The Thirty-Ninth AAAI Conference on Artificial Intelligence (AAAI-25)

Distances Between Top-Truncated Elections of Different Sizes

Piotr Faliszewski[1], **Jitka Mertlová**[2], **Pierre Nunn**[3], **Stanisław Szufa**[4], **Tomasz Wąs**[5]
[1]AGH University
[2]Czech Technical University in Prague
[3]Université de Rennes
[4]CNRS, LAMSADE, Université Paris Dauphine-PSL
[5]University of Oxford
faliszew@agh.edu.pl, mertlova.jita@gmail.com, pierre.nunn@ens-rennes.fr, s.szufa@gmail.com, tomasz.was@cs.ox.ac.uk

Abstract

The map of elections framework is a methodology for visualizing and analyzing election datasets. So far, the framework was restricted to elections that have equal numbers of candidates, equal numbers of voters, and where all the (ordinal) votes rank all the candidates. We extend it to the case of elections of different sizes, where the votes can be top-truncated. We use our results to present a visualization of a large fragment of the Preflib database.

Code — https://github.com/Project-PRAGMA/Map-Different-Sizes-AAAI-2025

1 Introduction

The map of elections framework is a methodology for analyzing and visualizing sets of elections, introduced by Szufa et al. (2020) and Boehmer et al. (2021). The basic idea is to depict (ordinal) elections as points on a plane, so that the closer two given points are, the more similar are their corresponding elections (see Figures 1a and 1b for examples, and Section 2 for definitions). To measure this similarity, one can either use the accurate but computationally challenging isomorphic swap distance (Faliszewski et al. 2019), or the less precise but efficiently computable positionwise distance (Szufa et al. 2020), or whatever other distance that is invariant to renaming the candidates and voters. Indeed, the elections may be unrelated to each other and we only care about their structural similarities. Unfortunately, the distances used so far are restricted to elections with the same numbers of candidates and voters, and require the votes to be complete, i.e., to rank all the candidates. We aim to rectify these two issues.

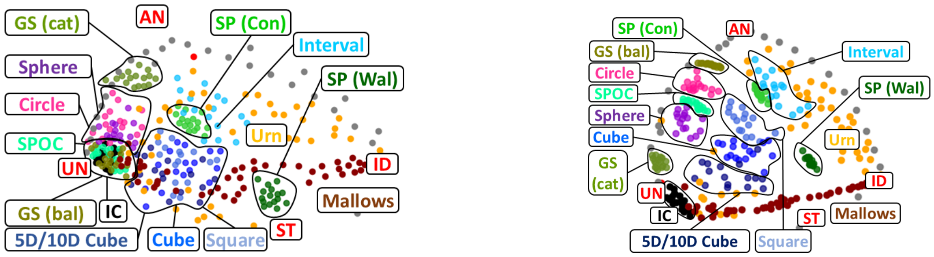
The original motivation behind the map framework was to better understand relations between various statistical cultures (i.e., models of generating random elections) and to present experimental results in a nonaggregate way (Szufa et al. 2020). For this, the restriction to particular election sizes and complete votes is natural, as we have full control over the data. However, maps are also useful for studying real-life elections. For example, we can get better insight into the nature of real-life elections by analyzing their positions on the map (Boehmer et al. 2021; Boehmer and Schaar 2023), or check which statistical cultures yield similar elections (Boehmer et al. 2021, 2022a). Yet, real-life data poses some issues, as seen in the next two datasets from Preflib (Mattei and Walsh 2013):

Irish General Elections. Preflib includes the ordinal votes cast in several constituencies in Ireland during the 2002 general election. Naturally, each of them had a different number of candidates (between 9 and 14) and a different number of voters (between around 30 000 and 64 000), most of whom ranked some top candidates only.

Formula 1. Here, each election is a Formula 1 season, where each vote is a race ranking the drivers (i.e., the candidates) in the order in which they finished. There are around 20 races and 20 drivers in each season, but these numbers vary, and not all drivers complete each race.

So far, to put such elections on a map, one had to fill in the missing parts of the preference orders, restrict the number of candidates to some common value (e.g., by deleting the worst-performing ones), and sample a fixed-size set of votes. We propose distances that can deal with different-sized/top-truncated elections natively, without preprocessing.

Ideally, we would like to extend the isomorphic swap and positionwise distances so that different-sized elections with "obviously identical" structure would be at distance zero. Unfortunately, this turns out to be impossible even for a rather minimal definition of an "obviously identical" structure (even without thinking of top-truncated votes). However, we do find two intuitively appealing distances that fulfill some of our requirements: One is an extension of the positionwise distance and one, which we call DAP, is based on analyzing diversity, agreement, and polarization within votes. In the spirit of our desiderata, DAP turns out to be strongly correlated with the isomorphic swap distance and, so, can serve as its replacement. We first describe our distances and then analyze their properties, both theoretically and experimentally. In the experiments, we largely focus on synthetic elections, as such data is easier to control, but our main motivation is to visualize real-life data. Hence, in Section 5 we form a map of a large fragment of Preflib, a database of real-life elections (see Figure 4). The overarching goal of the paper is to explain how we obtained this map, why the choices made on the way are justified, and what we can learn from it.



(a) Positionwise Distance.

(b) Swap Distance.

Figure 1: Maps of elections created using the (a) isomorphic swap and (b) positionwise distances. Each point corresponds to an election; the color represents the statistical culture it comes from: ID, UN, and AN refer to identity, uniformity, and antagonism elections; IC, Mallows, and urn mean impartial culture, the normalized Mallows model, and the Pólya-Eggenberger urn model, respectively; Interval, Square, Cube, 5D/10D Cube, Circle, and Sphere refer to Euclidean models; SP stands for single-peaked (in the Conitzer or Walsh models), SPOC for single-peaked on a circle, and GS for group-separable (caterpillar or balanced).

2 Preliminaries

For a positive integer k, we let $[k] = \{1, \dots, k\}$, and we write $\mathbb{R}_+$ to denote the set of nonnegative real numbers. For two equal-sized sets A and B, $\Pi(A, B)$ is the set of bijections between A and B.

