



**HAL**  
open science

## A simulation-based study of proximity between voting rules

Jean-Baptiste Aubin, Irène Gannaz, Samuela Leoni-Aubin, Antoine Rolland

► **To cite this version:**

Jean-Baptiste Aubin, Irène Gannaz, Samuela Leoni-Aubin, Antoine Rolland. A simulation-based study of proximity between voting rules. 2024. hal-04631154

**HAL Id: hal-04631154**

**<https://hal.science/hal-04631154v1>**

Preprint submitted on 1 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A simulation-based study of proximity between voting rules

Antoine Rolland<sup>1</sup>, Jean-Baptiste Aubin<sup>2</sup>, Irène Gannaz<sup>3</sup>, and Samuela Leoni<sup>2</sup>

<sup>1</sup>ERIC EA 3083, Université de Lyon, Université Lumière Lyon 2,  
5 Pierre Mendès France, 69596 Bron Cedex, France

<sup>2</sup>Univ Lyon, INSA Lyon, UJM, UCBL, ECL, ICJ, UMR5208,  
69621 Villeurbanne, France

<sup>3</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP,  
38000 Grenoble, France

June 2024

## Abstract

Our research program proposes to analyze the proximity between voting rules considering either their respective winners are identical or not. The more often two different voting rules lead to the same winner, the more similar they should be considered. Our methodology consists in simulating a large enough amount of elections, and analyzing the distance between each pair of voting rules based on the agreement/disagreement of the election winner. We analyze 15 voting rules, including several evaluation-based voting rules. The experimental design includes numbers of voters varying from 9 to 1001 and numbers of alternatives varying from 3 to 14. Some designs with correlations or different distributions are also explored. The study analyzes the influence of all these variables on the proximity between voting rules. Several proximity maps are presented. Unsurprisingly, voting methods based on the same principles (plurality rules, Condorcet like rules, evaluation-based rules) generally conduct to a proximity of winner. However, these proximities can change regarding the control variables. Moreover, some specific rules appear to be a kind of link or hyphen between two families of voting methods.

**Keywords:** Voting rules, Evaluation-based voting rules, Simulation, IC model.

# 1 Introduction

A voting rule is a process which, for all set of voters and alternatives, gives a winner of the election based on the voters' opinions on alternatives. These opinions can be expressed either by a ranking on the set of alternatives, or by evaluations given to each alternative. For example, Borda count is based on preference orders while majority judgment or approval voting are based on evaluations. Note that partial orders are used by some “weaker” (due to the lack of information used) voting rules as First-past-the-post voting, for example. Our objective is to compare both ranking-based and evaluation-based voting methods.

An axiomatic approach of comparison of two voting methods needs a common axiomatic framework. As ranking-based and evaluation-based voting methods do not use the same conceptual framework, we propose in this paper to compare voting rules through an empirical approach, i.e. looking at the result of the voting scheme rather than its properties. Our empirical approach consists in simulating several voting situations and deducing similarities based on the voting rules: the more often two different voting rules lead to the same winner, the more similar they should be considered, whatever their methodology or properties.

We restrict our framework to the continuous case for the evaluations scale. The discrete case, which implies specific tie-break rules, is not considered. In a precedent paper, [Rolland et al. \(2024\)](#) suggest several models to simulate evaluation-based voting data. In this paper, we mainly use the so-called IID models to explore the similarities between several voting rules, establishing an experimental setting where the number of voters, the number of alternatives and the distributions of the evaluations vary. This yields Impartial Culture modeling on rankings. This framework is the most usual in literature ([Boehmer et al., 2024](#)). A perspective is to extend the study to non IID (resp. non IC) models. We simulated few examples in this framework, and a more thorough simulation design may be the object of future work.

Several studies have already been proposed dealing with the similarities between voting rules. [Richelson \(1979\)](#), [Felsenthal and Machover \(2012\)](#) or [Kacprzyk et al. \(2020\)](#) compare ranking-based voting rules with respect to their satisfied properties. [Nurmi \(1987\)](#) does the same including evaluation-based methods, which are approval voting and median voting rule. [Conitzer and Sandholm \(2012\)](#) classify the common ranking-based voting rules with respect to a maximum likelihood model: it is also an axiomatic approach. [Tideman \(2023\)](#) propose some principles to choose among a large variety of voting processes, based on a large view of how do the voting rules work.

Using real-world elections to compare different voting rules has been proposed for example in [Mattei et al. \(2012\)](#) and [Song \(2023\)](#). [Tideman and Plassmann \(2013\)](#) propose a statistical modeling, generating voting

situations, which parameters are fitted using real-world elections. In their study, a best alternative is defined in the simulation process and their objective is to study which voting rule recover this alternative. [De Almeida et al. \(2019\)](#) propose a specific method to choose the best voting rule in a specific context.

Probabilistic studies have been proposed to study voting rules: [Nurmi \(1987\)](#) examines the probability that two different voting rules select two different winners. [Merlin et al. \(2000\)](#) compute the probability that some decision rules select the same winner.

Actually, two papers are similar, in their goal, to our paper. They both use simulated voting situations to compare voting rules. [McCabe-Dansted and Slinko \(2006\)](#) explores the similarities between 26 voting rules in a scenario with 5 alternatives and 85 voters, examining how variations in the homogeneity of voting preferences affect outcomes using a Polya urn model. They employ Hierarchical Ascending Classification and Multidimensional Scaling to cluster voting processes and facilitate visualization. [Green-Armytage et al. \(2016\)](#) compare 54 voting rules in a scenario with 3 alternatives and 99 voters. Their study deals with IC and IAC simulations, spatial models, and real datasets to evaluate the utilitarian efficiency and strategy-proofness of these rules, aiming to identify the “best voting rule” rather than classifying the studied rules. Their conclusions indicate that the Condorcet-Hare method represents a good compromise across the studied criteria, while the plurality method is clearly outperformed by numerous other methods in both dimensions. Our approach is more similar than [McCabe-Dansted and Slinko \(2006\)](#) in the sense that we observe the similarities between voting rules, and not their behavior with respect to given properties (which is done in [Green-Armytage et al. \(2016\)](#)). Additionally, [McCabe-Dansted and Slinko \(2006\)](#) consider fixed number of voters and alternatives, while we want to study their influence on voting procedures. We also include evaluation-based procedures in the study, which were lacking in [McCabe-Dansted and Slinko \(2006\)](#).

In this article, we aim to expand upon these previous studies by considering different scenarios, particularly regarding the number of voters, the number of alternatives, and the generation of voter evaluations (both in terms of distribution and dependencies). The article is organized as follows. Section 2 presents the selected voting rules. It also introduces the methodology and the experimental design of the study. Section 3 presents a global analysis of the simulation results, whereas Section 4 explores the influence of the simulation setting. Section 5 gives an insight on non IC frameworks, as a perspective of this article.

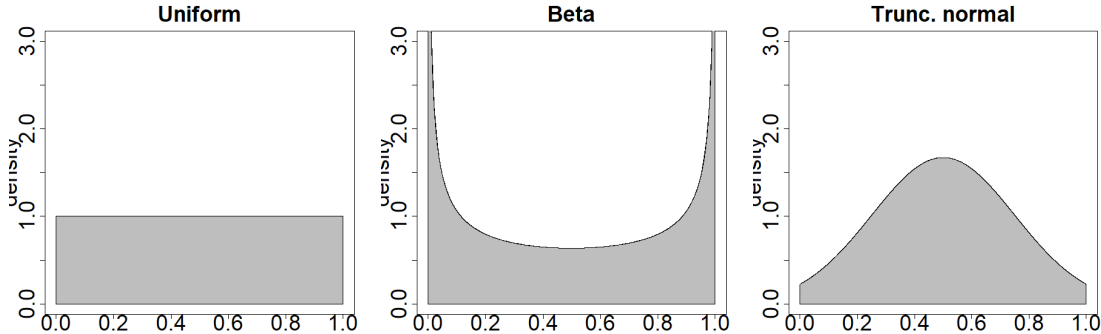


Figure 1: Uniform (left), Beta(0.5, 0.5) (center) and Truncated Normal(0.5, 0.25) (right) distributions

## 2 Hypothesis and Methodology

The goal of this paper is to find proximity relations between voting rules considering the winner of the election. The more often two different voting rules lead to the same winner, the more similar they should be considered. Our methodology consists in simulating a large enough amount of election situations, determining the winner for each election and each considered voting rule, and analysing the distance between each pair of voting rules based on the agreement/disagreement of the election winner.

