Advances in Data Analysis and Classification
https://doi.org/10.1007/s11634-025-00663-4

REGULAR ARTICLE

A distance-based aggregation method for finding consensus in preference-approvals

Alessandro Albano[1,2] · **Maurizio Romano**[3]

Received: 17 February 2024 / Revised: 19 December 2024 / Accepted: 4 December 2025

Abstract

This paper proposes a distance-based aggregation and consensus method for preference-approvals with potential applications in group decision-making, recommendation systems, and social choice theory. Preference-approvals are a type of preference data where individuals provide a list of approved alternatives in addition to a strict ranking. The proposed method aims to synthesise individual preference-approvals into a unified consensus representing the group's collective view. The consensus is the preference-approval, which minimises the average distance within the whole set of voters. The performance of the proposed method is investigated through a sensitivity analysis on the weighting parameter, two simulation studies, and two real-data applications.

Keywords Preference-approvals · Consensus · Preference aggregation · Distance-based method

Alessandro Albano and Maurizio Romano have equally contributed to this work.

✉ Maurizio Romano
romano.maurizio@unica.it

Alessandro Albano
alessandro.albano@unipa.it

1 Department of Economics, Business and Statistics, University of Palermo, Palermo, Italy

2 Sustainable Mobility Center (Centro Nazionale per la Mobilità Sostenibile—CNMS), Milan, Italy

3 Department of Business and Economics, University of Cagliari, Viale Fra Ignazio 17, Cagliari 09123, Italy

Published online: 06 January 2026

1 Introduction

Preference data have gathered significant attention owing to their applications in Group Decision-Making (GDM), recommendation systems, social choice theory, and beyond (see for example Yazidi et al. 2022). Within this context, preference aggregation stands out as a critical task for gaining information from preference data. Specifically, preference aggregation plays a crucial role in synthesising different individual preferences into a collective decision in GDM problems (Liu et al. 2019). Additionally, recommendation systems have embraced preference aggregation techniques to enhance user satisfaction and engagement (Pujahari and Sisodia 2020; Neve and Palomares 2019; Ismailoglu 2022).

Social choice theory, a field deeply rooted in understanding the aggregation of individual preferences to reach a *consensus*, has found new dimensions with advancements in preference aggregation methods (List 2022). The exploration of diverse mathematical models and aggregation algorithms has enriched the theoretical foundations of social choice theory, paving the way for robust analyses (Patty and Penn 2015).

Preference data manifest in various types, and one increasingly popular form is preference-approvals (see Brams 2008, Chapter 3, (Brams and Sanver 2009) and (Sanver 2010)).

Preference-approval structures combine the preference information of both preference rankings and approval voting, extending the classical ordinal preference model. These structures are particularly useful in different contexts. For instance, consider a scenario where a group of stakeholders is tasked with selecting a design for a new product. Instead of merely ranking the available designs, participants can provide a list of approved designs, indicating which options they find suitable. This dual information, comprising both ranking and approval, forms a preference-approval structure.

In the literature, preference-approvals are studied using diverse methods, including statistical techniques like clustering, distance-based approaches, and consensus measurement, along with conceptual and theoretical approaches. For instance, Kruger and Sanver (2021) investigated preference-approvals and noted challenges in combining ranking and approval information. They demonstrated that aggregating preference-approvals by decomposing rankings and approvals might be dictatorial, suggesting that a single individual or a small group could disproportionately influence the resulting decision. Kamwa (2023) studied preference-approvals and fallback voting by exploring some other properties and comparing these two voting systems. Barokas and Sprumont (2021) introduced a social choice rule *majority approval* and compared it with other rules. Additionally, Barokas (2022) developed an axiomatic approach to allocation rules, which is mathematically equivalent to preference-approvals but distinct from voting rules. Ye et al. (2024) extended classic preference-approvals proposing a novel structure based on a three-way group consensus decision-making approach. They also constructed a non-additive model over the framework of granular computing.

In bridging the gap to statistical methodologies, Albano et al. (2024) proposed novel pseudometrics for clustering alternatives in preference-approvals. Erdamar et al. (2014) focused on measuring consensus in groups of agents when they express their preferences over a fixed set of alternatives. They introduced a novel metric between preference-approvals and offered an axiomatic characterization.

In recent years, there has been a growing interest in preference-approval aggregation. Nevertheless, it remains an open and challenging research area.

Liang et al. (2018) proposed a method to fuse individual preference-approval structures in GDM. They focus on the weight of various preferences in the decision-making process, using prospect theory as a basis for their approach. Dong et al. (2021) presented axioms that imply the existence of a unique distance function of preference-approval structures. They further study a preference aggregation model in the group decision-making context based on these axioms. Escobedo et al. (2022) introduced a multimodal data aggregation methodology featuringoptimisation models and algorithms for jointly aggregating heterogeneous ordinal evaluations. Their approach is tailored for use with axiomatic distances rooted in social choice theory. Liu et al. (2023) deal with a multi-criteria group decision-making problem with preference approval structures. They propose a method that takes into account individual semantics and consensus in large-scale group decision-making. Albano et al. (2023) introduced a family of distances for preference-approvals. They also proposed a method for aggregating voters' preferences that relies on the Borda count extended to weak orders, paired with an average approval vector. The resulting aggregated structures are then called *Representative Preference-Approvals* (here abbreviated as RPA).

In this paper, we propose a novel distance-based aggregation and consensus method for preference-approvals, called DIVA (Divide and Conquer for Preference Approval). The foundation of DIVA draws inspiration from the classic axiomatic Kemeny approach, the most famous method for aggregating preference rankings. That is, the *consensus* is defined as the preference-approval that minimises the average distance across the entire set of voters. To quantify the disagreement between preference-approvals, we employ a family of distances defined by Erdamar et al. (2014), which is rooted in the well-established Kemeny distance. In this way, our approach leverages the benefits of widely recognised techniques for Kemeny ranking exploration while innovatively extending its capabilities to optimise for the approval search. Further, the DIVA algorithm is designed to consistently yield admissible solutions. This ensures the compatibility of the ranking and approval components.

The paper is structured as follows. In Sect. 2, we delve into preference-approval structures, elucidating their mathematical properties, codification methods, and the metric employed to quantify disagreement. Section 3 introduces the DIVA algorithm, our novel approach for aggregating a set of preference-approvals to establish a consensus. In Sect. 4, we present empirical studies to assess the flexibility of our method in assigning different weights to the ranking and approval components. Moving on to Sect. 5, we present two simulation studies that aim to evaluate the accuracy, reliability, and velocity of the proposed method. Additionally, Sect. 6 shows the application of the DIVA method in two real-world scenarios: the French Presidential Election (2002) and the Formula 1 World Championship (1950). Finally, Sect. 7 is dedicated to the discussion and conclusions drawn from our findings.

2 Notation

Formally, given a finite set of alternatives $X = \{x_1, \ldots, x_n\}$, the ranking π is a mapping from X to the set of ranks $\{1, \ldots, n\}$, endowed with the natural ordering of integers.

If the *n* items are ranked in *n* distinguishable ranks, a *linear order* is achieved (Cook 2006), $\pi \in L(X)$ where $L(X)$ denotes the set of linear orders on X.

When some items receive the same preference, then a *weak order* is obtained: $\pi \in W(X)$ where $W(X)$ denotes the set of weak orders on X. A weak order on X is also defined as a complete and transitive binary relation on X.

Let us denote with $\succ$ and $\sim$ the asymmetric, and the symmetric parts of π, respectively: $x_i \succ x_j$ if x_i is preferred to x_j, and $x_i \sim x_j$ if they are equally likeable. Given the set X, with $\mathcal{P}(X)$ we denote its power set, i.e., $I \in \mathcal{P}(X) \Leftrightarrow I \subseteq Y$.

2.1 Preference-approval

We have to consider that a set of voters, $V = \{v_1, \ldots, v_m\}$, with $m \geq 2$, have to express their opinions over $X = \{x_1, \ldots, x_n\}$ a finite set of alternatives, with $n \geq 2$.

Therefore, we assume that each voter ranks the alternatives in X by means of a weak order and, additionally, assesses each alternative as either acceptable or unacceptable by partitioning X into A, the set of *acceptable* alternatives, and $U = X \setminus A$, the set of *unacceptable* alternatives, where A and U can be empty sets.

We also assume the following generic consistency condition: if x_i is preferred to x_j in π, this preference should also be reflected in their approvals, and vice versa. In other words, ranking and approval must be concordant.

Definition 1 A *preference-approval* on X is a pair $(\pi, A) \in W(X) \times \mathcal{P}(X)$ satisfying the following conditions:

$$\forall x_i, x_j \in X \ \big((x_i \succ x_j \text{ and } x_j \in A) \Rightarrow x_i \in A\big) \tag{1}$$

$$\forall x_i, x_j \in X \ \big((x_i \succ x_j \text{ and } x_i \in U) \Rightarrow x_j \in U\big) \tag{2}$$

$$\forall x_i, x_j \in X \ \big((x_i \sim x_j \text{ and } x_i \in A) \Rightarrow x_j \in A\big) \tag{3}$$

$$\forall x_i, x_j \in X \ \big((x_i \sim x_j \text{ and } x_i \in U) \Rightarrow x_j \in U\big) \tag{4}$$

$$\forall x_i, x_j \in X \ \big((x_i \in A \text{ and } x_j \in U) \Rightarrow x_i \succ x_j\big) \tag{5}$$

$$\forall x_i, x_j \in X \ \big((x_i \in U \text{ and } x_j \in A) \Rightarrow x_i \prec x_j\big) \tag{6}$$

With $\mathcal{R}(X)$ we denote the set of preference-approvals on X. We now illustrate preference-approval structures through the following example.

Table 1 Formulas to derive the cardinality of the universe, $\Omega(n)$

Structure	$\Omega(n)$
Approvals	2^n
Linear orders	$n!$
Weak orders	$n!(\log_2 e)^{n+1}/2$
Preference-approvals	$\sum_{r=0}^{n}(r+1)!\,\boldsymbol{S}_n^{(r)}$

Table 2 Number of approvals, linear orders, weak orders, and preference-approvals

n	Approvals	Linear orders	Weak orders	Preference-approvals
2	4	2	3	8
5	32	120	541	2 612
10	1 024	3 628 800	102 247 563	862 440 068
20	1 048 576	$\sim 2.4 \cdot 10^{18}$	$\sim 2.6 \cdot 10^{21}$	$\sim 4.2 \cdot 10^{22}$

Example 1 Consider the preference-approval pair $(\pi, A) \in \mathcal{R}(\{x_1, \ldots, x_8\})$ represented by

$$\begin{array}{ccc} & x_4 & \\ x_1 & & x_6 \\ & x_2 & \\ \hline \\ & x_3 & \\ x_5 & x_7 & x_8 \end{array}$$

This representation indicates that alternatives in the same row are considered indifferent, while alternatives in the upper rows are preferred to those in the lower rows. Alternatives above the line are acceptable (i.e., $A = \{x_1, x_2, x_4, x_6\}$), while those below the line are deemed unacceptable (i.e., $U = \{x_3, x_5, x_7, x_8\}$).