Elections. An *election* is a pair $E = (C, V)$, where C is a set of m *candidates* and V is a collection of n *votes*, i.e., weak orders over C; we will also sometimes refer to members of V as *voters* (i.e., the agents who cast the respective votes). By a *size* of an election, we mean the pair (m, n). For two candidates $a, b \in C$ and voter $v \in V$, we write $a \succ_v b$ if v strictly prefers a over b and $a \sim_v b$ if v is indifferent between them. We write $a \succeq_v b$ if $a \succ_v b$ or $a \sim_v b$. By $N_E(a \succ b)$ we mean the number of voters in election E that strictly prefer a to b, and by $N_E(a \sim b)$ we mean the number of those indifferent between these candidates.

Consider vote v over candidate set C. We say that v is *complete* if it is a strict linear order over C, and that it is *top-truncated* if C can be partitioned into sets $C_v^{\uparrow}$ and $C_v^{\downarrow}$, so that v is a strict linear order over $C_v^{\uparrow}$ and for every $c \in C_v^{\uparrow}$ and $a, b \in C_v^{\downarrow}$ we have $c \succ_v a \sim_v b$. $C_v^{\uparrow}$ is the top part of v and $C_v^{\downarrow}$ is its truncated part (so a complete vote is top-truncated with empty truncated part.) An election is complete (top-truncated) if all its votes are complete (top-truncated).

Top-truncated elections are quite common in real-life election data (see e.g., *Preflib* database (Mattei and Walsh 2013)) and have been extensively studied in terms of their prevalence (Kilgour, Grégoire, and Foley 2020), efficiency (Borodin et al. 2022), effect on voting rules (Tomlinson, Ugander, and Kleinberg 2023), or a possibility of manipulation (Baumeister et al. 2012).

Special Elections. Following Boehmer et al. (2021), we consider the following families of characteristic (complete) elections (we consider m candidates and n voters): (i) In an *identity election* ($\mathrm{ID}_{m,n}$) all n votes are identical; (ii) an *antagonism election* ($\mathrm{AN}_{m,n}$) is a concatenation of two identity ones, where the voters in one have opposite preference orders to those in the other; and (iii) in a *uniformity election* ($\mathrm{UN}_{m,n}$) all possible votes appear equal number of times. The number of voters must be even in an antagonism election and must be a multiple of $m!$ in a uniformity one. We drop the number of voters from notation when it is irrelevant.

Distances. A *pseudodistance* over a set U is a function $d\colon U \times U \to \mathbb{R}_+$ such that for each $a, b, c \in U$ we have (1) $d(a, a) = 0$, (2) $d(a, b) = d(b, a)$, and (3) $d(a, b) + d(b, c) \geq d(a, c)$. We often drop "pseudo" prefix.

(Isomorphic) Swap Distance. Given two votes, u, v, over the same candidate set, we define their swap distance as $\mathrm{swap}(u, v) = \frac{1}{2}\sum_{a,b \in C}([a \succ_u b \wedge b \succeq_v a] + [a \succeq_u b \wedge b \succ_v a])$, where the expressions in square brackets evaluate to 1 when they are true and to 0 otherwise. That is, we add 1 for each pair of candidates for which both voters have a strict, opposite preference, and we add $1/2$ for each pair of candidates for which exactly one of the voters is indifferent.

For a pair of equal-sized/complete elections $E = (C, V)$ and $F = (D, U)$, such that $|C| = |D| = m$, $V = \{v_1, \dots, v_n\}$ and $U = \{u_1, \dots, u_n\}$, their *(isomorphic) swap distance* is the sum of the swap distances between the votes, under optimal matchings of the candidates and voters. Formally, $d_{\mathrm{swap}}(E, F)$ is equal to:

$$\min_{\substack{\pi \in \Pi([n],[n]) \\ \sigma \in \Pi(C,D)}} \textstyle\sum_{i \in [n]} \mathrm{swap}(\sigma(v_i), u_{\pi(i)}) / \left(\frac{1}{4}n(m^2 - m)\right),$$

where $\sigma(v_i)$ means vote v_i with each candidate $c \in C$ replaced with $\sigma(c) \in D$. By definition, d_{swap} is invariant to renaming the candidates and reordering the voters, and its value is zero if and only if the given elections are isomorphic (i.e., are identical up to renaming the candidates and reordering the voters). This distance was introduced by Faliszewski et al. (2019). Dividing by $\frac{1}{4}n(m^2 - m)$ ensures normalization of the largest distance to exactly 1 (for the case where n is even (Boehmer et al. 2022b)).

Wasserstein Distance. Given a vector $\vec{a} = (a_1, \dots, a_m) \in \mathbb{R}_+^m$, we identify it with a stepwise function $a\colon [0, 1] \to \mathbb{R}_+$, such that $a(0) = 0$ and for all $i \in [m]$ and x in $(\frac{i-1}{m}, \frac{i}{m}]$ we have $a(x) = (m \cdot a_i)/(a_1 + \cdots + a_m)$. Further, for each $x \in [0, 1]$ we set $A(x) = \int_0^x a(y)dy$. Note that the normalization in function $a(\cdot)$ ensures that $A(1) = 1$. For two vectors $\vec{a} \in \mathbb{R}^s$ and $\vec{b} \in \mathbb{R}^t$ of possibly different dimensions, we define their *Wasserstein distance* as $W(\vec{a}, \vec{b}) = \int_0^1 |A(x) - B(x)| dx$. As $a(x)$ and $b(x)$ are stepwise and $A(x)$ and $B(x)$ are piecewise linear, $W(\vec{a}, \vec{b})$ can be computed in polynomial time (assuming we can perform arithmetic operations in polynomial time).