### 2.1 Methodology

In the following, we consider situations with  $n$  voters and  $m$  alternatives. Each voter associates a grade from the continuous set  $[0, 1]$  to each alternative. These grades are simulated following three different methods: Uniform distribution, Beta distribution or truncated Normal distribution (see [Rolland et al. \(2024\)](#) for details and [Figure 1](#) for illustration). Preference rankings on alternatives are derived from the grades given by the voters to the alternatives: the greater the evaluation, the better the alternative should be ranked. The chosen experimental design controls the parameters as indicated in [Table 1](#).

voters	$n = 9, 15, 21, 51, 101, 1001$
alternatives	$m = 3, 4, 5, 6, 7, 9, 14$
Distributions	Uniform, Beta(0.5, 0.5), Truncated Normal(0.5, 0.25)

Table 1: Experimental design parameters

For each of the  $6 \times 7 \times 3 = 126$  combination of voters, alternatives and simulation processes, 1000 iterations have been run. A matrix of dissimilarity has been produced between the voting processes. The distance between two voting rules is the proportion of situations where the two rules lead to the same winner. Two voting rules are considered identical if their winner is always the same. The distance between them is then equal to 0. Two voting rules are considered totally dissimilar if they always choose two different winners. The distance between them is then equal to 1. Others distance scores should be possible (see below). The obtained matrix for a given combination of voters, alternatives and simulation process is therefore a matrix of dissimilarity with scores between 0 (if two rules always agree on the same elected alternative) and 1 (if two rules always disagree on the elected alternative). Therefore our study concerns 126 dissimilarity matrices. Therefore 126.000 simulations have been computed and combined to explore the influence of the number of voters, the number of alternatives or the evaluation simulation processes on the similarity of voting rules.

### **Tie-breaking and distance between voting-rules**

Whatever the voting rule, ties are a matter of concern, especially when the number of voters is small (9 or 15): several voting situations combined with voting rules lead to two or more alternatives to be ex-æquo. Our decision is not to break ties, and consider all the alternatives that are equally ranked first as equally winners of the election. Therefore, the result of a voting situation for a specific voting rules is a set of alternatives, or a 0/1  $m$ -vector with 1 if the alternative is in the set of winners and 0 if not. The distance between two voting rules is measured through the binary distance (as known as Jaccard distance) between the two vectors of winners. If  $\text{Win}_i$  is the set of winners given by voting rule  $i$ , the binary distance between voting rules  $i$  and  $j$  is given by

$$d(i, j) = 1 - \frac{|\text{Win}_i \cap \text{Win}_j|}{|\text{Win}_i \cup \text{Win}_j|}.$$

## **2.2 Voting rules**

Several authors have already presented a large list of voting rules, for instance [Felsenthal and Machover \(2012\)](#); [Durand \(2015\)](#), or more recently to [Tideman \(2023\)](#). We refer to these authors for more details about the large variety of voting rules. Among this large variety, we selected 16 voting rules using either preference rankings or evaluations for our study.

### 2.2.1 Plurality-based voting rules

These five first voting rules are based on the plurality concept: the alternative to be chosen should be the one supported by a majority of voters. Several meaning of the “majority” concept lead to several voting rules.

- Plurality rule (PLU1) - each voter selects its preferred alternative. The alternative with the greatest plurality score is the winner.
- Plurality two-round runoff procedure (PLU2) - in the first round, voters select their preferred alternative as in PLU1. If no alternative gets a majority, there is a second round involving only the top 2 alternatives of the first round.
- Hare procedure (HARE) - also known as alternative vote - each voter votes for its preferred alternative (in the set of remaining alternatives). The alternative with the smallest plurality score is eliminated and the procedure goes on until at least one alternative has a majority, which is the winner.
- Bucklin procedure (BUCK) - each voter votes for its preferred alternative. If an alternative reaches the majority, it is the winner. If not, each voter votes for its two preferred alternatives. The procedure goes on with the three preferred alternatives, etc. until at least one alternative reaches the majority. If there is only one alternative reaching the majority, it is the winner. If two or more alternatives reach the majority, the one with the greatest number of voters is the winner.

### 2.2.2 Condorcet-based voting rules

The Condorcet principle supposes that if there is an alternative preferred by a majority of voters to any other alternative in head-to-head comparisons then this alternative should be chosen. Such an alternative is called the Condorcet winner. It is often argued that the Condorcet winner should be elected.

- Condorcet winner rule (COND)- if a Condorcet winner exists, it is the winner of the election. Else, no alternative is considered to be the winner.

Nevertheless it is well-known that there are voting situations without any Condorcet winner ([Gehrlein, 1999](#)). Several voting rules have been proposed to determine, in that case, which alternative should be

considered as the nearest alternative to the Condorcet winner position. We introduce hereafter two voting rules based on the Condorcet principle.

- Minimax (MMAX) - if a Condorcet winner exists, it is the winner of the election. If not, each alternative is being affected its plurality score in its worst duel against any other alternative. Then the alternative with the highest score is chosen.
- Copeland (COPE) - each alternative gets 1 point for each won duel against each other alternative, 0.5 points if there is a tie and 0 points for each lost duel. The alternative with the best score wins the election.

### 2.2.3 Borda like voting rules

Borda like voting rules are based on average ranks of each alternative.

- Borda count (BORD) - the Borda count voting rule computes the average rank of preference for each alternative on the set of voters. The alternative with the best average rank wins the election.
- Nanson (NAN) - the Nanson rule is an iterative version of Borda rule. On each step, alternatives with an average rank equal or greater than the average rank ( $m/2$  for  $m$  alternatives) are cancelled, and the rule goes on with the remained alternatives until only one remains.

### 2.2.4 Evaluation-based voting rules

Evaluation-based voting rules are not using ranks of preferences but evaluations of voters on alternatives to compute the winner of the vote.

- Approval voting (AP) consists in a range voting with binary (0/1) evaluations for each alternative. In the simulation process, it is necessary to define the number of evaluations equal to 1 with respect to the continuous evaluations. Three different versions are considered:
  - Approval with threshold (AP-T) - any score greater than 0.7 corresponds to an approval (1) and any score less than 0.7 to an disapproval (0). Note that, in order to avoid trivial situations, for any voter its alternative with the maximum evaluation is always considered to be approved and its alternative with the minimum evaluation is always considered to be disapproved.



- Approval with Uniform distribution of approved alternatives (AP\_U) - each voter approves between 1 and  $m - 1$  alternatives, with  $m$  the total number of alternatives. The number of approved alternatives follows an Uniform distribution.
- Approval of half the alternatives (AP\_H) - each voter approves half of the alternatives, the ones with the best evaluations.
- Majority judgement (MJ) - the alternative with the best median evaluation wins the election. Note that the tie-break rule is not a matter of interest here as the evaluations are on a continuous scale, contrary to the original rule proposed by [Balinski and Laraki \(2010\)](#).
- Range Voting (RV) - the alternative with the best average evaluation wins the election.
- Weighted  $L^4$  Deepest Voting Rule (L4DV) - the winner of the election is determined by the choice of the innermost voter in the set of voters, which is determined through the use of the weighted  $L^4$  depth function (see [Aubin et al. \(2022\)](#) for details).

All these procedures can be seen as Weighted  $L^p$  Deepest Voting Rule, as defined in [Aubin et al. \(2022\)](#). AP, MJ, RV and L4DV correspond to Weighted  $L^p$  deepest voting with respectively  $p = 0, 1, 2, 4$ . AP was not considered in the former article, but it can be shown that it corresponds to the case  $p = 0^1$ .

