. 2016). Moreover, from a statistical perspective, the optimal solution to the Kemeny rank aggregation has an interpretation as a maximum likelihood estimator for a probabilistic model in which individual rankings are noisy estimates of an underlying true ranking (Young 1995). These axiomatic grounds and beneficial aggregation properties provide compelling theoretical reasons for using the Kemeny-Snell distance for rank aggregation.

Furthermore, d_{KS} is closely linked with the τ_x correlation coefficient, introduced by Emond and Mason (2002) and defined as:

$$\tau_x(\boldsymbol{a},\boldsymbol{b}) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij}}{n(n-1)},$$

where a_{ij} (and b_{ij}) is the (i,j)-element of the *ranking-matrix* of $\boldsymbol{a}$ (and $\boldsymbol{b}$), $[a_{ij}]$, given by:

$$a_{ij} = \begin{cases} 1 & \text{if } a_i \leq a_j, \\ -1 & \text{if } a_i > a_j, \\ 0 & \text{if } i = j. \end{cases} \tag{2}$$

Respective matrix representations for incomplete rankings are defined in Yoo et al. (2020). While the original τ coefficient (and the corresponding representation for d_{KS}) treats a tie as an expression of indifference by assigning it a value of 0 (see Emond and Mason (2002)), τ_x treats a tie as an expression of positive agreement by assigning it a value of 1 in the ranking-matrix. Although τ_x and d_{KS} have different treatments of ties, Emond and Mason (2002) proved that τ_x is connected to d_{KS} via the following equation:

$$\tau_x(\boldsymbol{a},\boldsymbol{b}) = 1 - \frac{2\,d_{KS}(\boldsymbol{a},\boldsymbol{b})}{n(n-1)}. \tag{3}$$

The above equation illustrates the connection between the measure of agreement (the correlation coefficient) and the measure of disagreement (the distance). To recognize this, it is important to explain that τ_x achieves values of 1 and -1 when there is complete agreement and complete disagreement, respectively, between two rankings $\boldsymbol{a}$ and $\boldsymbol{b}$. Hence, it can be interpreted that the expression of $\tau_x(\boldsymbol{a},\boldsymbol{b})$ starts from a default assumption of perfect agreement between $\boldsymbol{a}$ and $\boldsymbol{b}$ (i.e., a correlation value of 1), and then it subtracts any disagreements between $\boldsymbol{a}$ and $\boldsymbol{b}$ from this perfect agreement, as quantified by $d_{KS}(\boldsymbol{a},\boldsymbol{b})$. In the case when i and j are tied, d_{KS} subtracts 0 (i.e., indifference), as expected; however, in doing so, τ_x keeps the default assumption of agreement between i and j (for a detailed proof of Equation (3), see Emond and Mason (2002)).

From this connection, the respective problems give equivalent optimal solutions, that is,

$$\arg\min_{\boldsymbol{r}} \sum_{\ell=1}^{|L|} d_{KS}(\boldsymbol{a}^\ell,\boldsymbol{r}) = \arg\max_{\boldsymbol{r}} \sum_{\ell=1}^{|L|} \tau_x(\boldsymbol{a}^\ell,\boldsymbol{r}), \tag{4}$$

where $\boldsymbol{r}$ is a complete ranking and $\boldsymbol{a}^\ell$ is the evaluation from judge $\ell \in L$.

Another distance metric referenced in this paper is the Kendall distance, which is adapted from the Kendall-τ correlation coefficient (Kendall 1938), defined as:

$$d_\tau(\boldsymbol{a},\boldsymbol{b}) = \sum_{1\leq i<j\leq n} \mathbb{1}_{[(a_i-a_j)(b_i-b_j)<0]}.$$

This distance counts the number of the pairwise inversions between $\boldsymbol{a}$ and $\boldsymbol{b}$ and is equivalent to d_{KS} when the rankings are strict, although the distances are scaled differently. Specifically, when one pair of items has (strict) opposing preferences, d_{KS} accrues a value of 2 (based on one of the Kemeny-Snelly axioms), while d_τ accrues a distance a value of 1. Hence, the distances are related by the equation $d_{KS}(\boldsymbol{a},\boldsymbol{b}) = 2d_\tau(\boldsymbol{a},\boldsymbol{b})$. Since the original Kendall-τ distance is defined only for strict rankings, Brancotte et al. (2015) introduced the generalized Kendall-τ distance to handle non-strict rankings, which is defined as follows:

$$d_{\tau'}(\boldsymbol{a},\boldsymbol{b}) = \sum_{1\le i<j\le n} \mathbb{1}_{((a_i<a_j)\cap(b_i>b_j))\cup((a_i>a_j)\cap(b_i<b_j))\cup((a_i=a_j)\cap(b_i\neq b_j))\cup((a_i\neq a_j)\cap(b_i=b_j))}.$$

The main difference between the generalized Kendall-τ distance and d_{KS} is that when one ranking ties two specific alternatives and the other ranking does not, the generalized Kendall-τ distance returns the same distance as when the two rankings have opposite strict preferences. Conversely, the d_{KS} distance returns half of the distance value in the former case relative to the latter case.

3. Exact mathematical programming formulations for Kemeny aggregation

It is known that finding the consensus ranking is NP-hard (Bartholdi et al. 1989, Good 1975), even when there are only four complete rankings to be aggregated. Considering incomplete-ranking inputs exacerbates these computational difficulties. When solving the problem via the standard branch and bound algorithm, incompleteness increases solution symmetry, which is defined as a permutation of the values of the variables that preserves the set of solutions (Cohen et al. 2005, Liberti 2008). This has the effect of slowing down pruning of nodes and, consequently, leads to a larger branch and bound tree (Sherali and Smith 2001). Moreover, incomplete-ranking instances may yield a higher number of alternative optimal solutions than complete-ranking instances, which could also lead to less decisive outcomes (Yoo et al. 2020). Hence, the vast majority of works to date have been able to solve only small instances of non-strict complete rank aggregation problems exactly or they have primarily focused on (meta)heuristics (Davenport and Kalagnanam 2004, Amodio et al. 2016, D'Ambrosio et al. 2017, Mandal and Mukhopadhyay 2017) and approximation algorithms (i.e., heuristics with provable performance guarantees) (Fagin et al. 2003, Kenyon-Mathieu and Schudy 2007, Ailon et al. 2008, Ailon 2010). As a notable example, D'Ambrosio et al. (2017) developed a differential evolution algorithm for Kemeny rank aggregation, which does consider non-strict and incomplete rankings but is not guaranteed to return optimal solutions.

In the remainder of this section, we derive a new exact binary programming formulation that is applicable to non-strict complete and incomplete rankings. Beforehand, we introduce some existing integer programming formulations for distance-based rank aggregation.

3.1. Integer programming formulations for the generalized Kendall-τ distance

Cook et al. (2007) developed a binary programming formulation related to the Kemeny-Snell distance, and Conitzer et al. (2006) developed an integer programming formulation related to the Kendall-τ distance. Because these formulations do not deal with non-strict rankings, Brancotte et al. (2015) provided a revised integer programming formulation for the generalized Kendall-τ distance, given as follows:

$$\begin{aligned}
\underset{\boldsymbol{x}}{\text{minimize}} \quad & \sum_{\{v_i,v_j\}\subseteq V} (w_{j\leq i}x_{i<j} + w_{i\leq j}x_{j<i} + (w_{i<j}+w_{j<i})x_{i=j}) && && \text{(5a)}\\
\text{subject to} \quad & x_{i<j} + x_{j<i} + x_{i=j} = 1 && \forall\{v_i,v_j\}\in V && \text{(5b)}\\
& x_{i<k} - x_{i<j} - x_{j<k} \geq -1 && \forall\{v_i,v_j,v_k\}\in V && \text{(5c)}\\
& 2x_{i<j} + 2x_{j<i} + 2x_{j<k} + 2x_{k<j} - x_{i<k} - x_{k<i} \geq 0 && \forall\{v_i,v_j,v_k\}\in V && \text{(5d)}\\
& x_{i<j}, x_{j<i}, x_{i=j} \in \mathbb{B} && \forall\{v_i,v_j\}\in V. && \text{(5e)}
\end{aligned}$$

Constraint (5b) ensures that all pairs of alternatives are assigned exactly one of the three possible relative ordinal positions: preferred, dispreferred, or tied (the first two are strict and the third is non-strict). Constraints (5c) and (5d) enforce transitivity of strict and non-strict ordinal relationships. We note that this formulation's objective function does not align with the definition of d_{KS} (see Equation (1)), based on the different treatment of ties of $d_{\tau'}$. Using d_{KS}, when there are two rankings, where one ties v_i and v_j and the other one does not, this should return half the distance compared to when the ordinal relationships strictly oppose each other. For example, with $\boldsymbol{a}^1 = (1,2), \boldsymbol{a}^2 = (1,1), \boldsymbol{a}^3 = (2,1)$, this yields $d_{KS}(\boldsymbol{a}^1,\boldsymbol{a}^2) = 1$, while $d_{KS}(\boldsymbol{a}^1,\boldsymbol{a}^3) = 2$. To solve the Kemeny rank aggregation problem using Brancotte et al. (2015)'s formulation, it is necessary to modify its objective function to match the treatment of ties of d_{KS}. This is done in the following proposition.

Proposition 3.1 *To adopt the treatment of ties of d_{KS}, Brancotte et al. (2015)'s formulation is modified as follows:*

$$\begin{aligned}
&\underset{\boldsymbol{x}}{\textit{minimize}} \sum_{\{v_i,v_j\}\subseteq V} (w_{j<i} + \frac{1}{2}w_{j=i})x_{i<j} + (w_{i<j} + \frac{1}{2}w_{i=j})x_{j<i} + \frac{1}{2}(w_{i<j}+w_{j<i})x_{i=j}\\
&\textit{subject to constraints } (5b) - (5e)
\end{aligned}$$

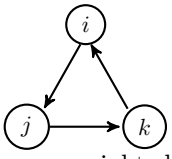
This modified version is used throughout the experiments to allow for fair comparison.

3.2. Deriving a new integer programming formulation

In this section, we leverage the correlation coefficient interpretation of Kemeny rank aggregation to derive a new integer programming formulation that is applicable to non-strict complete and incomplete rankings. The key to this formulation relies on devising a constraint set that ensures that the values of matrix induce a complete and consistent set of preferences, i.e., a complete and non-strict ranking. To this end, we develop a graph-based representation of the ranking-matrix (see Equation (2)) of a non-strict complete ranking.

Definition 3.2 *Let $G=(V,E)$ be unweighted directed graph for representing a non-strict complete ranking $\boldsymbol{r}$ as follows: V is the set of nodes (alternatives) and an each pair of nodes is connected by one or two directed edges $E \subseteq V \times V$ according to the preference relationship between each pair of alternatives. It includes the directed edge (i,j) if $r_i < r_j$ (i.e., $v_i \succ v_j$) and it includes the directed edges (i,j) and (j,i) if $r_i = r_j$ (i.e., $v_i \approx v_j$).*

From Definition 3.2, given a non-strict complete ranking, it is straightforward to construct its digraph (or matrix) representation. However, not every unweighted digraph will correspond to a complete and consistent set of preferences, since certain ones can induce preference cycles. For example, Figure (1a) can be represented via the matrix in Figure (1b), but these representations do not yield a non-strict complete ranking due to the preferential cycle ($v_i \succ v_j$, $v_j \succ v_k$, and $v_k \succ v_i$).



