

Diversity of Structured Domains via k -Kemeny Scores

Piotr Faliszewski,¹ Krzysztof Sornat,¹ Stanisław Szufa,^{1,2} Tomasz Waś³

¹AGH University, Poland

²CNRS, LAMSADE, Université Paris Dauphine – PSL, France

³Univeristy of Oxford, United Kingdom

Abstract

In the k -KEMENY problem, we are given an ordinal election, i.e., a collection of votes ranking the candidates from best to worst, and we seek the smallest number of swaps of adjacent candidates that ensure that the election has at most k different rankings. We study this problem for a number of structured domains, including the single-peaked, single-crossing, group-separable, and Euclidean ones. We obtain two kinds of results: (1) We show that k -KEMENY remains intractable under most of these domains, even for $k = 2$, and (2) we use k -KEMENY to rank these domains in terms of their diversity.

Code — www.github.com/Project-PRAGMA/kKemeny-Diversity-of-Domains-AAAI-2026

Extended version — www.arxiv.org/pdf/2509.15812

1 Introduction

An *ordinal election* consists of a set of candidates and a collection of votes, ranking these candidates from the most to the least desirable one, where each vote comes from a given *domain*. We study the diversity of structured domains, such as the single-peaked (Black 1958), single-crossing (Mirrlees 1971; Roberts 1977), group-separable (Inada 1964, 1969), and Euclidean ones (Enelow and Hinich 1984, 1990), as well as the diversity of elections with votes from these domains (see Section 2 for detailed definitions). In essence, structured domains restrict possible votes to those that are somehow *reasonable*; for example, in the single-peaked domain over the standard political left-right axis, one could not rank the extreme left-wing and right-wing candidates on the two top positions. To capture diversity, we employ a technique based on solving the k -Kemeny problem of Faliszewski et al. (2023). The idea is that an election—or, a structured domain—is diverse if it includes many very different votes that cannot be easily grouped into (a small number of) clusters. Our results come in two main flavors. First, we establish the computational complexity of the k -Kemeny problem across our domains. Second, we rank these domains—as well as several statistical cultures used to sample elections from them—with respect to their diversity.

Studying the diversity of structured domains and structured elections is important for several reasons:

1. Diversity is a fundamental property of elections and domains, yet formally capturing this intuitive notion is challenging (see the discussion of the related work below). Hence, it is valuable to both develop tools for analyzing diversity and to use these tools to gain better insights into various domains and elections.
2. Faliszewski et al. (2023) argued that one can understand the nature of elections—specifically, their locations on the map of elections (Szufa et al. 2025; Faliszewski et al. 2023; Boehmer et al. 2022b)—by analyzing their three natural properties, including diversity (the other two are polarization and agreement).
3. When designing numerical experiments on elections, one may wish to consider synthetic datasets with different levels of diversity. In particular, one may choose various statistical cultures—i.e., probabilistic distributions over votes—based on the diversity of the data they produce.

The classic Kemeny score of an election is the smallest number of swaps of adjacent candidates required to ensure that all votes are identical (in Section 2 we give a different but equivalent definition; this one follows the distance-rationalization framework (Baigent 1987; Meskanen and Nurmi 2008; Elkind, Faliszewski, and Slinko 2015)). Similarly, the k -Kemeny score is the smallest number of swaps, which ensure that the election consists of at most k different votes. Intuitively, to compute the k -Kemeny score we need to find k groups of similar votes (thereby solving a clustering problem) and compute their Kemeny scores separately. Faliszewski et al. (2023) argued that a weighted sum of k -Kemeny scores for different values of k gives a good measure of diversity; the same approach, albeit with different weights, was taken by Faliszewski et al. (2025a).

1.1 Our Contributions

We obtain the following sets of results. First, we find a polynomial-time algorithm for computing the k -Kemeny score of single-crossing elections, but show NP-hardness for single-peaked and group-separable ones, already for $k = 2$. For Euclidean elections, the results are more varied and depend on the exact assumptions.

Second, using k -Kemeny scores, we rank our structured domains from the most to the least diverse one. Surprisingly, the caterpillar group-separable domain turns out to be the

most diverse one, suggesting that it might deserve to be used in experiments more often.