Finally, we also consider

- Star rule (STAR) - first the average evaluation is calculated for each alternative and the best two are selected. Then a plurality voting is organized between these two alternatives and the winner wins the election.

## 3 Voting rules proximities

### 3.1 Analysis of distance between voting rules

Figure 2 shows the distance matrix between voting rules based on the 126.000 voting simulations: 7 alternatives levels  $\times$  6 voters levels  $\times$  3 simulation methods  $\times$  1000 simulations for each combination. It

<sup>1</sup>The  $L^0$  “norm” counts the number of non zero values. Taking  $L^0$ , the associated weighted deepest voting is defined as in [Aubin et al. \(2022\)](#), but it is not a deepest voting strictly speaking because taking  $p = 0$  does not lead to satisfying properties. Yet, the (weighted)  $L^0$  deepest voting can be seen as the procedure which elects the alternative with the highest non zero evaluations. This indeed corresponds to the AP procedure - with evaluations equal to 0 or 1.

has been obtained by averaging the distance matrices computed as described in Section 2.1. The distance between two voting rules is in the range  $[0, 1]$  and represents the proportion of voting situations where two procedures yield different winners. The average distance between two voting rules is 0.417. This value is very high, meaning that there are important differences between the voting rules. The minimum distance is between NAN and MMAX (0.123), and the maximum distance is between PLU1 and BUCK (0.61). The matrix shows that there is a structure relative to the voting rules. In particular, BORD, STAR, COPE, MMAX, NAN, COND and HARE form a relatively homogeneous group. These rules correspond to the Borda like and Condorcet like rules. It does not contain BUCK. BUCK can be seen as an intermediate rule between AP\_H and BORD, which seems natural by construction. Note also that PLU1 and PLU2 are far from the evaluation-based voting rules (AP\_U, AP\_T, AP\_H, MJ, RV, L4DV).

To improve the interpretation of the distance matrix, Figure 3 and Figure 4 show the associated multidimensional scaling (MDS) representation. The first three axes of the MDS represent 57% of the information contained in the distance matrix.

### 3.1.1 Discriminative axis analysis

The first axis clearly opposes, on one side, voting methods that rely on limited information -partial set of voters' preference orders- (PLU1, PLU2) to those that use much more comprehensive and nuanced information (evaluation-based voting methods such as MJ, RV, L4DV, but also BUCK), leaving in the middle the voting procedures based on the entire sets of voters' preference rankings (COND, MMAX, COP, etc). The second axis, on the other hand, opposes the Condorcet methods to all others, leading to a V-shaped structure with the plurality voting systems at the top left, the Condorcet-based (MMAX, COP) methods at the bottom, and the evaluation-based voting systems at the top right. The third axis discriminates between evaluation-based methods and approval-based methods. It shows that there is a graduation between APP rules. Next, MJ, RV and LP4D correspond to the  $L^p$  distance used in the deepest voting model introduced in Aubin et al. (2022); as explained previously, AP corresponds to (a weighted)  $L^0$  depth, MJ to (a weighted)  $L^1$  depth, RV to  $L^2$  depth and L4DV to  $L^4$  depth. These rules are, hence, sorted with respect to the exponent  $p$  of the corresponding weighted  $L^p$  depth. It corresponds to the sensibility of the rules to the extremes evaluations: the higher  $p$ , the more sensible to extremes evaluations. Finally, note that BUCK voting rule mainly corresponds to AP\_H, an approval voting with voters approving half of the alternatives -see Section 4 for a more in-depth analysis.

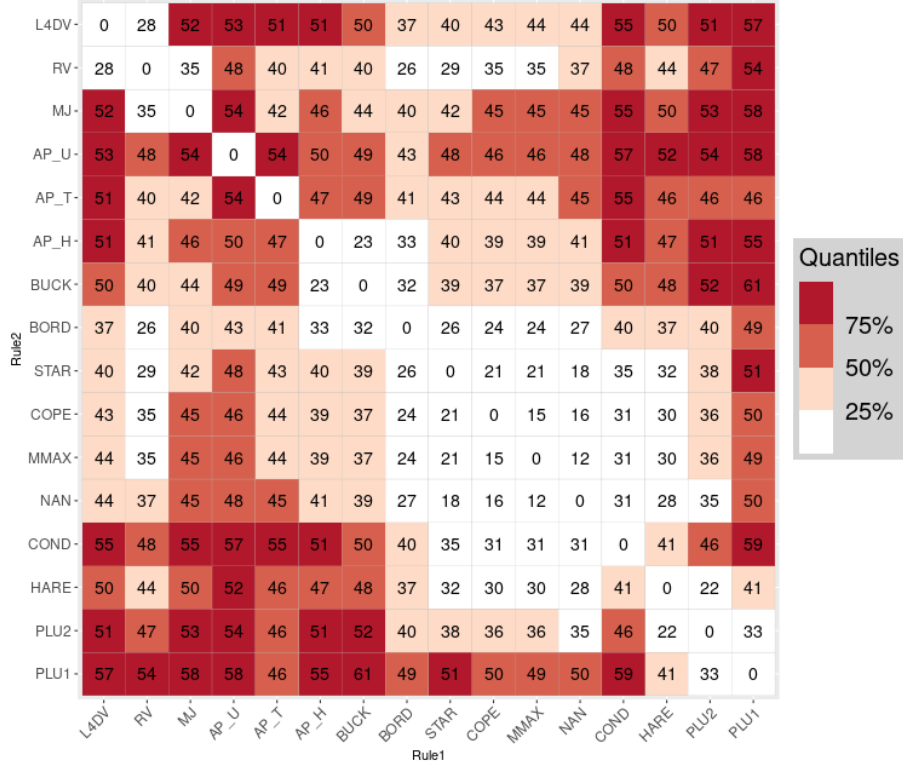


Figure 2: Distance matrix between voting rules. The displayed values are equal to 100 times the distance to increase the readability. The color of each entry corresponds to the membership of the inter-quartile intervals (respectively  $[0,0.357)$ ,  $[0.357, 0.437)$ ,  $[0.437,0.497)$ ,  $[0.497,1]$ ).

### 3.1.2 Family methods analysis

Considering the Plurality family (PLU1, PLU2, HARE and BUCK), PLU1 voting method stands out as the most characteristic: it is the furthest from the other voting methods (highest average distance to others). One possible explanation could be that it relies solely on the first preference of the voters, whereas all other voting methods are based on less partial information. PLU2 and HARE voting methods, iterative versions of PLU1 based on subsequent alternatives in the voters' preference rankings, exhibit

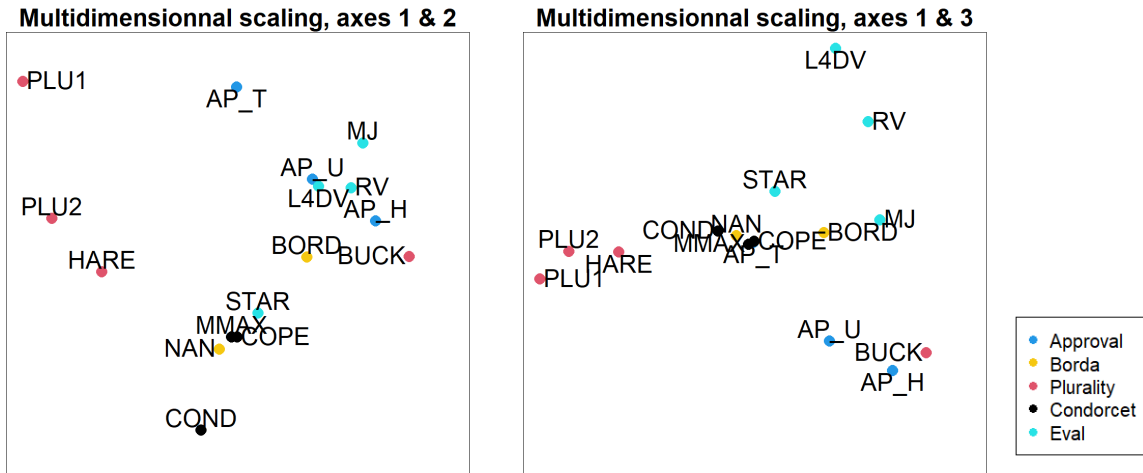


Figure 3: Multidimensional scaling representation of the distance matrix between voting rules - dimensions 1 and 2 (left) and dimensions 1 and 3 (right). Note that the goodness-of-fit of the 3 first axes are equal respectively to 0.24, 0.20 and 0.13.

similarities with PLU1 while being closer to the others. As mentioned above, BUCK, which is formally a plurality-based voting rule, is very different in its results from the other plurality rules, and acts as an intermediate between BORD and AP\_H. It is not surprising: BUCK is analogous in ranks to the MJ in evaluations (just as BORD is in relation to RV). The similarity between AP\_H and BUCK is easily explained by the fact that, when the number  $m$  of alternatives is great, BUCK selects a winner in  $m/2$  iterations, and then behaves exactly as an approval voting approving half of the alternatives.

