



Decision Analysis

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

A New Binary Programming Formulation and Social Choice Property for Kemeny Rank Aggregation

Yeawon Yoo , Adolfo R. Escobedo

To cite this article:

Yeawon Yoo , Adolfo R. Escobedo (2021) A New Binary Programming Formulation and Social Choice Property for Kemeny Rank Aggregation. Decision Analysis

Published online in Articles in Advance 16 Sep 2021

. <https://doi.org/10.1287/deca.2021.0433>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2021, INFORMS

Please scroll down for article—it is on subsequent pages





With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

A New Binary Programming Formulation and Social Choice Property for Kemeny Rank Aggregation

Yeawon Yoo,^a Adolfo R. Escobedo^b

^aDepartment of Applied Mathematics and Statistics, SNF Agora Institute, Johns Hopkins University, Baltimore, Maryland 21218; ^bSchool of Computing and Augmented Intelligence, Arizona State University, Tempe, Arizona 85281

Contact: yyoo9@jh.edu,  <https://orcid.org/0000-0001-9678-6607> (YY); adres@asu.edu,  <https://orcid.org/0000-0002-4843-3564> (ARE)

Received: May 12, 2020

Revised: October 26, 2020; March 9, 2021

Accepted: March 22, 2021

Published Online in Articles in Advance:
September 16, 2021

<https://doi.org/10.1287/deca.2021.0433>

Copyright: © 2021 INFORMS

Abstract. 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. Moreover, it leverages the equivalence of two ranking aggregation problems, namely, that of minimizing the Kemeny-Snell distance and of maximizing the Kendall- τ correlation, to compare the newly introduced binary programming formulation to a modified version of an existing integer programming formulation associated with the Kendall- τ distance. The new formulation has fewer variables and constraints, which leads to faster solution times. Moreover, we develop a new social choice property, the nonstrict 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 with benchmark instances from PrefLib, a library of preference data.

Funding: This work was supported by the U.S. Army Research Office [Grant W911NF1910260] and the National Science Foundation, Division of Information and Intelligent Systems [Grant 1850355].

Keywords: 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 (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 authors 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, rank aggregation has been widely used across a number of practical group decision-making settings. For instance, Fields et al. (2013) consider a healthcare problem of improving nurse triage and patient 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, Escobedo et al. 2021), and to improve the annual draft preparation decision-making process of Major League Baseball (Streib et al. 2012). Besides decision making, rank aggregation has extensive applicability to other fields, such as crowdsourcing and human computation (Steyvers et al. 2009; Mao et al. 2012, 2013; Kemmer et al. 2020), 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). For example, rank aggregation has been used within the field of crowdsourcing to derive robust collective estimates related to the order of historical events (Steyvers et al. 2009), puzzle games based on difficulty (Mao et al. 2012), and images containing dots based on the number of dots they contain (Kemmer et al. 2020). These and other 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—that is, 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, whereas 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 problem based on the Kemeny-Snell distance has competitive advantages over other aggregation frameworks. Known widely as *Kemeny (rank) aggregation*, the objective of this problem is to find a consensus ranking solution, which is defined as a ranking with the minimum number of pairwise reversals to the set of input rankings (see Section 2). Assuming that there are no cycles in the majority’s pairwise preferences and that the input rankings are complete, Kemeny aggregation is guaranteed to return the ranking solution that reflects the majority’s pairwise preferences. On the other hand, scoring methods are not guaranteed to do so; for instance, the consensus ranking solution may not place the *Condorcet winner* (see Section 5.1) in first place. Kemeny aggregation has been repeatedly demonstrated to be less vulnerable to manipulation than scoring methods and more robust to outliers (Feld and Grofman 1988, Favardin et al. 2002, Endriss et al. 2016). The satisfaction of this and other key social choice properties is a fundamental reason that Kemeny aggregation is preferred over various voting methods. 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 and various extensions of the Kemeny-Snell distance have been developed and applied in the area of decision analysis (e.g., Dwork et al. 2001, Cook 2006, Moreno-Centeno and Escobedo 2016) to reflect 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., Kendall 1938, Ahlgren et al. 2003, 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 nonparametric correlation coefficient that measures the agreement among *strict rankings* (i.e., rankings that do not allow ties); it was extended to handle *nonstrict rankings* (i.e., rankings that allow ties) in Kendall (1948). Emond and Mason (2002) provided another version of Kendall- τ correlation coefficient for nonstrict rankings and demonstrated that their τ -extended coefficient, τ_{N} , 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 National Science Foundation's Astronomy and Astrophysics Research Grants (AAG) program was 731 in 2014. As a second example from portfolio decision analysis, Keisler (2004) and Schilling et al. (2007) considered 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 nonhuman 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 2010a,b; Kolde et al. 2012; Marbach et al. 2012; Wald et al. 2012; Mandal and Mukhopadhyay 2017). Rank aggregation can be also used in a meta-search, 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 nonstrict 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 nonstrict 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 Kendall- τ distance. Section 5 introduces a social choice property that is applicable to nonstrict 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. Kemeny and Snell (1962) introduced an axiomatic distance for nonstrict complete rank aggregation, which is denoted as d_{KS} and defined with respect to ranking vectors \mathbf{a} and \mathbf{b} as follows:

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

In essence, distance d_{KS} counts the pairwise inversions between rankings \mathbf{a} and \mathbf{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

Table 1. Symbols and Notations

V	A set of alternatives (i.e., $V = \{v_1, v_2, v_3, \dots, v_n\}$), 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)
\mathbf{a}^ℓ	The ranking from judge ℓ ($\ell = 1, 2, 3, \dots, L $), where $\mathbf{a}^\ell \in A$
a_i^ℓ	Rank position of v_i in the evaluation from judge ℓ , where $\ell = 1, 2, 3, \dots, L $
$v_i > v_j$	v_i is preferred over v_j (i.e., $a_i < a_j$ for some ranking \mathbf{a})
$v_i \approx v_j$	v_i is tied with v_j (i.e., $a_i = a_j$ for some ranking \mathbf{a})
$v_i \geq v_j$	v_i is preferred over or tied with v_j (i.e., $a_i \leq a_j$ for some ranking \mathbf{a})
p_{ij}	The number of judges who prefer v_i over v_j (i.e., $ \{\mathbf{a}^\ell \in A : a_i^\ell < a_j^\ell\} $)
t_{ij}	The number of judges who tie v_i and v_j (i.e., $ \{\mathbf{a}^\ell \in A : a_i^\ell = a_j^\ell\} $)
$v_i >_M^m v_j$	A majority of judges prefers, rather than disprefers, v_i over v_j (i.e., $p_{ij} > p_{ji}$)
$v_i \approx_M^m v_j$	No majority of judges prefers or disprefers v_i over v_j (i.e., $p_{ij} = p_{ji}$)
$v_i >_M v_j$	A decisive majority of judges prefers v_i over v_j (i.e., $p_{ij} > p_{ji} + t_{ij}$)
$v_i \approx_M v_j$	No decisive majority of judges prefers v_i over v_j , or vice versa (i.e., $t_{ij} \geq p_{ij} - p_{ji} $)

advantages over other rank aggregation methods. Specifically, d_{KS} uniquely satisfies a set of intuitive axioms: nonnegativity, commutativity (symmetry), triangular inequality, anonymity, and 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, averaging 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 Kemeny 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

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

where α_{ij} (and β_{ij}) is the (i, j) -element of the *ranking-matrix* α_{ij} of a (and b), given by

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

Respective matrix representations for incomplete rankings are defined in Yoo et al. (2020). Whereas 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:

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

Equation (3) 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 a and b . Hence, it can be interpreted that the expression of $\tau_x(a, b)$ starts from a default assumption of perfect agreement between a and b (i.e., a correlation value of 1), and then it subtracts any disagreements between a and b from this perfect agreement, as quantified by $d_{KS}(a, b)$. 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,

$$\arg \min_r \sum_{\ell=1}^{|L|} d_{KS}(a^\ell, r) = \arg \max_r \sum_{\ell=1}^{|L|} \tau_x(a^\ell, r), \quad (4)$$

where r is a complete ranking and a^ℓ 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

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

This distance counts the number of pairwise inversions between a and 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-Snell axioms), whereas d_τ accrues a distance value of 1. Hence, the distances are related by the equation $d_{KS}(a, b) = 2d_\tau(a, b)$. Since the original Kendall- τ distance is defined only for strict rankings, Brancotte et al. (2015) redesigned the Kendall- τ distance for nonstrict rankings, which is defined as follows:

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

The main difference between d_{KS} and the Kendall- τ distance for nonstrict rankings is that when one ranking ties two specific alternatives and the other ranking does not, the Kendall- τ distance for nonstrict rankings 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 (Good 1975, Bartholdi et al. 1989), 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 the 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 nonstrict 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 aggregation, which does consider nonstrict 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 nonstrict complete and incomplete rankings. Beforehand, we introduce some existing integer programming formulations for distance-based rank aggregation.