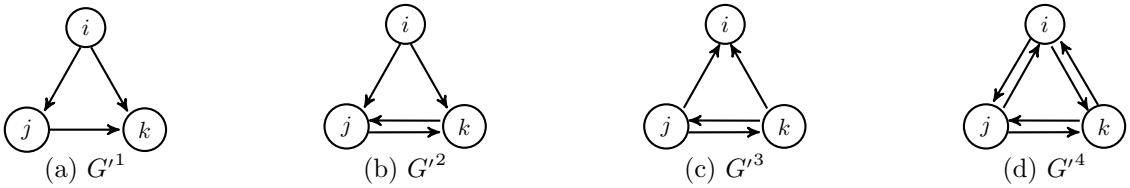
(a) Example of an unweighted directed graph

$$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

(b) Matrix representation of the digraph

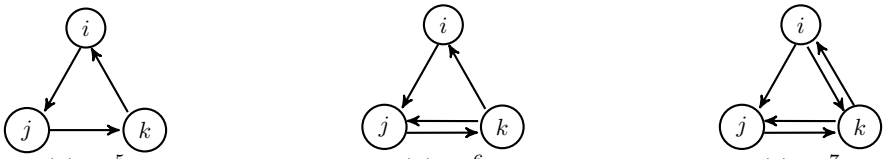
Figure 1 Not every unweighted digraph (or its matrix representation) yields a complete and consistent set of preferences

Hereafter, we define a *ranking-matrix graph* as an unweighted directed graph that induces a non-strict complete ranking (i.e., it does not create any preferential cycles). To identify a ranking-matrix graph structure, certain conditions are needed. For starters, since the solution must be a complete ranking, each pair of nodes in G must be connected by at least one directed edge. The following theorem specifies the remaining conditions for an arbitrary unweighted digraph to be a ranking-matrix graph. To this end, *a uni-cycle* is defined as a simple path that starts and ends on the same vertex in *one* direction *but not in the reverse direction. A bi-cycle* is defined as a simple path that starts and ends on the same vertex and can be traversed in *both* directions. According to these definitions, a bi-cycle and a uni-cycle are mutually exclusive. Additionally, it is not possible to have a uni-cycle of size 2 because, if there exists a directed edge from i to j and a directed edge from j to i, this creates a bi-cycle. The focus of the theorem and proof is to prevent graphs with uni-cycles since such structures can be associated with inconsistent sets of preferences (i.e., non-transitivity). To be more succinct and precise, we denote a graph without uni-cycles as a *uni-cycle-free graph* and a graph with at least one uni-cycle as a *unicyclic graph*. Figures 2 and 3 show the possible uni-cycle-free graphs and unicyclic graphs, respectively, over three alternatives.



(a) G'^1 (b) G'^2 (c) G'^3 (d) G'^4

Figure 2 Unicycle-free graphs



(a) G'^5 (b) G'^6 (c) G'^7

Figure 3 Unicyclic graphs

Theorem 3.3 *Let $G=(V,E)$ be an unweighted directed graph for representing a non-strict complete ranking $\boldsymbol{r}$ as follows: V is the set of nodes (alternatives) and an each pair of nodes is connected by one directed edge, that is, either $(i,j)\in E$ or $(j,i)\in E$ if $r_i<r_j$ (i.e., $v_i\succ v_j$) or two directed edges (i,j) and (j,i) if $r_i=r_j$ (i.e., $v_i\approx v_j$). Graph G is a ranking-matrix graph if and only if it does not contain uni-cycles (i.e., it is a unicycle-free graph).*

Proof. See Appendix A.

The results of this theorem and the following corollary will be used to derive a new integer programming formulation for Kemeny aggregation.

Corollary 3.4 *The ranking-matrix $\mathbf{S}\in\mathbb{Z}^{n\times n}$ induces a complete and consistent set of preferences of n alternatives (i.e., a non-strict complete ranking) if the following constraints are satisfied:*

$$s_{ij}-s_{kj}-s_{ik}\geq -1 \quad i,j,k=1,...,n; \quad i\neq j\neq k\neq i \tag{6a}$$
$$s_{ij}+s_{ji}\geq 0 \quad i,j=1,...,n; \quad i<j \tag{6b}$$
$$s_{ii}=0 \quad i=1,...,n; \tag{6c}$$
$$s_{ij}-2y_{ij}=-1 \quad i,j=1,...,n; \quad i=j \tag{6d}$$
$$s_{ij}\in\mathbb{Z}, y_{ij}\in\mathbb{B} \quad i,j=1,...,n; \tag{6e}$$

Proof. For ranking-matrix $\mathbf{S}=[s_{ij}]$ to represent a complete and consistent set of preferences, the following conditions must be met—note that the definition of s_{ij} is exactly the same as the definition of a_{ij} given by Equation (2). First, the diagonal elements must be set to 0, that is, $s_{ii}=0$, which is represented by constraint (6c). The off-diagonal elements must be non-zero values, specifically, they must be equal to 1 or -1. This is enforced via auxiliary binary variables y_{ij} in constraint (6d). Moreover, s_{ij} and s_{ji} cannot both be negative. Hence, constraint (6b) restricts at least one of s_{ij} and s_{ji} to be positive, when $i\neq j$. Constraint (6e) explicitly states the respective domains of s_{ij} and y_{ij}.

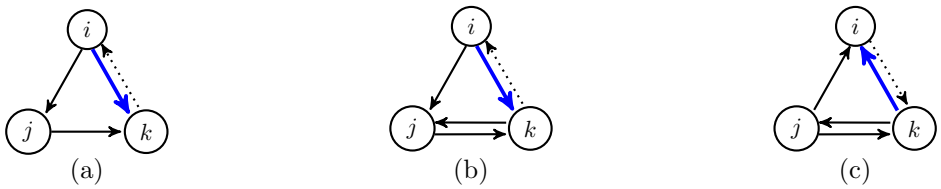


Figure 4 The dotted directed edge creates a uni-cycle

The proof of Theorem 3.3 explains that to check whether a uni-cycle of any length exists within a ranking-matrix digraph, it is sufficient to verify that no uni-cycles of length 3 exist. The possible unicycle-free and unicyclic and digraphs over three alternatives are shown in Figures 2 and 3. Each unicyclic graph of size 3 can be obtained by adding a particular directed edge of one of the unicycle-free graphs shown in Figure 2 or by replacing a particular directed edge with its reverse directed edge. This is depicted Figure 4. Specifically, adding the dotted directed edge or replacing the thick (blue) edge with the dotted directed edge creates a uni-cycle. For example, replacing the thick (blue) edge with the dotted directed edge in Figure (4a) and Figure (4b) yields Figure (3a) and Figure (3b), respectively. Also, adding the dotted directed edge in Figure (4b) yields Figure (3c), which is also a unicyclic graph.

This implies that the unicyclic graphs can be avoided by eliminating these additional or replacement edges from occurring. More specifically, as in Figures (4a) and (4b), whenever there is a directed edge from i to j, but not one from j to i, which gives that $s_{ij}=1$, $s_{ji}=-1$, and a directed edge from j to k (with or without one from k to j), which gives that $s_{jk}=1$ (here, the value of s_{kj} does not matter), the edge between i and k should be directed from i to k (and not in the opposite direction), which gives that $s_{ki}=-1$. This condition can be written as:

$$s_{ji}=-1, s_{jk}=1 \Longrightarrow s_{ki}=-1. \tag{7}$$

Moreover, as in Figure (4c), whenever there is a directed edge from j to i, but not one from i to j, which gives that $s_{ji}=1, s_{ij}=-1$, and a directed edge from k to j (with or without one from j to k), which gives that $s_{kj}=1$ (here, the value of s_{jk} does not matter), the edge between k and i should be directed from k to i (and not in the opposite direction), which gives that $s_{ik}=-1$. This condition can be written as:

$$s_{ij}=-1, s_{kj}=1 \Longrightarrow s_{ik}=-1. \tag{8}$$

Conditions (7) and (8) can be equivalently satisfied via the following linear constraints:

$$s_{ji}-s_{jk} \geq s_{ki}-1 \tag{9}$$

$$s_{ij}-s_{kj} \geq s_{ik}-1. \tag{10}$$

In fact, the formulation can be further reduced. By swapping labels i and j in constraint (10) (since it holds for any permutation of the labels), we can derive constraint (9) (i.e., it is redundant). Therefore, constraints (6a)-(6e) provide the full set of constraints. ■

3.3. Deriving a new binary programming formulation

We develop the first exact integer programming formulation for generalized Kemeny aggregation—that is, for input rankings that are complete and incomplete, with and without ties. The formulation is given by:

$$\begin{aligned} \underset{\mathbf{S}}{\text{maximize}} \quad & \sum_i \sum_j c_{ij} s_{ij} \\ \text{subject to} \quad & \text{constraints (6a)-(6e)} \end{aligned}$$

where $[c_{ij}] \in \mathbb{Z}^{n \times n}$ is the cumulative ranking-matrix of the input rankings, defined as $c_{ij} = \sum_{\ell=1}^{|L|} a_{ij}^{\ell}$, for $i, j = 1, ..., n$ (see Equation (2)), when the rankings are complete. When they are incomplete, the cumulative ranking-matrices defined by Yoo et al. (2020) can be utilized, which correspond to the incomplete-ranking distances introduced in Dwork et al. (2001) and Moreno-Centeno and Escobedo (2016). For the former it is defined as $c_{ij} = \sum_{l=1}^{|L|} = \frac{a_{ij}^{l}}{n(n-1)}$, where n is the total number of alternatives, and for the latter it is defined as $c_{ij} = \sum_{\ell=1}^{|L|} \frac{a_{ij}^{\ell}}{n^{\ell}(n^{\ell}-1)}$, where n^{ℓ} is the number of each alternative evaluated by judge ℓ. In effect, these expressions normalize the ranking-matrix values of each judge according to the total number of alternatives or to the number of alternatives evaluated by each judge. The formulation can be adapted for other incomplete-ranking measures (distances or correlation coefficients) that can be summarized via a respective ranking-matrix.

The Generalized Kemeny-aggregation Binary Programming formulation (GKBP) is obtained by substituting for s_{ij} with $(2y_{ij} - 1)$ in the integer programming formulation, which gives:

$$\begin{aligned} \underset{\mathbf{y}}{\text{maximize}} \quad & \sum_i \sum_j c_{ij}(2y_{ij} - 1) && && \text{(11a)} \\ \text{subject to} \quad & y_{ij} - y_{kj} - y_{ik} \geq -1 && i, j, k = 1, ..., n; \quad i \neq j \neq k \neq i && \text{(11b)} \\ & y_{ij} + y_{ji} \geq 1 && i, j = 1, ..., n; \quad i < j && \text{(11c)} \\ & y_{ij} \in \mathbb{B} && i, j = 1, ..., n; \quad i \neq j. && \text{(11d)} \end{aligned}$$

Combining Equation (6d) with the definition of ranking-matrix $[s_{ij}]$ gives the implicit definition of y_{ij}, which represents the ordinal relationship between alternatives v_i, v_j. Upon inspection, GKBP has $n^2 - n$ variables while Brancotte et al. (2015)'s formulation has $\frac{3}{2}n^2 - 2n$ variables; additionally, the new formulation has $n(n–1)(n–2)/2$ fewer constraints. Hence, GKBP has $O(n^2)$-fewer variables and $O(n^3)$-fewer constraints than Brancotte et al. (2015)'s formulation.

On a more fundamental level, GKBP can be connected to the theory of order polytopes. An order polytope $\mathbf{P_O^n}$ is the convex hull of vertices that represent the possible members of a specific type of binary relation on n alternatives. Notable examples are the linear order polytope $\mathbf{P_{LO}^n}$—the convex hull of binary relations that are total, irreflexive, and transitive—and the weak order polytope $\mathbf{P_{WO}^n}$—the convex hull of binary relations that are total, reflexive, and transitive—since their vertices correspond to strict complete rankings and non-strict complete rankings, respectively.

Previous works formulated Kemeny aggregation for strict complete rankings as a special case of the formulation of the linear ordering problem, whose aim is to find a linear ordering that maximizes the sum of weights c_{ij} in a weighted directed graph (Newman and Vempala 2001, Martí and Reinelt 2011). The ensuing theorem makes an analogous connection between GKBP and the weak order polytope. To the best of our knowledge, this is the first work to establish such a connection.

Theorem 3.5 *The GKBP constraints provide a logically equivalent formulation of the weak order polytope.*

Proof. See Appendix B.

From this theorem, the underpinnings of the GKBP formulation are strengthened through their connection with the theory of order polytopes. In fact, the GKBP constraints are equivalent to the basic family of facet-defining inequalities (see Fiorini and Fishburn (2004)). This connection gives the formulation inherent computational advantages since the facet-defining inequalities of $\mathbf{P^n_{WO}}$ could help obtain tighter lower bounds for the Kemeny aggregation problem within the branch and bound algorithm, thereby expediting solution times (Nemhauser and Wolsey 1988).

4. Computational studies of GKBP

The computational studies compare the performance of two formulations for rank aggregation: GKBP (formulation (11) in Section 3) and the modified version of Brancotte et al. (2015) stated in Proposition 3.1. The test datasets consist of probabilistic instances constructed based on the concept of Mallows distribution (see Section 4.1) and benchmark instances from PrefLib, a library of preference data (Mattei and Walsh 2013). Prior to describing the experiments, recall that GKBP finds a ranking that maximizes agreement quantified according to the Kendall τ-extended correlation coefficient and the modified Brancotte et al. (2015) model finds a ranking that minimizes disagreement quantified according to d_{KS}. Due to the connection of this distance-correlation coefficient pairing (see Equation (4)), the two respective problems are equivalent, which allows for a fair comparison of their performance.