Frequency Matrices and Positionwise Distance. Consider an election $E = (C, V)$, where $C = \{c_1, \dots, c_m\}$ and $V = (v_1, \dots, v_n)$. A *frequency matrix* of a top-truncated vote $v \in V$ is an $m \times m$ matrix $\mathrm{freq}(v)$ where the entry in the i-th row and j-th column is:

$$\mathrm{freq}(v)_{i,j} = \begin{cases} 1 & \text{if } c_j \in C_v^{\uparrow} \text{ and } v \text{ ranks it } i\text{-th there,} \\ 1/|C_v^{\downarrow}| & \text{if } c_j \in C_v^{\downarrow} \text{ and } i > |C^{\uparrow}|, \\ 0 & \text{otherwise.} \end{cases}$$

A frequency matrix of election E is the average of the frequency matrices of its votes, i.e., $\mathrm{freq}(E) = {}^1/_n \sum_{v_i \in V} \mathrm{freq}(v_i)$. Intuitively, $\mathrm{freq}(E)_{i,j}$ is the probability that in a randomly selected vote from E the j-th candidate is ranked in the i-th position (where being in the truncated part of a vote means having equal chance of being ranked in any of the truncated positions). Frequency matrices are normalized variants of *position matrices* of Szufa et al. (2020); see also the works of Boehmer et al. (2022a, 2023) (these authors did not consider top-truncated votes; we added this extension). Frequency matrices are bistochastic, i.e., each of their columns and each of their rows sums up to 1.

The Wasserstein distance between two bistochastic matrices, $X = [\vec{x}_1, \dots, \vec{x}_m]$ and $Y = [\vec{y}_1, \dots, \vec{y}_m]$, with m column vectors each, is the sum of the Wasserstein distances between their (optimally matched) column vectors. Formally:

$$d_W(X, Y) = \min_{\sigma \in \Pi([m],[m])} \tfrac{1}{m} \textstyle\sum_{i \in [m]} W(\vec{x}_i, \vec{y}_{\sigma(i)}).$$

Then, the positionwise distance of two equal-sized elections, E and F, is the Wasserstein distance of their frequency matrices, $d_{\mathrm{pos}}(E, F) = d_W(\mathrm{freq}(E), \mathrm{freq}(F))$. This distance was originally defined by Szufa et al. (2020) using the classic earth mover's distance (EMD) of Rubner, Tomasi, and Guibas (2000), defined for vectors of the same dimension that sum up to the same value. We prefer the Wasserstein distance because it applies to vectors of different dimensions, while being closely related to EMD (in the proposition below, $\mathrm{emd}(\vec{a}, \vec{b})$ is the EMD distance between vectors and $\vec{a}$ and $\vec{b}$, and is known to be between 0 and $m-1$; hence $mW(\vec{a}, \vec{b})$ is its good approximation both for very similar and quite far-off vectors).

Proposition 2.1. *For each two m-dimensional vectors $\vec{a}$ and $\vec{b}$ with nonnegative entries that sum up to 1, it holds that* $\max(\frac{1}{2}\mathrm{emd}(\vec{a}, \vec{b}), \mathrm{emd}(\vec{a}, \vec{b}) - 1) \leq m W(\vec{a}, \vec{b}) \leq \mathrm{emd}(\vec{a}, \vec{b})$.

Like the isomorphic swap distance, the positionwise distance is invariant to renaming the candidates and reordering the voters, but it differs in that it can assume value 0 even for pairs of nonisomorphic elections. Positionwise distance is normalized in the sense that its largest value is about $1/3$ (we omit exact calculations).

Maps of Elections. A map of elections is a set of elections with distances between each pair of them (called *original distances*). We associate each election with a point in a 2D space so that the Euclidean distances between the points resemble the original ones, using either multidimensional scaling (MDS (Kruskal 1964)) or Kamada-Kawai (KK (Kamada and Kawai 1989)) embedding algorithms.

3 Search for More Versatile Distances

Our first goal is to find a satisfying distance $\hat{d}$ over different-sized, complete elections (as a convention, we will be marking distances that can handle different-sized elections with a "hat" on top). Ideally, we would like $\hat{d}$ to be an extension of an appealing distance among equal-sized elections, and to put "obviously" identical elections at distance 0. Below we express these requirements formally:

1. We say that $\hat{d}$ is a *swap extension* if it is a distance and for each two equal-sized/complete elections E_1 and E_2 with n voters and m candidates, $\hat{d}(E_1, E_2) = d_{\mathrm{swap}}(E_1, E_2)$. We define a *positionwise extension* analogously.
2. We say that $\hat{d}$ is ID-*consistent* if for all $p, r, s, q \in \mathbb{N}$, $p, r \geq 3$, we have $\hat{d}(\mathrm{ID}_{p,q}, \mathrm{ID}_{r,s}) = 0$. We define AN-*consistency* and UN-*consistency* analogously. (We consider at least 3 candidates as $\mathrm{AN}_{2,2} = \mathrm{UN}_{2,2}$, so without this restriction assuming both AN- and UN-consistency would put all AN and UN elections at distance zero.)

We would like a swap extension or a positionwise extension that is ID-, AN-, and UN-consistent. Naturally, one could think of further conditions, but these are both natural and quite minimal. Yet, even this is too much to ask for.

Proposition 3.1. *There is no swap extension nor positionwise extension that simultaneously satisfies* ID-, AN- *and* UN-*consistency.*

This means that either we have to give up on at least one of the consistency properties, or on looking for swap/positionwise extensions. We explore both of these possibilities.

3.1 Intuitive Swap/Positionwise Extensions

As far as swap extensions go, one natural idea is that given two elections $E_1 = (C_1, V_1)$ and $E_2 = (C_2, V_2)$, where $|C_1| < |C_2|$, we define $\hat{d}^{\mathrm{tr}}_{\mathrm{swap}}(E_1, E_2)$ as $d_{\mathrm{swap}}(E'_1, E_2)$, where E'_1 is equal to E_1 with additional $|C_2| - |C_1|$ candidates that all voters have in their truncated parts. Unfortunately, this leads to maps where elections with fewer candidates are clustered together, so we lose significant amount of information about them.