Formulas to compute the cardinality of the universe, $\Omega(n)$, of approvals, linear orders, weak orders, and preference-approvals are observable in Table 1. Moreover, the formulas for computing the number of weak orders and preference-approvals are due to Good (1980) and Albano et al. (2023), respectively. Specifically, $S_n^{(r)}$ is a Stirling integer (number) of the second kind defined by David and Barton 1962, Abramowitz and Stegun 1964, and more thoroughly explained in Fisher and Yates 1953, while *r* denotes the number of distinct positions in a weak order on *n* alternatives, also known as *buckets* (Fürnkranz and Hüllermeier 2011).

Table 2 and Fig. 1 show the combinatorial explosion that occurs as the number of alternatives increases. Table 2 includes the cardinality of the universe, $\Omega(n)$, of approvals, linear orders, weak orders, and preference-approvals when the number of alternatives is $n = \{2, 5, 10, 20\}$. While Fig. 1 shows the $\Omega(n)$ in the log scale.

It is evident that the intricate nature of the preference-approval space, characterized by its rapidly increasing cardinality as the number of alternatives grows, presents an even harder challenge in the development of robust algorithms for effective preference aggregation.

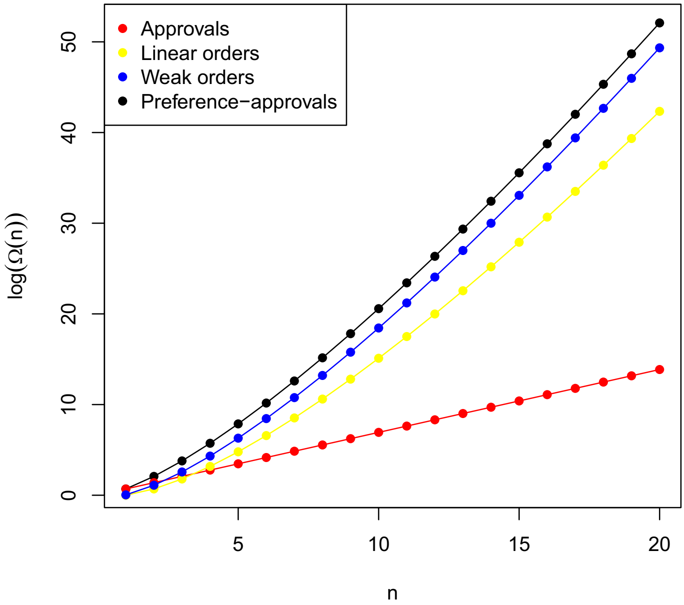

Fig. 1 Comparison of cardinalities of the universes in the log scale

2.2 Codifications

For a given linear order $\pi \in L(X)$, the position of each alternative $x_i \in X$ in π is determined by the mapping $P_\pi : X \longrightarrow \{1, \dots, n\}$, where each alternative is assigned a position based on its ranking. Specifically, the alternative ranked first is assigned position 1, the second-ranked alternative is assigned position 2, and so forth.

In contrast, various methods exist for assigning positions to alternatives in weak orders. For instance, when two alternatives share the first place, they may be assigned position 1, 2, or even 1.5.

Another strategy, employed by García-Lapresta and Pérez-Román (2011), draws inspiration from the works of Smith (1973), Black (1976), and Cook and Seiford (1982). For a given weak order $\pi \in W(X)$, the position of $x_i \in X$ in π is determined by the mapping $P_\pi : X \longrightarrow [1, n]$ s.t.:

$$P_\pi(x_i) = n - \#\{x_k \in X \mid x_i \succ x_k\} - \frac{1}{2} \cdot \#\{x_k \in X \setminus \{x_i\} \mid x_i \sim x_k\}. \quad (7)$$

Given $A \subseteq X$, the *indicator function* (or *characteristic function*) of A, $I_A : X \longrightarrow \{0, 1\}$, is defined as:

$$I_A(x_i) = \begin{cases} 1, & \text{if } x_i \in A, \\ 0, & \text{if } x_i \in X \setminus A. \end{cases} \quad (8)$$

Remark 1 Every preference-approval $(\pi, A) \in \mathcal{R}(\{x_1, \dots, x_n\})$ can be codified in terms of $P_\pi(x_i)$ (Eq. (7)) and $I_A(x_i)$ (Eq. (8)) as follows:

$$\big(P_\pi(x_1), P_\pi(x_2), \dots, P_\pi(x_n)\big) \;\&\; \big(I_A(x_1), I_A(x_2), \dots, I_A(x_n)\big). \tag{9}$$

Example 2 Consider the preference-approval $(\pi, A) \in \mathcal{R}(\{x_1, x_2, x_3, x_4\})$ represented by

$$\begin{array}{c} x_1 \\ \hline \\ x_3 \\ x_2 \end{array}$$

Following Eq. (9), (π, A) is codified as $(1, 3, 2) \& (1, 0, 0)$.

2.3 Axiomatic distance for preference-approvals

For a given pair of preference-approvals $\big((\pi_1, A_1), (\pi_2, A_2)\big) \in \mathcal{R}(X)$, Erdamar et al. (2014) introduced a distance metric, $d_\lambda\big((\pi_1, A_1), (\pi_2, A_2)\big)$, which combines ranking and approval distances through a convex combination. The Kemeny metric is used to measure ranking disagreement $d_K(\pi_1, \pi_2)$:

$$d_K(\pi_1, \pi_2) = \sum_{\substack{i,j=1 \\ \forall i<j}}^{n} \mid \text{sgn}\ \left(P_{\pi_1}(x_j) - P_{\pi_1}(x_i)\right) - \ \text{sgn}\ \left(P_{\pi_2}(x_j) - P_{\pi_2}(x_i)\right) \mid,$$

where the *sgn function*, returns 1 for positive inputs, -1 for negative inputs, and 0 for zero.

Equivalently, taking into account the *score matrix* $O_R(x_i, x_j)$ (May (1952) and Fishburn (1973), Kemeny and Snell 1962, p. 11, (Emond and Mason 2002)):

$$d_K(\pi_1, \pi_2) = \sum_{\substack{i,j=1 \\ \forall i<j}}^{n} |O_{\pi_1}(x_i, x_j) - O_{\pi_2}(x_i, x_j)|, \tag{10}$$

where the *score matrix* $O_R(x_i, x_j)$ codifies the order between two alternatives in a weak order $\pi \in W(X)$:

$$O_R(x_i, x_j) = \begin{cases} 1, & \text{if } x_i \succ x_j \\ 0, & \text{if } x_i \sim x_j \\ -1, & \text{if } x_j \succ x_i \end{cases} \tag{11}$$

Note that $d_K(\pi_1, \pi_2) \in [0, n \cdot (n-1)]$.

The approval disagreement is measured by the Hamming metric $d_H(A_1, A_2)$:

$$d_H(A_1, A_2) = \sum_{i=1}^{n} |I_{A_1}(x_i) - I_{A_2}(x_i)|. \tag{12}$$

Note that $d_H(A_1, A_2) \in [0, n]$.

To aggregate d_K and d_H into a global distance, both metrics are normalized to the same range $[0, 1]$ by dividing by their maximum distances.

The mappings $d_R : \mathcal{R}(X) \times \mathcal{R}(X) \longrightarrow [0, 1]$ and $d_A : \mathcal{R}(X) \times \mathcal{R}(X) \longrightarrow [0, 1]$ are defined s.t.:

$$d_R\big((\pi_1, A_1), (\pi_2, A_2)\big) = \frac{d_K(\pi_1, \pi_2)}{n \cdot (n-1)} \tag{13}$$

$$d_A\big((\pi_1, A_1), (\pi_2, A_2)\big) = \frac{d_H(A_1, A_2)}{n} \tag{14}$$

The normalized distances are then aggregated into a final preference-approval distance, $d_\lambda : \mathcal{R}(X) \times \mathcal{R}(X) \longrightarrow [0, 1]$, defined as

$$\begin{aligned} &d_\lambda\big((\pi_1, A_1), (\pi_2, A_2)\big) = \\ &\lambda \cdot d_R\big((\pi_1, A_1), (\pi_2, A_2)\big) + (1-\lambda) \cdot d_A\big((\pi_1, A_1), (\pi_2, A_2)\big), \end{aligned} \tag{15}$$

where $\lambda \in [0, 1]$ is a parameter controlling the relative relevance of the two components. Considering Eqs. (10) and (12), Eq. (15) can be rewritten as

$$\begin{aligned} &d_\lambda\big((\pi_1, A_1), (\pi_2, A_2)\big) = \\ &\frac{\lambda}{n \cdot (n-1)} \cdot \sum_{\substack{i, j = 1 \\ i < j}}^{n} |O_{\pi_1}(x_i, x_j) - O_{\pi_2}(x_i, x_j)| + \\ &\frac{1-\lambda}{n} \cdot \sum_{i=1}^{n} |I_{A_1}(x_i) - I_{A_2}(x_i)|. \end{aligned} \tag{16}$$

The distance metric d_λ, originally introduced by Erdamar et al. (2014), was later extended and generalized by Albano et al. (2023). They computed the disagreement between each pair of alternatives and subsequently combined the outcomes through a power mean with an exponent *r*. This extended metric was denoted as D_λ. The authors demonstrated that setting the exponent *r* to 1 yields D_λ equivalent to the original d_λ.

In this paper, we adopt the simplicity of the d_λ proposal by Erdamar et al. (2014) (equivalent to D_λ with $r = 1$). However, it's worth noting that our algorithm is versatile and can seamlessly accommodate the more general D_λ with any specified exponent r.

3 The proposal

Given a $m \times 2n$ matrix Π, whose l-th row represents the preference-approval associated with the l-th judge, the consensus preference-approval $(\hat{\pi}, \hat{A})$ is found by minimizing the average distance function d_λ for fixed λ:

$$(\hat{\pi}, \hat{A})_\lambda = \underset{(\pi,A)\in R^{\Omega(n)}}{\arg\min} \sum_{l=1}^{m} d_\lambda((\pi^{(l)}, A^{(l)}), (\pi, A)), \tag{17}$$

where $R^{\Omega(n)}$ is the universe of all preference-approvals with n objects.