3.1. Integer Programming Formulations for the 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 nonstrict rankings, Brancotte et al. (2015) provided a revised integer programming formulation for the Kendall- τ distance, given as follows:

$$\text{minimize}_x \sum_{\{v_i, v_j\} \subseteq V} (w_{j \leq i} x_{i < j} + w_{i \leq j} x_{j < i} + (w_{i < j} + w_{j < i}) x_{i=j}) \quad (5a)$$

$$\text{subject to } x_{i < j} + x_{j < i} + x_{i=j} = 1 \quad \forall \{v_i, v_j\} \in V, \quad (5b)$$

$$x_{i < k} - x_{i < j} - x_{j < k} \geq -1 \quad \forall \{v_i, v_j, v_k\} \in V, \quad (5c)$$

$$2x_{i < j} + 2x_{j < i} + 2x_{j < k} + 2x_{k < j} - x_{i < k} - x_{k < i} \geq 0 \quad \forall \{v_i, v_j, v_k\} \in V, \quad (5d)$$

$$x_{i < j}, x_{j < i}, x_{i=j} \in \{0, 1\} \quad \forall \{v_i, v_j\} \in V, \quad (5e)$$

where $w_{i \leq j}$ denotes the number of rankings with $v_i \geq v_j$ and $w_{i < j}$ denotes the number of rankings with $v_i > v_j$. 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 nonstrict). Constraints (5c) and (5d) enforce transitivity of strict and nonstrict 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_{τ} . 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 with when the ordinal relationships strictly oppose each other. For example, with $a^1 = (1, 2)$, $a^2 = (1, 1)$, $a^3 = (2, 1)$, this yields $d_{KS}(a^1, a^2) = 1$, whereas $d_{KS}(a^1, a^3) = 2$. To solve the Kemeny rank aggregation problem using the formulation of Brancotte et al. (2015), 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} , the formulation of Brancotte et al. (2015) is modified as follows:*

$$\begin{aligned} &\text{minimize}_x \sum_{\{v_i, v_j\} \subseteq V} \left(w_{j < i} + \frac{1}{2} w_{j=i} \right) x_{i < j} + \left(w_{i < j} + \frac{1}{2} w_{i=j} \right) x_{j < i} \\ &\quad + \frac{1}{2} (w_{i < j} + w_{j < i}) x_{i=j} \\ &\text{subject to constraints (5b) – (5e).} \end{aligned}$$

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 aggregation to derive a new integer programming formulation that is applicable to nonstrict complete and incomplete rankings. The key to this formulation relies on devising a constraint set that ensures that the values of a matrix induce a complete and consistent set of preferences, that is, a complete and nonstrict ranking. To this end, we develop a graph-based representation of the ranking-matrix (see Equation (2)) of a nonstrict complete ranking.

Definition 3.2. Let $G = (V, E)$ be an unweighted directed graph for representing a nonstrict complete ranking r as follows: V is the set of nodes (alternatives), and 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., $v_i > v_j$), 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 nonstrict 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 1(a) can be represented via the matrix in Figure 1(b), but these representations do not yield a nonstrict complete ranking due to the preferential cycle $v_i > v_j$, $v_j > v_k$, and $v_k > v_i$.

Hereafter, we define a *ranking-matrix graph* as an unweighted directed graph that induces a nonstrict 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 *unicycle* is defined as a simple path that starts and ends on the same vertex in *one* direction but *not*

in the reverse direction. A *bicycle* 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 bicycle and a unicycle are mutually exclusive. Additionally, it is not possible to have a unicycle of size 2, because, if there exists a directed edge from i to j and a directed edge from j to i , then this creates a bicycle. The focus of the theorem and proof is to prevent graphs with unicyles, 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 unicyles as a *unicycle-free graph* and a graph with at least one unicycle as a *unicyclic graph*. Figures 2 and 3 show the possible unicycle-free graphs and unicyclic graphs, respectively, over three alternatives.

Theorem 3.3. Let $G = (V, E)$ be an unweighted directed graph for representing a nonstrict complete ranking r as follows: V is the set of nodes (alternatives) and each pair of nodes $(i, j) \in V$ is connected by one or two directed edges; that is, either $(i, j) \in E$ or $(j, i) \in E$ if $r_i < r_j$ (i.e., $v_i > v_j$) or $(i, j) \in E$ and $(j, i) \in E$ 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 unicyles (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 $S \in \mathbb{Z}^{n \times n}$, along with corresponding auxiliary binary variables $Y \in \mathbb{B}^{n \times n}$, induce a complete and consistent set of preferences of n alternatives (i.e., a nonstrict complete ranking) if the following constraints are satisfied for some setting of S and Y :

$$s_{ij} - s_{kj} - s_{ik} \geq -1 \quad i, j, k = 1, \dots, n; \quad i \neq j \neq k \neq i \quad (6a)$$

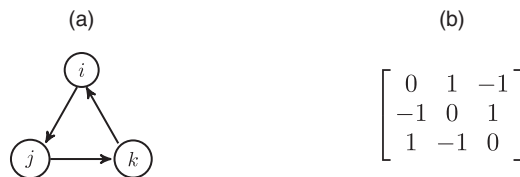
$$s_{ij} + s_{ji} \geq 0 \quad i, j = 1, \dots, n; \quad i < j \quad (6b)$$

$$s_{ii} = 0 \quad i = 1, \dots, n; \quad (6c)$$

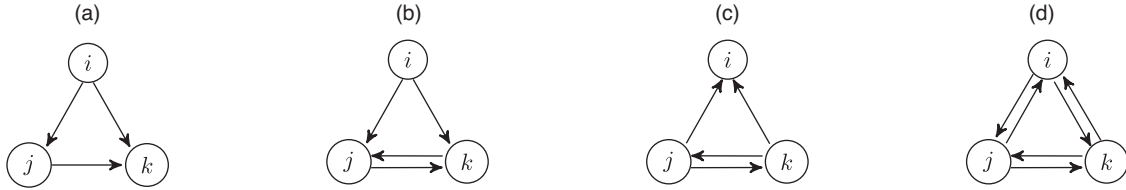
$$s_{ij} - 2y_{ij} = -1 \quad i, j = 1, \dots, n; \quad i = j \quad (6d)$$

$$s_{ij} \in \{-1, 0, 1\}, y_{ij} \in \{0, 1\} \quad i, j = 1, \dots, n. \quad (6e)$$

Figure 1. (a) Example of an Unweighted Directed Graph; (b) Matrix Representation of the Digraph



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

Figure 2. Unicycle-Free Graphs: (a) G'^1 ; (b) G'^2 ; (c) G'^3 ; and (d) G'^4 

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 α_{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 nonzero 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} .

The proof of Theorem 3.3 explains that to check whether a unicycle of any length exists within a ranking-matrix digraph, it is sufficient to verify that no unicycles 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 in Figure 4. Specifically, adding the dotted directed edge or replacing the thick (blue) edge with the dotted directed edge creates a unicycle. For example, replacing the thick (blue) edge with the dotted directed edge in Figure 4(a) and Figure 4(b) yields Figure 3(a) and Figure 3(b), respectively. Also, adding the dotted directed edge in Figure 4(b) yields Figure 3(c), 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 Figure 4(a) and (b), 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

$$s_{ji} = -1, s_{jk} = 1 \Rightarrow s_{ki} = -1. \quad (7)$$

Moreover, as in Figure 4(c), 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

$$s_{ji} = -1, s_{kj} = 1 \Rightarrow s_{ik} = -1. \quad (8)$$

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

$$s_{ji} - s_{jk} \geq s_{ki} - 1, \quad (9)$$

$$s_{ij} - s_{kj} \geq s_{ik} - 1. \quad (10)$$

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

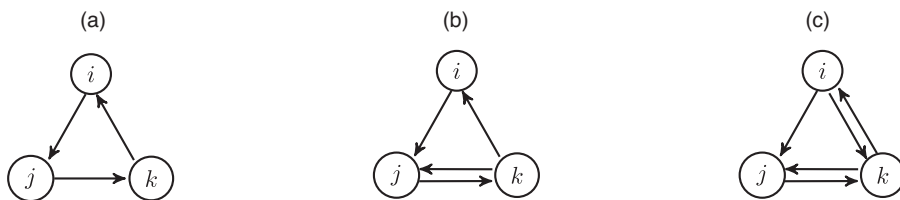
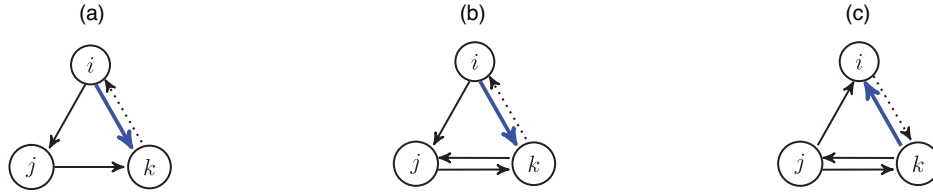
Figure 3. Unicyclic Graphs: (a) G'^5 ; (b) G'^6 ; and (c) G'^7 

Figure 4. (Color online) The Dotted Directed Edge Creates a Unicycle



Constraint (9) (i.e., it is redundant). Therefore, Constraints (6a)–(6e) provide the full set of constraints. \square