Proposition 3.2. $\hat{d}^{\mathrm{tr}}_{\mathrm{swap}}$ *is a swap extension that is neither* ID- *nor* AN- *nor* UN-*consistent.*

Alternatively, we could define $\hat{d}^{\mathrm{del}}_{\mathrm{swap}}(E_1, E_2)$ to be equal to the expected isomorphic swap distance of elections E_1 and E'_2, where E'_2 is obtained by deleting $|C_2| - |C_1|$ candidates from E_2 uniformly at random. However, this is not even a distance as it fails the triangle inequality.

Proposition 3.3. $\hat{d}^{\mathrm{del}}_{\mathrm{swap}}$ *fails triangle inequality.*

One could hope that violations of the triangle inequality are rare and, hence, could be ignored. Unfortunately, based on the experiments, this is not the case, and violations are common. All in all, we were not able to find a satisfying swap extension and we leave looking for one as an open problem. On the positive side, we do identify a natural, UN-consistent positionwise extension that not only applies to different-sized elections, but also seamlessly handles top-truncated votes.

For a matrix $X = [\vec{x}_1, \dots, \vec{x}_m]$ and an integer $k \geq 1$, by $\mathrm{str}_{mk}(X)$ we denote the matrix obtained from X by *stretching*, i.e., copying each of its component vectors k times, so that for each $i \in [mk]$, the i-th vector of $\mathrm{str}_{mk}(X)$ is $\vec{x}_{\lceil i/k \rceil}$. Then, we define the positionwise distance between elections E and F—with possibly different numbers of candidates and voters—as $\hat{d}_{\mathrm{pos}}(E, F) = d_W(\mathrm{str}_s(\mathrm{freq}(E)), \mathrm{str}_s(\mathrm{freq}(F)))$, where s is the least common multiple of the numbers of candidates in E and F. That is, to obtain $\hat{d}_{\mathrm{pos}}(E, F)$ we first compute the frequency matrices of E and F, then we duplicate their vectors so that both matrices end up with an equal number of columns, and, finally, we match these columns and sum up their Wasserstein distances (recall the definition of the positionwise distance from Section 2; note that here the vectors may be of different dimensions, which is why we chose to use the Wasserstein distance instead of EMD). Truncated votes are handled seamlessly because they are encoded in the frequency matrices. We show that $\hat{d}_{\mathrm{pos}}$ is indeed a positionwise extension and that it is UN-consistent.

Theorem 3.4. *$\hat{d}_{\mathrm{pos}}$ is a positionwise extension that is* UN-*consistent, but not* ID- *nor* AN-*consistent.*

Note that Proposition 3.1 does not not preclude the existence of a positionwise extension satisfying UN-consistency and either ID- or AN-consistency. Thus, our positionwise distance is not "optimal" in terms of the axioms it satisfies.

3.2 Feature Distance and ID/AN/UN-Consistency

In principle, Proposition 3.1 might hold simply because no distance is simultaneously ID-, AN-, and UN-consistent, irrespective of being a swap or a positionwise extensions. We show that this is not the case and at least one such distance exists. To this end, let $f = (f_1, \dots, f_k)$ be a collection of *features*, i.e., functions that given an election output a value between 0 and 1. For an election E, its feature vector is $f(E) = (f_1(E), \dots, f_k(E))$, and the *feature distance* $\hat{d}_f$ between elections E and F is $d_f(E, F) = \ell_2(f(E), f(F))$; naturally, one could also use ℓ_1 or some other distances.[1] For every collection of features f, $\hat{d}_f$ is indeed a distance over different-sized elections, and if the features are defined for top-truncated elections, so is $\hat{d}_f$.

Let $\mathrm{id}(E)$, $\mathrm{an}(E)$, and $\mathrm{un}(E)$ be three features such that their value is 0 if the input election is isomorphic to, respectively, some ID, AN, or UN election, and it is 1 otherwise.

[1]For two vectors, $\vec{a} = (a_1, \dots, a_k)$ and $\vec{b} = (b_1, \dots, b_k)$, we define $\ell_2(\vec{a}, \vec{b}) = \left(\sum_{i=1}^{k}(a_i - b_i)^2\right)^{1/2}$.

Proposition 3.5. *The distance defined by collection* $(\mathrm{id}, \mathrm{an}, \mathrm{un})$ *of features is* ID-, AN-, *and* UN-*consistent.*

While this distance has all the desired consistency properties, it puts all elections that are not isomorphic to ID, AN, or UN at distance zero and, hence, is not very useful. Next we develop a feature distance that implements a similar idea, but in a more sophisticated way. To this end, we focus on evaluating diversity, agreement, and polarization among the votes (indeed, UN captures ideal diversity, ID ideal agreement, and AN—polarization). Let $E = (C, V)$ be an election. The agreement for candidates $a, b \in C$ is:

$\alpha(a, b) = \max(|N_E(a \succ b) - N_E(b \succ a)|, N_E(a \sim b))/|V|$.

Intuitively, $|N_E(a \succ b) - N_E(b \succ a)|/|V|$ captures the agreement when voters lean toward strict preference over a and b, and $N_E(a \sim b)/|V|$ reflects the agreement when most of the voters put both a and b in the truncated parts of their votes. The agreement index of an election $E = (C, V)$ is an average of agreements for all pairs of candidates, i.e., $A(E) = \sum_{\{a,b\} \subseteq C} \alpha(a, b)/\binom{|C|}{2}$. This definition is a natural extension of the agreement index studied for complete elections by Alcalde-Unzu and Vorsatz (2013), Hashemi and Endriss (2014), and Can, Ozkes, and Storcken (2015). Note that these papers use other names and interpretations for this index—we follow the approach of Faliszewski et al. (2023).

Regarding diversity and polarization, we also follow Faliszewski et al. (2023), but with a few changes. For each integer i, let the empirical i-Kemeny score of election E be:

$\mathrm{emk}_i(E) = \min_{v_1, \dots, v_i \in V}\left(\sum_{v \in V}\left(\min_{j \in [i]} \mathrm{swap}(v, v_j)\right)\right)$.