The experiments were performed on machines equipped with 36GB of RAM memory shared by two Intel Xeon E5-2680 processors running at 2.40 GHz; code was written in Python and the formulations were solved using CPLEX solver version 12.8.0 (IBM Knowledge Center 2017).

4.1. Instances from probabilistic distributions

The formulations are first tested on randomized instances constructed from rankings sampled from a probabilistic distribution with an underlying ground truth and an adjustable level of noise/error. This choice allows for the generation of instances with differing levels of difficulty, thereby enabling a systematic comparison of the formulations. Among the existing options for generating randomized

rankings, the Mallows-ϕ model (Mallows 1957, Diaconis 1988, Marden 1996, Critchlow 2012) is the most popular and has been used similarly in other works—e.g., (Lu and Boutilier 2014, Betzler et al. 2014, Asfaw et al. 2017, Crispino et al. 2019, Yoo et al. 2020).

The Mallows-ϕ model is a Kendall-τ distance-based model (i.e., the Kemeny-Snell distance-based model when the rankings are strict and complete), which is parameterized by a "ground truth" (or reference) ranking $\underline{\boldsymbol{a}}$ and "dispersion" $\phi \in (0, 1]$. These parameters are used to quantify the probability of obtaining a complete ranking $\boldsymbol{a}$ as:

$$P(\boldsymbol{a}) = P(\boldsymbol{a}|\underline{\boldsymbol{a}}, \phi) = \frac{\phi^{d_\tau(\boldsymbol{a},\underline{\boldsymbol{a}})}}{\sum_{\boldsymbol{a}\in\Omega_C} \phi^{d_\tau(\boldsymbol{a},\underline{\boldsymbol{a}})}},$$

where Ω_C is the space of complete rankings. When sampling from this distribution, as ϕ gets closer to 0, the generated ranking converges to $\underline{\boldsymbol{a}}$; as ϕ gets closer to 1, any complete ranking has equal probability of occurring (i.e., this becomes the uniform distribution). Prior works have developed efficient algorithms for sampling rankings from the Mallows-ϕ model (Doignon et al. 2004, Ceberio et al. 2015, Irurozki et al. 2016). We use a slightly modified version of the repeated insertion model of Doignon et al. (2004), which was originally designed for strict complete rankings. The next paragraph describes how ties and incompleteness are added to the rankings generated by the repeated insertion model so as to provide suitable instances for testing the featured formulations.

To generate non-strict rankings, a random number u is repeatedly drawn from a discrete uniform distribution $U(1, h-1)$, where h is the highest-valued (worst) rank position in the current ranking. The alternative in rank position u is tied with the alternative in the next rank position higher than u. Ties are repeatedly inserted until the number of tied alternatives reaches a specified threshold, which is set to $0.5n$. For example, let $\boldsymbol{a} = (1, 2, 3, 3, 5)$ and $u = 3$. The next rank position higher than u is 5, and v_5 is the alternative with this rank. Therefore, $\boldsymbol{a}$ becomes $(1, 2, 3, 3, 3)$, and the process stops because the number of tied alternatives reaches the threshold (i.e., $3 > 0.5 \cdot 5 = 2.5$).

To generate incomplete rankings, we utilize the repeated insertion model extension introduced in (Yoo et al. 2020). This model applies the repeated insertion model on a subset $V' \subset V$ and then marks the alternatives $V \backslash V'$ as unranked. Ties are inserted to incomplete rankings using the same procedure as with complete rankings but restricted to the alternatives in V'. The pseudocode for generating non-strict and incomplete ranking instances can be found in Appendix D.

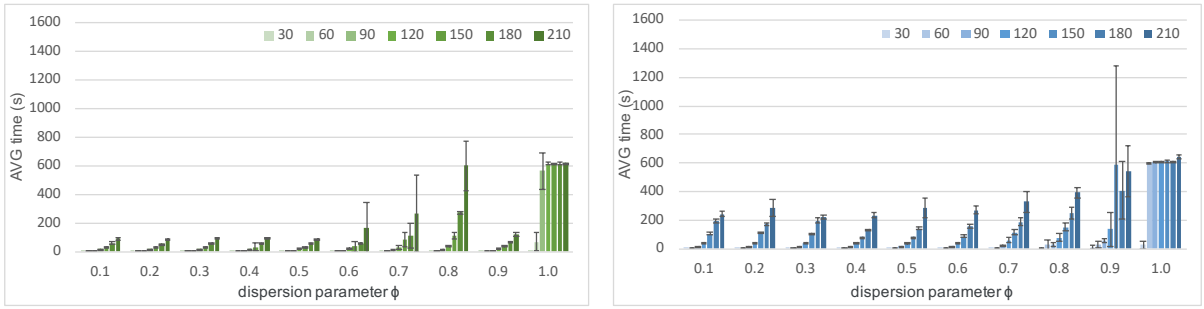
4.2. Configurations of probabilistic distribution experiments

The first set of instances is constructed and guided based on the concept of Mallows distribution; specifically, instances are obtained by sampling complete rankings from the Mallows-ϕ distribution and inserting ties and/or incompleteness, as described in the preceding subsection. We first investigate the effect of varying the dispersion parameter, $\phi \in \{0.1, 0.2, ..., 0.9, 1.0\}$, and the number of alternatives, $n \in \{30, 60, 90, ..., 210\}$; the number of judges is fixed to 50.

For each of the parameter configurations detailed above, in all upcoming experiments, the computing times of each formulation are individually recorded for 10 corresponding instances, which are summarized via average (AVG) and standard deviation (SD) values (represented via error bars). When a formulation cannot return an optimal solution within a 600-second (10-minute) time limit, the relative optimality gap is recorded (a solution is considered as optimal when the relative optimality gap is less than equal to 0.0001). Note that the relative optimality gap for a maximization problem is defined as:

$$\text{Relative optimality gap} = \frac{\text{best relaxation bound} - \text{objective function value for best integer solution}}{\text{objective function value for best integer solution+1e-10}}$$

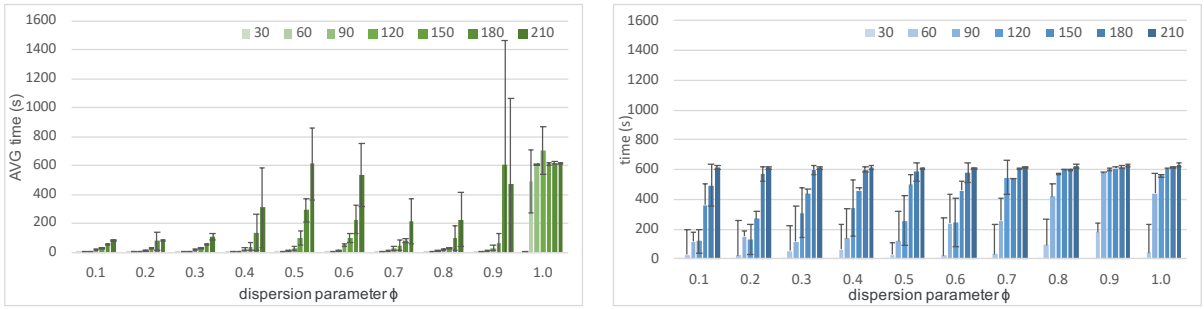
For example, for an instance solved with one of the featured formulations, the objective function value of the current best integer solution was 0.312 and the relative optimality gap was 0.842; this indicates that the objective function value of the optimal solution could be as high as 0.575. As shown in Figure 5 and 6, the computing times for some instances exceed 600 seconds; this occurs because CPLEX can be slow to terminate when a new incumbent solution is found close to the time limit (IBM Support 2019).



(a) Computing times via GKBP (b) Computing times via Brancotte et al. (2015)

Figure 5 **Non-strict complete rankings: GKBP solves all probabilistic instances within the time limit and significantly faster, even for larger n**

As shown in Figure 5, for non-strict complete ranking instances, GKBP finds the optimal solution in less time than Brancotte et al. (2015)'s formulation for most values of n and ϕ. In general, for both formulations, computing times increase with the value of ϕ and n. GKBP returns the optimal solution for all instances within the time limit, except for $\phi = 1.0$ with $n \geq 90$, while Brancotte et al. (2015) is not able to solve some instances with $\phi = 0.7$ with $n = 150$, $\phi = 0.8$ with $n = 210$, and most instances with $\phi \geq 0.9$. Despite the fact that GKBP found an optimal solution faster than Brancotte's model for most of the tested instances, when both reached the time limit without an optimal solution, the optimality gaps of GKBP were at times larger. For example, the average relative optimality gaps over the 10 instances with $\phi = 1.0$ and $n = 90$ were 2.99 for GKBP and 1.00 for Brancotte et al. (2015)'s model.



(a) Computing times via GKBP (b) Computing times via Brancotte et al. (2015)

Figure 6 **Non-strict incomplete rankings: Brancotte et al. (2015)'s formulation cannot reach optimality within the time limit for most instances with higher ϕ and n**

Figure 6 displays the computing times for non-strict incomplete ranking instances. Compared to complete rankings, it takes longer to reach optimality for most values of n and ϕ. Similar to the prior results, GKBP reaches optimality in a shorter amount of time than Brancotte et al. (2015). GKBP attains the optimal solution except four instances with $\phi = 0.9$ and $n = 180, 210$ and most instances with $\phi = 1.0$ and $n \geq 60$, while Brancotte et al. (2015) cannot solve most instances, except those with $n = 30, 60$ for all ϕ. For example, for an instance with $\phi = 0.9$ and $n = 210$, the relative optimality gap of GKBP is 0.0001, which is considered as optimal, while that of Brancotte et al. (2015) is 1.00.

4.3. Instances from Preflib benchmark dataset

The second set of instances is selected from the library of preference data Preflib (Mattei and Walsh 2013), specifically the "Order with Ties - Complete List (TOC)" dataset. This benchmark dataset consists of 378 instances with differing numbers of alternatives and rankings (i.e., judges) obtained from various domains, and they include real-world data (e.g., figure skating competitions, crosscountry skiing and ski jump championship results). The instances include results from Formula One racing and human computation activities, which tend to be relatively less subjective and possess a higher level of collective similarity, as well as data from elections and pure preferences (e.g., the Sushi data set), which tend to be more subjective and possess a lower level of collective similarity. For instance, Milosz and Hamel (2018) estimated the Mallows dispersion parameter ϕ of the "Websearch" instance to be 0.0265, which is a relatively low value. Although we do not have information on the specific dispersion values apart from this single instance, Preflib instances encompass a wide range of subjectivity, meaning the inputs are expected to have varying degrees of collective similarity, as suggested by this discussion.

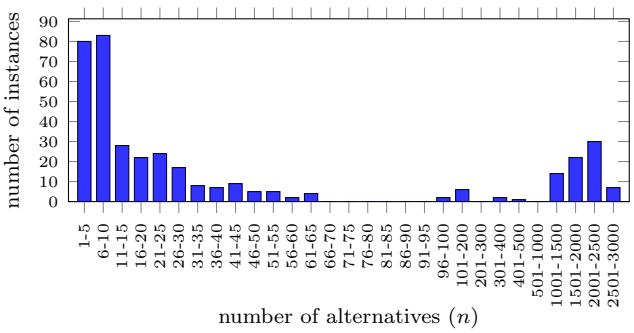


Figure 7 **The number of alternatives (n) of most instances in the Preflib dataset ranges between 3 and 65**

Figure 7 summarizes the distribution of the instances according to ranges of n (number of alternatives). As shown in the figure, most instances have $n \leq 65$, but there are a few instances with $n \geq 1000$.

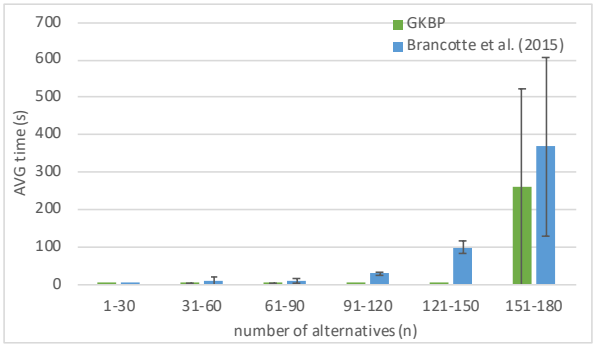


Figure 8 **GKBP is faster, on average, than Brancotte et al. (2015)'s formulation over all instances with $n \leq 300$, grouped according to the ranges of n**

For this experiment, only the instances with $n < 300$ are considered; many instances with $n > 300$ resulted in termination likely due to insufficient memory. In all, there are 302 instances with $3 \leq n \leq 170$; for clarity, the instances are grouped by intervals of 30 over the range of n in Figure 8. The longest computing time for GKBP is 338.94 seconds when there are 170 alternatives, while the time limit is reached with an relative optimality gap of 0.46, on average, for Brancotte et al. (2015). From this analysis, it is evident that GKBP can also solve benchmark problems noticeably faster.