The voting methods in the center of the MDS representation come from Condorcet (COND, COPE, MMAX) and Borda (BORD, NAN) families, adding STAR method. Not surprisingly, voting methods inspired by the idea of Condorcet (MMAX and COPE) are the closest from COND. A bit further but still at low distances, BORD and NAN, as well as STAR, complete the group by being closer to COND than the voting methods of other groups. This can be understood because STAR is a hybrid voting method that initially relies on the averages of evaluations to select two alternatives (similar to RV), then elects the alternative receiving the most votes among the two finalists (similar to PLU1 or PLU2). STAR is then more likely to choose a Condorcet winner than just RV. It is then placed between the Condorcet-based and the evaluation-based families. Next, NAN is close to the Condorcet like group. It is known that NAN elects a Condorcet winner if one exists, explaining this observation. Finally, BORD is not far from voting methods based on evaluations. This can be understood through an analogy between BORD and

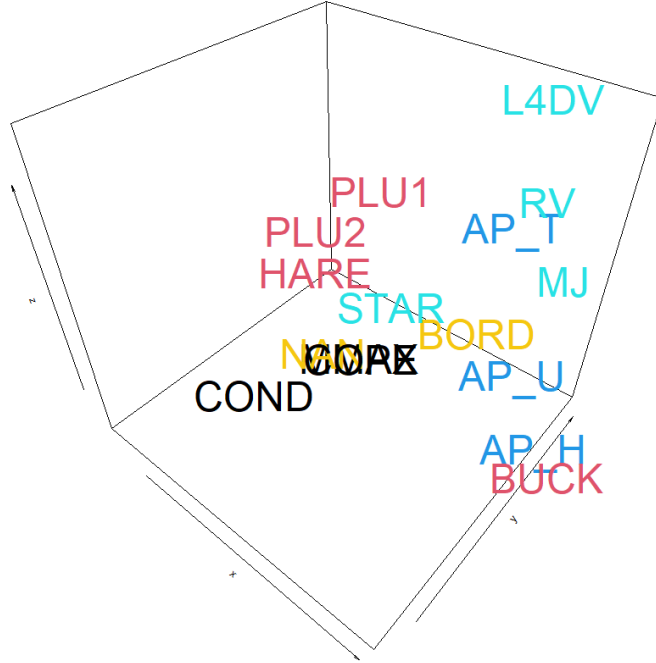


Figure 4: Multidimensional scaling representation of the distance matrix between voting rules - three dimensions.

RV: RV elects the alternative whose average evaluation is the highest; BORD elects the alternative whose average ranks is the best. Thus, in the case where the scores are inverted ranks (0 for last, 1 for second to last, etc.), the two voting methods coincide.

The last family consists in voting methods based on evaluations, and includes AP\_H, AP\_U, AP\_T, MJ, RV, and L4DV. Several elements stand out in this group. For instance, L4DV is atypical: it is notably distant from the voting methods in its class based on the  $L^1$  (resp.  $L^0$ ) distance such as MJ (resp. AP). However, it aligns more closely with RV (and BORD). This can be explained by the nature of winner selection, which is more sensitive to extreme evaluations than others (as discussed earlier). Similarly, AP\_T lies between evaluation-based voting rules and PLU1, as in several cases approving all

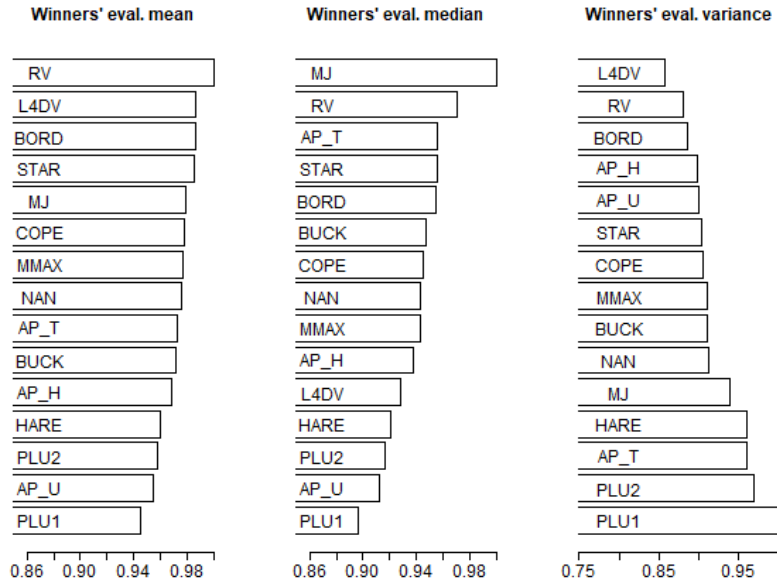


Figure 5: Average mean, median and variance of winner's evaluations.

the alternatives with an evaluation greater than the threshold consists in approving only one alternative, i.e. it has the same behavior as PLU1.

### 3.2 Comparison of winners' characteristics

It is clear that, depending on the voting system adopted, the characteristics of the elected alternative will differ. To explore this, we evaluate some characteristics of the elected alternatives. Given a voting procedure, we calculate for the elected alternative on each voting situation the mean, the median and the variance of the evaluations (resp. the rankings). The mean of these statistical characteristics on all the voting situations are given respectively in Figure 5 for evaluations and Figure 6 for rankings. As Condorcet's procedure does not always select a winner, we choose here not to display its characteristics.

Of course, by definition, an alternative elected by the RV will have the highest possible average rating. Similarly, an alternative elected by BORD will have the highest average rank. All the same, an alternative elected using MJ will have the highest median of its ratings, and an alternative elected using BUCK will have the highest median of its ranks.

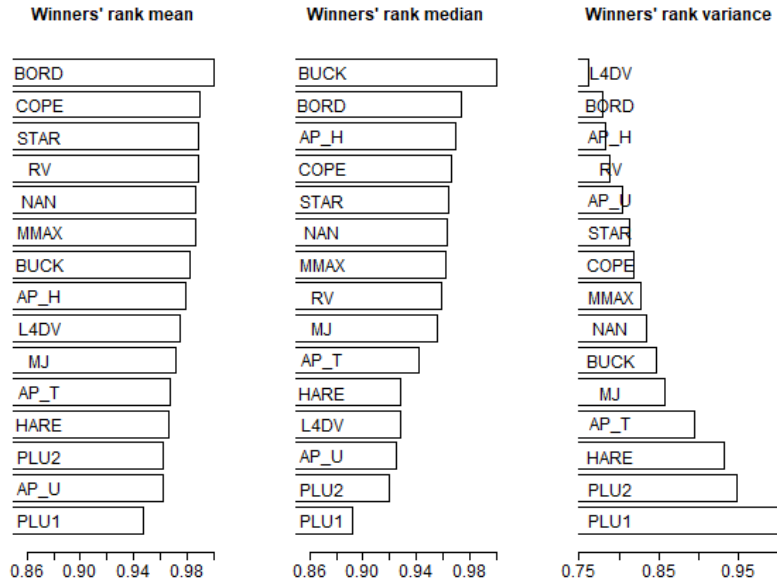


Figure 6: Average mean, median and variance of winner’s rankings.