That is, to compute $\mathrm{emk}_i(E)$, we seek i votes from the election so that the sum of the swap distances of each vote in E to the closest selected one is minimized. E.g., $\mathrm{emk}_1(E)$ is an approximation of the classic Kemeny score (Kemeny 1959). To normalize $\mathrm{emk}_i(E)$, we divide by its maximal possible value, i.e., $1/2 \cdot |V| \cdot \binom{|C|}{2}$. We define the diversity and polarization indices as:

$$D(E) = \tfrac{2}{5}\textstyle\sum_{i=1}^{5} \mathrm{emk}_i(E)/(|V| \cdot \binom{|C|}{2}),$$

$$P(E) = 2 \cdot (\mathrm{emk}_1(E) - \mathrm{emk}_2(E))/(|V| \cdot \binom{|C|}{2}).$$

The latter is simply an approximation of the polarization index introduced by Faliszewski et al. (2023) and the former is a heuristic built on top of their diversity index; we chose the constant 5 as our initial experiments have shown that it captures the same notion as their approach, and it allows for fast computation; in general, computing $\mathrm{emk}_i(E)$ is intractable (Faliszewski et al. 2023), so we approximate it using their local search approach.

We prove that for every i, the value of the empirical i-Kemeny score of the uniformity election converges to its maximal value as the number of candidates grows.

Proposition 3.6. *For every constant $i \in \mathbb{N}$, it holds that* $\lim_{m \to \infty} \mathrm{emk}_i(\mathrm{UN}_{m,m!}) = 1/2 \cdot m! \cdot \binom{m}{2}$.

This gives the following sanity check (arrows mean convergence as the number of candidates grows):

$$\begin{aligned} &D(\mathrm{ID}) = 0, && A(\mathrm{ID}) = 1, && P(\mathrm{ID}) = 0, \\ &D(\mathrm{UN}) \to 1, && A(\mathrm{UN}) = 0, && P(\mathrm{UN}) \to 0, \\ &D(\mathrm{AN}) = 1/5, && A(\mathrm{AN}) = 0, && P(\mathrm{AN}) = 1. \end{aligned}$$

These results are intuitive as in ID all the voters agree, UN is most diverse, and AN is most polarized. The DAP distance, denoted $\hat{d}_{\mathrm{dap}}$, is a feature distance that uses D, A, and P as the features. We see that DAP is ID-consistent and AN-consistent, but not UN-consistent (but it satisfies this property in an approximate sense—the distance between two UN elections with different numbers of candidates gets smaller and smaller as the numbers of candidates in these elections increase).

Proposition 3.7. $\hat{d}_{\mathrm{dap}}$ *is* ID-*consistent and* AN-*consistent, but is not* UN-*consistent.*

While our diversity feature may appear somewhat ad hoc, we believe that it captures the intuitive notion of diversity among votes and our results are robust to tweaking it. Indeed, in Section 4.2 we find strong correlation between the DAP distance and the swap one (which implicitly relies on diversity analysis (Faliszewski et al. 2023)).

3.3 Positionwise Distance Versus DAP

Positionwise distance and DAP vary in several significant ways. Foremost, the positionwise distance deals with different numbers of candidates by, effectively, creating their virtual copies, so that the elections it analyzes look as if they were equal-sized (this process is hidden in computing the Wasserstein distance and in stretching the matrices). On the other hand, DAP identifies structural properties of the elections (diversity, agreement, and polarization) and rescales them to a common denominator, so that different-sized elections can be compared on common grounds.

Second, the two distances treat top-truncated votes differently. DAP inherits its approach from the swap distance that its features are based on (the agreement index is similar in this respect): If a voter puts some candidates in the truncated part, then DAP assumes that he or she sees them as equally bad. In contrast, the positionwise distance assumes that the voter can rank them, but chose not to report it, so the distance assumes a uniform distribution over possible completions.

4 Maps of Synthetic Datasets and Evaluation

Our next goal is to understand how the positionwise and DAP distances behave in practice. To this end, we form and analyze maps of synthetic elections, and we evaluate how the distances between elections change as we vary either their size or their level of truncation. First, we describe our data.