By construction, the minimization of d_λ entails the simultaneous minimization of both rank and approval distances. Therefore, a first attempt to reduce the problem is to find marginal solutions $\hat{\pi}_M$ and $\hat{A}_M$ such that:

$$\left(\hat{\pi}_M = \underset{\pi\in S^n}{\arg\min} \sum_{l=1}^{m} d_K(\pi^{(l)}, \pi), \; \hat{A}_M = \underset{A\in\{0,1\}^n}{\arg\min} \sum_{l=1}^{m} d_H(A^{(l)}, A) \right), \tag{18}$$

where S^n ($\{0,1\}^n$) is the universe of weak orders (approvals) of n elements. Moreover, $d_H(A^{(l)}, A)$ and $d_K(\pi^{(l)}, \pi)$ are respectively the Hamming and the Kemeny distance between the preference and the approval part of the l-th row and the candidate consensus.

Unfortunately, if the two optimisations are carried out separately, non-admissible solutions may be generated. That is, the marginal consensus ranking $\hat{\pi}_M$ may not be compatible with the the consensus approval $\hat{A}_M$, violating the conditions in the Eqs. 1, 2, 3, 4, 5 and 6.

Example 3 Let us consider the following set of voters, Table 3.

In this case, the consensus ranking found using the D'Ambrosio et al. (2015) method $\hat{\pi}_M = (1,1,2)$ and the consensus approval $\hat{A}_M = (1,0,0)$ found by calcu-

Table 3 Preference-approval data for Example 3

Preference Approval votes	π		A
(π_1,A_1)	(1, 1, 2)	&	(1,1,0)
(π_2,A_2)	(1, 2, 1)	&	(1,1,1)
(π_3,A_3)	(2, 1, 3)	&	(0,0,0)
(π_4,A_4)	(1, 2, 2)	&	(0,0,0)
(π_5,A_5)	(1, 1, 2)	&	(1,1,0)
(π_6,A_6)	(1, 2, 3)	&	(1,0,0)
(π_7,A_7)	(2, 2, 1)	&	(0,0,0)

lating the element-wise median of the binary approval matrices for all judges, are incompatible. As a matter of fact, the pair $(\hat{\pi}_M = (1, 1, 2), \hat{A}_M = (1, 0, 0))$ violates the condition in Eq. 1.

To cope with this problem, the Divide and Conquer for Preference Approval (DIVA) algorithm, designed to minimise both the ranking and the approval distances whilst avoiding non-admissible solutions, is hereby proposed in Alg. 1. After checking the compatibility of the marginal solutions, the DIVA algorithm works through two phases: *from rank to approval* and *from approval to rank*. Each phase provides a distinct set of solutions while considering, for any possible admitted value, to have a certain fixed ranking/approval.

First, the data matrix Π is split in half to define the two submatrixes that compose it $\Pi = (\Pi^r, \Pi^a)$, respectively, containing the ranking and the approval component. Next, the algorithm starts the first phase, *from rank to approval*. Therefore, it uses a consensus ranking function $f_c(\cdot)$ to estimate it (π_1) considering the ranking data matrix Π^r, and initialises the first approval candidate A' as the rounded by-column average Π^a. Thus, all the possible combinations of compatible approvals are evaluated and compared by keeping fixed π_1 while using the averaged d_λ s.t.:

$$r_1 = \left(\pi_1, \arg\min_{A} \left(\bar{d_\lambda}((\pi_1, A), \Pi)\right)\right), \qquad \text{with } A \in \{0, 1\}^n. \tag{19}$$

In this way, the first set of possible solutions r_1 is composed by the combination of the fixed π_1 and those approvals that minimise the considered distance.

The algorithm now starts with the last phase, *from approval to rank*. Following the same philosophy as the previous one, here, the set of solutions is found by keeping fixed an approval $A \in \{0, 1\}^n$. By construction, all the approved items (Π_1^r) must be in the top position of the estimated consensus ranking; on the other side, those that are not approved (Π_0^r) must be at the bottom. This information provides an interesting and helpful insight to estimate a consensus ranking that will be the best one across all the compatible with *A*:

$$\pi' = [f_c(\Pi_1^r), f_c(\Pi_0^r)] \tag{20}$$

with

$$\Pi_1^r = [\ \Pi^r[:;j]\ ] \qquad \forall j = 1 \cdot a_1, \ldots, n \cdot a_n \qquad \forall a \in A$$

$$\Pi_0^r = [\ \Pi^r[:;j]\ ] \qquad \forall j = 1 \cdot \neg a_1, \ldots, n \cdot \neg a_n \qquad \forall a \in A$$

Therefore, the second set of possible solutions r_2 is composed by the combination of each fixed approval $A \in \{0, 1\}^n$ with the associated consensus ranking π' that minimise the considered distance:

$$r_2 = \arg\min_{(\pi', A)} \left(\bar{d_\lambda}((\pi', A), \Pi)\right) \tag{21}$$

Thus, the final solution $\hat{r} \in \{r_1 \cup r_2\}$ is the one that minimises the average distance $\hat{r}_\lambda = d_\lambda(\hat{r}, \Pi)$ between all those founds in the two phases:

$$\hat{r} = \arg\min_r \left(\bar{d_\lambda}(r, \Pi)\right) \qquad \text{with } \; r \in \{r_1 \cup r_2\} \tag{22}$$

Input: Π; $f_c(\cdot)$;

1: **Define:**
2: $\Pi^r = \Pi[:;1,\ldots,n]$
3: $\Pi^a = \Pi[:;(n+1),\ldots,2n]$
4: $\pi_1 = f_c(\Pi^r)$
5: $A' = \left[\ \overline{\Pi^a[:;1]},\ \ldots,\ \overline{\Pi^a[:;n]}\ \right]$
6: **if** $r = (\pi_1, A')$ not violates Eqs. 1-6 **then** [Eq. 18]
7: $\quad r = (\pi_1, A')$
8: $\quad$ GOTO line 25
9: **end if**
10: **From rank to approval** Given a fixed rank find the best approval
11: **for each** $A \in \{0,1\}^n$ **do**
12: $\quad$ **if** $\bar{d_\lambda}((\pi_1, A), \Pi) < \bar{d_\lambda}((\pi_1, A'), \Pi)$ **then**
13: $\quad\quad A' = A$
14: $\quad$ **end if**
15: **end for**
16: $r = (\pi_1, A')$ [Eq. 19]
17: **From approval to rank** Given a fixed approval find the best rank
18: **for each** $A \in \{0,1\}^n$ **do**
19: $\quad \Pi^r_1 = [\ \Pi^r[:;j]\ ] \qquad \forall j = 1 \cdot a_1, \ldots, n \cdot a_n \qquad \forall a \in A$
20: $\quad \Pi^r_0 = [\ \Pi^r[:;j]\ ] \qquad \forall j = 1 \cdot \neg a_1, \ldots, n \cdot \neg a_n \qquad \forall a \in A$
21: $\quad \pi' = [f_c(\Pi^r_1), f_c(\Pi^r_0)]$ [Eq. 20]
22: $\quad$ **if** $\bar{d_\lambda}((\pi', A), \Pi) < \bar{d_\lambda}(r, \Pi)$ **then**
23: $\quad\quad r = (\pi', A')$ [Eq. 21]
24: $\quad$ **end if**
25: **end for**

Output: $\{\hat{r} = r\}$; $\{\hat{r}_\lambda = \bar{d_\lambda}(r, \Pi)\}$ [Eq. 22]

Algorithm 1 Divide and conquer for preference approval (DIVA)

The DIVA algorithm has been implemented in `R` (R Core Team 2023) and is currently available at the following GitHub repository: urlr.me/pPRmY4. Furthermore, a dedicated package will soon be released. It's worth noting that, despite being able to work with any type of algorithm for finding the consensus ranking $f_c(\cdot)$, by considering the state of the art our default implementation relies on the `ConsRank` `R` package (D'Ambrosio 2023), which implements many algorithms for finding the consensus ranking (D'Ambrosio et al. 2015, 2017; Albano and Plaia 2021).

4 Impact of λ on the DIVA procedure

The parameter λ plays an important role in the DIVA algorithm, as it influences the results in two specific ways: (i) it can modify the consensus preference-approval solution $\hat{r}$ produced by the algorithm, and (ii) it affects the quality of the solution, $\hat{r}_\lambda = \bar{d}_\lambda(r, \Pi)$.

4.1 Impact of λ on the DIVA solution $\hat{r}$

The influence of λ on the consensus preference-approval arises only in specific cases. To illustrate this, we note that the DIVA algorithm begins by separately optimizing the ranking and approval components, as described in Eq. 18. The first attempt checks whether the marginal consensus ranking π_1, obtained by minimizing the Kemeny distance, is compatible with the marginal consensus approval A', obtained by minimizing the Hamming distance d_H, (see Step 7 on the Algorithm 1).

If π_1 and A' are compatible, the parameter λ does not affect the final consensus preference-approval. In this scenario, the same solution is obtained regardless of the value of λ, which coincides with the solution derived from computing the marginal consensus ranking and marginal consensus approval (Eq. 18).

In contrast, when π_1 and A' are not compatible, the algorithm proceeds to its two phases, *from rank to approval* and *from approval to rank*. Here, the solutions obtained in each phase are evaluated and compared using the average distance function d_λ, which explicitly depends on λ (see Eq. 19 and Eq. 21). In this case, λ directly influences the choice of the final consensus preference-approval by determining the weight assigned to the ranking and approval distances. Specifically, higher values of λ place greater emphasis on minimizing the ranking distance, while lower values prioritize the approval distance.

To empirically assess the extent to which λ influences the consensus solution, we performed simulation studies inspired by the approach proposed in Dong et al. (2021). The steps of our simulations are as follows. Firstly, we generate preference-approvals and calculate the marginal consensus ranking $\hat{\pi}_M$ and marginal consensus approval $\hat{A}_M$ as described in Eq. 18. Then, by applying the DIVA algorithm to the whole set of preference-approvals we obtain the joint solution $\hat{r} = (\hat{\pi}, \hat{A})$. Finally, we study how the ranking and approval component of the DIVA solution deviates from $\hat{\pi}_M$ and $\hat{A}_M$ as a function of λ. To quantify this difference, we introduce two indices SCR and SMA, which are based respectively on d_R and d_A defined in Eqs. 13 and 14.

$$\mathrm{SCR}(\lambda) = d_R(\hat{\pi}_M, \hat{\pi}_{(\lambda=\lambda_0)}), \tag{23}$$

$$\mathrm{SMA}(\lambda) = d_A(\hat{A}_M, \hat{A}_{(\lambda=\lambda_0)}). \tag{24}$$

To generate preference-approvals, a vector ($\boldsymbol{\mu}$) is populated by random means that are drawn from a normal distribution: $\mu \sim \mathcal{N}(10, 1)$, whilst a $n \times n$ identity matrix is used as covariance matrix (Σ) s.t.:

$$\Sigma = \begin{bmatrix} 10 & 0 & \dots & 0 & 0 \\ 0 & 10 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 10 & 0 \\ 0 & 0 & \dots & 0 & 10_n \end{bmatrix}.$$

Thus, true ranking configurations Γ are sampled from a multivariate normal distribution: $\Gamma \sim \text{MVN}(\boldsymbol{\mu}, \Sigma)$. To derive preference-approvals, preference rankings are extracted from Γ, while the number of approved alternatives is generated from a Poisson distribution $\mathcal{P}(\alpha)$. Where, $\alpha = \{5, 10, 25\}$ respectively for $m = n = \{10, 20, 50\}$.