5. A structural social choice property for complete rankings with ties

To further expedite the solution process of Kemeny rank aggregation, this section will devise a structural social choice property, which enables the decomposition certain large problems into a collection of smaller subproblems.

5.1. The Condorcet criterion

Condorcet (1785) proposed a social choice property, which has come to be aptly known as the Condorcet Criterion (CC), that states that if a majority of voters prefers one alternative in pairwise fashion over all others, that alternative alone should obtain the best position in the voting outcome. Formally, recalling that p_{ij} is the number of judges who prefer alternative v_i over alternative v_j, CC can be written as:

$$\text{If } \exists v_i \in V, \text{ s.t. } p_{ij} > p_{ji} \text{ (i.e., } v_i \overset{m}{\succ} v_j) \ \forall v_j \in V \backslash \{v_i\} \implies v_i^* \succ v_j^*, \text{ or equivalently, } \boldsymbol{r}_i^* < \boldsymbol{r}_j^*,$$

where $\boldsymbol{r}^*$ is the final aggregate ranking and $v_i^* \succ v_j^*$ indicates that v_i is ranked strictly better than v_j in $\boldsymbol{r}^*$. CC provides limited usefulness for decision-making since it only applies to the winning alternative or to the losing alternative (i.e., it specifies a rule to select the Condorcet winner/loser). An extended version of the Condorcet winner is the Smith set; the winning (losing) Smith set is the smallest non-empty set of alternatives that defeats (is defeated by, resp.) every alternative outside the set in a pairwise election (Smith 1973). Truchon et al. (1998) provided another natural extension of CC, called the Extended Condorcet Criterion (XCC). This property requires that if V can be organized into a partition $\mathcal{V} := \{V^1, ..., V^K\} \in \mathcal{P}(V)$, such that all alternatives in subset $V^k \in \mathcal{V}$ are pairwise preferred over all alternatives in subset $V^{k'} \in \mathcal{V}$ by a majority (i.e., $V^k \overset{m}{\succ} V^{k'}$), where $k < k'$, then the alternatives in V^k must be ranked strictly better than all alternatives in $V^{k'}$ in the optimal ranking. Formally, this property can be written as:

$$\begin{aligned}&\text{If } \exists \mathcal{V} := \{V^1, V^2, ..., V^K\} \in \mathcal{P}(V), \text{ s.t. } V^k \overset{m}{\succ} V^{k'}, \text{ for } 1 \le k < k' \le K \\ &\implies v_i^* \succ v_j^*, \text{ or equivalently, } \boldsymbol{r}_i^* < \boldsymbol{r}_j^*, \ \forall v_i \in V^k, \ \forall v_j \in V^{k'}.\end{aligned}$$

Table 2 illustrates how XCC can be applied to the Kemeny aggregation problem. In the example, since $\mathcal{V}^{\text{XCC}} := \{\{v_1, v_2\}, \{v_3\}, \{v_4\}\}$ is a partition satisfying XCC, the Kemeny optimal ranking is expected to place v_1 and v_2 ahead of v_3 and v_4, and to place v_3 ahead of v_4 (the optimal ordering between v_1 and v_2 cannot be determined from the application of this property alone). As shown in the table, the Kemeny optimal rankings (three in this case) are all consistent with XCC. Note that Kemeny rank aggregation (as well as other distance-based methods) may yield more than one optimal solution (Young and Levenglick 1978, Dwork et al. 2001). Indeed, Muravyov (2014) explained that the number of optimal solutions in Kemeny aggregation can at times be greater

									Kemeny optimal rankings		
	$\boldsymbol{a}^1$	$\boldsymbol{a}^2$	$\boldsymbol{a}^3$	$\boldsymbol{a}^4$	$\boldsymbol{a}^5$	$\boldsymbol{a}^6$	$\boldsymbol{a}^7$	$\boldsymbol{a}^8$			
v_1	1	2	1	2	1	2	1	2	1	1	2
v_2	2	1	2	1	2	1	2	1	1	2	1
v_3	3	3	3	4	4	4	3	3	3	3	3
v_4	4	4	4	3	3	3	4	4	4	4	4

Table 2 The Kemeny optimal solutions are consistent with XCC

than the number of input rankings and that these solution rankings may rank the alternatives in significantly different ways, which can lead to ambiguity—this is called *Paradox of Kemeny*. This is explained to a great extent by the fact that the Kemeny rank aggregation problem can be characterized as finding the median ranking among the given set of rankings. Medians do not need to be unique and, therefore, Kemeny optimal rankings are not guaranteed to be unique as well (Kemeny and Snell 1962).

	$\boldsymbol{a}^1$	$\boldsymbol{a}^2$	$\boldsymbol{a}^3$	$\boldsymbol{a}^4$	$\boldsymbol{a}^5$	$\boldsymbol{a}^6$	$\boldsymbol{a}^7$	$\boldsymbol{a}^8$	$\boldsymbol{a}^9$	Kemeny optimal ranking	XCC solution
v_1	1	1	1	1	1	1	2	2	2	1	1
v_2	1	3	3	3	2	2	1	1	1	1	2
v_3	3	1	1	1	3	2	2	2	2	1	3

Table 3 The optimal solution is not consistent with CC and XCC

Although both CC and XCC have been implemented to refine the complexity of Kemeny aggregation (i.e., providing parameterized complexity with respect to the number of subsets and the size of subsets in $\mathcal{V}^{\text{XCC}}$), they are not appropriate for non-strict rankings. In the example showcased in Table 3, we have that $p_{12}=5$, $p_{21}=3$, $t_{12}=1$, $p_{23}=5$, $p_{32}=3$, $t_{23}=1$, $p_{13}=3$, $p_{31}=0$ and $t_{13}=6$. According to the definition of XCC, the final optimal solution is expected to be $(1,2,3)$, since $\mathcal{V}^{\text{XCC}} := \{\{v_1\},\{v_2\},\{v_3\}\}$ is a partition satisfying XCC—that is, $p_{ij} > p_{ji}$ for $1 \le i < j \le 3$. However, v_1, v_2, and v_3 are tied in the aggregate ranking when optimizing with the Kemeny-Snell distance and allowing ties. Effectively, this implies that Kemeny aggregation for non-strict rankings is not consistent with XCC. In order to overcome this inadequacy, we define a new social choice property in the ensuing subsection.

5.2. The Non-strict extended Condorcet criterion and its structural decomposition

To introduce a substitute to XCC that is suitable for non-strict rankings, we first define an important concept. We say that *a decisive majority* prefers an alternative v_i over an alternative v_j, written as $v_i \overset{M}{\succ} v_j$, if $p_{ij} > p_{ji} + t_{ij}$, that is, the number of people who prefer v_i over v_j is greater than the number of people who prefer v_j over v_i plus those who tie them. If neither $v_i \overset{M}{\succ} v_j$ nor $v_j \overset{M}{\succ} v_i$, there is no decisive majority that prefers v_i over v_j, and vice versa, written as $v_i \overset{M}{\approx} v_j$. Similarly, we say that a decisive majority prefers all alternatives in the partition V^k over all alternatives in the partition $V^{k'}$, written as $V^k \overset{M}{\succ} V^{k'}$, if $p_{ij} > p_{ji} + t_{ij}$, $\forall v_i \in V^k, \forall v_j \in V^{k'}$. If neither $V^k \overset{M}{\succ} V^{k'}$ nor $V^{k'} \overset{M}{\succ} V^k$, there is no decisive majority that prefers V^k over $V^{k'}$, and vice versa, written as $V^k \overset{M}{\approx} V^{k'}$.

Definition 5.1 *Let $\mathcal{V} := \{V^1, V^2, ...V^K\}$ s.t. $V^k \overset{M}{\succ} V^{k'}$ for $1 \le k < k' \le K$. The Non-strict Extended Condorcet Criterion (NXCC) requires that all $v_i \in V^k$ must precede all $v_j \in V^{k'}$ in the final ranking. That is,*

$$\begin{aligned} &\textit{If } \exists \mathcal{V} := \{V^1, V^2, ..., V^K\} \in \mathcal{P}(V),\ \textit{s.t. } V^k \overset{M}{\succ} V^{k'},\ \textit{ for } 1 \le k < k' \le K \\ &\quad \Longrightarrow v_i^* \succ v_j^* \textit{ or equivalently, } \boldsymbol{r}_i^* < \boldsymbol{r}_j^*,\ \forall v_i \in V^k,\ \forall v_j \in V^{k'}. \end{aligned}$$

A basic implication of this property is that, the more subsets the partition has, the smaller the sizes of the subproblems that need to be solved (since each subset will tend to have fewer alternatives). Note that, when all rankings are strict and complete, NXCC is exactly XCC, because $t_{ij} = 0$.

To apply NXCC, it is necessary to determine the ordered partition of subsets of alternatives—in which lower-indexed subsets are each preferred over higher-indexed subsets by a decisive majority—from the data. This can be done via Algorithm 3 in Appendix E, which has a worst-case complexity of $O(n^2)$, where n is the number of alternatives.

The main difference between NXCC and XCC is that ties are or are not considered, respectively, to determine the majority's strict pairwise preferences. Specifically, XCC does not consider ties to be relevant to the conclusion that $v_i \in V^k$ should be strictly preferred over all $v_j \in V^k$. On the other hand, NXCC requires that to arrive at this conclusion, the number of judges who strictly prefer v_i over v_j should be greater than those who do not—which includes those who tie them or who strictly prefer v_j over v_i. Table 3 illustrates that the outcome of XCC decomposition is not consistent with Kemeny aggregation for non-strict rankings. Therein, since $t_{13} > p_{13} - p_{31}$ (the number of judges who tie v_1 and v_3 is greater than the net difference between the number of judges who have a strict preference), it cannot be concluded that v_1 should be ahead of v_3 in the final optimal ranking. Hence, these two alternatives and every other alternative between them cannot be ordered a priori into separate subsets, which is the outcome obtained when NXCC is applied to the example. The ensuing paragraphs formally prove that the Kemeny optimal solution satisfies NXCC when the rankings are non-strict and complete. This is done through Lemma 5.2 and Theorem 5.3. Beforehand, it is useful to define some additional notation.

Notation 1 *The reduced instance associated with two subsets V^k, $V^{k'} \subset V$, written as $A_{[k\cup k']} = A_{[V^k \cup V^{k'}]}$ is the submatrix induced by rows $V^k \cup V^{k'}$ of A. Similarly, $\boldsymbol{a}^{\ell}_{[k\cup k']} = \boldsymbol{a}^{\ell}_{[V^k\cup V^{k'}]}$ and $\boldsymbol{r}^{*}_{[k\cup k']} = \boldsymbol{r}^{*}_{[V^k\cup V^{k'}]}$ are the reduced evaluation from judge ℓ and the optimal reduced ranking with respect to $V^k \cup V^{k'}$, respectively.*

When exactly two alternatives are considered, this notation is modified as follows.

Notation 2 *The reduced instance $A_{\{i,j\}} = A_{\{v_i,v_j\}}$ is the submatrix induced by alternatives v_i and v_j. Similarly, $\boldsymbol{a}^{\ell}_{\{i,j\}} = \boldsymbol{a}^{\ell}_{\{v_i,v_j\}}$ and $\boldsymbol{r}^{*}_{\{i,j\}} = \boldsymbol{r}^{*}_{\{v_i,v_j\}}$ are the reduced evaluation from judge ℓ and the optimal reduced ranking with respect to v_i and v_j, respectively.*

Furthermore, the notation is combined to specify the ranking position of single alternative within a reduced problem space.

Notation 3 *The ordinal position of v_i in the reduced evaluation from judge $\boldsymbol{a}^{\ell}$ with respect to $V^k \cup V^{k'}$ is denoted as $a^{\ell}_{i|[k\cup k']}$.*

Using this notation, the cumulative distance between $\{\boldsymbol{a}^\ell\}_{\ell=1}^{|L|}$ and $\boldsymbol{r}$, accrued by only alternatives v_i and v_j, can be written as $\sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}}, \boldsymbol{r}_{\{i,j\}})$. Similarly, the distance between $\{\boldsymbol{a}^\ell\}_{\ell=1}^{|L|}$ and $\boldsymbol{r}$, accrued by all alternatives in V^k and all alternatives in $V^{k'}$, can be written as $\sum_{\ell\in L} d(\boldsymbol{a}^\ell_{[k\cup k']}, \boldsymbol{r}_{[k\cup k']})$.