Additionally, with regard to the different characteristics, we can observe the following:

- With regard to the mean of the evaluations, after RV, three voting procedures have a high mean for the winner: BORD, based on average on ranks; L4DV, which is very close to RV; STAR, which elects one of the two alternatives with the highest average ratings. On the opposite, HARE and AP\_U elect alternatives whose average ratings are low, but even higher than PLU2 and especially PLU1, which has the lowest values.
- With regard to the median of the evaluations, MJ and RV have the highest values. This is consistent since these evaluation-based methods are both designed to elect the alternatives with the highest grades. The four lowest values are obtained for the same voting systems as before, with PLU1 far lower than the others.
- With regard to the variance of the evaluations, the lowest variance is obtained with L4DV, followed by RV and BORD. Note that a low variance alternative can be interpreted as a consensual alternative. On the opposite, PLU2 and PLU1 lead to the election of polarizing alternatives (like AP\_T and HARE to a lesser extent).

- With regard to the mean of the rankings, after BORD, procedures COPE, STAR and RV are also convincing. As usual, on the opposite, we find HARE, AP\_U, PLU2 and PLU1 electing alternatives with the lowest rankings' average.
- With regard to the median of the rankings, after BUCK, BORD procedure is still the best while L4DV joins HARE, AP\_U, PLU2 and PLU1 at the bottom of the table.
- With regard to the variance of the rankings, L4DV is the procedure electing the alternative with the least variance of obtained rankings. It is followed by BORD, AP\_H and RV. On the other side, PLU1, PLU2, HARE and AP\_T elect the alternative with the highest variance of rankings, that is, an alternative that is less consensual than the alternative elected by L4DV, BORD, AP\_H and RV.

If the objective is to elect the alternative with the highest and least dispersed ratings (resp. rankings), then BORD and RV should be preferred, and PLU1 and PLU2 should be excluded.

The differences between the Approval voting procedures can be explained by the influence of the number of approvals. In AP\_T, evaluations higher than 0.7 lead to an approval and evaluations lower than 0.7 are rejected. Then, the higher the mean and the higher the variance, the higher the probability that the evaluation is higher than 0.7. AP\_U behaves differently, since the number of approvals is random. Then, the alternative chosen by PLU1, which has the highest probability to be evaluated higher than others, has the highest probability to be approved. This explains the proximity between PLU1 and AP\_U for the mean and the median (of rankings and on evaluations). The variance is nevertheless lower for AP\_U since, as the number of approvals is random, the variance has less effect on the fact that an alternative is approved. In AP\_H, each voter approves half of the alternatives. It is intermediate between AP\_U and AP\_T.

## 4 Parameters' influence

To complete the overall study carried out so far, we now study the influence of the parameters of the voting situations on the voting rules behavior. As explained in Section 2, the two main parameters are the number of alternatives and the number of voters. The evaluations distributions may also influence evaluation-based voting rules, but not rank-based ones. Observe that McCabe-Dansted and Slinko (2006) considered 85 voters and 5 alternatives and Green-Armytage et al. (2016) considered 99 voters and 3 alternatives.



## 4.1 Influence of the number of alternatives

Figure 7 displays the distance between voting rules for various number of alternatives (with all cases of voters). The distance between voting rules increases as the number of alternatives increases. If two methods elect a winner randomly, then there would be a 1-in-3 chance that the two methods produce the same winner with 3 alternatives, giving a theoretical distance between two different voting rules equal to 0.66 (2/3). Similarly, a theoretical distance of 0.93 (13/14) would be observed in a 14-alternative ballot. In our experiment, the median distances observed are equal to 0.27 for 3 alternatives and 0.60 for 14 alternatives. The first conclusion is that voting systems do not elect winners at random, since the observed median is 0.35 lower than the one with random procedures.

### 3 alternatives

Actually, very small number of alternatives lead to particular cases where two different voting rules always give the same winner. Note that, for 3 alternatives, AP\_H and PLU1 are similar, the same holds for PLU2 and HARE. To a lesser extent, this is also the case between AP\_T and PLU1 (and AP\_H). In fact, with 3 alternatives (resp. 14 alternatives), the probability of a voter approving only one alternative is around 78% (resp. 1.4 %), and in this case, the two voting methods are identical. Moreover, as it is very likely to have a Condorcet winner with 3 alternatives (94% chance, [Gehrlein \(1999\)](#)), all the voting methods electing a Condorcet winner will be close. This is indeed what we observe with MMAX and NAN, COPE or CONDO on Figure 8, which is the multidimensional scaling representation of distances between voting rules for  $m = 3$ .

### More than 3 alternatives

As stated before, distances between voting rules are growing with the number of alternatives. From 4 to 14 alternatives, Figure 9 shows a juxtaposition of voting rules positions on the first plane of the multidimensional scaling derived from the distance matrices (presented in Figure 7). Each curve corresponds to a voting rule, from its position for 4 alternatives to its position for 14 alternatives. Three main phenomena can be observed. First, it is clear that distances between voting rules increase with the number of alternatives, especially for the Plurality family rules which are more and more far from the other rules, as they use very few information compared to the other rules. Second, COND also moves away from most procedures, because there are fewer and fewer Condorcet winners. Third, it is worth noting that as the