Third, we find that the typical way of sampling Euclidean elections, used in most computational social choice papers that employ these models—see the survey of Boehmer et al. (2024)—by necessity does not generate a sizable fraction of votes that belong to this domain. We show how this affects the diversity of generated elections, as compared to sampling votes from these domains uniformly at random.

We also make multiple remarks about our domains, showing their various quirks. Proofs are in the extended version.

1.2 Related Work

The two most related papers are those of Faliszewski et al. (2023, 2025a), where the authors introduce and use measures of diversity based on computing weighted sums of k -Kemeny scores. The former work also argues that many previously studied notions of diversity—e.g., those considered by Alcalde-Unzu and Vorsatz (2013, 2016), Can, Ozkes, and Storcken (2015, 2017) and Hashemi and Endriss (2014)—capture (dis)agreement among the votes rather than diversity, and conflate the notions of diversity and polarization.

Recently, Ammann and Puppe (2025) developed a number of diversity notions for preference domains, based on the multi-attribute approach of Nehring and Puppe (2002). The main idea is to count how many “attributes” its votes have, where an attribute can be a property such as “candidate c is ranked on top.” In similar spirit, Karpov et al. (2024) analyze how many different rankings of at least s candidates appear in a given Condorcet domain. While we focus on several mainstream domains, Karpov et al. (2024) consider numerous special ones (and Ammann and Puppe (2025) did not apply their measures to any particular domain). Overall, these approaches count occurrences of particular structures in the votes of a domain, whereas our k -Kemeny-based approach analyzes interrelations between these votes.

The problem of computing the Kemeny score was shown to be NP-hard by Bartholdi, Tovey, and Trick (1989), and its exact complexity was established by Hemaspaandra, Spakowski, and Vogel (2005). There are various ways of circumventing these intractability results, ranging from approximation algorithms (Ailon, Charikar, and Newman 2008), through parameterized approaches (Betzler et al. 2009; Betzler, Bredebeck, and Niedermeier 2014), and heuristics (Conitzer, Davenport, and Kalagnanam 2006). While it is well-known that Kemeny score can be computed in polynomial-time for Condorcet domains (Barbut 1980; Truchon 1998), doing so is also possible if the input election is, in a certain formal sense, close to being in some such domains (Cornaz, Galand, and Spanjaard 2012, 2013). Researchers also considered the complexity of computing Kemeny scores in Euclidean elections (Escoffier, Spanjaard, and Tydrichová 2022; Hamm, Lackner, and Rapberger 2021). Furthermore, De et al. (2024) analyzed (parameterized) complexity of finding all 1-Kemeny rankings, and Arrighi et al. (2021) looked for a set of good approximations of 1-Kemeny rankings that are dissimilar from each other. Computational aspects of k -Kemeny scores were, so far, considered only by Faliszewski et al. (2023).

2 Preliminaries

Given a set of candidates $C = \{c_1, \dots, c_m\}$, we write $\mathcal{L}(C)$ to denote the set of all strict rankings (linear orders) over C , and we refer to $\mathcal{L}(C)$ as the *full domain* over C . We often focus on various other domains \mathcal{D} that are subsets of $\mathcal{L}(C)$, typically referred to as *structured domains*. For a ranking v and candidates a and b , we write $v: a \succ b$ to indicate that v ranks a higher than b . If A and B are two disjoint subsets of candidates, then by $v: A \succ B$ we mean that v prefers each member of A to each member of B . The *swap distance* of two rankings $u, v \in \mathcal{L}(C)$ is the number of swaps of adjacent candidates needed to transform u into v (equal to the number of inversions, i.e., the number of pairs of candidates that are ranked differently in u and v).

Elections. An *election* $E = (C, V)$ consists of a set C of candidates and a collection V of voters, where every voter has a vote from $\mathcal{L}(C)$, ranking the candidates from the most to the least desirable one. To streamline our discussion, we use the same symbols to refer to both the voters and their votes, with the exact meaning clear from the context. We often focus on elections where the votes are restricted to belong to some structured domain, rather than the full one.