Lemma 5.2 *Let $\{V^1, V^2\} \in \mathcal{P}(V')$ (i.e., $\{V^1, V^2\}$ is a bipartition of V'). Consider the reduced aggregation problem consisting of input rankings $A_{[1\cup 2]}$, which include only the evaluations over $V^1 \cup V^2$. If $V^1 \overset{M}{\succ} V^2$, every alternative $v_i \in V^1$ should obtain a better position than every alternative $v_j \in V^2$ in the optimal solution to the reduced problem, that is, $r^*_{i|[1\cup 2]} < r^*_{j|[1\cup 2]}$, where $\boldsymbol{r}^*_{[1\cup 2]} := \underset{\boldsymbol{r}_{[1\cup 2]}}{\arg\min} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{[1\cup 2]}, \boldsymbol{r}_{[1\cup 2]})$.*

Proof. See Appendix C.

Theorem 5.3 *Kemeny aggregation satisfies NXCC when inputs are non-strict and complete.*

Proof. Define the partition $\mathcal{V} = \{V^1, V^2, ..., V^K\}$, where $V^k = \{v^k_1, v^k_2, ..., v^k_{|V^k|}\}$ and assume $V^k \overset{M}{\succ} V^{k'}$ for every k, k', where $1 \le k < k' \le K$. Let $\boldsymbol{r}^*$ be a Kemeny optimal ranking. In order to prove that Kemeny aggregation satisfies NXCC when rankings are non-strict and complete—which means that all elements in V^k should precede all elements in $V^{k'}$ in $\boldsymbol{r}^*$, the Kemeny-Snell distance between V^k and $V^{k'}$ can be calculated for all $k < k'$. To this end, the cumulative Kemeny-Snell distance is expanded as follows:

$$\begin{aligned}
\sum_{\ell\in L} d(\boldsymbol{a}^\ell, \boldsymbol{r}^*) &= \sum_{i=1}^{|V|-1} \sum_{j=i+1}^{|V|} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}}, \boldsymbol{r}^*_{\{i,j\}}) \\
&= \sum_{k=1}^{K} \sum_{k'=1}^{K} \sum_{i=1}^{|V^k|} \sum_{j=1}^{|V^{k'}|} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}}, \boldsymbol{r}^*_{\{i,j\}}) \\
&= \underbrace{\sum_{k=1}^{K} \sum_{i=1}^{|V^k|-1} \sum_{j=i+1}^{|V^k|} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}}, \boldsymbol{r}^*_{\{i,j\}})}_{\text{within subset } V^k \text{ (intrasubset distances)}} + \underbrace{\sum_{k=1}^{K-1} \sum_{k'=k+1}^{K} \sum_{i=1}^{|V^k|} \sum_{j=i}^{|V^{k'}|} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}}, \boldsymbol{r}^*_{\{i,j\}})}_{\text{between subsets } V^k \text{ and } V^{k'} \text{ (intersubset distances)}} .
\end{aligned}$$

Note that if all $v_i \in V^k$ can be proved to be strictly preferred or strictly dispreferred over all $v_j \in V^{k'}$ in the optimal solution, for all $k < k'$, this guarantees an optimal ordered partition $\mathcal{V} = \{V^1, V^2, ..., V^K\}$, with $V^1 \succ V^2 \succ ... \succ V^K$, that allows us to ignore or postpone ordering the alternatives within each subset (i.e., between $v_i, v_{i'} \in V^k$). This is the focus of the remainder of the proof. Now, we can obtain the following bound involving the optimal cumulative intersubset distance:

$$\min_{\boldsymbol{r}^*} \sum_{k=1}^{K-1} \sum_{k'=k+1}^{K} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{[k\cup k']}, \boldsymbol{r}^*) \ge \sum_{k=1}^{K-1} \sum_{k'=k+1}^{K} \min_{\boldsymbol{r}^*_{[k\cup k']}} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{[k\cup k']}, \boldsymbol{r}^*_{[k\cup k']}). \tag{12}$$

Notice the corresponding optimal solutions of the right-hand side of inequality (12) give all 2-subset optimal reduced orderings from $\{V^1, V^2, ..., V^K\} \in \mathcal{V}$, which can produce preference cycles (i.e., contradictions) when combined. For example, assuming $v_i \in V^1, v_j \in V^2, v_k \in V^3$, the optimal solutions to the reduced problems can be $r^*_{i|[1\cup2]} < r^*_{j|[1\cup2]}$, $r^*_{j|[2\cup3]} < r^*_{k|[2\cup3]}$ and $r^*_{k|[1\cup3]} < r^*_{i|[1\cup3]}$. On the other hand, the corresponding optimal solution of the left-hand side of inequality (12) gives a Kemeny optimal ordering of the K subsets. In other words, the collection of right-hand side subproblems can be interpreted as a relaxed version of the left-hand side rank aggregation problem. In particular, the relaxed problem does not enforce the preference transitivity between all subsets orderings. However, when the optimal solutions to the 2-subset optimal reduced orderings yield a combined solution that is feasible to the left-hand side problem, an optimal solution for the original rank aggregation problem has been found, since the objective values will also be equal.

By Lemma 5.2, when $V^k \overset{M}{\succ} V^{k'}$, we know that the corresponding optimal solution to the reduced problem induced by each pair V^k and $V^{k'}$ places every alternative $v_i \in V^k$ ahead of every alternative $v_j \in V^{k'}$, for all $k < k'$. Hence, the individual 2-subset orderings can be combined into a K-subset partial ordering $\mathcal{V}^* = \{V^1, V^2, ..., V^K\}$ with $V^1 \succ V^2 \succ ... \succ V^K$, without contradictions or preferences cycles. In particular, the combined solution to the reduced subproblems is feasible to the original problem and $r^*_i < r^*_j$, where $v_i \in V^k$, $v_j \in V^{k'}$, for all $k < k'$, giving an optimal partial ranking of all alternatives in V. Hence, Kemeny aggregation satisfies NXCC when the rankings are non-strict and complete. ■

Besides offering significant potential computational advantages, NXCC brings various practical benefits for decision-making. This includes the ability to focus on the most relevant alternatives. It is usually not known a priori which alternatives will be the ones to occupy the top, middle, or bottom positions in the consensus ranking; the exact positions are ultimately revealed through the aggregation process. One exception by which it may be possible to know such information ahead of time is through the application of the NXCC property developed in this work. In particular, NXCC takes advantage of the pairwise comparison information to determine whether there exist subsets of alternatives that will always be preferred over other subsets of alternatives. By determining the NXCC partition, decision-makers can focus on the exact ordering of the alternatives contained in just the top and/or bottom subsets. That is, alternatives that belong to subsets in the middle of the partition can be dropped from consideration while formally guaranteeing that the relative ordering of the remaining alternatives will not be affected. An additional related benefit is the ability to certifiably rule out certain outcomes even when the consensus ranking is not unique. That is, even if an instance may have multiple alternative optimal solutions but only one is obtained by the exact solution method, the NXCC decomposition would guarantee that alternatives in higher-indexed subsets will never be ranked ahead of alternatives in lower-indexed subsets. Put otherwise, it could help determine that many alternatives will never occupy the top positions of any optimal ranking.

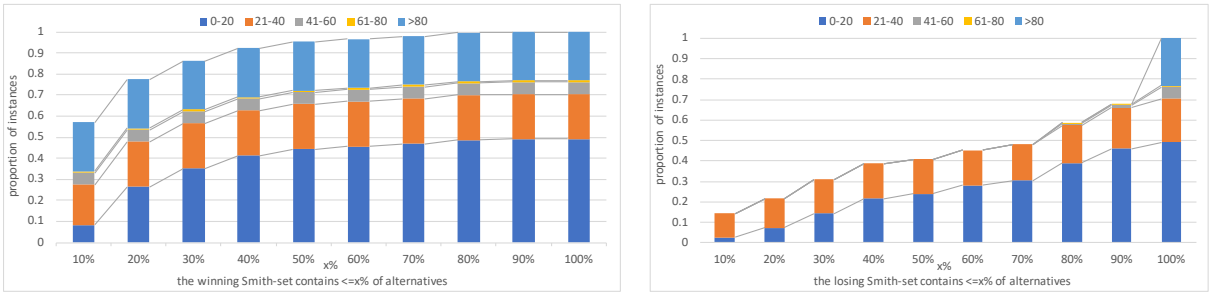
6. Computational studies of NXCC decomposition

The purpose of this experiment is to provide an estimate of how NXCC could improve the computing time of Kemeny aggregation via partitioning and bring about other practical benefits. To highlight the practicality of the property, this work experiments on the Preflib benchmark dataset consisting of 378 instances described in Section 4.3. During the experiment, the following information is recorded: (1) the existence of a Condorcet winner and loser, (2) whether the partition is non-trivial after the decomposition (i.e., $|\mathcal{V}| \geq 2$, where $|\mathcal{V}|$ is the number of subsets in the partition), and (3) the size of the winning Smith set and the losing Smith set (see Section 5.1); Table 4 and Figure 9 display this information. Checking the existence of a Condorcet winner and the size of the winning and losing Smith set are two indicators of the effectiveness of NXCC. In particular, a smaller winning Smith set helps to narrow down the winners or most relevant candidates. Likewise, a larger losing Smith set helps decision-makers rule out many irrelevant alternatives, since the complement of the losing Smith set becomes smaller. In addition, practitioners are often most interested in the alternatives that obtain the top positions. For instance, in recommendation system, it is more important to suggest a set of items that is most likely to be preferred (i.e., top alternative sets), rather than suggesting middle or the least preferred items (Davidson et al. 2010). The following analysis helps better understand and quantify the practical benefits of NXCC decomposition.

Key Instance Characteristics	Proportion of Instances
Condorcet winner	191 out of 378 (50.53%)
Condorcet loser	62 out of 378 (16.40%)
Non-trivial partition ($\|\mathcal{V}\| \geq 2$)	230 out of 378 (60.85%)

Table 4 NXCC helps identify the most relevant candidates in the instances from the Preflib dataset

More detailed information about the Smith sets is visualized in Figure 9, which shows the cumulative distribution of the proportion of instances in the dataset with respect to the percentage of alternatives in the winning Smith set and the losing Smith set. Specifically, each bar graph represents the number of instances whose Smith set contains at most $x\%$ of the total number of alternatives. For example, Figure (9a) shows that approximately 58% of instances have a small winning Smith set (i.e., at most $.1n$ alternatives in the winning Smith set); in particular, at most 10% of the alternatives (i.e., $.1n$ alternatives) are contained in the winning (top) Smith set for the majority of the instances with more than 80 alternatives (colored in light blue in the graph). Moreover, Figure (9b) shows that approximately 60% of the instances have a large losing Smith set; in particular, more than 90% of the alternatives are contained in the losing (bottom) Smith set for instances with more than 80 alternatives (colored in light blue in the graph). From these observations, we conclude that NXCC decomposition can significantly simplify the identification of relevant candidates from these benchmark instances.