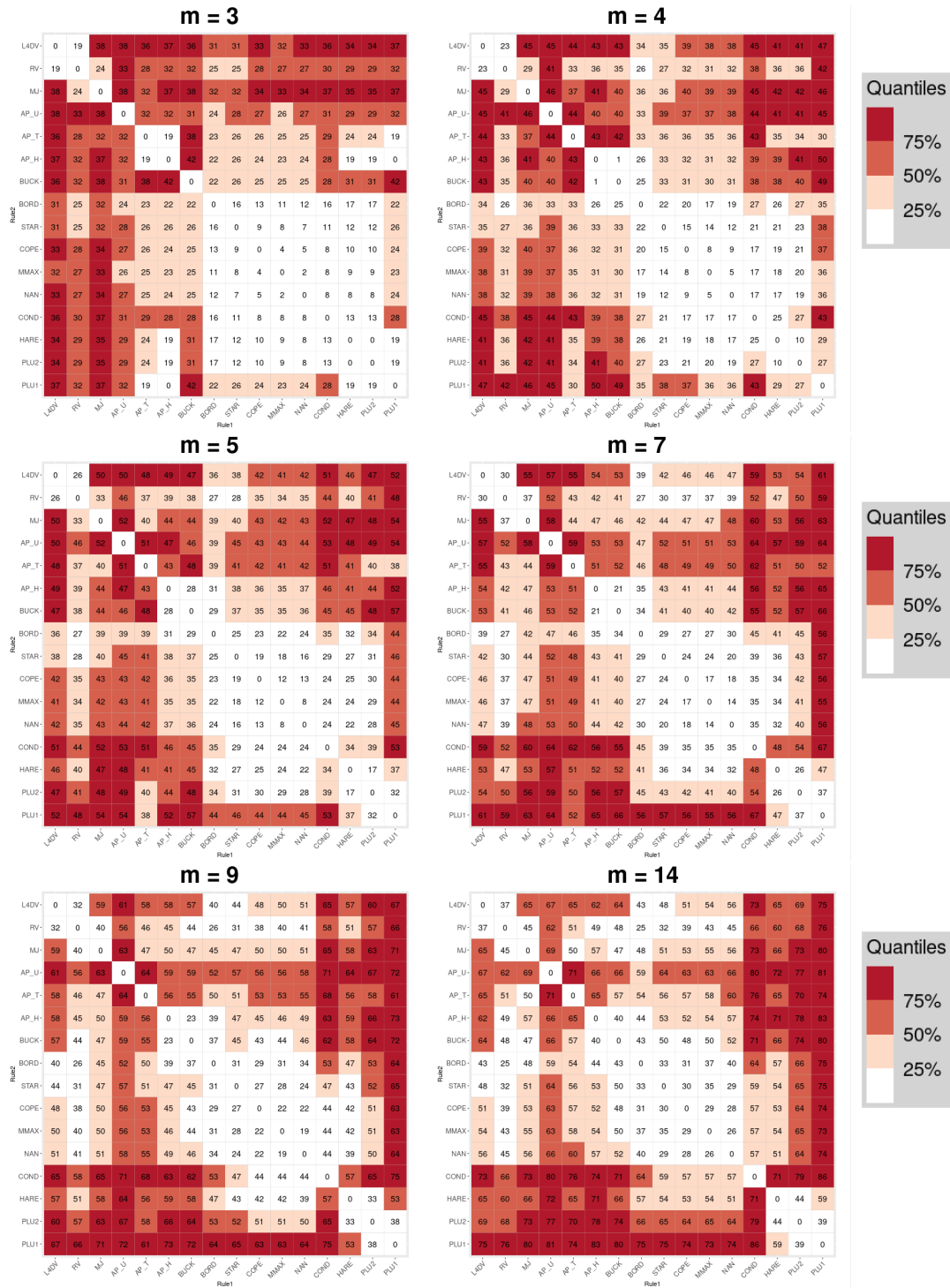


Figure 7: Distance matrices between voting rules respectively for  $m = 3, 4, 5, 7, 9,$  and  $14$  alternatives. The displayed values are equal to 100 times the distance to increase the readability. The color of each entry corresponds to the membership of the inter-quartile intervals, calculated separately for each matrix. The medians are respectively equal to 27, 37, 41, 48, 52 and 60.

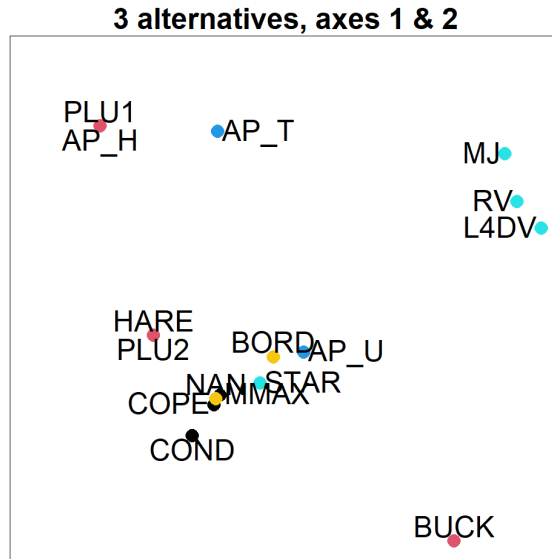


Figure 8: MDS for 3 alternatives. The goodness-of-fit equals 0.58.

number of alternatives increases, the gap of information in ranks and in scores decreases. Therefore, the evaluation-based and ranking-based voting methods using similar mechanism are getting closer, as RV and BORD (using evaluation mean or rank mean), and MJ and BUCK (using evaluation median or rank median). As a consequence, evaluation-based voting rules become “closer” to the other voting rules.

## 4.2 Influence of the number of voters

Distance matrices depending on the number of voters are given in Figure 10. It is obvious that the differences between these distance matrices are negligible: therefore, we can say that the number of voters has no influence on the proximity of two different voting rules. Especially, there appears to be no difference between the 101 voters distance matrix and the 1001 voters distance matrix. This point is a matter of importance, as it implies that a relatively small number of voters (let say around 100) is enough to study the empirical results of voting rules without lack of generality. For example, McCabe-Dansted and Slinko (2006) (resp. Green-Armytage et al. (2016)) present results obtained with 85 sample of voters (resp. 99), which is good enough to conclude.

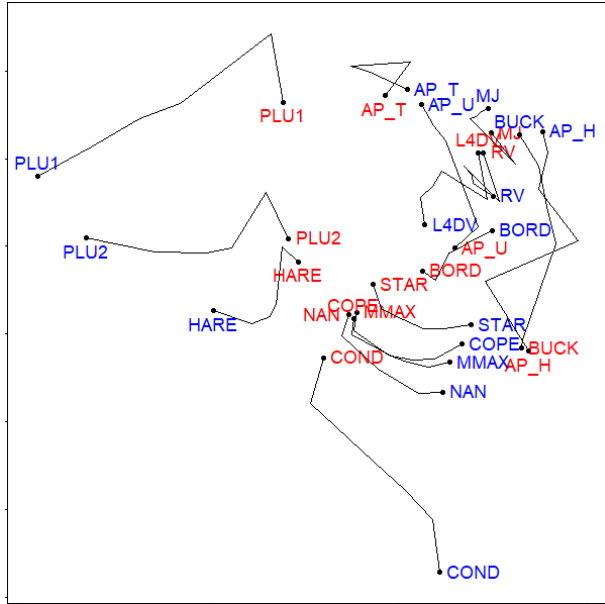


Figure 9: MDS from 4 to 14 alternatives. Red labels correspond to the case  $m = 4$  alternatives and blue labels to the case  $m = 14$  alternatives. Black lines give the position of the corresponding voting procedures when the number of alternatives varies.

### 4.3 Influence of the distributions on evaluations

Changing the law of evaluations obviously doesn't change the laws of ranks, since we're still in an IC framework, so the distances between ranking-based voting methods must remain unchanged, regardless of the distribution chosen. On the other hand, the distances between evaluation-based voting procedures (as L4DV, RV, MJ, AP) may be modified fairly slight: distances from MJ do not change, as the median is the same for the three distributions; distances from RV change a bit more, as the mean is more sensitive to the extreme values than the mode. The voting rule which is most affected by the distribution is clearly L4DV, whose behavior is strongly impacted by extreme values, which are more common with the beta distribution rather than the uniform or truncated normal distribution. Approval voting rules are also affected by the distributions: with the truncated normal distribution, approval voting with a threshold of 0.7 becomes close to first-past-the-post, as with parameters  $\mu = 0.5$  and  $\sigma = 0.25$ , there are 20% of values above 0.7, i.e. only one to three alternatives on average. Note that AP\_U (resp. AP\_H) is not affected by the distribution, as the number of alternatives is selected at random (resp. fixed), with not link to the value of evaluations. They can therefore be considered more as ranking-based than evaluation-based