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

$$\begin{aligned} & \text{maximize}_s \quad \sum_i \sum_j c_{ij} s_{ij} \\ & \text{subject to} \quad \text{constraints (6a) – (6e),} \end{aligned}$$

where $[c_{ij}] \in \mathbb{Z}^{n \times n}$ is the cumulative ranking-matrix of the input rankings, defined as $c_{ij} = \sum_{\ell=1}^{|I|} \alpha_{ij}^\ell$, for $i, j = 1, \dots, n$ (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 by Dwork et al. (2001) and Moreno-Centeno and Escobedo (2016). For the for-

mer, it is defined as $c_{ij} = \sum_{\ell=1}^{|I|} \frac{\alpha_{ij}^\ell}{n(n-1)}$, where n is the total number of alternatives, and, for the latter, it is defined as $c_{ij} = \sum_{\ell=1}^{|I|} \frac{\alpha_{ij}^\ell}{n^\ell(n^\ell-1)}$, where n^ℓ is the number of alternatives 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

$$\text{maximize}_y \quad \sum_i \sum_j c_{ij} (2y_{ij} - 1) \quad (11a)$$

$$\text{subject to} \quad y_{ij} - y_{kj} - y_{ik} \geq -1 \quad i, j, k = 1, \dots, n; \quad i \neq j \neq k \neq i \quad (11b)$$

$$y_{ij} + y_{ji} \geq 1 \quad i, j = 1, \dots, n; \quad i < j \quad (11c)$$

$$y_{ij} \in \{0, 1\} \quad i, j = 1, \dots, n; \quad i \neq j. \quad (11d)$$

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 and v_j . Upon inspection, GKBP has $n^2 - n$ variables, whereas the formulation of Brancotte et al. (2015) 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 the formulation of Brancotte et al. (2015).

On a more fundamental level, GKBP can be connected to the theory of order polytopes. An order polytope 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 P_{LO}^n —the convex hull of binary relations that are total, irreflexive, and transitive—and the weak order polytope P_{WO}^n —the convex hull of binary relations that are total, reflexive, and transitive—since their vertices correspond to strict complete rankings and nonstrict 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 P_{WO}^n 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 data sets 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 36 gigabytes of RAM memory shared by two Intel Xeon E5-2680 processors running at 2.40 gigahertz (note that all tests are run on the ASU High Performance Computing Agave Cluster); 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., Betzler et al. 2014, Lu and Boutilier 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{a} and “dispersion” $\phi \in (0, 1]$. These parameters are used to quantify the probability of obtaining a complete ranking \mathbf{a} as

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

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{a} ; as ϕ gets closer to 1, any complete ranking has equal probability of occurring (i.e., this becomes the uniform distribution). Note that the Mallows’ model can be used to generate a set of noisy rankings from a ground truth and a given dispersion parameter that is shared by all of the rankings. On the other hand, Kemeny aggregation returns the maximum likelihood estimator of a model in which each judge provides a noisy estimate of one ground-truth ranking, with each judge possessing the same dispersion or noise parameter ϕ . Hence, each process can be interpreted as being the inverse of the other. 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 nonstrict 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, say a . The alternative in rank position u is tied with the alternative in the next rank position higher (i.e., worse) than u . Ties are repeatedly inserted until the number of tied alternatives reaches or first exceeds a specified threshold, which is set to $0.5n$. For example, let $a = (1, 2, 3, 3, 5)$ and $u = 3$. The next rank position higher than u is 5, and v_5 is the alternative with this rank. Therefore, 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 $V \setminus V'$ 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 nonstrict 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;

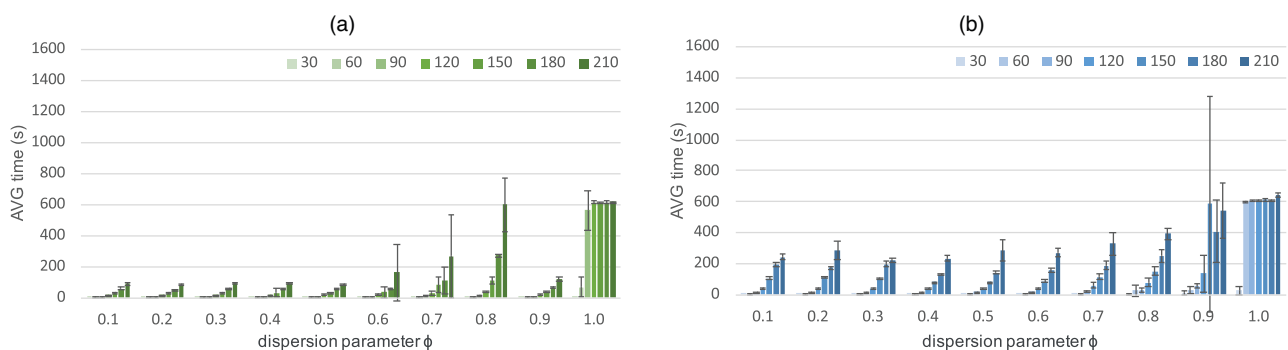
specifically, instances are obtained by sampling complete rankings from the Mallows- ϕ distribution and inserting ties and/or incompleteness, as described in the preceding subsection. We first investigate the effect of varying the dispersion parameter, $\phi \in \{0.1, 0.2, \dots, 0.9, 1.0\}$, and the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

For each of the aforementioned parameter configurations, 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 or 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 Figures 5 and

Figure 5. (Color online) Nonstrict Complete Rankings: (a) Computing Times via GKBP; (b) Computing Times via Brancotte et al. (2015)



Note. GKBP solves nearly all probabilistic instances within the time limit and significantly faster, even for larger n (note that shading in these charts denotes the number of alternatives in the instances).

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).

As shown in Figure 5, for nonstrict complete ranking instances, GKBP finds the optimal solution in less time than the formulation of Brancotte et al. (2015) 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 $n \geq 90$, whereas Brancotte et al. (2015) are not able to solve some instances for $\phi = 0.7$ with $n = 150$, for $\phi = 0.8$ with $n = 210$, and most instances for $\phi \geq 0.9$. Despite the fact that GKBP found an optimal solution faster than the model of Brancotte et al. 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 the model of Brancotte et al. (2015).

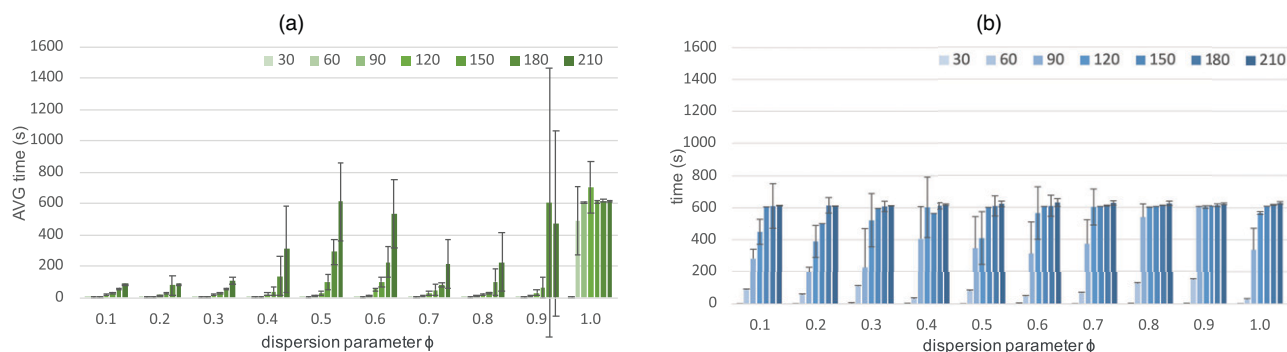
Figure 6 displays the computing times for nonstrict incomplete ranking instances. Before explaining the results, we note that the model of Brancotte et al. (2015) is not originally designed to handle incomplete rankings, whereas GKBP can handle incomplete rankings using the cumulative ranking-matrices defined by Yoo et al. (2020). To compare the models under the same treatment of incomplete rankings, we

normalize $d_{\tau'}$ with the same normalization factor as d_{NPKS} (see Section 3.3). Compared with 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 the model of Brancotte et al. (2015). GKBP attains the optimal solution except for four instances with $\phi = 0.9$ and $n = 180$ and 210 , and most instances with $\phi = 1.0$ and $n \geq 60$, whereas Brancotte et al. (2015) cannot solve most instances, except those for all ϕ with $n = 30, 60, 90$, and 120 . 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, whereas that of Brancotte et al. (2015) is 1.00. We remark that the performance of the model of Brancotte et al. (2015) over these instances is worse without the inclusion of the normalization factor.