(a) Cumulative distribution - winning Smith set size (b) Cumulative distribution - losing Smith set size

Figure 9 NXCC yields small winning Smith sets and large losing Smith sets for more than 50% of the instances

Next, we test the computational benefits of the decomposition. To do so, this study compares the computing times of solving the full (non-decomposed) instances (see the computing time of solving these instances in Figure 8) and solving the decomposed instances. The latter include the partitioning time and the time of solving the subproblem for each subset in sequential manner. The experiment was conducted on the PrefLib instances with a non-trivial partition (i.e., $|\mathcal{V}| \geq 2$) and $n \leq 300$ (because CPLEX could not solve non-decomposed instances with more alternatives); the number of instances that meet these conditions is 177. Table 5 shows the *relative improvement* in computing time, which is defined as follows:

$$\frac{(\text{time to solve full (non-decomposed) problem}) - \sum(\text{time to solve decomposed subproblems})}{\sum(\text{time to solve decomposed subproblems})} \times 100$$

$\lvert\mathcal{V}\rvert$	2	3	4	5-10	11-20	21-30
number of instances	72	33	20	21	12	19
relative improvement	25%	44%	72%	65%	67%	96%

Table 5 Applying GKBP for each subset after partitioning reduces the computing time by at least 25%

As shown in Table 5, the computing time is reduced when a higher number of subsets is obtained after the decomposition. For example, when $|\mathcal{V}| = 4$, the computing time is reduced by 72%, on average, whereas the computing time is reduced by 25% on average when $|\mathcal{V}| = 2$. If each subset were solved by multiple processors at the same time, the computing time could be further reduced. Hence, using distributed computing resources, the more finely decomposed $\mathcal{V}$ is, the faster that large instances could be solved with the combination of GKBP and NXCC. It is pertinent to point out, however, that most of the benchmark instances with less than 150 alternatives can be solved within a minute using GKBP (without decomposition). To further assess the potential computational improvements of NXCC, we also apply the decomposition algorithm to non-strict complete ranking instances generated using the procedure described in Section 4.1. Here, we restrict n to larger values, specifically $n \in \{90, 120, 150, 180, 210\}$, and generate 10 instances for each value. Moreover, we fix $\phi = 0.8$, since this is the setting after which certain instances cannot be solved

to optimality by GKBP within ten minutes (see Figure (5a)). In other words, such instances are relatively difficult but they can still be solved by GKBP within a reasonable amount of time.

Table 6 shows the (absolute) improvement in computing time, which is calculated as the difference between the time to solve full (non-decomposed) problem and the cumulative time to solve the decomposed subproblems (i.e., the numerator in the relative improvement calculation).

n	90	120	150	180	210
(absolute) improvement (s)	12.12	47.31	125.84	269.35	$\geq$ 560.42

Table 6 The effectiveness of NXCC is more prominent on larger, more difficult instances

As shown in Table 6, after applying NXCC, computing times improve significantly for every n; most strikingly, decomposed instances with 210 alternatives are solved within few seconds after decomposition, but the original non-decomposed instances were not solved within 600 seconds (note that non-decomposed instances with 210 alternatives are not solved within the time limit, meaning the absolute improvement is at least what is shown on the table). We surmise that NXCC is more effective in these instances because they are moderately difficult to solve.

7. Conclusion

This paper develops a new binary programming formulation and a new social choice property for generalized Kemeny rank aggregation. The major benefit of the generalized Kemeny-aggregation binary programming formulation is that it is applicable to a wide variety of ordinal preferences including complete and incomplete rankings, with and without ties. Additionally, it has fewer variables and constraints compared to a closely related integer programming formulation for the generalized Kendall-τ distance, leading to computational savings, as demonstrated by the featured set of experiments. Furthermore, the paper introduced a computationally expedient social choice property, which unlike the original Condorcet criterion and the extended Condorcet criterion, aligns with the Kemeny consensus ranking when there are ties in the inputs.

The connection to the facet-defining inequalities of the weak order polytope of the binary programming formulation and the structural decomposition enabled by the novel social choice property demonstrate the theoretical rigor of the proposed exact approaches. Moreover, results of the featured experiments on randomized instances and benchmark data show their substantial computational advantages. Through their combination, certain instances that could only be solved approximately can now be solved exactly, even when the input evaluations are relative non-cohesive and/or contain hundreds of alternatives.

The proposed methodologies for rank aggregation can be applied to many practical settings where high-dimensional rankings that may be strict or non-strict and complete or incomplete may be encountered. For example, they could be used to aggregate evaluations of research proposals,

entries into a paper competition, university rankings, and many other applications that require the evaluation of a long list of alternatives. In such cases, the input evaluations are likely to be non-strict and/or incomplete. In addition to group decision-making situations, the presented methodologies could be beneficial for various non-human decision-making contexts that seek to reconcile long heterogeneous ordered lists, such as in metasearch and bioinformatics. Using the social-choice property inspired decomposition along with the GKBP binary programming formulation, decision-makers may be able to solve certain of these large rank aggregation problems exactly.

The generalized Kemeny-aggregation binary programming formulation can be applied to complete and incomplete rankings with ties; however, the Non-strict Extended Condorcet criterion is only applicable to complete rankings with ties. For example, when rankings include too much incomplete information (i.e, many pairs of alternatives are not compared), it is not reasonable to use the current definition of a decisive majority since, for it to be useful, it would require one same preference relation to be made by more than half of judges for each pair of items. If, for example, only two of ten judges evaluate two specific alternatives and their preferences agree on these two alternatives, it is hard to say that there exists a decisive majority because more than three-fourths of judges do not evaluate those two alternatives. Thus, in testing the newly defined social choice property along with the decisive majority, the paper only focused on complete rankings. Future work will involve more computational experiments to test the performance of a suitable social choice property for incomplete rankings. Moreover, to enhance the computational speed of Kemeny aggregation, parallel programming and/or valid inequalities (see Nemhauser and Wolsey (1988)) will be further explored.

Acknowledgments

The authors gratefully acknowledge funding support from the National Science Foundation (Award 1850355) and the Army Research Office (Award 74113NSII).

Appendices

A. Proof of Theorem 3.3

Theorem 3.3 Let $G=(V,E)$ be an unweighted directed graph for representing a non-strict complete ranking $\boldsymbol{r}$ as follows: V is the set of nodes (alternatives) and an each pair of nodes is connected by one directed edge, that is, either $(i,j)\in E$ or $(j,i)\in E$ if $r_i<r_j$ (i.e., $v_i\succ v_j$) or two directed edges (i,j) and (j,i) if $r_i=r_j$ (i.e., $v_i\approx v_j$). Graph G is a ranking-matrix graph if and only if it does not contain uni-cycles (i.e., it is a unicycle-free graph).

Proof. Let $G'=(V',E')$ be a subgraph of G, that is, $V'\subseteq V$, $E'=(V'\times V')\cap E$. A bi-cycle exists whenever, for every adjacent pair of nodes i_k, i_{k+1} in a path $i_1,i_2,i_3,...,i_p$, we have $(i_k,i_{k+1})\in E'$ and $(i_{k+1},i_k)\in E'$, for $1\le k\le p-1<|V'|$. Clearly, when there exists a bi-cycle in G', it means every alternative included in the bi-cycle is tied. That is, when there exist directed cycles $i_1,i_2,...,i_{p-1},i_p,i_1$ and $i_1,i_p,i_{p-1},...,i_2,i_1$ with distinct nodes $i_1,i_2,...,i_{p-1},i_p\in V'$, then $i_1\approx i_2\approx ...\approx i_p$. Hence, whenever a graph of size 3 has a bi-cycle, it leads to a valid setting for the corresponding ranking-matrix entries. Specifically, this gives that $a_{i_\ell,i_{\ell'}}=1$, $a_{i_{\ell'},i_\ell}=1$ for all i_ℓ, $i_{\ell'}\in\{i_1,i_2,...,i_p\}$, where $i_\ell\neq i_{\ell'}$.

Recall that we can focus on a graph of length 3 or greater since a graph of size 2 cannot contain uni-cycle. Moreover, to check if these cycles exist in a digraph having at least one directed edge between every pair of nodes, it is sufficient to concentrate on finding unicyclic triads (i.e., uni-cycles of size 3), as a preference graph without unicyclic triads cannot have any higher-order uni-cycles (Gass 1998). Hence, $|V'|=3$ from this point, without loss of generality. To continue, Figure 2 lists distinct isomorphic classes of unicycle-free digraphs of size 3, while Figure 3 lists distinct isomorphic classes of unicyclic digraphs of size 3. Even though more unicyclic and unicycle-free graphs of size 3 are possible, it is only necessary to consider three respective isomorphic classes given in Figures 2 and 3; all other graphs can be represented by permuting the labels of the given graphs.

Figure 10 **Isomorphically equivalent digraphs of G'^6**

For Figure (3a), because $v_i\succ v_j$, $v_j\succ v_k$, and $v_k\succ v_i$, the preference relation includes the cycle $v_i\succ v_j\succ v_k\succ v_i$, which does not yield a proper ranking of three alternatives). For Figure (3b), because $v_i\succ v_j$, $v_j\approx v_k$, and $v_k\succ v_i$, the preference relation includes the cycle $v_i\succ v_j\approx v_k\succ v_i$, which also does not yield a complete and consistent set of preferences. For Figure (3c), because $v_i\succ v_j$, $v_j\approx v_k$, and $v_k\approx v_i$, the preference relation $v_i\succ v_j\approx v_k\approx v_i$ is not a consistent set of preferences. Therefore, the graphs with a uni-cycle do not yield a complete and consistent set of preferences, meaning they cannot correspond to a ranking. Using proof by exhaustion, we have demonstrated that graph G is a ranking-matrix graph if and only if it is a unicycle-free graph. ■

B. Proof of Theorem 3.5

Theorem 3.5 The GKBP constraints provide a logically equivalent formulation of the weak order polytope.

Proof. Fiorini and Fishburn (2004) provides a binary programming formulation for $\mathbf{P^n_{WO}}$. For $i, j \in \{1...n\}$, decision variable x_{ij} is defined as:

$$x_{ij} = \begin{cases} 1 & \text{if } i \preceq j \text{ (i.e., } i \text{ is not preferred to } j) \\ 0 & \text{otherwise.} \end{cases}$$

These decision variables have an oppositely aligned interpretation as those of GKBP in terms of preferential information. Specifically, when $y_{ij} = 1$ in GKBP, this indicates that the solution regards i as being preferred over j or tied with j (i.e., an expression of agreement). On the other hand, when $x_{ij} = 1$ in $\mathbf{P^n_{WO}}$, this indicates that the solution regards i as not being preferred to j, which implicitly indicates a dispreference and/or ties (i.e., seemingly both an expression of agreement and disagreement). When $y_{ij} = 0$ in GKBP, this indicates that the solution regards i as being dispreferred over j (i.e., an expression of disagreement). On the other hand, when $x_{ij} = 0$ in $\mathbf{P^n_{WO}}$, this indicates that the solution regards i as being preferred to j (i.e., an expression of agreement). The decision variables in the weak order polytope can be also represented via a ranking-matrix graph. When $x_{ij} = 1$, there is one directed edge (j, i) or there are two directed edges (i, j) and (j, i); there is one when $x_{ji} = 0$, and there are two when $x_{ji} = 1$. The constraints for $\mathbf{P^n_{WO}}$ are given as:

$$x_{ij} \leq 1 \tag{13a}$$
$$x_{ij} + x_{ji} \geq 1 \tag{13b}$$
$$x_{ik} - x_{ij} - x_{jk} \geq -1 \tag{13c}$$

By substituting x_{ij} with y_{ji}, the constraints become:

$$y_{ji} \leq 1 \tag{14a}$$
$$y_{ji} + y_{ij} \geq 1 \tag{14b}$$
$$y_{ki} - y_{ji} - y_{kj} \geq -1. \tag{14c}$$

Notice that the combination of Equation (14a) and (14b) gives the domain of y_{ij}. Equation (14b) matches with constraint (11c) in GKBP. Furthermore, the above constraints hold for any permutation of the labels; therefore, changing i to j, j to k, and k to i yields:

$$y_{ij} - y_{kj} - y_{ik} \geq -1. \tag{15}$$

Equation (15) is the same as constraint (11b) in GKBP. Therefore, the constraints of GKBP provide a logically equivalent formulation of the weak order polytope. ■

C. Proof of Lemma 5.2

Lemma 5.2 Let $\{V^1, V^2\} \in \mathcal{P}(V')$ (i.e., $\{V^1, V^2\}$ is a bipartition of V'). Consider the reduced aggregation problem consisting of input rankings $A_{[1\cup2]}$, which include only the evaluations over $V^1 \cup V^2$. If $V^1 \overset{M}{\succ} V^2$, every alternative $v_i \in V^1$ should obtain a better position than every alternative $v_j \in V^2$ in the optimal solution to the reduced problem, that is, $r^*_{i|[1\cup2]} < r^*_{j|[1\cup2]}$, where $\boldsymbol{r}^*_{[1\cup2]} := \underset{\boldsymbol{r}_{[1\cup2]}}{\arg\min} \sum_{\ell\in L} d(\boldsymbol{a}^\ell_{[1\cup2]}, \boldsymbol{r}_{[1\cup2]})$.

Proof. We prove this by contradiction. Let $\boldsymbol{r}^*_{[1\cup2]}$ be a Kemeny optimal ranking to the reduced problem involving only V^1 and V^2 and assume that $r^*_{i|[1\cup2]} \geq r^*_{j|[1\cup2]}$, for at least one alternative pair v_i, v_j, where $v_i \in V^1, v_j \in V^2$—i.e., there exists at least one alternative in V^1 which is tied or dispreferred over at least one alternative in V^2. Additionally, denote the ranking where all alternatives in V^1 are preferred over all alternatives in V^2 by $\bar{\boldsymbol{r}}^*_{[1\cup2]}$.