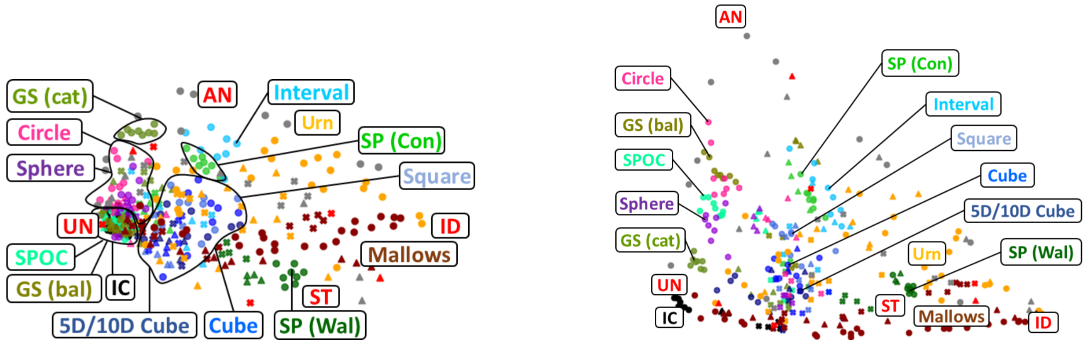
4.1 Datasets

Our *basic dataset* consists of 326 complete elections with 8 candidates and 96 voters each, and is nearly identical to the one used by Faliszewski et al. (2023) (in particular, we chose the same numbers of candidates and voters as they did). The main part of the dataset consists of elections generated according to the impartial culture (IC), normalized Mallows (Mallows 1957; Boehmer et al. 2021), Pólya-Eggenberger urn (Berg 1985; McCabe-Dansted, Pritchard, and Slinko 2008), and Euclidean models (see, e.g., the work of Enelow and Hinich (1984)). Under impartial culture, we draw each vote uniformly at random. The normalized Mallows model is similar but the votes are clustered around a given central one (the strength of this clustering is controlled by parameter norm-$\phi \in [0, 1]$, the higher the value the less concentrated are the votes; see the work of Boehmer, Faliszewski, and Kraiczy (2023) for a discussion of this model). The urn model generates elections with clusters of identical votes (the larger its parameter of contagion $\alpha \geq 0$ is, the fewer clusters there are, each containing more votes (Faliszewski et al. 2023)). In the Euclidean models, candidates and voters are points in some Euclidean space and voters prefer the closer candidates (we draw the points uniformly from a unit hypercube or hypersphere of a given dimension; for 1-, 2-, and 3- dimensional hypercubes we refer to the models as Interval, Square, and Cube; for 1- and 2-dimensional hyperspheres we refer to them as Circle and Sphere). The dataset also includes elections generated using statistical cultures that yield single-peaked (Black 1958), SPOC (Peters and Lackner 2020), and group-separable (Inada 1964, 1969) elections. Finally, we add ID, AN, and an approximation of UN elections together with artificial elections forming paths between these three on our maps (which include ST election in which each voter ranks the same half of the candidates on top, but otherwise the votes are chosen uniformly at random).

We generated the *size-oriented dataset* in the same way as the basic one, except that for each culture we partitioned its elections into four groups, with either 8 or 16 candidates and either 96 or 192 voters. We obtain top-truncated elections from complete ones by using the following methods:

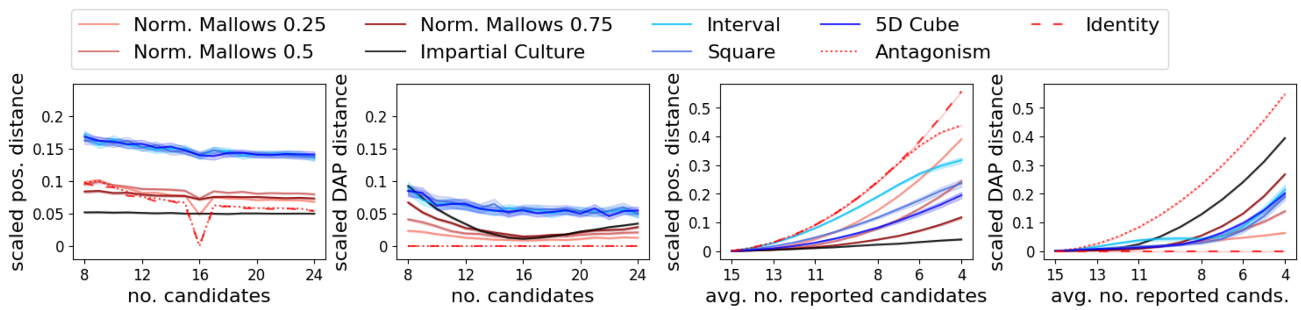
1. Top-k truncation removes from each vote the candidates below position k. Such data appears, e.g., in Preflib in the sushi dataset, where people rank their top 10 sushi types out of 100 available ones (Kamishima 2003) (see Preflib file `00014-00000002.soi`).
2. Random cut truncation is parameterized by probability p. For each vote, we consider its candidates from top to bottom and with probability $1 - p$ we stop the process, truncating the vote after the current candidate. Some of the political elections follow a similar pattern, e.g., UK Labour Party leadership election (Riley, Ryan, and Smith 2010) (see Preflib file `00030-00000001.soi`).
3. Random drop truncation moves each candidate in each vote to the truncated part, independently, with probability p. This imitates sport elections—such as those for Formula 1—where each player fails to finish a given competition with some probability.

To form the *comprehensive dataset*, we took the size-oriented one and for each group of elections (of a given size, generated using a given statistical culture) we left the first half of the elections in the group intact, we applied the top-k truncation to the next quarter of them, and we applied random cut truncation to the last quarter. We chose the truncation parameters so that, in expectation, each voter ranked half of the candidates. We generated the *truncation-oriented dataset* in the same way, but starting from the basic dataset. We also generated a *random drop dataset*, but we omitted it from the experiments as its elections have very different nature (we analyze them in the extended version of the paper).



(a) Positionwise Distance Map. (b) DAP Distance Map.

Figure 2: Maps of elections created using the positionwise and DAP distances, for the truncation-oriented datasets. Top-k truncated elections are marked with triangles, random-cut truncated ones with crosses, and complete ones with circles.



(a) Distances between complete elections of different sizes. (b) Distances between elections and their top-k truncations.

Figure 3: Plots (a) show average distance from size-16 elections to different-sized, complete elections from the same culture, as a fraction of the maximum distance in our dataset (positionwise distance on the left, DAP distance on the right). Plots (b) show average distance from a complete election to its top-k truncation, as a fraction of the maximum distance in our dataset (positionwise distance on the left, DAP distance on the right).

4.2 Maps of Elections

We show maps for the truncation-based dataset obtained using our two distances in Figure 2, with the KK embedding (Kamada and Kawai 1989) used for the positionwise distance, and MDS (Kruskal 1964) used for DAP. The maps for size-based and comprehensive datasets can be found in the extended version of the paper.

The map obtained using the positionwise distance is similar to that of the basic dataset (shown in Figure 1a) and the map obtained using the DAP distance resembles the map of the basic dataset obtained using the swap distance (Figure 1b), albeit with some level of degradation. For the positionwise map, the truncated elections tend to be closer to UN than their complete counterparts and for DAP we see a cluster of random-cut elections one-third of a way between UN and ID elections. Unfortunately, if we used random drop truncation (with dropping probability of 0.5) then all thus-generated elections would form a single cluster in the vicinity of UN. One explanation for this is that random drop affects the internal structure of elections very strongly. For example, in the AN_8 election only two candidates are ever ranked first, but with random drop truncation every candidate has a nonnegligible probability of being ranked first. Thus, analyzing random-drop elections is particularly difficult. Still, the maps indicate that, on the high level, both distances give intuitive, reasonable results.