(k -)Kemeny Rankings. Let us fix an election $E = (C, V)$ and let r be some ranking from $\mathcal{L}(C)$. Its *Kemeny score* is:

$$\text{Kemeny}_E(r) = \sum_{v \in V} \text{swap}(v, r).$$

A ranking with the lowest Kemeny score is known as a *Kemeny ranking* (Kemeny 1959). Similarly, the k -Kemeny score of a set $R = \{r_1, \dots, r_k\}$ of k rankings is:

$$k\text{-Kemeny}_E(R) = \sum_{v \in V} \min_{i \in [k]} \text{swap}(v, r_i).$$

We refer to the set that minimizes the k -Kemeny score as the *k -Kemeny set* and to its members as *k -Kemeny rankings* (Faliszewski et al. 2023). We study the problem below.

Definition 2.1. *In the k -KEMENY problem we are given an election E as well as integers k and q , and we ask if there is a set R of k rankings such that $k\text{-Kemeny}_E(R) \leq q$.*

Structured Domains. In addition to the full domain, we also consider its various structured variants defined below:

Single-Peaked Domain (SP). Let \triangleleft be some ranking from $\mathcal{L}(C)$, referred to as an axis. The *single-peaked domain (SP)* for \triangleleft consists of all rankings $v \in \mathcal{L}(C)$ that satisfy the following condition: For every $t \in [|C|]$ the top t candidates in v form an interval within \triangleleft . This domain was introduced by Black (1958).

Single-Crossing Domains (SC). A subset of $\mathcal{L}(C)$ is *single-crossing (SC)* if it is possible to order its members as (v_1, \dots, v_n) , so that for each pair of candidates $a, b \in C$ there is a number $t_{ab} \in [n]$ such that voters $v_1, \dots, v_{t_{ab}}$ rank a and b in one way, and the remaining ones rank them in the opposite way. These domains were introduced by Mirrlees (1971) and Roberts (1977).

Group-Separable Domains (GS). Let \mathcal{T} be a rooted, ordered tree, where each leaf is labeled with a unique candidate from C and each internal node has at least two children. A vote $v \in \mathcal{L}(C)$ is *consistent* with \mathcal{T} if we

can obtain it by reading the labels of the leaves from left to right, after possibly reversing the order of some nodes' children. A *group-separable domain* (GS) for tree \mathcal{T} contains all votes consistent with \mathcal{T} . The notion of group-separability is due to Inada (1964, 1969), but we follow the equivalent definition of Karpov (2019). Whenever we speak of GS, we mean its variant for a given binary tree.

It is well-known that all single-peaked domains for candidate sets of the same cardinality are isomorphic, so we typically speak of *the* single-peaked domain. On the other hand, if there are at least three candidates then there are many different single-crossing domains; these observations are made explicitly, e.g., by Faliszewski et al. (2025b). Group-separable domains are isomorphic, provided that the underlying trees are isomorphic. We are particularly interested in:

Balanced Group-Separable Domain (GS/bal). A GS domain is *balanced* if its underlying tree is a binary tree where each level—except possibly the last one—is completely filled (we refer to such trees as *balanced*)

Caterpillar Group-Separable Domain (GS/cat). A GS domain is *caterpillar* if its underlying tree is binary, where each internal node has at least one leaf as a child (we refer to such trees as *caterpillar*).

A domain \mathcal{D} is a *Condorcet domain* if for each election in \mathcal{D} there is a ranking r , called *Condorcet ranking*, such that for each two candidates a and b , $r: a \succ b$ implies that at least half of the voters prefer a to b . All the above domains are Condorcet. Other domains we consider are:

Single-Peaked on a Graph Domains (SP/G). For a graph G where each vertex is labeled with a unique candidate, a vote $v \in \mathcal{L}(C)$ is single-peaked with respect to G if for each $t \in [|C|]$ the graph induced by the top t candidates is connected (so, for a path we get the classic single-peaked domain). Similarly to GS, in each SP/G domain we assume a specific connected graph G for each candidate set, which we can compute in polynomial-time. This domain was mentioned, e.g., by Elkind, Lackner, and Peters (2017); its variant for trees is due to Demange (1982).