This process is repeated 500 times, and the average results are reported in Fig. 2.

The results show that SCR and SMA follow an inverse pattern. Specifically, SCR decreases with increasing λ, suggesting that placing greater emphasis on rankings causes the DIVA solution's ranking component to move closer to the marginal consensus ranking $\hat{\pi}_M$. Conversely, SMA increases with λ, reflecting that as less weight is given to the approval component, the DIVA solution's approval component diverges further from the marginal consensus approval $\hat{A}_M$. Notably, there is an

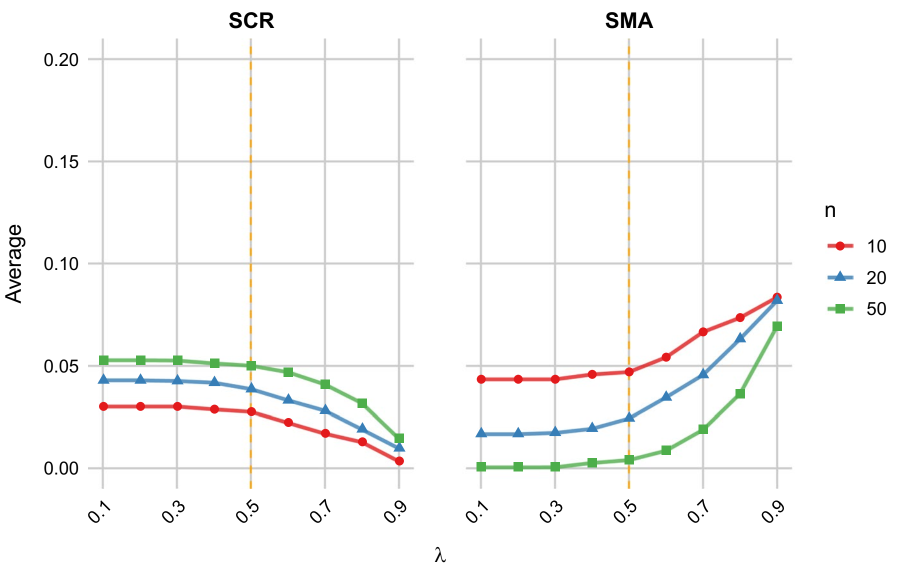

Fig. 2 Average values of the indices SCR and SMA as functions of λ, calculated over 500 simulation runs, for different scenarios ($m = n = \{10, 20, 50\}$). The dashed vertical line represents $\lambda = 0.5$, where approvals and rankings are equally important

inflection point around $\lambda = 0.5$, where the decrease in SCR and the increase in SMA accelerates. Even though trends are observed, we highlight that the observed changes are moderate. The maximum value of SMA remains just below 0.1, while SCR does not exceed 0.05, underscoring the robustness of the DIVA solution to variations in λ.

The values of SCR and SMA are comparable but not perfectly symmetrical. This pattern arises because rankings and approvals are distinct structures, and different distance metrics are used. Specifically, even after normalization to the $[0, 1]$ range, the Kemeny and Hamming distances exhibit slightly different behaviors. For example, the Hamming distance tends to show larger jumps since it takes fewer values in the $[0, 1]$ interval, whereas the Kemeny distance demonstrates more gradual changes, as it can take more intermediate values. Finally, the behavior across different sample sizes ($n = 10, 20, 50$) does not show significant differences in the general trends of SCR and SMA.

4.2 Impact of λ on the quality of the DIVA solution $\hat{r}_\lambda$

The purpose of this section is to analyze how λ affects the quality of the DIVA solution under varying levels of noise in the ranking and approval components. To evaluate the quality of the DIVA solution, we compute the average distance between the consensus solution and the input preference-approvals provided by all voters:

$$\bar{d_\lambda}((\hat{\pi}, \hat{A}), \Pi) = \frac{1}{m} \sum_{l=1}^{m} d_\lambda \left((\pi^{(l)}, A^{(l)}), (\hat{\pi}, \hat{A}) \right). \tag{25}$$

To systematically study the impact of λ, we generate synthetic preference-approvals where the noise levels in the ranking and approval components can be controlled separately. The noise in the ranking components is quantified by computing the average scaled Kemeny distance between all pairs of rankings provided by the voters:

$$\frac{1}{\binom{m}{2}} \sum_{i<j} d_R(\pi_i, \pi_j). \tag{26}$$

Similarly, the noise in the approval components is measured using the average scaled Hamming distance between all pairs of approval sets:

$$\frac{1}{\binom{m}{2}} \sum_{i<j} d_A(A_i, A_j). \tag{27}$$

To assess the effect of λ on the quality of the DIVA solution, we design four experimental settings with varying levels of noise in the ranking and approval components. The settings are defined as follows:

Setting 1: This setting models a scenario where preference-approvals are highly consistent across voters. Both rankings and approval components exhibit strong agreement, as indicated by low Kemeny and Hamming distances.

Setting 2: In this setting, variability is introduced in voters' preference-approvals. Both rankings and approval sets show moderate agreement, reflected in Kemeny and Hamming distances.

Setting 3: This setting models situations where voters agree on the relative ordering of alternatives but differ significantly in their binary approvals. That is, rankings remain relatively consistent (low Kemeny distance), while approval sets show higher variability (moderate Hamming distance).

Setting 4: This scenario simulates situations where voters agree on which alternatives to approve but differ in their preferences over the relative order. Specifically, approval sets are highly consistent (low Hamming distance), while rankings exhibit moderate variability (higher Kemeny distance).

In each setting, the number of alternatives n was set to 10, 20, and 30, with the number of judges (m) also equal to n. For each setting, we performed 100 iterations, generating preference-approval profiles according to the distributions described in Table 4. Only profiles satisfying the range constraints specified in the table (in columns $\bar{d}_R$ and $\bar{d}_A$) were retained. The generation process began by constructing a vector $\boldsymbol{\mu}$, where each element represents a random mean sampled from a normal distribution $\mathcal{N}(10, \sigma)$. To generate preference-approvals in each scenario, we begin by creating a vector $\boldsymbol{\mu}$, where its elements are random means drawn from a normal distribution: $\mu \sim \mathcal{N}(10, \sigma)$. The covariance matrix Σ is chosen to be an identity matrix of size

Table 4 Summary of simulation settings, including the average noise values for both ranking and approval components across the different configurations

Setting	*n*	Ranking generation	Approval generation	$\bar{d}_R$	$\bar{d}_A$
1	10	$\mu = \mathcal{N}(10, 5.5)$,$\Sigma = I_n$	*Pois*(0.6)	[0.070, 0.090]	[0.070, 0.090]
1	20	$\mu = \mathcal{N}(10, 5.5)$,$\Sigma = I_n$	*Pois*(1.8)	[0.070, 0.090]	[0.070, 0.090]
1	50	$\mu = \mathcal{N}(10, 5.5)$,$\Sigma = I_n$	*Pois*(8.5)	[0.070, 0.090]	[0.070, 0.090]
2	10	$\mu = \mathcal{N}(10, 1)$,$\Sigma = I_n$	*Pois*(3.1)	[0.325,0.345]	[0.325,0.345]
2	20	$\mu = \mathcal{N}(10, 1)$,$\Sigma = I_n$	*Pois*(7)	[0.325,0.345]	[0.325,0.345]
2	50	$\mu = \mathcal{N}(10, 1)$,$\Sigma = I_n$	*Pois*(20)	[0.325,0.345]	[0.325,0.345]
3	10	$\mu = \mathcal{N}(10, 10.5)$,$\Sigma = I_n$	*Uniform*(0, *n*)	[0.030, 0.050]	[0.325,0.345]
3	20	$\mu = \mathcal{N}(10, 10.5)$,$\Sigma = I_n$	*Uniform*(0, *n*)	[0.030, 0.050]	[0.325,0.345]
3	50	$\mu = \mathcal{N}(10, 12)$,$\Sigma = I_n$	*Uniform*(0, *n*)	[0.030, 0.050]	[0.325,0.345]
4	10	$\mu = \mathcal{N}(10, 1)$,$\Sigma = I_n$	$\text{Binomial}(p = 0.975)$	[0.325,0.345]	[0.025,0.055]
4	20	$\mu = \mathcal{N}(10, 1)$,$\Sigma = I_n$	$\text{Binomial}(p = 0.975)$	[0.325,0.345]	[0.025,0.055]
4	50	$\mu = \mathcal{N}(10, 1)$,$\Sigma = I_n$	$\text{Binomial}(p = 0.975)$	[0.325,0.345]	[0.025,0.055]

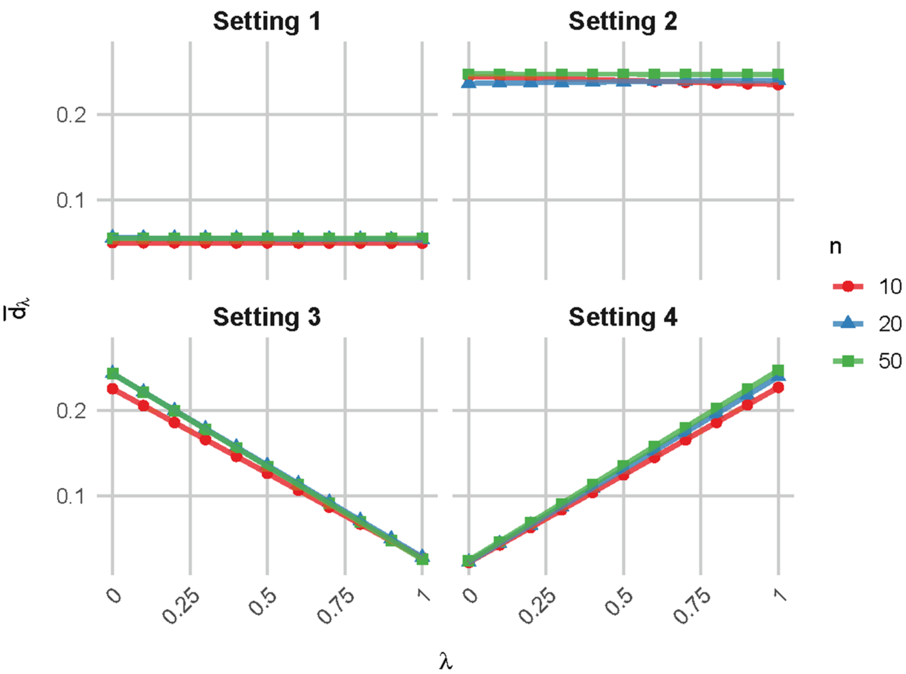

Fig. 3 Goodness of the DIVA solution as a function of λ for different settings

$n \times n$. True ranking configurations Γ are then sampled from a multivariate normal distribution: $\Gamma \sim \mathrm{MVN}(\boldsymbol{\mu}, \Sigma)$. The number of approved alternatives is determined by either a Poisson, Binomial, or Uniform distribution, depending on the specific setting (Table 4). The specific parameters for the distributions used to generate the preference-approvals in each setting were selected based on a sensitivity analysis. The goal was to generate configurations that exhibited similar levels of noise in both the ranking and approval components across different values of n, while also capturing the intended characteristics of each setting.