Initially, choose an arbitrary alternative $v_i \in V^1$. Then, the Kemeny-Snell distance between v_i and every alternative $v_j \in V^2$ is calculated as:

$$\sum_{v_j\in V^2}\sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}|[1\cup2]}, \boldsymbol{r}^*_{\{i,j\}|[1\cup2]}) = \frac{1}{2}\sum_{v_j\in V^2}\sum_{\ell\in L} |\operatorname{sign}(a^\ell_{i|[1\cup2]} - a^\ell_{j|[1\cup2]}) - \operatorname{sign}(r^*_{i|[1\cup2]} - r^*_{j|[1\cup2]})|.$$

Without loss of generality, a coefficient $1/2$ can be ignored in the remainder of the proof because it is a constant term which only affects the objective-value scaling. Moreover, since there are three possible ordinal relationships between v_i and v_j in $\boldsymbol{r}^*$, we have that,

$$\sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}}, \boldsymbol{r}^*_{\{i,j\}}) = \begin{cases} \sum_{\ell\in L} |\operatorname{sign}(a^\ell_i - a^\ell_j) - (-1)| = 2p_{ji} + t_{ij} & \text{if } r^*_i < r^*_j \\ \sum_{\ell\in L} |\operatorname{sign}(a^\ell_i - a^\ell_j) - 0| = p_{ij} + p_{ji} & \text{if } r^*_i = r^*_j \\ \sum_{\ell\in L} |\operatorname{sign}(a^\ell_i - a^\ell_j) - 1| = 2p_{ij} + t_{ij} & \text{if } r^*_i > r^*_j \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the distance between v_i and all $v_j \in V^2$ can be factored as,

$$\sum_{v_j\in V^2}\sum_{\ell\in L} d(\boldsymbol{a}^\ell_{\{i,j\}|[1\cup2]}, \boldsymbol{r}^*_{\{i,j\}|[1\cup2]}) = \sum_{\substack{v_j\in V^{k'} \\ \text{s.t. } r^*_{i|[1\cup2]} > r^*_{j|[1\cup2]}}} (2p_{ij} + t_{ij}) + \sum_{\substack{v_j\in V^{k'} \\ \text{s.t. } r^*_{i|[1\cup2]} = r^*_{j|[1\cup2]}}} (p_{ij} + p_{ji}) + \sum_{\substack{v_j\in V^{k'} \\ \text{s.t. } r^*_{i|[1\cup2]} < r^*_{j|[1\cup2]}}} (2p_{ji} + t_{ij}).$$

From the assumption that $V^1 \overset{M}{\succ} V^2$ for all $v_i \in V^1$ and $v_j \in V^2$, because $p_{ij} > p_{ji} + t_{ij}$, we can derive the following inequalities involving the second ($r^*_i = r^*_j$) and third ($r^*_i > r^*_j$) cases above:

$$\begin{aligned} 2p_{ji} + t_{ij} &< p_{ij} + p_{ji} \\ 2p_{ji} + t_{ij} &< 2p_{ij} + t_{ij}, \end{aligned}$$

which results in the following relationship:

$$\sum_{\substack{v_j\in V^{k'} \\ \text{s.t. } r^*_{i|[1\cup2]} > r^*_{j|[1\cup2]}}} (2p_{ij} + t_{ij}) + \sum_{\substack{v_j\in V^{k'} \\ \text{s.t. } r^*_{i|[1\cup2]} = r^*_{j|[1\cup2]}}} (p_{ij} + p_{ji}) + \sum_{\substack{v_j\in V^{k'} \\ \text{s.t. } r^*_{i|[1\cup2]} < r^*_{j|[1\cup2]}}} (2p_{ji} + t_{ij}) > \sum_{v_j\in V^2} (2p_{ji} + t_{ij}).$$

That is,

$$\sum_{v_j \in V^2} \sum_{\ell \in L} d(\boldsymbol{a}^{\ell}_{\{i,j\}|[1\cup 2]}, \boldsymbol{r}^{*}_{\{i,j\}|[1\cup 2]}) > \sum_{v_j \in V^2} \sum_{\ell \in L} d(\boldsymbol{a}^{\ell}, \bar{\boldsymbol{r}}^{*}_{\{i,j\}|[1\cup 2]})$$

which means the optimal ranking $\boldsymbol{r}^*$ where v_i is tied or dispreferred over some alternatives in V^2 returns a longer cumulative distance than the ranking $\bar{\boldsymbol{r}}^*$ where v_i is preferred over all $v_j \in V^2$. This contradicts the assumption that $\boldsymbol{r}^*$ is the optimal ranking, since it does not return the shortest cumulative distance. Therefore, if $V^1 \overset{M}{\succ} V^2$, assigning v_i with a better ranking position than all alternatives $v_j \in V^2$ returns a strictly shorter distance than a ranking where alternative v_i is tied with or dispreferred over an arbitrary alternative v_j, that is $r^*_{i|[1\cup 2]} < r^*_{j|[1\cup 2]}$, $\forall v_j \in V^2$. Since v_i was chosen arbitrarily, this holds for all alternatives in V^1. Hence, if $V^1 \overset{M}{\succ} V^2$, every alternative $v_i \in V^1$ should obtain a better position than every alternative $v_j \in V^2$ in the optimal solution. ■

D. Pseudocode of generating non-strict complete and incomplete rankings

Algorithm 1 and 2 describe how non-strict complete and incomplete ranking instances are constructed and guided based on the concept of Mallows distribution. Before introducing the algorithms, we provide needed definitions. Let $\boldsymbol{a}^{-1}$ be an *alternative-ordering* induced from rankings by sorting the alternatives from best to worst, according to their ranks. For example, for $\boldsymbol{a} = (1,5,2,4,3)$, $\boldsymbol{a}^{-1} = (v_1, v_3, v_5, v_4, v_2)$. Extending this notation, $\boldsymbol{a}^{-1}(i)$ specifies the ith-highest ranked alternative in $\boldsymbol{a}$ (Doignon et al. 2004); in the aforementioned example, $\boldsymbol{a}^{-1}(3) = v_5$. When $\boldsymbol{a}$ is a non-strict ranking, the alternative-ordering is obtained by putting alternatives with the same rank position into preference equivalence classes; for example, for $\boldsymbol{a} = (1,3,3,1,5)$, $\boldsymbol{a}^{-1} = (\langle v_1, v_4\rangle, \langle v_2, v_3\rangle, v_5)$. Additionally, $\underline{\boldsymbol{a}}^{-1}|_{V_{\boldsymbol{a}}}$ is an alternative ordering that is projected to the alternatives in $V_{\boldsymbol{a}}$, where $V_{\boldsymbol{a}}$ is an alternative set which is evaluated by $\boldsymbol{a}$. Finally, UniformDist(L,U) denotes the discrete uniform distribution, where L and U are the minimum and maximum values of the distribution.

Algorithm 1 Generating non-strict complete rankings

Input: Dispersion: ϕ, reference alternative ordering: $\underline{\boldsymbol{a}}^{-1}$
Output: A set of non-strict complete rankings

for $i = 1, 2, ..., |V|$ **do**
  for $j = 1, 2, ..., i$ **do**
    $a^{-1}(j) \leftarrow \underline{a}^{-1}(i)$ with probability: $p_{ij} = \phi^{i-j}/(1 + \phi + \cdots + \phi^{i-1})$
while (the number of alternatives involved in ties) $\leq 0.5n$ **do**
  $u \leftarrow$ UniformDist$(1, h-1)$, where h is the worst (highest-ranked) position in $\boldsymbol{a}$
  $v \leftarrow a^{-1}(u+1)$, and then $a_v \leftarrow u$

Algorithm 2 Generating non-strict incomplete rankings

Input: Dispersion: ϕ, alternative set for $\boldsymbol{a}$: $V_{\boldsymbol{a}}$, projected reference alternative ordering: $\underline{\boldsymbol{a}}^{-1}|_{V_{\boldsymbol{a}}}$
Output: A set of non-strict incomplete rankings

for $i = 1, 2, ..., |V_a|$ **do**
  for $j = 1, 2, ..., i$ **do**
    $a^{-1}(j) \leftarrow \underline{a}^{-1}(i)|_{V_{\boldsymbol{a}}}$ with probability: $p_{ij} = \phi^{i-j}/(1 + \phi + \cdots + \phi^{i-1})$
for $i = 1, 2, ..., |V|$ **do**
  if $v_i \in V_{\boldsymbol{a}}$ **then**
    $a_i \leftarrow$ rank position of v_i in $\boldsymbol{a}^{-1}$
  else
    $a_i \leftarrow \bullet$
while (the number of alternatives involved in ties) $\leq 0.5n$ **do**
  $u \leftarrow$ UniformDist$(1, h-1)$, where h is the worst (highest-ranked) position in $\boldsymbol{a}$
  $v \leftarrow a^{-1}(u+1)$, and then $a_v \leftarrow u$

E. Pseudocode of decomposition algorithm

Algorithm 3 Decomposition Algorithm

Input: $\{p_{ij}\}$,$\{p_{ji}\}$,$\{t_{ij}\}$
Output: An ordered partition of subsets $\mathcal{V}=\{V^1,V^2,...,V^K\}$

$\mathcal{V}=\{\{v_1\}\}$
for $i=2,3,...,|V|$ **do**
  if $(\exists v_j \in V^1$ s.t., $t_{ij} \geq |p_{ij}-p_{ji}|)$, **or**
    $(\exists v_j \in V^1$ s.t., $p_{ij} > p_{ji}+t_{ij}$ **and** $\exists v_{j'} \in V^1 \backslash \{v_j\}$ s.t., $p_{j'i} > p_{ij'}+t_{ij'})$ **then**
    Put v_i in V^1 and $k \leftarrow 2$
  else if $\forall v_j \in V^1$ s.t., $p_{ij} > p_{ji}+t_{ij}$ **then**
    Insert v_i before V^1, increment the index of subsets after $V^{\sigma(i)}$ by 1, and $k \leftarrow 3$
  else if $\forall v_j \in V^1$ s.t., $p_{ji} > p_{ij}+t_{ij}$ **then**
    Insert v_i after V^1, increment the index of subsets after $V^{\sigma(i)}$ by 1, and $k \leftarrow 3$
  while $k \leq |\mathcal{V}|$ **do**
    if $(\exists v_j \in V^k$ s.t., $t_{ij} \geq |p_{ij}-p_{ji}|)$, **or**
      $(\exists v_j \in V^k$ s.t., $p_{ij} > p_{ji}+t_{ij}$ **and** $\exists v_{j'} \in V^k \backslash \{v_j\}$ s.t., $p_{j'i} > p_{ij'}+t_{ij'})$ **then**
      Merge subsets from $V^{\sigma(i)}$ to V^k
      Decrease the index of subsets after V^k by $k-\sigma(i)$ and $k \leftarrow \sigma(i)+1$
    else if $\forall v_j \in V^k$ s.t., $p_{ij} > p_{ji}+t_{ij}$ **then**
      $k \leftarrow k+1$
    else if $\forall v_j \in V^k$ s.t., $p_{ji} > p_{ij}+t_{ij}$ **then**
      if $|\sigma(i)-k|=1$ and $|V^{\sigma(i)}|=1$ **then**
        Move $V^{\sigma(i)}$ after V^k and increment the index of subsets after $V^{\sigma(i)}$ by 1
      else
        Merge subsets from $V^{\sigma(i)}$ to V^k
        Decrease the index of subsets after V^k by $k-\sigma(i)$ and $k \leftarrow \sigma(i)+1$

* $\sigma(i)$ is the index of the subset containing v_i.

Note that this decomposition algorithm has a worst-case complexity of $O(n^2)$, which is explained as follows. At iteration $i-1$, the algorithm inserts v_i before, within, or after the existing subsets in the working partition $\mathcal{V}=\{V^1,V^2,...,V^{\kappa(i)}\}$, where $\kappa(i)$ is the number of subsets prior to the iteration. To determine its precise point of insertion, v_i is compared at most to the alternatives in all subsets from V^1 to $V^{\kappa(i)}$, which takes at most $(i-1)$ comparisons, each of which is assumed to take constant time. This has to be done for $i=2,\ldots,n$. Therefore, the algorithm requires at most $n(n-1)/2$ such comparisons, resulting in a worst-case complexity of $O(n^2)$.