Single-Peaked on a Circle Domain (SPOC). This is the SP/G domain for the case where graph G is a cycle; it is due to Peters and Lackner (2020).

d -Euclidean Domains. Let d be a positive integer and let $x: C \rightarrow \mathbb{R}^d$ be a function that associates each candidate with a point in the Euclidean space (called *embedding function*). A vote $v \in \mathcal{L}(C)$ belongs to the domain induced by x if there is a point $x_v \in \mathbb{R}^d$ such that for each two candidates $a, b \in C$, if $v: a \succ b$ then the Euclidean distance between x_v and $x(a)$ is smaller than between x_v and $x(b)$. These domains are discussed, e.g., by Enelow and Hinich (1984, 1990).

Statistical Cultures. Fix a candidate set C and some domain $\mathcal{D} \subseteq \mathcal{L}(C)$. A statistical culture is a distribution over the votes from \mathcal{D} . In particular, by *impartial culture over* \mathcal{D} we mean the uniform distribution over \mathcal{D} . If we omit \mathcal{D} , then we mean impartial culture of $\mathcal{L}(C)$. We introduce further statistical cultures in Section 4.1.

3 Computational Complexity of k -KEMENY

The k -KEMENY problem is NP-hard even if $k = 1$ and $n = 4$ voters (Dwork et al. 2001; Biedl, Brandenburg, and Deng 2009). However, for Condorcet domains and $k = 1$ it can be solved trivially (it suffices to compute the Kemeny score of the Condorcet ranking, which is guaranteed to be optimal). We show that for some prominent Condorcet domains, including SP, GS/bal, and GS/cat, as well as for SP/G domains, k -KEMENY becomes NP-hard already for $k = 2$. We also discuss the complexity of k -KEMENY on Euclidean domains. We supplement these results with a general FPT algorithm for Condorcet domains, parameterized by the number of voters, and an outright polynomial-time algorithm for SC elections.

3.1 Intractability Results for Structured Domains

The key idea of our hardness proofs is to give reductions from the HYPERCUBE 2-SEGMENTATION (H2S) problem. To define this problem, we need some additional notation. For a binary string x , we write $x[j]$ to refer to its j -th symbol. *Hamming* distance between two equal-length strings x and y , denoted $\text{ham}(x, y)$, is the number of positions on which these two strings differ. For a sequence $S = (s_1, \dots, s_n)$ of binary strings, each of length m , the Hamming distance between S and another binary string r of length m is $\text{ham}(S, r) = \sum_{i=1}^n \text{ham}(s_i, r)$. A string that minimizes the Hamming distance to S is called *central* for S and for every position $i \in [m]$, has the same symbol on this position as at least half of the strings in S (in case of a tie, a central string can take either symbol). The Hamming distance of such a string to S is the *Hamming score* of S , denoted $\text{ham}(S)$.

Definition 3.1. *An instance of H2S consists of a sequence $S = (s_1, \dots, s_n)$ of binary strings and of an integer t . We ask if it is possible to partition S into two groups, such that the sum of their Hamming scores is at most t .*

The first NP-completeness claim for H2S appears in the conference paper of Kleinberg, Papadimitriou, and Raghavan (1998), but without a proof. Its journal version does not include a proof either (Kleinberg, Papadimitriou, and Raghavan 2004), but claims MAXSNP-hardness (also without proof). The NP-hardness proof was eventually presented 16 years later by Feige (2014), who also argued why the MAXSNP-hardness claim was incorrect.

Theorem 3.1. *k -KEMENY is NP-complete even for $k = 2$ and elections that are both SP and GS/bal.*

Proof sketch. To show NP-hardness, we reduce from H2S. Let the input instance consist of an integer t and a sequence $S = (s_1, \dots, s_n)$ of binary strings, each of length m . We form a set of candidates $C = \{a_1, \dots, a_m\} \cup \{b_1, \dots, b_m\}$ and we say that a preference order v is *aligned* if it is of the form: $v: \{a_1, b_1\} \succ \{a_2, b_2\} \succ \dots \succ \{a_m, b_m\}$. An aligned vote is consistent with length- m binary string x if for each $j \in [m]$, we have $v: a_j \succ b_j$ when $x[j] = 1$ and $v: b_j \succ a_j$ if $x[j] = 0$. We form an election $E = (C, V)$, where for each string s_i we have exactly one vote v_i , aligned and consistent with it. We form a k -KEMENY instance with

election E , $k = 2$, and $q = t$. We observe that election E is single-peaked with respect to societal axis:

$$a_m \triangleleft \cdots \triangleleft a_2 \triangleleft a_1 \triangleleft b_1 \triangleleft b_2 \triangleleft \cdots \triangleleft b_m.$$