4.3. Instances from PrefLib Benchmark Data Set

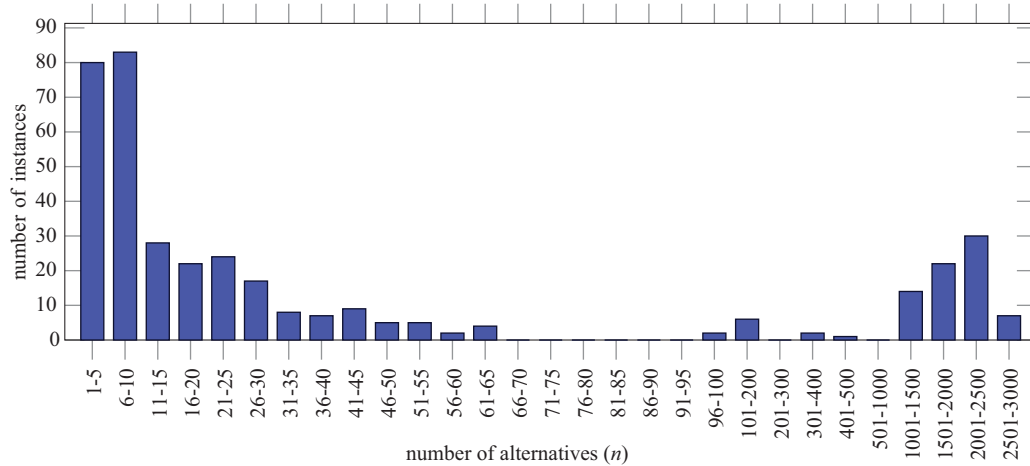
The second set of instances is selected from the library of preference data PrefLib (Mattei and Walsh, 2013), specifically the “Order with Ties—Complete List (TOC)” data set. This benchmark data set consists of 378 instances with differing numbers of alternatives and rankings (i.e., judges) obtained from various domains, and they include real-world data (e.g., figure skating competitions and cross-country skiing and ski jump championship results). The instances include results from Formula One racing and human computation activities, which tend to be relatively

Figure 6. (Color online) Nonstrict Incomplete Rankings: (a) Computing Times via GKBP; (b) Computing Times via Brancotte et al. (2015)



Note. The formulation of Brancotte et al. (2015) cannot reach optimality within the time limit for most instances with higher ϕ and n (note that the color of these charts denotes the number of alternatives in the instances).

Figure 7. (Color online) The Number of Alternatives (n) of Most Instances in the PrefLib Data Set Ranges Between 3 and 65



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 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 $n \geq 1,000$.

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, whereas the time limit is reached with a 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.

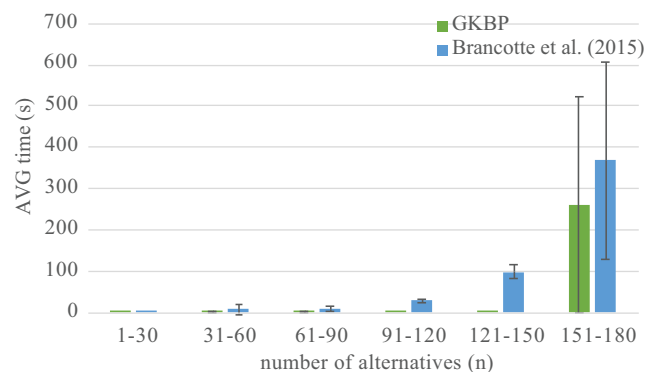
5. A Structural Social Choice Property for Complete Rankings with Ties

To further expedite the solution process of Kemeny aggregation, this section will devise a structural social choice property, which enables the decomposition of 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

Figure 8. (Color online) GKBP Is Faster, on Average, Than the Formulation of Brancotte et al. (2015) over All Instances with $n \leq 300$, Grouped According to the Ranges of n



others, then 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

$$\begin{aligned} &\text{If } \exists v_i \in V, \text{ s.t. } p_{ij} > p_{ji} \text{ (i.e., } v_i \succ^m v_j) \quad \forall v_j \in V \setminus \{v_i\} \\ &\Rightarrow v_i^* > v_j^*, \text{ or equivalently, } r_i^* < r_j^*, \end{aligned}$$

where r^* is the final aggregate ranking and $v_i^* > v_j^*$ indicates that v_i is ranked strictly better than v_j in r^* . At the preliminary screening stage of decision making, keeping a diverse and large set of candidates, rather than selecting few candidates, provides decision makers with broader options. For this reason, CC provides limited usefulness for decision making, since it can only identify one winning alternative (i.e., the Condorcet winner) or one losing alternative (i.e., the Condorcet loser), when it is satisfied. An extended version of the Condorcet winner is the Smith set; the winning (losing) Smith set is the smallest nonempty set of alternatives that defeats (is defeated by, respectively) every alternative outside the set in a pairwise election (Smith 1973). Truchon (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 $\mathcal{V} := \{V^1, \dots, V^K\} \in \mathcal{P}(V)$, 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., $V^k \succ^m V^{k'}$), 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

$$\begin{aligned} &\text{If } \exists \mathcal{V} := \{V^1, V^2, \dots, V^K\} \in \mathcal{P}(V), \text{ s.t. } V^k \succ^m V^{k'}, \text{ for } 1 \leq k < k' \leq K \\ &\Rightarrow v_i^* > v_j^*, \text{ or equivalently, } r_i^* < r_j^*, \quad \forall v_i \in V^k, \quad \forall v_j \in V^{k'}. \end{aligned}$$

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 aggregation (as well as other distance-based methods) may yield more than

Table 2. The Kemeny Optimal Solutions Are Consistent with XCC

	a^1	a^2	a^3	a^4	a^5	a^6	a^7	a^8	Kemeny optimal rankings
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

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 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 the *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).

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}^{XCC}), they are not appropriate for nonstrict 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 $1 \leq i < j \leq 3$. 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 nonstrict rankings is not

Table 3. The Optimal Solution Is Not Consistent with CC and XCC

	a^1	a^2	a^3	a^4	a^5	a^6	a^7	a^8	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

consistent with XCC. In order to overcome this inadequacy, we define a new social choice property in the ensuing subsection.

5.2. The Nonstrict Extended Condorcet Criterion and Its Structural Decomposition

To introduce a substitute to XCC that is suitable for nonstrict 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 $v_i \succ^M v_j$, 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 $v_i \succ^M v_j$ nor $v_j \succ^M v_i$, then there is no decisive majority that prefers v_i over v_j , and vice versa, written as $v_i \approx^M v_j$. Similarly, we say that a decisive majority prefers all alternatives in the subset V^k over all alternatives in the subset $V^{k'}$, written as $V^k \succ^M V^{k'}$, if $p_{ij} > p_{ji} + t_{ij}$, $\forall v_i \in V^k, \forall v_j \in V^{k'}$. If neither $V^k \succ^M V^{k'}$ nor $V^{k'} \succ^M V^k$, then there is no decisive majority that prefers V^k over $V^{k'}$, and vice versa, written as $V^k \approx^M V^{k'}$.

Definition 5.1. Let $\mathcal{V} := \{V^1, V^2, \dots, V^K\}$ s.t. $V^k \approx^M V^{k'}$ for $1 \leq k < k' \leq K$. The nonstrict 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,

$$\text{if } \exists \mathcal{V} := \{V^1, V^2, \dots, V^K\} \in \mathcal{P}(V), \text{ s.t. } V^k \approx^M V^{k'}, \\ \text{for } 1 \leq k < k' \leq K$$

$$\Rightarrow v_i^* > v_j^* \text{ or equivalently, } r_i^* < r_j^*, \forall v_i \in V^k, \forall v_j \in V^{k'}.$$

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 E.1 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 nonstrict 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 nonstrict and complete. This is done through Lemma 5.2 and Theorem 5.3. Beforehand, it is useful to introduce some additional notation.

Notation 1. The reduced instance associated with two subsets of alternatives $V^k, V^{k'} \subset V$, 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, $a_{[k \cup k']}^\ell = a_{[V^k \cup V^{k'}]}^\ell$ and $r_{[k \cup k']}^* = r_{[V^k \cup V^{k'}]}^*$ 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, $a_{\{i,j\}}^\ell = a_{\{v_i, v_j\}}^\ell$ and $r_{\{i,j\}}^* = r_{\{v_i, v_j\}}^*$ are the reduced evaluation from judge ℓ and the optimal reduced ranking with respect to v_i and v_j , respectively.

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

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

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

Lemma 5.2. Let $V^1, V^2 \subset V$ with $V^1 \cap V^2 = \emptyset$. Consider the reduced aggregation problem consisting of input rankings $A_{[1 \cup 2]}$, that is, the part of the evaluations involving only $V^1 \cup V^2$. If $V^1 \succ^M V^2$, then 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 $r_{[1 \cup 2]}^* := \arg \min_{r_{[1 \cup 2]}} \sum_{\ell \in L} d(a_{[1 \cup 2]}^\ell, r_{[1 \cup 2]})$.

Proof. See Appendix C.

Theorem 5.3. Kemeny aggregation satisfies NXCC when the inputs are nonstrict and complete.