4.3 Varying Election Sizes Truncation Level

Next, we evaluate the robustness of our distances in a more quantitative way. First, we analyze their ability to recognize similar elections with different numbers of candidates. To this end, for each of several statistical cultures (IC, normalized Mallows model with norm-$\phi \in \{0.25, 0.5, 0.75\}$, and 1D/2D/5D Euclidean models with points distributed uniformly on unit hypercubes) and for each integer m between 8 and 24, we generated 100 pairs of elections, each with 192 voters, where one election in the pair had m candidates and the other one had 16, and we computed their average distance (normalized by the largest distance that occurred in the datasets from Section 4.1, i.e., an approximate diameter of the election space; consequently, the results for positionwise and DAP are on the same scale). We report the results in Figure 3(a), the shaded areas show 95% confidence intervals (we also included ID and AN elections in the plot; IC can be seen as an approximation of UN). Ideally, we would like our plots to consist of flat lines, close to zero. This would mean that a given distance can recognize structural similarities between elections generated in the same way, even if these elections have different numbers of candidates. Hence, in our view DAP performs better as its scaled values are significantly smaller, even if the results for Mallows are less flat (as the Mallows and IC lines are increasing, the reader may worry what happens for even more candidates; in short, they increase slowly, with IC reaching about 10% of the diameter for the case of 100 candidates). It is reassuring that for DAP the plots for ID and AN are flat at value zero (as DAP is ID- and AN-consistent), and for positionwise the plot for IC is flat and close to zero (as positionwise is UN-consistent and IC elections approximate UN).

To analyze the influence of truncation, for each of the cultures from the previous experiment we generated 25 elections with 16 candidates and 192 voters, and applied each of our truncation methods to each of them changing the parameter to obtain varying levels of vote completeness. For each truncated election, we computed its distance from the original, complete one (and normalized it as previously). In Figure 3(b) we plot the average of these values, for the case of top-k truncation; shaded area shows 95% confidence interval. We see that if the voters rank at least half of the candidates, then the average distance of the truncated election from its complete variant is (a) less than 20% of the diameter for the positionwise distance (indeed, even less than 10% for most of the cultures), and (b) less than 5% of the diamater for DAP (except for AN and IC elections, where it is $\leq 30\%$ and $\leq 15\%$ of the diameter, respectively). Thus, both distances handle truncated data well, but DAP has an advantage (however, for the other truncation types DAP and positionwise perform similarly to each other). Overall, the experiments point to DAP.

5 Map of Preflib

Last but not least, we present the *Map of Preflib*, obtained using the DAP distance (we offer more in-depth analysis in the extended version of the paper). Boehmer et al. (2021) already tried putting Preflib elections on the map, but their approach was limited to elections with 10 candidates and 100 complete votes, obtained by involved preprocessing. We use raw, unprocessed elections from Preflib, with the number of candidates ranging from 3 to $\approx 2\,600$ and the number of voters from 4 to $\approx 64\,000$. We also use more elections, as Preflib was extended since their work.

We show our map in Figure 4: Each black dot represents a single election from Preflib, while large pale discs represent elections generated synthetically (much more detailed analysis is available in the extended version of the paper). We find that most of the Preflib elections fall between Euclidean ones (with uniform distribution of candidate and voter points inside a unit hypercube of between 2 and 10 dimensions), Mallows elections (with norm-ϕ values in the range $[0.2, 0.7]$), and urn elections (with contagion parameter in the range $[0.2, 0.5]$, or $[0.2, 2]$ if one wants to include elections with large clusters of identical votes; similar elections happen in Preflib, but are more rare). This motivates the use of these models and parameter ranges to generate realistic-looking data using only the few best-known statistical cultures. Some previous papers, including those of Boehmer et al. (2021) and Faliszewski et al. (2023), argued that the urn model does not generate realistic elections but our results counter this. That said, we also see an area with many Preflib elections, but no elections from our statistical cultures (black dots over white background). It would be interesting to find statistical cultures that do cover this space.

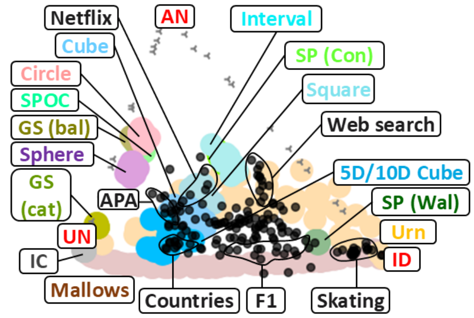


Figure 4: Map of Preflib elections (black dots) in addition to the synthetic ones (large pale discs), obtained using the DAP distance.

Next, let us anlyze the locations of the Preflib elections with respect to ID, UN, and AN. Foremost, we see that the area near AN is empty and, hence, Preflib elections are not strongly polarized (the same effect was visible in the map of Boehmer et al. (2021)). Some elections that (weakly) stand out in this respect include those in the WebSearch dataset (but they include only 4 votes each and, hence, are very particular), some Netflix elections (where each voter ranks the same 3 or 4 movies, so some level of disagreement is expected) and some APA elections (which involve 5 candidates for the president of the American Psychological Association, with some tension between "academics" and "clinicians"). Elections from the (Figure) Skating, Formula 1, and Countries datasets represent the other extreme and are located between ID and UN, close to the Mallows elections. These locations are quite natural: Indeed, we expect the judges in figure skating, who evaluate the same performances, conducted on the same day, to be correlated, and we expect Formula 1 races that happen over the course of a season to be more varied.

6 Summary

We found that the DAP distance is an interpretable, scalable way of assessing similarity between different-sized/top-truncated elections, suitable to form a map of (a fragment of) Preflib. By analyzing this map, we were able to draw a number of conclusions, including the parameters for Mallows, urn, and Euclidean models that yield realistic elections.