In each scenario, we evaluate how the goodness of the DIVA solution varies as a function of λ. Higher values of $\bar{d}_\lambda$ indicate a poorer quality of the consensus. The results are summarized in Fig. 3.

In Setting 1, where both ranking and approval components have minimal noise, the average distance $\bar{d}_\lambda$ remains constant across all values of λ and is consistently low. This behavior suggests that the consensus solution is stable and unaffected by changes in λ, as there is little disagreement among the voters' preference-approvals. In Setting 2, where moderate noise is present in both components, $\bar{d}_\lambda$ also remains constant but at higher values compared to Setting 1. This result highlights that when noise is balanced across ranking and approval parts, tuning λ has no significant impact on the solution quality, as the noise is uniformly distributed.

In Setting 3, where the ranking component has minimal noise and the approval component exhibits significant noise, $\bar{d}_\lambda$ decreases monotonically as λ increases.

This indicates that giving more weight to the ranking component effectively reduces the discrepancy caused by the noisy approvals, leading to an improved consensus solution. Conversely, in Setting 4, where the ranking component has moderate noise and the approval component has minimal noise, $\bar{d}_\lambda$ increases as λ grows. In this case, placing more weight on the ranking part amplifies the effect of noise in the ranking component, resulting in a poorer-quality consensus.

Overall, while λ has little effect under minimal or balanced noise conditions (Settings 1 and 2), it becomes crucial when the noise levels are asymmetric (Settings 3 and 4).

5 Simulations

The aim of this set of simulations is to analyse the efficacy of the proposed algorithm concerning its ability to yield valid solutions (i.e. consistent preference-approvals), processing times, and solution quality. In the first simulation scenario (Model I, Sec. 5.1), the DIVA algorithm is evaluated by comparing it against the Representative Preference-Approvals (RPA) solution (Albano et al. 2023).

In the simulation under Model I, synthetic preferences are generated from a multivariate normal distribution, and approvals are drawn from a uniform distribution. In the second simulation scenario (Model II, Sec. 5.2), an extended Mallow model tailored to preference-approvals is introduced. This model is employed to evaluate the accuracy of the proposed solution.

In the experimental evaluations, we set the weighting parameter λ to 0.5, assuming equal relevance between the ranking and approval components. The simulations are carried out using `R`.

5.1 Model I

Two distinct simulation settings have been formulated for Model I. In the initial setting, preference-approvals are generated without admissible ties, implying that all rankings are linear orders. In the following setting, an arbitrary number of ties is introduced into the ranking component.

Inspired by D'Ambrosio et al. (2022) for the two simulation settings, nine distinct scenarios have been designed, varying the number of alternatives across three values $n = \{10, 20, 50\}$, and adjusting the number of judges $m = \{n, 2n, 3n\}$ to three different multiples of n. Each unique scenario is subjected to $B = 100$ iterations for a comprehensive evaluation.

In each iteration, a vector $(\boldsymbol{\mu})$ is populated by random means that are drawn from a normal distribution: $\mu \sim \mathcal{N}(0, 1)$, whilst a $n \times n$ identity matrix is used as covariance matrix (Σ) s.t.:

$$\Sigma = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1_n \end{bmatrix}$$

Moreover, true ranking configurations Γ are sampled from a multivariate normal distribution: $\Gamma \sim \mathrm{MVN}(\boldsymbol{\mu}, \Sigma)$. Therefore, to derive preference-approvals, preference rankings are extracted from Γ, and approval thresholds are generated following a discrete Uniform $\mathcal{U}(0, n)$.

In the first simulation setting, only strict rankings are derived from Γ, while in the second simulation setting a number of tied alternatives $g \leq n$ is added by sampling from a Poisson distribution, $g \sim Pois(0.4n)$.

Three indices are employed for evaluation purposes:

1. computational times in seconds;
2. percentage of instances where the methods return a valid solution;
3. Minimum Distance Index (MDI). This performance metric compares the proposed solution and the RPA solution in terms of the average distance from the set of voters.

With respect to the third performance metric, MDI, averaged over the iterations, is defined s.t.:

$$\mathrm{MDI} = \frac{1}{B} \sum_{b=1}^{B} \mathrm{sgn} \left(\bar{d}_\lambda(\mathrm{RPA}, \Pi)^{(b)} - \bar{d}_\lambda(\mathrm{DIVA}, \Pi)^{(b)} \right), \tag{28}$$

where B represents the number of iterations, $\bar{d}_\lambda(\mathrm{RPA}, \Pi)$ and $\bar{d}_\lambda(\mathrm{DIVA}, \Pi)$ denotes the average distance of the corresponding solution to the whole set of generated preference-approvals in the b-th replication.

Table 5 Performances metrics in the first setting with no ties

n	m	Time(s)		% Admissible solutions		MDI
		DIVA	RPA	DIVA	RPA	
10	10	0.198	< 0.01	100	67	0.81
	20	0.245	< 0.01	100	73	0.77
	30	0.246	< 0.01	100	85	0.70
20	20	2.175	< 0.01	100	40	1.00
	40	2.478	< 0.01	100	53	0.99
	60	2.303	< 0.01	100	64	0.98
50	50	27.092	< 0.01	100	23	1.00
	100	34.402	< 0.01	100	27	1.00
	150	44.886	< 0.01	100	36	1.00

Table 6 Performances metrics in the second setting with tied alternatives

n	m	Time(s)		% Admissible solutions		MDI
		DIVA	RPA	DIVA	RPA	
10	10	0.238	< 0.01	100	61	1.00
	20	0.305	< 0.01	100	65	1.00
	30	0.291	< 0.01	100	78	1.00
20	20	1.157	< 0.01	100	39	1.00
	40	1.457	< 0.01	100	42	1.00
	60	1.607	< 0.01	100	63	1.00
50	50	11.378	< 0.01	100	20	1.00
	100	14.367	< 0.01	100	32	1.00
	150	17.388	< 0.01	100	33	1.00

Basically, MDI counts 1 each time the proposed solution is closer to the set of votes than the RPA solution, or when the RPA method does not return an admissible solution. It counts 0 when both solutions share the same distance, and -1 otherwise. The cumulative contributions are summed and then divided by the number of replications B. An $\text{MDI} \approxeq 0$ suggests an equivalence between the methods, while $\text{MDI} \approxeq 1$ ($\text{MDI} \approxeq -1$) signify superior solutions in favor of the DIVA (RPA) approach.

Tables 5 and 6 report the results of the simulation studies under the two settings.

Despite the superior computational values of the RPA solution, the proposed algorithm remains competitive, delivering efficient processing times even in scenarios with increased complexity, such as $n = 50$. Additionally, it is noteworthy that our solution improves execution times in scenarios with ties, contributing to its adaptability and efficiency.

Notably, the proposed method consistently delivers admissible solutions across all scenarios in the two simulation models. This underscores its robust capability to find rankings and approvals that are both consistent. In stark contrast, the RPA solution exhibits a decline in its capacity to provide admissible solutions as the number of alternatives n increases. Moreover, its performance slightly degrades further when ties are introduced.

Finally, the MDI value exhibits high values even for smaller values of n, quickly reaching 1 as n increases. While ties are present, the MDI value remains steadfastly at 1 across all scenarios. This shows the superiority of the DIVA over the RPA approach, particularly in scenarios characterised by numerous or tied preferences.

5.2 Model II

The simulation under Model II is specifically designed to evaluate the DIVA's ability to produce solutions closely aligned with the preferences of the designated voter set. To accomplish this goal, this article introduces, for the first time, an extended version of the Mallow model tailored to preference approvals generation.

The Mallows Model (Mallows 1957) stands as one of the earliest proposed probability models for rankings, and its relevance is evident in both theoretical and applied research. This exponential model is characterised by a central permutation π_0 and a dispersion parameter θ. When $\theta \neq 0$, π_0 represents the mode of the distribution, that is the ranking with the highest probability of being generated. As the distance from the central permutation increases, the probability of any other ranking exponentially decreases. The dispersion parameter θ regulates the rate of this decline. The probability for a generic ranking π is formulated as a function of θ and is expressed as follows:

$$Pr(\pi) = \frac{\exp(-\theta d(\pi, \pi_0))}{\psi(\theta)}, \tag{29}$$

where d denotes a ranking distance measure, and $\psi(\theta)$ serves as a normalisation constant.

In this paper, we introduce a modified version of the Mallows model specifically designed to deal with preference-approvals. The extension of the model to preference approvals entails the introduction of a novel definition of the mode. In this context, the mode represents a valid preference-approvals $(\pi_0,\ A_0)$. Another necessary adaptation involves the choice of distance metric for the model, here we employ d_λ to quantify the dissimilarity between preference-approval configurations. The functional form of which is specified as follows:

$$Pr\left((\pi, A)\right) = \frac{\exp\left(-\alpha\theta d_\lambda\left((\pi, A), (\pi_0, A_0)\right)\right)}{\psi(\alpha\theta)}, \tag{30}$$

where d_λ is a preference-approval distance measure, defined in Eq. 16, and $\psi(\alpha\theta)$ is a normalisation constant. Furthermore, in contrast to the traditional Mallows model, the dispersion parameter θ is multiplied by an integer α to account for the range of d_λ, which inherently ranges between 0 and 1. This adjustment ensures a more effective representation of the dispersion parameter within the context of the preference-approval model.