References

Ahlgren P, Jarneving B, Rousseau R (2003) Requirements for a cocitation similarity measure, with special reference to pearson's correlation coefficient. *Journal of the Association for Information Science and Technology* 54(6):550–560.

Ailon N (2010) Aggregation of partial rankings, p-ratings and top-m lists. *Algorithmica* 57(2):284–300.

Ailon N, Charikar M, Newman A (2008) Aggregating inconsistent information: ranking and clustering. *Journal of the ACM (JACM)* 55(5):23.

Amodio S, DAmbrosio A, Siciliano R (2016) Accurate algorithms for identifying the median ranking when dealing with weak and partial rankings under the kemeny axiomatic approach. *European Journal of Operational Research* 249(2):667–676.

Arrow KJ, et al. (1951) Social choice and individual values .

Asfaw D, Vitelli V, Sørensen Ø, Arjas E, Frigessi A (2017) Time-varying rankings with the bayesian mallows model. *Stat* 6(1):14–30.

Bartholdi J, Tovey CA, Trick MA (1989) Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare* 6:157–165.

Betzler N, Bredereck R, Niedermeier R (2014) Theoretical and empirical evaluation of data reduction for exact kemeny rank aggregation. *Autonomous Agents and Multi-Agent Systems* 28(5):721–748.

Brancotte B, Yang B, Blin G, Cohen-Boulakia S, Denise A, Hamel S (2015) Rank aggregation with ties: Experiments and analysis. *Proceedings of the VLDB Endowment* 8(11):1202–1213.

Brandt F, Conitzer V, Endriss U, Lang J, Procaccia AD (2016) *Handbook of computational social choice* (Cambridge University Press).

Ceberio J, Irurozki E, Mendiburu A, Lozano JA (2015) A review of distances for the mallows and generalized mallows estimation of distribution algorithms. *Computational Optimization and Applications* 62(2):545–564.

Cohen D, Jeavons P, Jefferson C, Petrie KE, Smith BM (2005) Symmetry definitions for constraint satisfaction problems. *International Conference on Principles and Practice of Constraint Programming*, 17–31 (Springer).

Condorcet Md (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* .

Conitzer V, Davenport A, Kalagnanam J (2006) Improved bounds for computing kemeny rankings. *AAAI*, volume 6, 620–626.

Cook WD (2006) Distance-based and ad hoc consensus models in ordinal preference ranking. *European Journal of Operational Research* 172(2):369–385.

Cook WD, Golany B, Penn M, Raviv T (2007) Creating a consensus ranking of proposals from reviewers' partial ordinal rankings. *Computers & Operations Research* 34(4):954–965.

Cook WD, Kress M (1985) Ordinal ranking with intensity of preference. *Management science* 31(1):26–32.

Cook WD, Seiford LM (1978) Priority ranking and consensus formation. *Management Science* 24(16):1721–1732.

Crispino M, Arjas E, Vitelli V, Barrett N, Frigessi A, et al. (2019) A bayesian mallows approach to nontransitive pair comparison data: How human are sounds? *The Annals of Applied Statistics* 13(1):492–519.

Critchlow DE (2012) *Metric methods for analyzing partially ranked data*, volume 34 (Springer Science & Business Media).

Cushman P, Hoeksema JT, Kouveliotou C, Lowenthal J, Peterson B, Stassun KG, von Hippel T (2015) Impact of declining proposal success rates on scientific productivity. *arXiv preprint arXiv:1510.01647* .

Davenport A, Kalagnanam J (2004) A computational study of the kemeny rule for preference aggregation. *AAAI*, volume 4, 697–702.

Davidson J, Liebald B, Liu J, Nandy P, Van Vleet T, Gargi U, Gupta S, He Y, Lambert M, Livingston B, et al. (2010) The youtube video recommendation system. *Proceedings of the fourth ACM conference on Recommender systems*, 293–296.

de Borda JC (1781) Mémoire sur les élections au scrutin, histoire de l?académie royale des sciences. *Paris, France* .

Desarkar MS, Sarkar S, Mitra P (2016) Preference relations based unsupervised rank aggregation for metasearch. *Expert Systems with Applications* 49:86–98.

Diaconis P (1988) Group representations in probability and statistics. *Lecture notes-monograph series* 11:i–192.

Doignon JP, Pekeč A, Regenwetter M (2004) The repeated insertion model for rankings: Missing link between two subset choice models. *Psychometrika* 69(1):33–54.

Dummett M (1998) The borda count and agenda manipulation. *Social Choice and Welfare* 15(2):289–296.

Dwork C, Kumar R, Naor M, Sivakumar D (2001) Rank aggregation methods for the web. *Proceedings of the 10th international conference on the World Wide Web*, 613–622 (New York, NY, USA: ACM), ISBN 1-58113-348-0.

D'Ambrosio A, Mazzeo G, Iorio C, Siciliano R (2017) A differential evolution algorithm for finding the median ranking under the kemeny axiomatic approach. *Computers & Operations Research* 82:126–138.

Emond EJ, Mason DW (2002) A new rank correlation coefficient with application to the consensus ranking problem. *Journal of Multi-Criteria Decision Analysis* 11(1):17–28.

Endriss U, Obraztsova S, Polukarov M, Rosenschein JS (2016) Strategic voting with incomplete information .

Fagin R, Kumar R, Mahdian M, Sivakumar D, Vee E (2004) Comparing and aggregating rankings with ties. *Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, 47–58.

Fagin R, Kumar R, Sivakumar D (2003) Efficient similarity search and classification via rank aggregation. *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, 301–312 (ACM).

Farah M, Vanderpooten D (2007) An outranking approach for rank aggregation in information retrieval. *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, 591–598 (ACM).

Favardin P, Lepelley D, Serais J (2002) Borda rule, copeland method and strategic manipulation. *Review of Economic Design* 7(2):213–228.

Feld SL, Grofman B (1988) The borda count in n-dimensional issue space. *Public Choice* 59(2):167–176.

Fields EB, Okudan GE, Ashour OM (2013) Rank aggregation methods comparison: A case for triage prioritization. *Expert Systems with Applications* 40(4):1305–1311.

Fiorini S, Fishburn PC (2004) Weak order polytopes. *Discrete mathematics* 275(1-3):111–127.

Gao Y, Xu K (2019) prankaggreg: A fast clustering based partial rank aggregation. *Information Sciences* 478:408–421.

Gass S (1998) Tournaments, transitivity and pairwise comparison matrices. *Journal of the Operational Research Society* 49(6):616–624.

Good I (1975) The number of orderings of n candidates when ties are permitted. *Fib. Quart* 13:11–18.

Goodin RE, List C (2006) A conditional defense of plurality rule: generalizing may's theorem in a restricted informational environment. *American Journal of Political Science* 50(4):940–949.

Heiser WJ (2004) Geometric representation of association between categories. *Psychometrika* 69(4):513–545.

Hochbaum DS, Levin A (2006) Methodologies and algorithms for group-rankings decision. *Management Science* 52(9):1394–1408.

IBM Knowledge Center (2017) Ibm ilog cplex optimization studio v12.8.0 documentation .

IBM Support (2019) URL `https://www.ibm.com/support/pages/apar/RS03137`.

Irurozki E, Calvo B, Lozano JA, et al. (2016) Permallows: An r package for mallows and generalized mallows models. *Journal of Statistical Software* 71(12):1–30.

Keisler J (2004) Value of information in portfolio decision analysis. *Decision analysis* 1(3):177–189.

Kemeny JG, Snell LJ (1962) Preference ranking: An axiomatic approach. *Mathematical Models in Social Science*, 9–23 (Boston: Ginn).

Kendall MG (1938) A new measure of rank correlation. *Biometrika* 30(1/2):81–93.

Kendall MG (1945) The treatment of ties in ranking problems. *Biometrika* 239–251.

Kendall MG (1948) Rank correlation methods .

Kenyon-Mathieu C, Schudy W (2007) How to rank with few errors. *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, 95–103 (ACM).

Kolde R, Laur S, Adler P, Vilo J (2012) Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* 28(4):573–580.

Krylovas A, Zavadskas EK, Kosareva N, Dadelo S (2014) New kemira method for determining criteria priority and weights in solving mcdm problem. *International Journal of Information Technology & Decision Making* 13(06):1119–1133.

Liberti L (2008) Automatic generation of symmetry-breaking constraints. *International Conference on Combinatorial Optimization and Applications*, 328–338 (Springer).

Lin S (2010a) Rank aggregation methods. *Wiley Interdisciplinary Reviews: Computational Statistics* 2(5):555–570.

Lin S (2010b) Space oriented rank-based data integration. *Statistical Applications in Genetics and Molecular Biology* 9(1).

Lu T, Boutilier C (2014) Effective sampling and learning for mallows models with pairwise-preference data. *The Journal of Machine Learning Research* 15(1):3783–3829.

Mallows CL (1957) Non-null ranking models. i. *Biometrika* 44(1/2):114–130.

Mandal M, Mukhopadhyay A (2017) Multiobjective pso-based rank aggregation: Application in gene ranking from microarray data. *Information Sciences* 385:55–75.

Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, Allison KR, Aderhold A, Bonneau R, Chen Y, et al. (2012) Wisdom of crowds for robust gene network inference. *Nature methods* 9(8):796.

Marden JI (1996) *Analyzing and modeling rank data* (CRC Press).

Martí R, Reinelt G (2011) *The Linear Ordering Problem* (Springer).

Mattei N, Walsh T (2013) Preflib: A library for preferences http://www. preflib. org. *International Conference on Algorithmic DecisionTheory*, 259–270 (Springer).

Miller GA (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* 63(2):81–97, URL `http://www.musanim.com/miller1956/`.

Milosz R, Hamel S (2018) Exploring the median of permutations problem. *Journal of Discrete Algorithms* 52:92–111.

Mohammadi M, Rezaei J (2020) Ensemble ranking: Aggregation of rankings produced by different multi-criteria decision-making methods. *Omega* 102254.

Moreno-Centeno E, Escobedo AR (2016) Axiomatic aggregation of incomplete rankings. *IIE Transactions* 48(6):475–488.

Muravyov SV (2014) Dealing with chaotic results of kemeny ranking determination. *Measurement* 51:328–334.

Nemhauser GL, Wolsey LA (1988) Integer programming and combinatorial optimization. *Wiley, Chichester. GL Nemhauser, MWP Savelsbergh, GS Sigismondi (1992). Constraint Classification for Mixed Integer Programming Formulations. COAL Bulletin* 20:8–12.

Newman A, Vempala S (2001) Fences are futile: On relaxations for the linear ordering problem. *International Conference on Integer Programming and Combinatorial Optimization*, 333–347 (Springer).

Nurmi H, Meskanen T (2000) Voting paradoxes and mcdm. *Group Decision and Negotiation* 9(4):297–313.

Schilling MS, Oeser N, Schaub C (2007) How effective are decision analyses? assessing decision process and group alignment effects. *Decision Analysis* 4(4):227–242.

Sherali HD, Smith JC (2001) Improving discrete model representations via symmetry considerations. *Management Science* 47(10):1396–1407.

Smith JH (1973) Aggregation of preferences with variable electorate. *Econometrica: Journal of the Econometric Society* 1027–1041.

Streib N, Young SJ, Sokol J (2012) A major league baseball team uses operations research to improve draft preparation. *Interfaces* 42(2):119–130.

Truchon M, et al. (1998) An extension of the condorcet criterion and kemeny orders. *Cahier* 9813.

Wald R, Khoshgoftaar TM, Dittman D, Awada W, Napolitano A (2012) An extensive comparison of feature ranking aggregation techniques in bioinformatics. *2012 IEEE 13th International Conference on Information Reuse & Integration (IRI)*, 377–384 (IEEE).

Ye M, Liang C, Yu Y, Wang Z, Leng Q, Xiao C, Chen J, Hu R (2016) Person reidentification via ranking aggregation of similarity pulling and dissimilarity pushing. *IEEE Transactions on Multimedia* 18(12):2553–2566.

Yilmaz E, Aslam JA, Robertson S (2008) A new rank correlation coefficient for information retrieval. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, 587–594 (ACM).

Yoo Y, Escobedo AR, Skolfield JK (2020) A new correlation coefficient for comparing and aggregating non-strict and incomplete rankings. *European Journal of Operational Research* 285(3):1025–1041.

Young HP, Levenglick A (1978) A consistent extension of condorcets election principle. *SIAM Journal on applied Mathematics* 35(2):285–300.

Young P (1995) Optimal voting rules. *Journal of Economic Perspectives* 9(1):51–64.