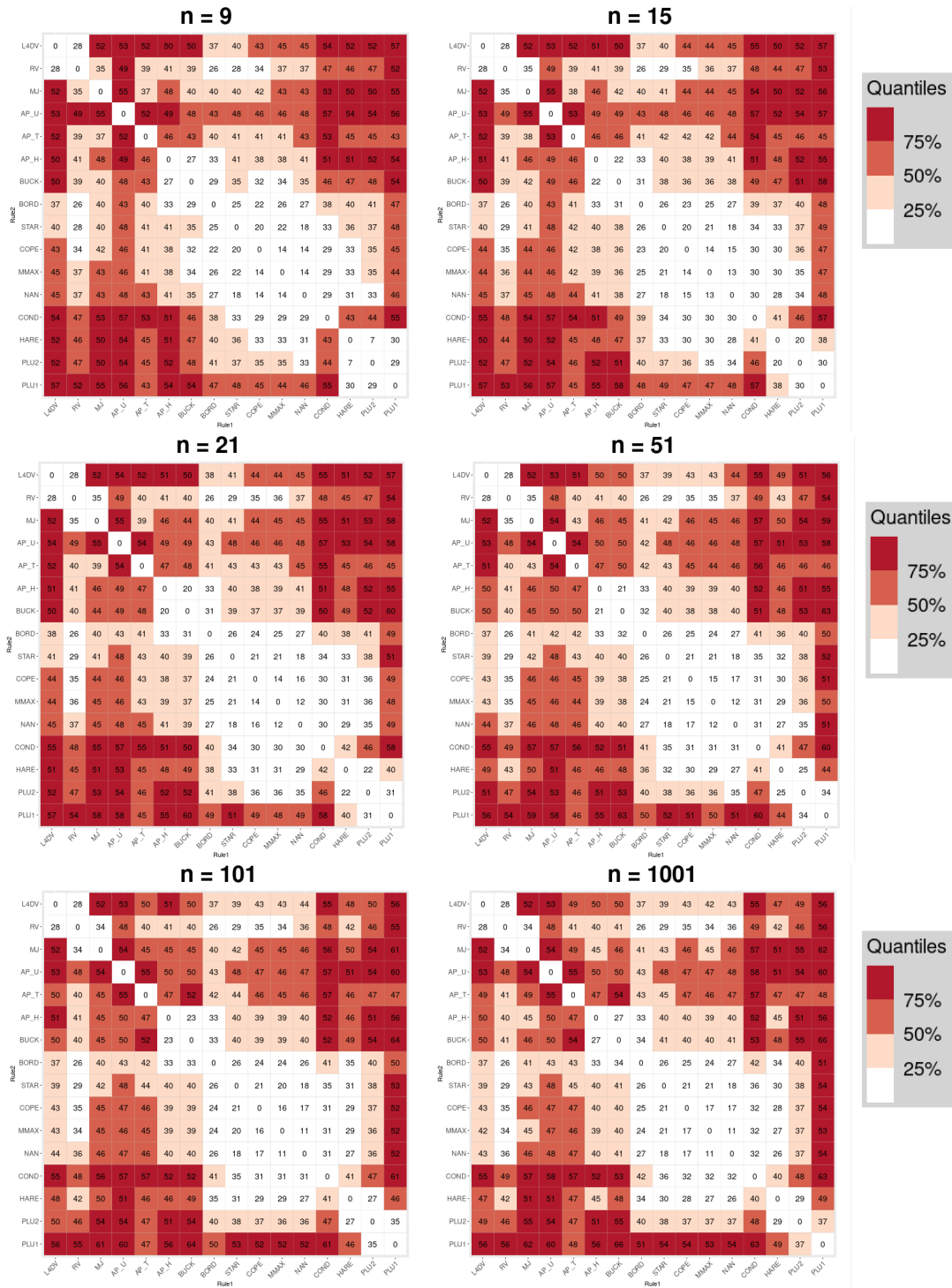


Figure 10: Distance matrices between voting rules respectively for  $n = 9, 15, 21, 51, 101,$  and  $1001$  voters. The displayed values are equal to 100 times the distance to increase the readability. The color of each entry corresponds to the membership of the inter-quartile intervals, calculated separately for each matrix.

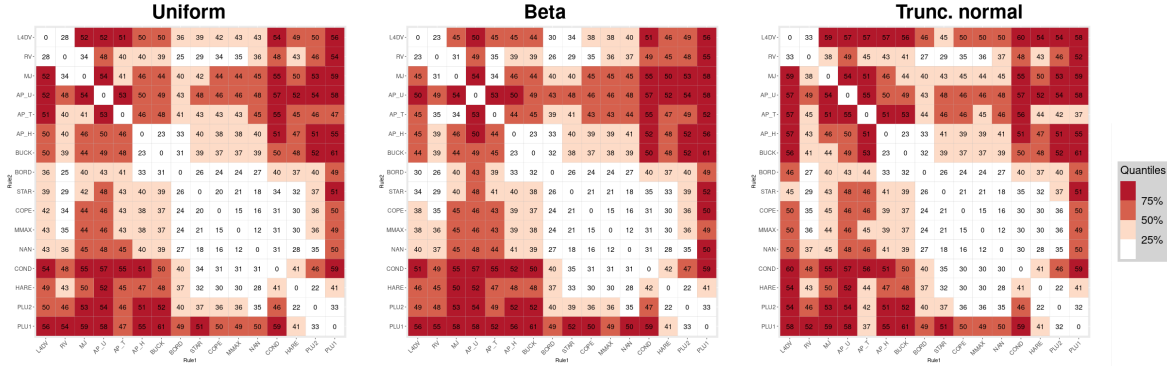


Figure 11: Distance matrices between voting rules respectively for uniform distributions, beta distributions and truncated normal distributions. The distributions are displayed in Figure 1. The displayed values are equal to 100 times the distance to increase the readability. The color of each entry corresponds to the membership of the inter-quartile intervals, calculated separately for each matrix.

voting rules.

## 5 Special cases

All the experiments above have been conducted based on independent and identically distributed evaluations of voters on alternatives. They lead to Impartial Cultures (IC) models on rankings. This framework can be enlarged to non IC models. McCabe-Dansted and Slinko (2006) use for example a Polya urn model to extend to non IC models. With evaluation based simulations, a perspective is to explore two directions:

- Special case 1: What if the evaluations distributions are not independent?
- Special case 2: What if the alternatives' evaluations distributions are not identically distributed?

### 5.1 Special case 1: correlated scores

In the following, we consider  $n = 101$  voters, and  $m = 4$  alternatives. Let us first consider a case where the evaluations for the alternatives are correlated. Denote  $E_A, E_B, E_C$  and  $E_D$  the evaluations of four alternatives. We consider a framework where  $E_A, E_B, E_C$  and  $E_D$  have uniform marginal distributions

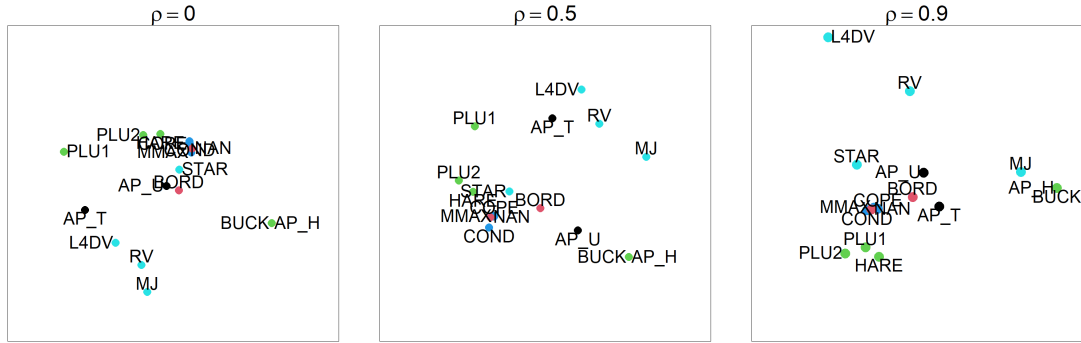


Figure 12: Multidimensional scaling representation of distance between voting rules for correlated data. On the left  $\rho = 0$ , on middle  $\rho = 0.5$  and on the right  $\rho = 0.9$ . The goodness-of-fit are respectively 0.46, 0.46, and 0.51.

and  $Cor(E_A, E_B) = Cor(E_C, E_D) = \rho$ ;  $Cor(E_A, E_C) = Cor(E_B, E_C) = Cor(E_A, E_D) = Cor(E_B, E_D) = -\rho$ , with  $\rho \in \{0, 0.5, 0.9\}$ .

The resulting MDS is displayed in Figure 12. Introducing correlations changes the similarities between voting procedures. We still recover a Condorcet like cluster, with MMAX, NAN, BORD and COND gathered. But STAR is further away the cluster under correlation. PLU2, PLU1 and HARE also gathered together and are nearer to the Condorcet family as the correlation increases. Evaluation based procedures, L4DV, MJ, RV, moved away from each others with correlation.

## 5.2 Special case 2: consensus or polarized alternative

Another extension is to consider voters with different and independent profiles of evaluations. We introduce three distributions, respectively corresponding to a neutral alternative, to a consensual alternative and to a polarizing alternative. The distributions are displayed in Figure 13.

We simulate three scenarios:

- Scenario 1: all three alternatives are neutral. This corresponds to the framework of Section 3 and Section 4.
- Scenario 2: two alternatives are neutral and one is consensual. In this case, the consensual alternative tends to have better evaluations.