Thus, the distances between the generated preferences and the DIVA solutions are examined in this simulation study, taking the central permutation of the Mallow model as a reference point for comparison.

To derive the probability of generating each preference-approval within the entire universe, in a brute-force approach, it would be necessary to compute the distance for each pair $d_\lambda\left((\pi_0, A_0), (\pi_i, A_i)\right)$, where $(\pi_i, A_i)\big) \in R^{\Omega(n)}$. However, as indicated in Table 2, this operation becomes computationally infeasible as *n* grows.To address this challenge, sub-universes were introduced from which preference-approvals are sampled. To generate sub-universes, we start with a modal preference-approval and introduce noise by perturbing both the ranking and approval components. Multiple perturbations are applied to ensure that the distances of the generated preference-

Table 7 DIVA performance under various settings

m	θ	$n=10$		$n=20$		$n=50$	
		$\Delta\bar{d}$	$\overline{\text{RDR}}$	$\Delta\bar{d}$	$\overline{\text{RDR}}$	$\Delta\bar{d}$	$\overline{\text{RDR}}$
$m=n$	0.01	0.137	27.497	0.181	35.847	0.244	46.529
	0.5	0.005	2.543	0.015	7.099	0.057	26.792
	1	0.007	7.838	0.005	5.043	0.035	32.914
$m=2n$	0.01	0.130	26.305	0.178	35.772	0.243	46.436
	0.5	0.004	2.021	0.012	6.250	0.056	26.273
	1	0.007	7.307	0.004	3.805	0.035	31.882
$m=3n$	0.01	0.128	25.887	0.179	35.977	0.242	46.385
	0.5	0.004	1.884	0.012	5.800	0.056	26.161
	1	0.007	7.167	0.003	3.376	0.035	31.619

approvals to the modal preference-approval, cover the entire 0-1 range. The provided pseudocode of Alg. 2 elucidates the extraction whole process.

Input: n; n_u; m; $R^{\Omega(n)}$; $\alpha = 10$; θ; (π_0, A_0); d_0

Generate a sub-universe $T^{(n)} \subset R^{\Omega(n)}$ *with* $|T^{(n)}| = n_u$

Generate an equidistributed sequence of n *elements*

Generate a distribution $\Phi(d_0) = \frac{exp(-\alpha\theta d_0)}{\psi(\alpha\theta)}$ *that map each sequence value*

Define:

$\Pi' = \{(\pi_1, A_1) = \text{NULL}, \ldots, (\pi_m, A_m) = \text{NULL}\}$

for $l = 1$ to m **do**

  Distance generator

  $d_l \sim \Phi(d_0)$

  Preference-Approval search

  $(\tilde{\pi}_l, \tilde{A}_l) = \arg\min_{(\pi,A)\in T^{(n)}} |(d_\lambda((\pi, A), (\pi_0, A_0)) - d_l|$

  $\Pi'_l = (\tilde{\pi}_l, \tilde{A}_l)$

end for

Output: $\{\Pi'\}$

Algorithm 2 Sample preference-approvals from mallows model (SPA-MM)

This strategic approach allows for efficient sampling from sub-universes, avoiding computational challenges and enabling a comprehensive exploration of preference-approvals across varying distances.

The metrics employed for assessing the algorithm's ability to approximate solutions aligned with the generated preference-approvals are the mean average difference ($\Delta\bar{d}$) and the Relative Difference Reduction ($\overline{\text{RDR}}$) averaged over B iterations, defined as:

$$\Delta\bar{d} = \frac{1}{B}\sum_{b=1}^{B}\left(\bar{d}_\lambda((\pi_0, A_0), \Pi')^{(b)} - \bar{d}_\lambda(\text{DIVA}, \Pi')^{(b)}\right), \tag{31}$$

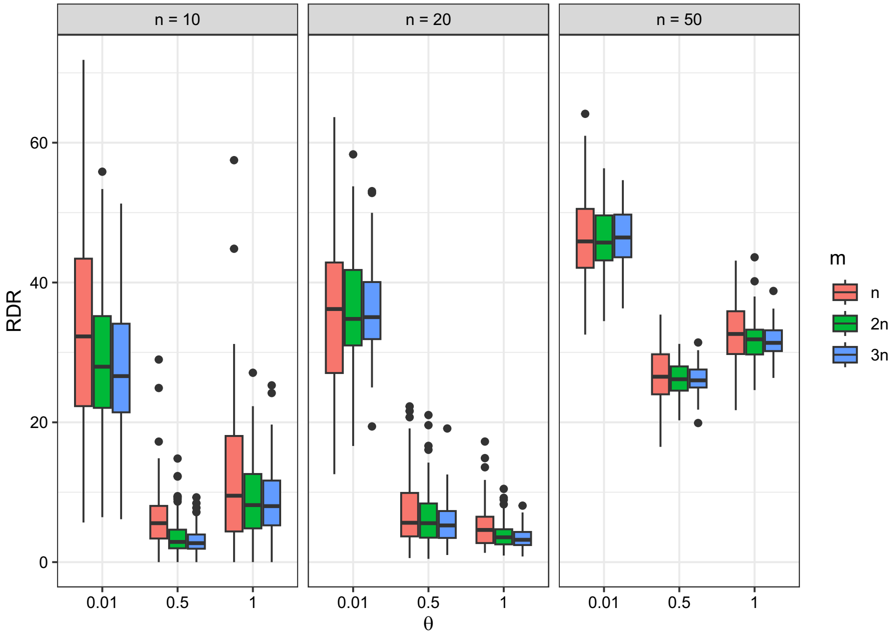

Fig. 4 RDR distributions across different θ values, grouped by n. Colors represent m levels

$$\overline{\mathrm{RDR}} = \frac{100}{B} \times \sum_{b=1}^{B} \frac{\left(\bar{d}_{\lambda}((\pi_0, A_0), \Pi')^{(b)} - \bar{d}_{\lambda}(\mathrm{DIVA}, \Pi')^{(b)}\right)}{\bar{d}_{\lambda}((\pi_0, A_0), \Pi')^{(b)}}. \tag{32}$$

Table 7 presents the DIVA's performance while being applied to datasets created by SPA-MM under the configurations delineated by the parameters $n = \{10, 20, 50\}$, $m = \{n, 2n, 3n\}$, and $\theta = \{0.01, 0.5, 1\}$, while α has been fixed to 10. Each unique scenario is subjected to $B = 100$ iterations for a comprehensive evaluation.

Table 7 reveals noteworthy trends in the algorithm's performance across different configurations. As n increases, there is a consistent improvement in $\Delta\bar{d}$ and $\overline{\mathrm{RDR}}$. This trend is particularly evident at $n = 50$, indicating a moderate gain in using the DIVA algorithm compared to the mode of the Mallow model (π_0, A_0). As θ increases, the noise in data generation reduces, resulting in a convergence of the DIVA algorithm's solutions with (π_0, A_0). Meanwhile, the impact of the number of voters m on the algorithm's agreement with voter preferences appears less pronounced compared to the influence of n and θ. Finally, the average RDR affirms the algorithm's consistent enhancement over its baseline, with significant improvements observed, such as approximately 30% decrease at $n = 50$ and $\theta = 1$, emphasizing its adaptability and effectiveness in diverse voter scenarios. Those variations are easier to observe in Fig. 4, where the whole distribution of the RDR index over the iterations is plotted.

Table 8 List of French politicians with their id

Politician	Id
Francois Bayrou	x_1
Olivier Besancenot	x_2
Christine Boutin	x_3
Jacques Cheminade	x_4
Jean-Pierre Chevenement	x_5
Jacques Chirac	x_6
Robert Hue	x_7
Lionel Jospin	x_8
Arlette Laguiller	x_9
Brice Lalonde	x_{10}
Corine Lepage	x_{11}
Jean-Marie Le Pen	x_{12}
Alain Madelin	x_{13}
Noel Mamere	x_{14}
Bruno Maigret	x_{15}

6 Case studies

This section presents two case studies using data from the Preference Learning and Preference Elicitation Repository[1] (PrefLib, Mattei and Walsh 2013). Specifically, the employed datasets are: an analog to the 2002 French Presidential Election Dataset (collected by Laslier and Straeten 2004) and the 1950 Formula 1 World Championship (provided within the collection made by Boehmer and Schaar 2023).

6.1 French presidential election (2002)

The dataset comprises 398 approval ballots and subjective ratings on a 20-point scale, collected from students at the *Institut d'Etudes Politiques de Paris*, focusing on potential candidates for the 2002 French presidential election. Both approval ballots and subjective ratings for each candidate by individual voters are present in the dataset.

The approval ballots are expressed as either 1.0 for approved or 0.0 for not approved. While subjective ratings are expressed using a 20-point scale, a score of -1.0 indicates no input, distinct from a rating of 0.0, which represents the lowest possible rating. Table 8 reports the list of French politicians with their identification.

To perform PA aggregation, the original data, which combines ratings and approvals, needs to be transformed into a set of admissible preference-approvals. To achieve this, a preference ranking is derived from each voter by arranging candidates in ascending order of their ratings. However, two challenges emerged: some voters have given incomplete ratings whilst some others expressed incoherent evaluations.

With respect to incomplete ratings, a common approach in the literature is to place all objects that are not evaluated at the bottom of the rankings, assuming that respon-

[1] Available for download at https://preflib.simonrey.fr/dataset/00029 and https://preflib.simonrey.fr/dataset/00052.

Table 9 Example of incompatible ratings and approval, voter number 10 $V^{(10)}$

Politician	Rating	Approval
Francois Bayrou	0	0
Olivier Besancenot	0	0
Christine Boutin	0	0
Jacques Cheminade	0	0
Jean-Pierre Chevenement	6	1
Jacques Chirac	5	0
Robert Hue	2	0
Lionel Jospin	8	0
Arlette Laguiller	0	0
Brice Lalonde	3	0
Corine Lepage	3	0
Jean-Marie Le Pen	0	0
Alain Madelin	0	0
Noel Mamere	7	1
Bruno Maigret	0	0

Table 10 DIVA solution French presidential election (2002) data

x_8
x_1 x_5 x_6 x_{14}
x_2 x_3 x_4 x_7 x_9 x_{10} x_{11} x_{12} x_{13} x_{15}

Table 11 RPA solution French presidential election (2002) data

x_8
x_5
x_{14}
x_6
x_1
x_7
x_9
x_{11}
x_{13}
x_{10}
x_2
x_3
x_4
x_{12}
x_{15}

dents have no preference among them (Li et al. 2017). Following this principle, not-evaluated objects were placed just above disapproved objects if approved, and at the very end if not approved. To clarify this approach, let us assume that the *l*-th voter

Table 12 Comparison of solutions French presidential election (2002) data

	$\bar{d_\lambda}$	Time (s)
DIVA	0.217	0.86
RPA	0.242	< 0.01

$V^{(l)}$ splits the alternatives into approved set $A_{V^{(l)}}$ and not approved set $U_{V^{(l)}}$. Let us also say that $V^{(l)}$ does not provide any rating for x_j. Then the rank x_j is found as:

$$P_{V^{(l)}}(x_j) = \begin{cases} max\big(P(x|\, x \in A_{V^{(l)}})\big) + 1 & \text{if } x_j \in A \\ max\big(P(x|\, x \in U_{V^{(l)}})\big) + 1 & \text{if } x_j \in U. \end{cases}$$

Note that the incomplete rankings are assigned while being careful not to violate the conditions in Eqs. 1, 2, 3, 4, 5 and 6, that is, $P\big(x|\, x \in A_{V^{(l)}}\big) < P\big(x|\, x \in U_{V^{(l)}}\big)\ \forall x \in X$. Thus, if $x_j \in A$ and its $P_{V^{(l)}}$ is assigned using this strategy, to avoid violating Eqs. 1, 2, 3, 4, 5 and 6 we also need to adjust the rankings of the not-approved items by adding 1 to all of them (i.e., shifting).