Proof. Define the partition $\mathcal{V} = \{V^1, V^2, \dots, V^K\}$, where $V^k = \{v_1^k, v_2^k, \dots, v_{|V^k|}^k\}$ and assume that $V^k \succ^M V^{k'}$ for every k, k' , where $1 \leq k < k' \leq K$. Let r^* be a Kemeny optimal ranking. In order to prove that Kemeny aggregation satisfies NXCC when rankings are nonstrict and complete—which means that all elements in V^k should precede all elements in $V^{k'}$ in 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:

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

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 $k < k'$, then this guarantees an optimal ordered partition $\mathcal{V} = \{V^1, V^2, \dots, V^K\}$, with $V^1 \succ V^2 \succ \dots \succ V^K$; the ordering of the alternatives within each subset (i.e., between $v_i, v_{i'} \in V^k$) can be performed in a subsequent step and is ignored for the rest of the proof. We can derive the following bound on the optimal cumulative intersubset distances:

$$\begin{aligned} \min_{r^*} \sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{\ell \in L} d(a_{[k \cup k']}^\ell, r_{[k \cup k']}^*) \\ \geq \sum_{k=1}^{K-1} \sum_{k'=k+1}^K \min_{r_{[k \cup k']}^*} \sum_{\ell \in L} d(a_{[k \cup k']}^\ell, r_{[k \cup k']}^*). \end{aligned} \quad (12)$$

The optimal solutions from the right-hand side of Inequality (12) correspond to the $K(K-1)/2$ optimal reduced orderings of all subset pairs from $\{V^1, V^2, \dots, V^K\} \in \mathcal{V}$, which can produce preference cycles (i.e., contradictions) when combined. For example, assuming that $v_i \in V^1, v_j \in V^2, v_k \in V^3$, the optimal solutions to the reduced problems can be $r_{i|[1 \cup 2]}^* < r_{j|[1 \cup 2]}^*, r_{j|[2 \cup 3]}^* < r_{k|[2 \cup 3]}^*$ and $r_{k|[1 \cup 3]}^* < r_{i|[1 \cup 3]}^*$. On the other hand, the optimal solution from 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 two-subset subproblems can be interpreted as a relaxed version of the left-hand-side rank aggregation problem. The relaxed problem does not enforce the preference transitivity between all subset orderings. However, when the optimal solutions to the two-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 $V^k \succ^M V^{k'}$, the 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 two-subset orderings can be combined into a K -subset partial ordering $\mathcal{V}^* = \{V^1, V^2, \dots, V^K\}$ with $V^1 \succ V^2 \succ \dots \succ V^K$, 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 nonstrict and complete. \square

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 data set 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 nontrivial 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 sets 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

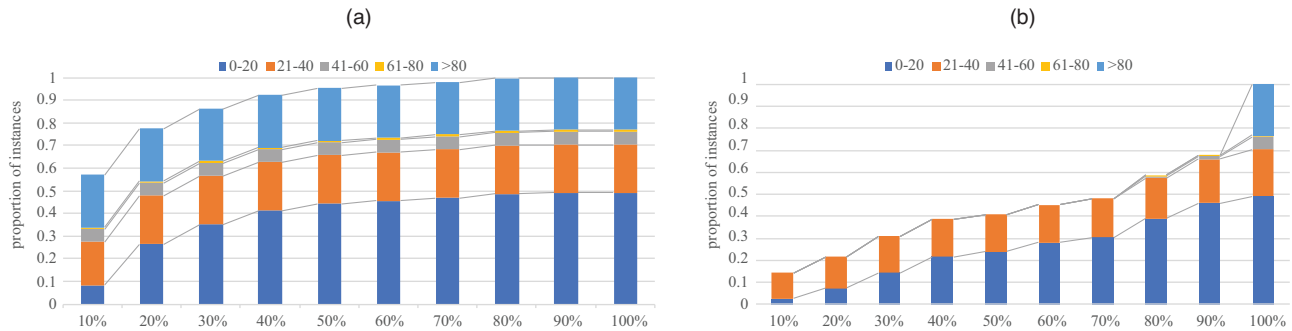
Smith set helps decision makers rule out many irrelevant alternatives, since the complement of the losing Smith set becomes smaller. In addition, practitioners are often most interested in the alternatives that obtain the top positions. For instance, in a recommendation system, it is more important to suggest a set of items that is most likely to be preferred (i.e., top alternative sets), rather than suggesting the middle or least-preferred items (Davidson et al. 2010). The following analysis helps better understand and quantify the practical benefits of NXCC decomposition.

More detailed information about the Smith sets is visualized in Figure 9, which shows the cumulative distribution of the proportion of instances in the data set 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 9(a) shows that approximately 58% of instances have a small winning Smith set (i.e., at most $0.1n$ alternatives in the winning Smith set); in particular, at most 10% of the alternatives (i.e., $0.1n$ 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 9(b) 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.

Next, we test the computational benefits of the decomposition. To do so, this study compares the computing times of solving the full (nondecomposed) 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 nontrivial partition (i.e., $|\mathcal{V}| \geq 2$) and $n \leq 300$ (because CPLEX could not solve nondecomposed instances with more alternatives); the number of instances that meet these conditions is 177. Table 5 shows the *relative improvement* in

Table 4. NXCC Helps Identify the Most Relevant Candidates in the Instances from the PrefLib Data Set

Key instance characteristics	Proportion of instances
Condorcet winner	191 out of 378 (50.53%)
Condorcet loser	62 out of 378 (16.40%)
Nontrivial partition ($ \mathcal{V} \geq 2$)	230 out of 378 (60.85%)

Figure 9. (Color online) Cumulative Distribution: (a) Winning Smith Set Size; (b) Losing Smith Set Size

Note. NXCC yields small winning Smith sets and large losing Smith sets for more than 50% of the instances (note that the color legend of these charts serves to group the tested instances according to the range of number of alternatives they consider).

computing time (i.e., reduction in computing time), which is defined as follows:

$$\frac{(\text{time to solve full (nondecomposed) problem}) - \sum (\text{time to solve decomposed subproblems})}{(\text{time to solve full (nondecomposed) problems})} \times 100\%.$$

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 was solved by multiple processors at the same time, then 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 nonstrict 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 10 minutes (see Figure 5(a)). 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 (nondecomposed) problems and the cumulative time to solve decomposed subproblems (i.e., the numerator in the relative improvement calculation).

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 a few seconds after decomposition, but the original nondecomposed instances were not solved within 600 seconds (note that nondecomposed 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.

Table 5. Applying GKBP for Each Subset After Partitioning Reduces the Computing Time by at Least 25%

$ \mathcal{V} $	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 6. The Effectiveness of NXCC Is More Prominent on Larger, More Difficult Instances

n	90	120	150	180	210
(Absolute) improvement(s)	12.12	47.31	125.84	269.35	560.42+
Standard deviation(s)	0.50	3.60	29.60	72.41	207.23

7. Conclusion

This paper develops a new binary programming formulation and a new social choice property for generalized Kemeny 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 with 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 introduces 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 relatively noncohesive 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 nonstrict and complete or incomplete may be encountered. For example, they could be used to aggregate evaluations of research proposals, entries into a paper competition, journal 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 nonstrict and/or incomplete (see, e.g., Cook et al. 2010, Escobedo et al. 2021). In addition to group decision-making situations, the presented methodologies could be beneficial for various nonhuman decision-making contexts that seek to reconcile long heterogeneous ordered lists, such as in metasearch and bioinformatics (see, e.g., Lin 2010b, Mandal and Mukhopadhyay

2017). Using the social choice property inspired decomposition along with the GKBP binary programming formulation, decision makers may be able to solve certain large rank aggregation problems exactly.

The generalized Kemeny-aggregation binary programming formulation can be applied to complete and incomplete rankings with ties; however, the nonstrict 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 2 of 10 judges evaluate two specific alternatives and their preferences agree on these two alternatives, then 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, this paper only focuses on complete rankings. Future work will involve more computational experiments to test the performance of a suitable social choice property for incomplete rankings encountered in real-world applications, including human computation (e.g., Mao et al. 2013, Kemmer et al. 2020) and wireless sensor networks (e.g., Fishbain and Moreno-Centeno 2016, Skolfield et al. 2020). Moreover, to enhance the computational speed of Kemeny aggregation, parallel programming and/or valid inequalities (e.g., Doignon and Fiorini 2001, Escobedo and Yasmin 2021) will be further explored.

Appendix A. Proof of Theorem 3.3

Theorem 3.3. Let $G = (V, E)$ be an unweighted directed graph for representing a nonstrict complete ranking \mathbf{r} as follows: V is the set of nodes (alternatives) and an each pair of nodes $i, j \in V$ is connected by one or two directed edges; that is, either $(i, j) \in E$ or $(j, i) \in E$, if $r_i < r_j$ (i.e., $v_i > v_j$), or $(i, j) \in E$ and $(j, i) \in E$, 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 unicycles (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 bicycle exists whenever, for every adjacent pair of nodes i_k, i_{k+1} in a path $i_1, i_2, i_3, \dots, i_p$, we have $(i_k, i_{k+1}) \in E'$ and $(i_{k+1}, i_k) \in E'$, for $1 \leq k \leq p-1 < |V'|$. Clearly, when there exists a bicycle in

Figure A.1. Isomorphically Equivalent Digraphs of G' 

G' , it means that every alternative included in the bicycle is tied. That is, when there exist directed cycles $i_1, i_2, \dots, i_{p-1}, i_p, i_1$ and $i_1, i_p, i_{p-1}, \dots, i_2, i_1$ with distinct nodes $i_1, i_2, \dots, i_{p-1}, i_p \in V'$, we have $i_1 \approx i_2 \approx \dots \approx i_p$. Hence, whenever a graph of size 3 has a bicycle, 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 $i_\ell, i_{\ell'} \in \{i_1, i_2, \dots, i_p\}$, 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 a unicycle. 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., unicycles of size 3), as a preference graph without unicyclic triads cannot have any higher-order unicycles (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, whereas 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 (see Figure A.1 as an example).