Nevertheless, we believe that there is no one-fit-all method for creating maps, and different applications may call for different approaches. Thus, considering different distances, including but not limited to feature distances with different set of features, and developing arguments for their usefulness might be a fruitful direction for future research.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101002854), from the French government under management of Agence Nationale de la Recherche as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute), and from the European Union under the project Robotics and advanced industrial production (reg. no. CZ.02.01.01/00/22_008/0004590). J. Mertlová acknowledges the additional support of the Faculty of Information Technology of the Czech Technical University in Prague via the Výzkumné léto (VýLeT) project. S. Szufa was supported by the Foundation for Polish Science (FNP). T. Wąs was partially supported by EPSRC under grant EP/X038548/.

References

Alcalde-Unzu, J.; and Vorsatz, M. 2013. Measuring the Cohesiveness of Preferences: An Axiomatic Analysis. *Social Choice and Welfare*, 41(4): 965–988.

Baumeister, D.; Faliszewski, P.; Lang, J.; and Rothe, J. 2012. Campaigns for Lazy Voters: Truncated Ballots. In *Proceedings of AAMAS-2012*, 577–584.

Berg, S. 1985. Paradox of Voting Under an Urn Model: The Effect of Homogeneity. *Public Choice*, 47(2): 377–387.

Black, D. 1958. *The Theory of Committees and Elections*. Cambridge University Press.

Boehmer, N.; Bredereck, R.; Elkind, E.; Faliszewski, P.; and Szufa, S. 2022a. Expected Frequency Matrices of Elections: Computation, Geometry, and Preference Learning. In *Proceedings of NeurIPS-2022*.

Boehmer, N.; Bredereck, R.; Faliszewski, P.; Niedermeier, R.; and Szufa, S. 2021. Putting a Compass on the Map of Elections. In *Proceedings of IJCAI-2021*, 59–65.

Boehmer, N.; Cai, J.-Y.; Faliszewski, P.; Fan, A.; Janeczko, L.; Kaczmarczyk, A.; and Wąs, T. 2023. Properties of Position Matrices and Their Elections. In *Proceedings of AAAI-2023*, 5507–5514.

Boehmer, N.; Faliszewski, P.; and Kraiczy, S. 2023. Properties of the Mallows Model Depending on the Number of Alternatives: A Warning for an Experimentalist. In *Proceedings of ICML-2023*.

Boehmer, N.; Faliszewski, P.; Niedermeier, R.; Szufa, S.; and Wąs, T. 2022b. Understanding Distance Measures Among Elections. In *Proceedings of IJCAI-2022*, 102–108.

Boehmer, N.; and Schaar, N. 2023. Collecting, Classifying, Analyzing, and Using Real-World Ranking Data. In *Proceedings of AAMAS-2023*, 1706–1715.

Borodin, A.; Halpern, D.; Latifian, M.; and Shah, N. 2022. Distortion in Voting with Top-t Preferences. In *Proceedings of IJCAI-2022*, 116–122.

Can, B.; Ozkes, A. I.; and Storcken, T. 2015. Measuring Polarization in Preferences. *Mathematical Social Sciences*, 78: 76–79.

Enelow, J.; and Hinich, M. 1984. *The Spatial Theory of Voting: An Introduction*. Cambridge University Press.

Faliszewski, P.; Kaczmarczyk, A.; Sornat, K.; Szufa, S.; and Wąs, T. 2023. Diversity, Agreement, and Polarization in Elections. In *Proceedings of IJCAI-2023*, 2684–2692.

Faliszewski, P.; Skowron, P.; Slinko, A.; Szufa, S.; and Talmon, N. 2019. How Similar Are Two Elections? In *Proceedings of AAAI-2019*, 1909–1916.

Hashemi, V.; and Endriss, U. 2014. Measuring Diversity of Preferences in a Group. In *Proceedings of ECAI-2014*, 423–428.

Inada, K. 1964. A Note on the Simple Majority Decision Rule. *Econometrica*, 32(32): 525–531.

Inada, K. 1969. The Simple Majority Decision Rule. *Econometrica*, 37(3): 490–506.

Kamada, T.; and Kawai, S. 1989. An Algorithm for Drawing General Undirected Graphs. *Information Processing Letters*, 31(1): 7–15.

Kamishima, T. 2003. Nantonac Collaborative Filtering: Recommendation Based on Order Responses. In *Proceedings of KDD-2003*, 583–588.

Kemeny, J. 1959. Mathematics Without Numbers. *Daedalus*, 88: 577–591.

Kilgour, D. M.; Grégoire, J.-C.; and Foley, A. M. 2020. The prevalence and consequences of ballot truncation in ranked-choice elections. *Public Choice*, 184: 197–218.

Kruskal, J. 1964. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. *Psychometrika*, 29(1): 1–27.

Mallows, C. 1957. Non-null ranking models. *Biometrica*, 44: 114–130.

Mattei, N.; and Walsh, T. 2013. PrefLib: A Library for Preferences. In *Proceedings of ADT-2013*, 259–270.

McCabe-Dansted, J.; Pritchard, G.; and Slinko, A. 2008. Approximability of Dodgson's Rule. *Social Choice and Welfare*, 31(2): 311–330.

Peters, D.; and Lackner, M. 2020. Preferences Single-Peaked on a Circle. *Journal of Artificial Intelligence Research*, 68: 463–502.

Riley, J.; Ryan, I.; and Smith, W. 2010. The Election, by Instant Runoff Voting, of UK Labour Party Leader in late-September 2010. https://rangevoting.org/LabourUK2010.html.

Rubner, Y.; Tomasi, C.; and Guibas, L. 2000. The Earth Mover's Distance as a Metric for Image Retrieval. *International Journal of Computer Vision*, 40(2): 99–121.

Szufa, S.; Faliszewski, P.; Skowron, P.; Slinko, A.; and Talmon, N. 2020. Drawing a Map of Elections in the Space of Statistical Cultures. In *Proceedings of AAMAS-2020*, 1341–1349.

Tomlinson, K.; Ugander, J.; and Kleinberg, J. 2023. Ballot length in instant runoff voting. In *Proceedings of AAAI-2023*, 5841–5849.