The second issue concerns voters who exhibited non-admissible preference approvals. For instance, consider the evaluations of voter number 10, $V^{(10)}$ in Table 9. The incoherence in $V^{(10)}$ arises from a discrepancy between the Rating and Approval columns for certain politicians, notably Lionel Jospin, Alain Madelin, and Jean-Pierre Chevenement. The specific issue is that Lionel Jospin has the highest rating (8), but he is marked as not approved (Approval: 0), while Alain Madelin and Jean-Pierre Chevenement, who have lower ratings (both with ratings of 7 and 6, respectively), are marked as approved (Approval: 1). Therefore, Lionel Jospin's rating does not align with the approval status, creating an inconsistency for $V^{(10)}$. Checking for inconsistencies in the whole dataset led to the exclusion of 90 incoherent preference approvals expressed by voters.

Tables 10, 11 and 12 report the DIVA and RPA solutions with the corresponding average d_λ and computing times (seconds). The weighting parameter λ has been set to 0.5, assuming that ranking and approval parts are equally important.

Both solutions are admissible and consistently place x_8 at the top, indicating a consensus on the highest-preferred candidate. However, differences in the ordering of intermediate and lower-tier candidates contribute to variations in the average distance measure (d_λ). Specifically, the DIVA solution achieves a lower average distance from the set of voters (0.217), keeping the execution time under 1 s. Furthermore, the DIVA takes into account the high presence of ties observed in the data that are reflected in the final solution. This underscores the sensitivity of the DIVA solution to the preferential arrangement of candidates, highlighting the potential impact on the overall reliability of the consensus.

6.2 Formula 1 world championship (1950)

The Formula 1 World Championship, which consists of a series of races, has been one of the premier forms of racing around the world since its inaugural season in

Table 13 List of 1950' drivers with their id

Racer Lastname	Id	Racer Lastname	Id	Racer Lastname	Id
Pozzi	x_1	Banks	x_{28}	Cantrell	x_{55}
Bettenhausen	x_2	Comotti	x_{29}	Levrett	x_{56}
Hampshire	x_3	Fangio	x_{30}	Chiron	x_{57}
Shawe Taylor	x_4	Peter Walker	x_{31}	Hanks	x_{58}
Rol	x_5	Rolt	x_{32}	Graffenried	x_{59}
Sanesi	x_6	Fohr	x_{33}	Schell	x_{60}
Fagioli	x_7	Davies	x_{34}	Hoyt	x_{61}
Crossley	x_8	Gerard	x_{35}	Leslie Johnson	x_{62}
Flaherty	x_9	Branca	x_{36}	Reg Parnell	x_{63}
Pagani	x_{10}	Bonetto	x_{37}	Levegh	x_{64}
Louveau	x_{11}	Guy Mairesse	x_{38}	Paul Russo	x_{65}
Martin	x_{12}	Whitehead	x_{39}	Schindler	x_{66}
Jackson	x_{13}	Faulkner	x_{40}	Rathmann	x_{67}
Serafini	x_{14}	Biondetti	x_{41}	Webb	x_{68}
Holland	x_{15}	Ader	x_{42}	Parsons	x_{69}
Rose	x_{16}	Dick Rathmann	x_{43}	Walt Brown	x_{70}
Kelly	x_{17}	Hartley	x_{44}	Ruttman	x_{71}
Sommer	x_{18}	Agabashian	x_{45}	Farina	x_{72}
Rosier	x_{19}	Bira	x_{46}	Dinsmore	x_{73}
Villoresi	x_{20}	Wallard	x_{47}	Manzon	x_{74}
Cabantous	x_{21}	Murray	x_{48}	George Connor	x_{75}
Fry	x_{22}	Taruffi	x_{49}	Pian	x_{76}
Darter	x_{23}	Claes	x_{50}	Mcgrath	x_{77}
Harrison	x_{24}	Chaboud	x_{51}	Hellings	x_{78}
Ascari	x_{25}	Pietsch	x_{52}	Trintignant	x_{79}
Holmes	x_{26}	Green	x_{53}	Etancelin	x_{80}
Mcdowell	x_{27}	Gonzalez	x_{54}	Chitwood	x_{81}

Table 14 DIVA solution Formula 1 (1950) data, Approved = Top 3

x_7
x_{72}
x_{19} x_{30}
x_{50} x_{80}
x_{18} x_{25} x_{57} x_{59}
x_{21} x_{46}
All the others ex-aequo...

1950. As stated by the authors of the dataset Boehmer and Schaar (2023), an election has been created considering the finishing times of all drivers in all laps of races in 1950-2020. Thus, each vote corresponds to a race of a specific year, and drivers are ranked by their total finishing time in this race. For each year, they have been made available one dataset. Within this section, we will focus on the first one, which consists of the inaugural season (1950).

Table 15 DIVA solution Formula 1 (1950) data, Approved = Top 5

x_{72}
x_{7}
x_{19}
———
x_{30}
x_{50} x_{80}
x_{18} x_{25} x_{57} x_{59}
x_{21} x_{46}
All the others ex-aequo…

Table 16 DIVA solution Formula 1 (1950) data, Approved = Top 10

x_{72}
x_{7}
x_{19}
x_{80}
———
x_{30}
x_{50}
x_{18} x_{25} x_{57} x_{59}
x_{21} x_{46}
All the others ex-aequo…

Table 17 RPA solution Formula 1 (1950) data, Approved = Top 3

x_{7}
———
x_{72}
x_{19}
x_{30}
x_{80}
x_{50}
x_{59}
x_{18} x_{57}
x_{46}
x_{25}
x_{24}
.
.
.

Therefore, the dataset comprises $m = 7$ judges (i.e. the races) and $n = 81$ items (i.e. the drivers). Table 13 reports the list of each driver with their identification.

To perform PA aggregation, the original data, which consists of incomplete rankings only, needs to be transformed into a set of admissible preference-approvals. To achieve this, we considered as *approved* those drivers ranked in the top 3, 5, 10

Table 18 RPA solution Formula 1 (1950) data, Approved = Top 5

x_7
x_{72}
x_{19}
x_{30}
x_{80}
x_{50}
x_{59}
x_{18} x_{57}
x_{46}
x_{25}
x_{24}
x_{39}
x_{21} x_{35} x_{69}
.
.
.

Table 19 RPA solution Formula 1 (1950) data, Approved = Top 10

x_7
x_{72}
x_{19}
.
.
.
.
.
.

Table 20 Comparison of achieved best solutions, Formula 1 (1950) data

	$\bar{d}_\lambda$	Time (s)	Consistent	Approval type
DIVA	0.106	8.55	Yes	Top 3
	0.117	7.71	Yes	Top 5
	0.145	7.85	Yes	Top 10
RPA	0.208	< 0.01	Yes	Top 3
	0.219	< 0.01	Yes	Top 5
	–	< 0.01	**No**	Top 10

(respectively) and *non-approved* the others. By construction, approvals that are made in this way do not violate Eqs. 1, 2, 3, 4, 5 and 6.

With respect to incomplete rankings, we follow the same approach used for the 2002 French Presidential Election dataset (Sec. 6.1): placing at the bottom of the rankings all objects that are not evaluated, assuming that respondents have no preference among them.

Tables 14, 15, and 16 report the DIVA solutions ($\lambda = 0.5$). It is possible to observe how, compared to the DIVA algorithm, RPA solutions are more likely to avoid ex-aequo expecially for the non-approved items (Tables 17 and 18) or are not consistent. In fact, partial results reported in Table 19 are provided by a solution that set for position number 4 a non-approved item whilst an approved one for position number 5.

Furthermore, how it is possible to observe in Table 20, the DIVA solution achieves a lower average distance from the set of voters in the three considered cases (Approval = Top 3, Top 5, and Top 10). Thus, despite of being slightly slower than RPA, DIVA takes into account the high presence of ties observed in the data that are reflected in the final solution.

7 Conclusions

This paper presents an innovative distance-based approach designed to aggregate different preferences within a group, establishing a consensus on preference-approvals. By proposing a solution for extracting a common preference, the methodology outlined herein contributes to the field of decision-making processes, providing a framework that enhances both effectiveness and efficiency in scenarios involving varied individual preferences. We highlight that the proposed algorithm has promising applications within preference learning algorithms, serving as a valuable input for predictive modeling. The incorporation of this distance-based methodology into preference learning frameworks opens avenues for more accurate predictions, especially in scenarios characterised by different individual preferences.

A comprehensive evaluation of the method was conducted, including a sensitivity analysis on the weighting parameter λ, two simulation studies, and two real-data applications.

The analysis of λ highlights its dual role in influencing the DIVA solution and its quality. We showed that as λ increases, the ranking component of the DIVA solution aligns more closely with the marginal consensus ranking ($\hat{\pi}_M$), while the approval component diverges from the marginal consensus approval ($\hat{A}_M$). Further insights were drawn from the behavior of the quality of the DIVA solution under different noise conditions. Under minimal or balanced noise, the quality of the DIVA solution stays constant across all λ values. However, in asymmetric noise conditions, λ plays a critical role in balancing the effects of noise between the ranking and approval components.

In addition to the sensitivity analysis, the simulation studies assessed DIVA's computational performance, solution quality, and alignment with voter preferences. DIVA demonstrates competitive computational performance, consistently delivering admissible solutions and outperforming existing methods, particularly in scenarios with tied preferences. Furthermore, we introduced an extended Mallows model tailored for preference approvals, demonstrating the ability of DIVA to align solutions with voter preferences. The adapted model, utilising a unique mode definition and distance metric, showed improved performance with an increasing voter count.