For Figure 3(a), because $v_i > v_j, v_j > v_k$, and $v_k > v_i$, the preference relation includes the cycle $v_i > v_j > v_k > v_i$, which does not yield a proper ranking of three alternatives). For Figure 3(b), because $v_i > v_j, v_j \approx v_k$, and $v_k > v_i$, the preference relation includes the cycle $v_i > v_j \approx v_k > v_i$, which also does not yield a complete and consistent set of preferences. For Figure 3(c), because $v_i > v_j, v_j \approx v_k$, and $v_k \approx v_i$, the preference relation $v_i > v_j \approx v_k \approx v_i$ is not a consistent set of preferences. Therefore, the graphs with a unicycle 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. \square

Appendix 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) provide a binary programming formulation for P_{wo}^n . For $i, j \in \{1 \dots n\}$, decision variable x_{ij} is defined as

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

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 P_{wo}^n , 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 P_{wo}^n , 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 P_{wo}^n are given as

$$x_{ij} \leq 1, \quad (\text{A.1a})$$

$$x_{ij} + x_{ji} \geq 1, \quad (\text{A.1b})$$

$$x_{ik} - x_{ij} - x_{jk} \geq -1. \quad (\text{A.1c})$$

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

$$y_{ji} \leq 1, \quad (\text{A.2a})$$

$$y_{ji} + y_{ij} \geq 1, \quad (\text{A.2b})$$

$$y_{ki} - y_{ji} - y_{kj} \geq -1. \quad (\text{A.2c})$$

Notice that the combination of Inequalities (A.2a) and (A.2b) gives the domain of y_{ij} . Inequality (A.2b) matches with Constraint (11c) in GKBP. Furthermore, Constraints (A.2a)–(A.2c) hold for any permutation of the labels; therefore, changing i to j , j to k , and k to i yields

$$y_{ij} - y_{kj} - y_{ik} \geq -1. \quad (\text{A.3})$$

Inequality (A.3) is the same as Constraint (11b) in GKBP. Therefore, the constraints of GKBP provide a logically equivalent formulation of the weak order polytope. \square

Appendix C. Proof of Lemma 5.2

Lemma 5.2. Let $V^1, V^2 \subset V$ with $V^1 \cap V^2 = \emptyset$. Consider the reduced aggregation problem consisting of input rankings $A_{[1 \cup 2]}$, that is, the part of the evaluations involving only $V^1 \cup V^2$. If $V^1 \succ^M V^2$, then 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 $r_{i[1 \cup 2]}^* := \arg \min_{r_{i[1 \cup 2]}} \sum_{\ell \in L} d(a_{i[1 \cup 2]}^\ell, r_{i[1 \cup 2]})$.

Proof. We let $r_{i[1 \cup 2]}^*$ be a Kemeny optimal ranking to the reduced problem involving only V^1 and V^2 , and assume that $r_{i[1 \cup 2]}^* \geq r_{j[1 \cup 2]}^*$, for at least one alternative pair v_i, v_j , where $v_i \in V^1, v_j \in V^2$ —that is, there exists at least one alternative in V^1 that 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 $\bar{r}_{i[1 \cup 2]}^*$.

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

$$\begin{aligned} & \sum_{v_j \in V^2} \sum_{\ell \in L} d(a_{i,j}^\ell, r_{i,j}^*) \\ &= \frac{1}{2} \sum_{v_j \in V^2} \sum_{\ell \in L} |\text{sign}(a_{i,j}^\ell - a_{j,i}^\ell) - \text{sign}(r_{i,j}^* - r_{j,i}^*)|. \end{aligned}$$

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

$$\sum_{\ell \in L} d(a_{i,j}^\ell, r_{i,j}^*) = \begin{cases} \sum_{\ell \in L} |\text{sign}(a_{i,j}^\ell - a_{j,i}^\ell) - (-1)| = 2p_{ji} + t_{ij} & \text{if } r_i^* < r_j^* \\ \sum_{\ell \in L} |\text{sign}(a_{i,j}^\ell - a_{j,i}^\ell) - 0| = p_{ij} + p_{ji} & \text{if } r_i^* = r_j^* \\ \sum_{\ell \in L} |\text{sign}(a_{i,j}^\ell - a_{j,i}^\ell) - 1| = 2p_{ij} + t_{ij} & \text{if } r_i^* > r_j^* \\ 0 & \text{otherwise.} \end{cases}$$

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

$$\begin{aligned} & \sum_{v_j \in V^2} \sum_{\ell \in L} d(a_{i,j}^\ell, r_{i,j}^*) \\ &= \sum_{\substack{v_j \in V^k \\ \text{s.t. } r_{i[1 \cup 2]}^* > r_{j[1 \cup 2]}^*}} (2p_{ij} + t_{ij}) + \sum_{\substack{v_j \in V^k \\ \text{s.t. } r_{i[1 \cup 2]}^* = r_{j[1 \cup 2]}^*}} (p_{ij} + p_{ji}) \\ &+ \sum_{\substack{v_j \in V^k \\ \text{s.t. } r_{i[1 \cup 2]}^* < r_{j[1 \cup 2]}^*}} (2p_{ji} + t_{ij}). \end{aligned}$$

From the assumption that $V^1 \succ^M V^2$ 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 ($r_i^* > r_j^*$) aforementioned cases:

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

which results in the following relationship:

$$\begin{aligned} & \sum_{\substack{v_j \in V^k \\ \text{s.t. } r_{i[1 \cup 2]}^* > r_{j[1 \cup 2]}^*}} (2p_{ij} + t_{ij}) + \sum_{\substack{v_j \in V^k \\ \text{s.t. } r_{i[1 \cup 2]}^* = r_{j[1 \cup 2]}^*}} (p_{ij} + p_{ji}) \\ &+ \sum_{\substack{v_j \in V^k \\ \text{s.t. } r_{i[1 \cup 2]}^* < r_{j[1 \cup 2]}^*}} (2p_{ji} + t_{ij}) > \sum_{v_j \in V^2} (2p_{ji} + t_{ij}). \end{aligned}$$

That is,

$$\sum_{v_j \in V^2} \sum_{\ell \in L} d(a_{i,j}^\ell, r_{i,j}^*) > \sum_{v_j \in V^2} \sum_{\ell \in L} d(a_{i,j}^\ell, \bar{r}_{i,j}^*),$$

which means that the optimal ranking r^* , where v_i is tied or dispreferred over some alternatives in V^2 , returns a longer cumulative distance than the ranking \bar{r}^* , where v_i is preferred over all $v_j \in V^2$. This contradicts the assumption that r^* is the optimal ranking, since it does not return the shortest cumulative distance. Therefore, if $V^1 \succ^M V^2$, then 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 $V^1 \succ^M V^2$, then every alternative $v_i \in V^1$ should obtain a better position than every alternative $v_j \in V^2$ in the optimal solution. \square

Appendix D. Pseudocode of Generating Nonstrict Complete and Incomplete Rankings

Algorithms D.1 and D.2 describe how nonstrict 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 a^{-1} be an *alternative ordering* induced from rankings by sorting the alternatives from best to worst, according to their ranks. For example, for $a = (1, 5, 2, 4, 3)$, $a^{-1} = (v_1, v_3, v_5, v_4, v_2)$. Extending this notation, $a^{-1}(i)$ specifies the i th highest ranked alternative in a (Doignon et al. 2004); in the aforementioned example, $a^{-1}(3) = v_5$. When a is a nonstrict ranking, the alternative ordering is obtained by putting alternatives with the same rank position into preference equivalence classes; for example, for $a = (1, 3, 3, 1, 5)$, $a^{-1} = (\langle v_1, v_4 \rangle, \langle v_2, v_3 \rangle, v_5)$. Additionally, $\underline{a}^{-1}|_{V_a}$ is an alternative ordering that is projected to the alternatives in V_a , where V_a is an alternative set that is evaluated by a . Finally, $\text{UniformDist}(L, U)$ denotes the discrete uniform distribution, where L and U are the minimum and maximum values of the distribution.