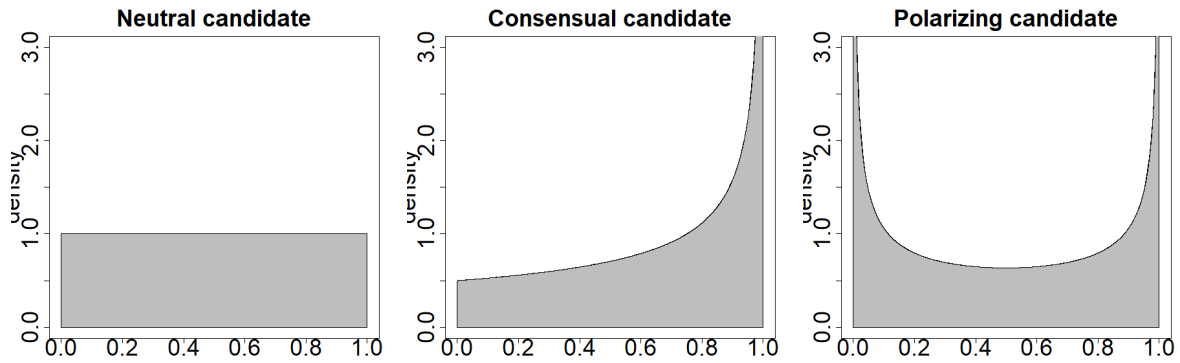


Figure 13: Neutral, consensual and polarized distributions of evaluations. Probability distribution functions of evaluations on one alternative: Uniform distribution (left), Beta distribution with  $\alpha = 1$  and  $\beta = 0.5$  (center) and Beta distribution with  $\alpha = 0.5$  and  $\beta = 0.5$  (right).

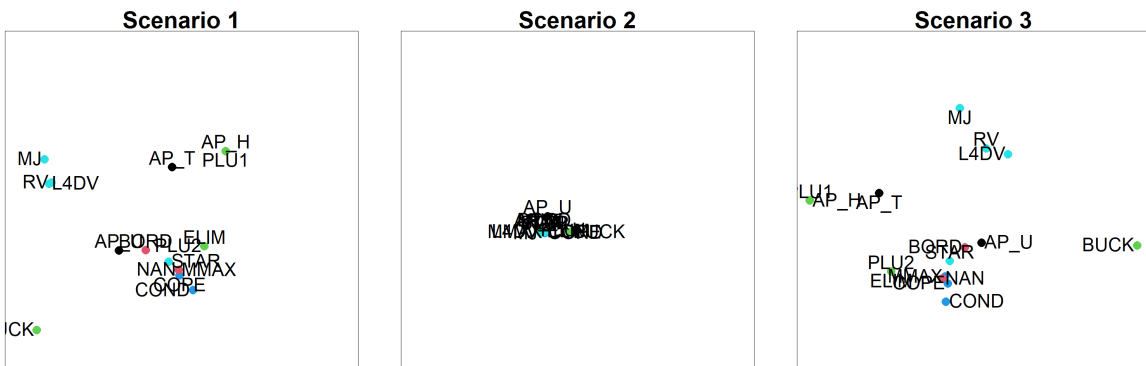


Figure 14: MDS for Scenario 1, Scenario 2 and Scenario 3. The three graphs have the same scale. The goodness-of-fit are respectively 0.60, 0.85, and 0.65.

- Scenario 3: two alternatives are neutral and one is polarizing. In this case, half of the voters may give a lower evaluation to the polarizing alternative and half may give a higher evaluation.

MDS for the three scenarios are given in Figure 14. They show that in Scenario 2, when there is a consensual alternative, the distance between the voting rules is very low. The procedures indeed globally agree to choose the consensual alternative. Next, no main differences can be seen between Scenario 1 and Scenario 3.



## Conclusion

This article investigates the empirical closeness of voting procedures on simulated data. An originality of our work is to consider evaluation-based methods, together with ranking-based ones. Our main study deals with independent and identically distributed evaluations, which yields Impartial Culture on rankings. We recover empirically the families of voting rules – Condorcet like, evaluation-based, plurality and approval voting, Bucklin. Additionally, our study gives an insight of the characteristics of the winners, depending of the voting procedure used. We show for example that (one-round or two-round) plurality voting selects winners with lower mean and higher variance (on rankings and on evaluations) than other procedures. We also highlight that the number of alternatives influences the proximity between voting rules, while it is not the case for the number of voters.

Future works may extend the study to non IC frameworks for ranking-based procedures, respectively not considering IID frameworks for evaluations. We already show in this article in simulated examples that releasing either assumptions of independence or of equal distributions modify the proximities between voting rules. Studying more extensively how releasing IC structure modify the voting rules, by introducing dependent and different distributions, is an interesting perspective.

## Acknowledgment

The authors thanks Nathan Grimault for his help on an earlier version of this work.

## References

- Aubin, J.-B., Gannaz, I., Leoni, S., and Rolland, A. (2022). Deepest voting: A new way of electing. *Mathematical Social Sciences*, 116:1–16.
- Balinski, M. and Laraki, R. (2010). *Majority Judgment: Measuring, Ranking, and Electing*. The MIT Press.
- Boehmer, N., Faliszewski, P., Janeczko, L., Kaczmarczyk, A., Lisowski, G., Pierczyński, G., Rey, S., Stolicki, D., Szufa, S., and Wąs, T. (2024). Guide to numerical experiments on elections in computational social choice. arXiv preprint arXiv:2402.11765.

- Conitzer, V. and Sandholm, T. (2012). Common voting rules as maximum likelihood estimators. *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence, UAI 2005*.
- De Almeida, A. T., Morais, D. C., and Nurmi, H. (2019). *Systems, procedures and voting rules in context: A primer for voting rule selection*, volume 9. Springer.
- Durand, F. (2015). *Towards less manipulable voting systems*. Ph.d., Université Pierre et Marie Curie - Paris VI.
- Felsenthal, D. S. and Machover, M. (2012). *Electoral Systems: Paradoxes, Assumptions, and Procedures*. Springer.
- Gehrlein, W. V. (1999). Approximating the probability that a condorcet winner exists. In *Proceedings of National Decision Sciences Institute*, pages 626–628.
- Green-Armytage, J., Tideman, T., and Cosman, R. (2016). Statistical evaluation of voting rules. *Social Choice and Welfare*, 46(1):183–212.
- Kacprzyk, J., Merigó, J. M., Nurmi, H., and Zadrozny, S. (2020). Multi-agent systems and voting: How similar are voting procedures. In *IPMU 2020, CCIS 1237*, pages 172–184.
- Mattei, N., Forshee, J., and Goldsmith, J. (2012). An empirical study of voting rules and manipulation with large datasets. In *Proceedings of COMSOC*, volume 59. Citeseer.
- McCabe-Dansted, J. and Slinko, A. (2006). Exploratory analysis of similarities between social choice rules. *Group Decision and Negotiation*, 15:77–107.
- Merlin, V., Tataru, M., and Valognes, F. (2000). On the probability the all decision rules select the same winner. *Journal of Mathematical Economics*, 33:183–207.
- Nurmi, H. (1987). *Comparing Voting Systems*. D. Reidel Publishing Company.
- Richelson, J. T. (1979). A comparative-analysis of social choice functions, I, II, III: A summary. *Behavioral Science*, 24(5):355–355.
- Rolland, A., Aubin, J.-B., Irène, G., and Leoni, S. (2024). Probabilistic models of profiles for voting by evaluation. *Social Choice and Welfare*. Forthcoming.
- Song, C. G. (2023). What difference does a voting rule make? *Constitutional Political Economy*, 34(3):275–285.
- Tideman, N. (2023). Choosing among the variety of proposed voting reforms. *Constitutional Political Economy*, 34(3):471–481.

Tideman, N. and Plassmann, F. (2013). Developing the aggregate empirical side of computational social choice. *Annals of Mathematics and Artificial Intelligence*, 68:31–64.