Finally, we present two case studies using real-world data analogous to the 2002 French Presidential Election Dataset and the 1950 Formula 1 World Championship, collected from the Preference Learning and Preference Elicitation Repository (PrefLib). Despite challenges such as incomplete ratings and inconsistent evaluations, DIVA successfully aggregates preference-approvals, providing a consensus that aligns well with the voters' preferences.

Moving forward, future research directions stemming from our study include investigating DIVA's adaptability to dynamic preferences and evolving decision contexts, such as online recommendation systems or real-time group decision-making processes. Additionally, we plan to address uncertainty in preference-approvals and develop methods to quantify and account for robustness in consensus formation. Forby, we plan to publish an R package containing our proposal. These avenues of exploration hold the potential to further enhance the applicability and effectiveness of DIVA in diverse decision-making scenarios.

Author Contributions All authors have contributed equally to the work.

Funding Open access funding provided by Università degli Studi di Cagliari within the CRUI-CARE Agreement. The research work of Alessandro Albano has been supported by the European Union - NextGenerationEU - National Sustainable Mobility Center CN00000023, Italian Ministry of University and Research Decree n. 1033— 17/06/2022, Spoke 2, CUP B73C2200076000 and Targeted Research Funds 2024 (FFR 2024) of the University of Palermo (Italy). Maurizio Romano acknowledges financial support under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5 - Call for tender No.3277 published on December 30, 2021 by the Italian Ministry of University and Research (MUR) funded by the European Union – Next Generation EU. Project Code ECS0000038 – Project Title eINS Ecosystem of Innovation for Next Generation Sardinia – CUP F53C22000430001- Grant Assignment Decree No. 1056 adopted on June 23, 2022 by the MUR.

Data availability The DIVA algorithm has been implemented in R and is currently available at the following GitHub repository, likewise all the data considered in this paper: urlr.me/pPRmY4.

Declarations

Conflict of interest The authors declare no Conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent For this type of study formal consent is not required.

References

Abramowitz M, Stegun IA (1964) Handbook of mathematical functions with formulas, graphs, and mathematical tables, vol 55. US Government printing office, Washington, D.C., USA

Albano A, Plaia A (2021) Element weighted Kemeny distance for ranking data. Electron J Appl Stat Anal 14(1):117–145. https://doi.org/10.1285/i20705948v14n1p117

Albano A, García-Lapresta JL, Plaia A et al (2023) A family of distances for preference-approvals. Ann Oper Res 323(1–2):1–29. https://doi.org/10.1007/s10479-022-05008-4

Albano A, García-Lapresta JL, Plaia A et al (2024) Clustering alternatives in preference-approvals via novel pseudometrics. Stat Methods Appl 33(1):61–87. https://doi.org/10.1007/s10260-023-00718-w

Barokas G (2022) Revealed desirability: a novel instrument for social welfare. Theor Decis 93(4):649–661

Barokas G, Sprumont Y (2021) The broken Borda rule and other refinements of approval ranking. Soc Choice Welfare 58(1):187–199

Black D (1976) Partial justification of the Borda count. Public Choice 28:1–15

Boehmer N, Schaar N (2023) Collecting, classifying, analyzing, and using real-world ranking data. In: Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, AAMAS '23, p 1706–1715

Brams SJ (2008) Mathematics and democracy: designing better voting and fair-division procedures. Math Comput Model 48(9):1666–1670

Brams SJ, Sanver MR (2009) Voting systems that combine approval and preference. In: Brams SJ, Gehrlein WV, Roberts FS (eds) The Mathematics of Preference, Choice and Order: Essays in Honor of Peter C. Fishburn. Springer Berlin Heidelberg, Berlin, Heidelberg, p 215–237, https://doi.org/10.1007/978-3-540-79128-7_12

Cook WD (2006) Distance-based and ad hoc consensus models in ordinal preference ranking. Eur J Oper Res 172(2):369–385.https://doi.org/10.1016/j.ejor.2005.03.048

Cook WD, Seiford LM (1982) On the Borda-Kendall consensus method for priority ranking problems. Manage Sci 28(6):621–637

D'Ambrosio A (2023) ConsRank: Compute the Median Ranking(s) According to the Kemeny's Axiomatic Approach. https://CRAN.R-project.org/package=ConsRank, R package version 2.1.4

D'Ambrosio A, Amodio S, Iorio C (2015) Two algorithms for finding optimal solutions of the Kemeny rank aggregation problem for full rankings. Electron J Appl Stat Anal 8(2):198–213. https://doi.org/10.1016/j.ejor.2015.08.048

D'Ambrosio A, Mazzeo G, Iorio C et al (2017) A differential evolution algorithm for finding the median ranking under the Kemeny axiomatic approach. Comput Oper Res 82:126–138

D'Ambrosio A, Vera JF, Heiser WJ (2022) Avoiding degeneracies in ordinal unfolding using Kemeny-equivalent dissimilarities for two-way two-mode preference rank data. Multivar Behav Res 57(4):679–699

David FN, Barton DE (1962) Combinatorial Chance. Hafner Publishing Co., New York, NY, USA

Dong Y, Li Y, He Y et al (2021) Preference-approval structures in group decision making: axiomatic distance and aggregation. Decis Anal 18(4):273–295

Emond EJ, Mason DW (2002) A new rank correlation coefficient with application to the consensus ranking problem. J Multi-Criteria Decis Anal 11(1):17–28

Erdamar B, García-Lapresta JL, Pérez-Román D et al (2014) Measuring consensus in a preference-approval context. Information Fusion 17:14–21

Escobedo AR, Moreno-Centeno E, Yasmin R (2022) An axiomatic distance methodology for aggregating multimodal evaluations. Inf Sci 590:322–345

Fishburn PC (1973) The Theory of Social Choice. Princeton University Press, Princeton, NJ, http://www.jstor.org/stable/j.ctt13x10fr

Fisher RA, Yates F (1953) Statistical tables for biological, agricultural and medical research. Hafner Publishing Co., New York, NY, USA

Fürnkranz J, Hüllermeier E (2011) Preference learning: an introduction. In: Fürnkranz J, Hüllermeier E (eds) Preference Learning. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 1–17. https://doi.org/10.1007/978-3-642-14125-6_1

García-Lapresta JL, Pérez-Román D (2011) Measuring consensus in weak orders. In: Herrera-Viedma E, García-Lapresta JL, Kacprzyk J et al (eds) Consensual Processes. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 213–234. https://doi.org/10.1007/978-3-642-20533-0_13

Good IJ (1980) C59 . The number of orderings of n candidates when ties and omissions are both allowed. J Stat Comput Simul 10(2):159–159. https://doi.org/10.1080/00949658008810357

Ismailoglu F (2022) Aggregating user preferences in group recommender systems: a crowdsourcing approach. Decis Support Syst 152:113663. https://doi.org/10.1016/j.dss.2021.113663

Kamwa E (2023) On two voting systems that combine approval and preferences: fallback voting and preference approval voting. Public Choice 196(1):169–205. https://doi.org/10.1007/s11127-023-01070-z
Kemeny JG, Snell J (1962) Mathematical models in the social sciences. Blaisdall Publishing Company, New York, NY, USA
Kruger J, Sanver MR (2021) An Arrovian impossibility in combining ranking and evaluation. Soc Choice Welfare 57:535–555
Laslier JF, Straeten KV (2004) Une expérience de vote par assentiment lors de l'élection présidentielle française de 2002. Revue Française de Science Politique 54(1):99–130. https://doi.org/10.3917/rfsp.541.0099
Li X, Wang X, Xiao G (2017) A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications. Brief Bioinform 20(1):178–189. https://doi.org/10.1093/bib/bbx101
Liang H, Xiong W, Dong Y (2018) A prospect theory-based method for fusing the individual preference-approval structures in group decision making. Comput Ind Eng 117:237–248. https://doi.org/10.1016/j.cie.2018.01.001
List C (2022) Social Choice Theory. In: Zalta EN, Nodelman U (eds) The Stanford Encyclopedia of Philosophy, Winter 2022 edn. Metaphysics Research Lab, Stanford University, https://plato.stanford.edu/archives/win2022/entries/social-choice/
Liu H, Xu Z, Jiang L et al (2023) Multi-criteria group decision making with preference approval structures: a personalized individual semantics approach. Inform Fusion 96:80–91. https://doi.org/10.1016/j.inffus.2023.03.009
Liu Q, Crispino M, Scheel I et al (2019) Model-based learning from preference data. Annu Rev Stat Appl 6:329–354. https://doi.org/10.1146/annurev-statistics-031017-100213
Mallows CL (1957) Non-null ranking models. I. Biometrika 44(1/2):114–130. https://doi.org/10.2307/2333244
Mattei N, Walsh T (2013) Preflib: A library for preferences HTTP://PREFLIB.ORG. In: Perny P, Pirlot M, Tsoukiàs A (eds) Algorithmic Decision Theory. Springer Berlin Heidelberg, Berlin, Heidelberg, Lecture Notes in Artificial Intelligence, pp 259–270, https://doi.org/10.1007/978-3-642-41575-3_20
May KO (1952) A set of independent necessary and sufficient conditions for simple majority decision. Econometrica 20(4):680–684. https://doi.org/10.2307/1907651
Neve J, Palomares I (2019) Aggregation strategies in user-to-user reciprocal recommender systems. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). IEEE, pp 4031–4036, https://doi.org/10.1109/SMC.2019.8914362
Patty JW, Penn EM (2015) Aggregation, evaluation, and social choice theory. The Good Society 24(1):49–72. https://doi.org/10.5325/goodsociety.24.1.0049
Pujahari A, Sisodia DS (2020) Aggregation of preference relations to enhance the ranking quality of collaborative filtering based group recommender system. Expert Syst Appl 156:113476. https://doi.org/10.1016/j.eswa.2020.113476
R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/
Sanver MR (2010) Approval as an intrinsic part of preference. In: Laslier JF, Sanver MR (eds) Handbook on approval voting. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 469–481. https://doi.org/10.1007/978-3-642-02839-7_20
Smith JH (1973) Aggregation of preferences with variable electorate. Econometrica 41(6):1027–1041
Yazidi A, Ivanovska M, Zennaro FM et al (2022) A new decision making model based on rank centrality for gdm with fuzzy preference relations. Eur J Oper Res 297(3):1030–1041. https://doi.org/10.1016/j.ejor.2021.05.030
Ye J, Sun B, Bai J et al (2024) A preference-approval structure-based non-additive three-way group consensus decision-making approach for medical diagnosis. Inform Fusion 101:102008. https://doi.org/10.1016/j.inffus.2023.102008