Algorithm D.1. (Generating Nonstrict Complete Rankings)**Input:** Dispersion: ϕ , reference alternative ordering: \underline{a}^{-1} **Output:** A set of nonstrict complete rankings

```

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

```

Algorithm D.2. (Generating Nonstrict Incomplete Rankings)**Input:** Dispersion: ϕ , alternative set for a : V_a , projected reference alternative ordering: $\underline{a}^{-1}|_{V_a}$ **Output:** A set of nonstrict incomplete rankings

```

1: for  $i = 1, 2, \dots, |V_a|$  do
2:   for  $j = 1, 2, \dots, i$  do
3:      $a^{-1}(j) \leftarrow \underline{a}^{-1}(i)|_{V_a}$  with probability:  $p_{ij} = \phi^{i-j} / (1 + \phi + \dots + \phi^{i-1})$ 
4: for  $i = 1, 2, \dots, |V|$  do
5:   if  $v_i \in V_a$  then
6:      $a_i \leftarrow$  rank position of  $v_i$  in  $a^{-1}$ 
7:   else
8:      $a_i \leftarrow \bullet$ 
9: while (the number of alternatives involved in ties)  $\leq 0.5n$  do
10:   $u \leftarrow \text{UniformDist}(1, h-1)$ , where  $h$  is the worst (highest-ranked) position in  $a$ 
11:   $v \leftarrow a^{-1}(u+1)$ , and then  $a_v \leftarrow u$ 

```

Appendix E. Pseudocode of Decomposition Algorithm**Algorithm E.1.** (Decomposition Algorithm)**Input:** $\{p_{ij}\}, \{p_{ji}\}, \{t_{ij}\}$ **Output:** An ordered partition of subsets $\mathcal{V} = \{V^1, V^2, \dots, V^K\}$

```

1:  $\mathcal{V} = \{\{v_1\}\}$ 
2: for  $i = 2, 3, \dots, |V|$  do
3:   if  $(\exists v_j \in V^1 \text{ s.t. } t_{ij} \geq |p_{ij} - p_{ji}|)$ , or
4:      $(\exists v_j \in V^1 \text{ s.t. } p_{ij} > p_{ji} + t_{ij} \text{ and } \exists v_{j'} \in V^1 \setminus \{v_j\} \text{ s.t. } p_{ji'} > p_{ij'} + t_{ij'})$  then
5:     Put  $v_i$  in  $V^1$  and  $k \leftarrow 2$ 
6:   else if  $\forall v_j \in V^1 \text{ s.t. } p_{ij} > p_{ji} + t_{ij}$  then
7:     Insert  $v_i$  before  $V^1$ , increment the index of subsets after  $V^{\sigma(i)}$  by 1, and  $k \leftarrow 3$ 
8:   else if  $\forall v_j \in V^1 \text{ s.t. } p_{ji} > p_{ij} + t_{ij}$  then
9:     Insert  $v_i$  after  $V^1$ , increment the index of subsets after  $V^{\sigma(i)}$  by 1, and  $k \leftarrow 3$ 
10:  while  $k \leq |\mathcal{V}|$  do
11:    if  $(\exists v_j \in V^k \text{ s.t. } t_{ij} \geq |p_{ij} - p_{ji}|)$ , or
12:       $(\exists v_j \in V^k \text{ s.t. } p_{ij} > p_{ji} + t_{ij} \text{ and } \exists v_{j'} \in V^k \setminus \{v_j\} \text{ s.t. } p_{ji'} > p_{ij'} + t_{ij'})$  then

```

```

13:      Merge subsets from  $V^{\sigma(i)}$  to  $V^k$ 
14:      Decrease the index of subsets after  $V^k$  by  $k - \sigma(i)$  and  $k \leftarrow \sigma(i) + 1$ 
15:    else if  $\forall v_j \in V^k \text{ s.t. } p_{ij} > p_{ji} + t_{ij}$  then
16:       $k \leftarrow k + 1$ 
17:    else if  $\forall v_j \in V^k \text{ s.t. } p_{ji} > p_{ij} + t_{ij}$  then
18:      if  $|\sigma(i) - k| = 1$  and  $|V^{\sigma(i)}| = 1$  then
19:        Move  $V^{\sigma(i)}$  after  $V^k$  and increment the index of subsets after  $V^{\sigma(i)}$  by 1
20:      else
21:        Merge subsets from  $V^{\sigma(i)}$  to  $V^k$ 
22:        Decrease the index of subsets after  $V^k$  by  $k - \sigma(i)$  and  $k \leftarrow \sigma(i) + 1$ 

```

Note. $\sigma(i)$ 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 $\mathcal{V} = \{V^1, V^2, \dots, V^{\kappa(i)}\}$, 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 $i = 2, \dots, n$. 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. *J. Assoc. Inform. Sci. Tech.* 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. *J. Assoc. Comput. Machinery* 55(5):23.
- Amodio S, D'Ambrosio A, Siciliano R (2016) Accurate algorithms for identifying the median ranking when dealing with weak and partial rankings under the Kemeny axiomatic approach. *European J. Oper. Res.* 249(2):667–676.
- Arrow KJ (1951) *Social Choice and Individual Values* (Wiley, New York).
- 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. *Soc. Choice 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 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. *Proc. 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, Cambridge, UK).
- Ceberio J, Irurrozki E, Mendiburu A, Lozano JA (2015) A review of distances for the Mallows and generalized Mallows estimation of distribution algorithms. *Comput. Optim. Appl.* 62(2):545–564.
- Cohen D, Jeavons P, Jefferson C, Petrie KE, Smith BM (2005) Symmetry definitions for constraint satisfaction problems. Van Beek P, ed. *Internat. Conf. Principles Practice Constraint Programming* (Springer, Berlin), 17–31.
- Condorcet JAN (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (L'Imprimerie Royale, Paris).
- Conitzer V, Davenport A, Kalagnanam J (2006) Improved bounds for computing Kemeny rankings. Cohn A, ed. *Proc. 21st National Conf. Artificial Intelligence*, vol. 1 (AAAI Press, Palo Alto, CA), 620–626.
- Cook WD (2006) Distance-based and ad hoc consensus models in ordinal preference ranking. *Eur. J. Oper. Res.* 172(2):369–385.
- Cook WD, Kress M (1985) Ordinal ranking with intensity of preference. *Management Sci.* 31(1):26–32.
- Cook WD, Seiford LM (1978) Priority ranking and consensus formation. *Management Sci.* 24(16):1721–1732.
- Cook WD, Raviv T, Richardson AJ (2010) Aggregating incomplete lists of journal rankings: An application to academic accounting journals. *Accounting Perspect.* 9(3):217–235.
- Cook WD, Golany B, Penn M, Raviv T (2007) Creating a consensus ranking of proposals from reviewers' partial ordinal rankings. *Comput. Oper. Res.* 34(4):954–965.
- Crispino M, Arjas E, Vitelli V, Barrett N, Frigessi A (2019) A Bayesian Mallows approach to nontransitive pair comparison data: How human are sounds? *Ann. Appl. Statist.* 13(1):492–519.
- Critchlow DE (2012) *Metric Methods for Analyzing Partially Ranked Data* (Springer, New York).
- 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. Preprint, submitted October 4, <https://arxiv.org/abs/1510.01647>.
- D'Ambrosio A, Mazzeo G, Iorio C, Siciliano R (2017) A differential evolution algorithm for finding the median ranking under the Kemeny axiomatic approach. *Comput. Oper. Res.* 82(C):126–138.
- Davenport A, Kalagnanam J (2004) A computational study of the Kemeny rule for preference aggregation. Cohn AG, ed. *Proc. 19th National Conf. Artificial Intelligence* (AAAI Press, Palo Alto, CA), 697–702.
- Davidson J, Liebal B, Liu J, Nandy P, Van Vleet T, Gargi U, Gupta S, et al. (2010) The YouTube video recommendation system. *Proc. 4th ACM Conf. Recommender Systems* (ACM, New York), 293–296.
- de Borda JC (1781) *Mémoire sur les élections au scrutin* (L'Académie royale des sciences, Paris).
- Desarkar MS, Sarkar S, Mitra P (2016) Preference relations based unsupervised rank aggregation for metasearch. *Expert Systems Appl.* 49(C):86–98.
- Diaconis P (1988) *Group Representations in Probability and Statistics* (Institute of Mathematical Statistics, Hayward, CA).
- Doignon JP, Fiorini S (2001) Facets of the weak order polytope derived from the induced partition projection. *SIAM J. Discrete Math.* 15(1):112–121.
- 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.
- Dummet M (1998) The Borda count and agenda manipulation. *Soc. Choice Welfare* 15(2):289–296.
- Dwork C, Kumar R, Naor M, Sivakumar D (2001) Rank aggregation methods for the web. *Proc. 10th Internat. Conf. World Wide Web* (ACM, New York), 613–622.
- Emond EJ, Mason DW (2002) A new rank correlation coefficient with application to the consensus ranking problem. *J. Multi-Criteria Decision Anal.* 11(1):17–28.
- Endriss U, Obraztsova S, Polukarov M, Rosenschein JS (2016) Strategic voting with incomplete information. Brewka G, ed. *Proc. 25th Internat. Joint Conf. Artificial Intelligence* (AAAI Press, Palo Alto, CA), 236–242.
- Escobedo AR, Yasmin R (2021) Derivations of large classes of facet-defining inequalities of the weak order polytope using ranking structures. Preprint, submitted March 25, <https://arxiv.org/abs/2008.03799>.
- Escobedo AR, Moreno-Centeno E, Yasmin R (2021) An axiomatic distance methodology for aggregating multimodal evaluations. Preprint, submitted January 23, http://www.optimization-online.org/DB_HTML/2021/01/8223.html.
- Fagin R, Kumar R, Sivakumar D (2003) Efficient similarity search and classification via rank aggregation. *Proc. 2003 ACM SIGMOD Internat. Conf. Management Data* (ACM, New York), 301–312.
- Fagin R, Kumar R, Mahdian M, Sivakumar D, Vee E (2004) Comparing and aggregating rankings with ties. *Proc. 23rd ACM SIGMOD-SIGACT-SIGART Sympos. Principles Database Systems* (ACM, New York), 47–58.
- Farah M, Vanderpooten D (2007) An outranking approach for rank aggregation in information retrieval. *Proc. 30th Annual Internat. ACM SIGIR Conf. Res. Development Inform. Retrieval* (ACM, New York), 591–598.
- Favardin P, Lepellet D, Serais J (2002) Borda rule, Copeland method and strategic manipulation. *Rev. Econom. 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 Appl.* 40(4):1305–1311.
- Fiorini S, Fishburn PC (2004) Weak order polytopes. *Discrete Math.* 275(1–3):111–127.
- Fishbain B, Moreno-Centeno E (2016) Self calibrated wireless distributed environmental sensory networks. *Sci. Rep.* 6:24382.
- Gao Y, Xu K (2019) pRankAggreg: A fast clustering based partial rank aggregation. *Inform. Sci.* 478:408–421.
- Gass S (1998) Tournaments, transitivity and pairwise comparison matrices. *J. Oper. Res. Soc.* 49(6):616–624.
- Good I (1975) The number of orderings of n candidates when ties are permitted. *Fibonacci Quart.* 13:11–18.
- Goodin RE, List C (2006) A conditional defense of plurality rule: Generalizing May's theorem in a restricted informational environment. *Amer. J. Political Sci.* 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 Sci.* 52(9):1394–1408.

- IBM Knowledge Center (2017) CPLEX optimization studio v12.8.0 documentation, <https://www.ibm.com/support/pages/cplex-optimization-studio-v128>.
- IBM Support (2019) RS03137: CPLEX may ignore time limits on highly symmetric models on which a new incumbent is found close to the time limit, <https://www.ibm.com/support/pages/apar/RS03137>.
- Irurozki E, Calvo B, Lozano JA (2016) PerMallows: An R package for Mallows and generalized Mallows models. *J. Statist. Software* 71(12):1–30.
- Keisler J (2004) Value of information in portfolio decision analysis. *Decision Anal.* 1(3):177–189.
- Kemeny JG, Snell LJ (1962) Preference ranking: An axiomatic approach. *Mathematical Models in Social Science* (Ginn, Boston), 9–23.
- Kemmer R, Yoo Y, Escobedo A, Maciejewski R (2020) Enhancing collective estimates by aggregating cardinal and ordinal inputs. Aroyo L, Simperl E, eds. *Proc. 8th AAAI Conf. Human Comput. Crowdsourcing* (AAAI Press, Palo Alto, CA), 73–82.
- 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* 33(3):239–251.
- Kendall MG (1948) *Rank Correlation Methods* (Griffin, London).
- Kenyon-Mathieu C, Schudy W (2007) How to rank with few errors. *Proc. 39th Annual ACM Sympos. Theory Comput.* (ACM, New York), 95–103.
- Kolde R, Laur S, Adler P, Vilo J (2012) Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* 28(4):573–580.
- Liberti L (2008) Automatic generation of symmetry-breaking constraints. Yang B, Du D-Z, Wang CA, eds. *Internat. Conf. Combin. Optim. Appl.* (Springer, Berlin), 328–338.
- Lin S (2010a) Rank aggregation methods. *WIREs Comput. Statist.* 2(5):555–570.
- Lin S (2010b) Space oriented rank-based data integration. *Statist. Appl. Genetics Molecular Biol.* 9(1):20.
- Lu T, Boutilier C (2014) Effective sampling and learning for Mallows models with pairwise-preference data. *J. Machine Learning Res.* 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. *Inform. Sci.* 385–386:55–75.
- Mao A, Procaccia AD, Chen Y (2012) Social choice for human computation. *Workshops 26th AAAI Conf. Artificial Intelligence* (AAAI Press, Palo Alto, CA), 137–142.
- Mao A, Procaccia AD, Chen Y (2013) Better human computation through principled voting. *Proc. 27th AAAI Conf. Artificial Intelligence* (AAAI Press, Palo Alto, CA), 1142–1148.
- Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, Allison KR, et al. (2012) Wisdom of crowds for robust gene network inference. *Nature Methods* 9(8):796–804.
- Marden JI (1996) *Analyzing and Modeling Rank Data* (CRC Press, Boca Raton, FL).
- Martí R, Reinelt G (2011) *The Linear Ordering Problem* (Springer, Berlin).
- Mattei N, Walsh T (2013) PrefLib: A library for preferences. Perny P, Pirlot M, Tsoukiàs A, eds. *Internat. Conf. Algorithmic Decision Theory* (Springer, Berlin), 259–270.
- Miller GA (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psych. Rev.* 63(2):81–97.
- Milosz R, Hamel S (2018) Exploring the median of permutations problem. *J. Discrete Algorithms* 52–53:92–111.
- Moreno-Centeno E, Escobedo AR (2016) Axiomatic aggregation of incomplete rankings. *IIE Trans.* 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, UK).
- Newman A, Vempala S (2001) Fences are futile: On relaxations for the linear ordering problem. Aardal K, Gerards B, eds. *Internat. Conf. Integer Programming Combin. Optim.* (Springer, Berlin), 333–347.
- Schilling MS, Oeser N, Schaub C (2007) How effective are decision analyses? Assessing decision process and group alignment effects. *Decision Anal.* 4(4):227–242.
- Sherali HD, Smith JC (2001) Improving discrete model representations via symmetry considerations. *Management Sci.* 47(10):1396–1407.
- Skolfield JK, Yasmin R, Escobedo AR, Huie LM (2020) A comparison of axiomatic distance-based collective intelligence methods for wireless sensor network state estimation in the presence of information injection. 6th World Forum Internet Things (WF-IoT) (IEEE, Piscataway, NJ), 1–6.
- Smith JH (1973) Aggregation of preferences with variable electorate. *Econometrica* 41(6):1027–1041.
- Steyvers M, Miller B, Hemmer P, Lee MD (2009) The wisdom of crowds in the recollection of order information. Bengio Y, Schuurmans D, Lafferty JD, Williams CKI, Culotta A, eds. *Proc. 22nd Internat. Conf. Neural Inform. Processing Systems* (Curran Associates, Red Hook, NY), 1785–1793.
- 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 (1998) An extension of the Condorcet criterion and Kemeny orders. Cahier 98-15, Centre de Recherche en Économie et Finance Appliquées, Université Laval, Québec, Canada.
- Wald R, Khoshgoftaar TM, Dittman D, Awada W, Napolitano A (2012) An extensive comparison of feature ranking aggregation techniques in bioinformatics. 13th Internat. Conf. Inform. Reuse Integration (IRI) (IEEE, Piscataway, NJ), 377–384.
- 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 Trans. Multimedia* 18(12):2553–2566.
- Yilmaz E, Aslam JA, Robertson S (2008) A new rank correlation coefficient for information retrieval. *Proc. 31st Annual Internat. ACM SIGIR Conf. Res. Development Inform. Retrieval* (ACM, New York), 587–594.
- Yoo Y, Escobedo AR, Skolfield JK (2020) A new correlation coefficient for comparing and aggregating non-strict and incomplete rankings. *Eur. J. Oper. Res.* 285(3):1025–1041.
- Young HP, Levenglick A (1978) A consistent extension of Condorcet's election principle. *SIAM J. Appl. Math.* 35(2):285–300.
- Young P (1995) Optimal voting rules. *J. Econom. Perspect.* 9(1):51–64.

Yeawon Yoo is a postdoctoral fellow in the Department of Applied Mathematics and Statistics and at the SNF Agora Institute at Johns Hopkins University. She earned her PhD in

Industrial Engineering from Arizona State University, Tempe, Arizona, in 2021. She graduated magna cum laude from Pohang University of Science and Technology (POSTECH) in South Korea with a BS in industrial and management engineering in 2016. She has previously worked as an intern at Los Alamos National Laboratory. Her research interests lie in a broad and dynamic set of problems related to group decision making, crowdsourcing, political science, and society.

Adolfo R. Escobedo is an assistant professor in the School of Computing and Augmented Intelligence, Arizona State University. He earned his PhD in industrial and systems engineering from Texas A&M University, College Station, Texas, in 2016. His research interests are in the theory and application of mathematical programming and computing. His applications of interest include the design and analysis of algorithms for group decision making, computational social choice, and crowdsourcing.