It is also balanced group-separable, as witnesses by balanced binary tree \mathcal{T} with $2m$ leaves, where reading the labels of the leaves from left to right gives order $a_1 \succ b_1 \succ a_2 \succ b_2 \succ \cdots \succ a_m \succ b_m$. We ask if there are two rankings r' and r'' such that k -Kemeny $_E(\{r', r''\}) \leq t$. \square

Before we discuss an analogous result for caterpillar group-separable elections, we make an observation about how one can generate GS/cat votes, or verify that votes are GS/cat. Take an axis $c_1 \triangleleft \cdots \triangleleft c_m$, corresponding to a caterpillar tree (the leaves, read from left to right, are labeled with c_1, \dots, c_m). To form a GS/cat vote for this axis, we consider the candidates in the order c_1, c_2, \dots, c_m and for each c_i we choose whether to place it on the highest or the lowest still available position. We refer to this as *caterpillar vote construction (CVC)*. The appeal is that if we sample decisions in CVC uniformly at random, it is very similar to uniform sampling of SP votes: There, we always place the considered candidate in the lowest available position, randomizing between selecting the top- or bottom not-yet-ranked candidate from the axis (Walsh 2015); a similar relation between GS/cat and SP was already noted by Boehmer et al. (2022a).

Theorem 3.2. *k -KEMENY is NP-complete even for $k = 2$ and elections that are caterpillar group-separable.*

Proof sketch. We sketch a reduction from H2S, similar in spirit to the one for Theorem 3.1, but with a different representation of the votes. Consider an H2S instance with sequence $S = (s_1, \dots, s_n)$ of binary strings of length- m each, and an integer t . Let $M = m^{10}n^{10}$.

We form a candidate set $C = A \cup B \cup X$, where $A = \{a_1, \dots, a_m\}$, $B = \{b_1, \dots, b_m\}$ and $X = \{x_1, \dots, x_M\}$, where X is a large set of dummy candidates. For each binary string z of length m and each position $j \in [m]$, we let $c_z(j)$ be a_j if $z[j] = 1$, and we let $c_z(j)$ be b_j if $z[j] = 0$. By $\bar{c}_z(j)$ we mean the unique candidate in $\{a_j, b_j\} \setminus \{c_z(j)\}$. For each binary string z of length m , we define ranking $r(z)$:

$$r(z): c_z(1) \succ \cdots \succ c_z(m) \succ X \succ \bar{c}_z(m) \succ \cdots \succ \bar{c}_z(1),$$

where by X we mean listing members of X from x_1 to x_M . For example, if z has prefix 101, then $r(z)$ is of the form:

$$r(101\dots): a_1 \succ b_2 \succ a_3 \succ \cdots \succ X \succ \cdots \succ b_3 \succ a_2 \succ b_1.$$

We form an election E with candidate set C and a single vote $r(s_i)$ for each string $s_i \in S$. Note that E is GS/cat for axis $a_1 \triangleleft b_1 \triangleleft a_2 \triangleleft b_2 \triangleleft \cdots \triangleleft a_m \triangleleft b_m \triangleleft x_1 \triangleleft \cdots \triangleleft x_M$.

Our k -KEMENY instance consists of election E , $k = 2$, and $q = 2Mt + 2nm^2$. Intuitively, whenever some symbol differs between a string from the input and the central one for its group, this corresponds to, at most, $2M + 2m^2$ swaps within a corresponding vote ($2M$ to swap the respective members of A and B between the sides of X , and the remaining ones to arrange their final positions). \square

We can also extend Theorem 3.1 to show NP-hardness for all SP/G domains, which, e.g., include the SPOC domain.

Theorem 3.3. *For every SP/G domain, k -KEMENY is NP-complete even for $k = 2$ and elections from this domain.*

For the case of 2-Euclidean elections (and, naturally, higher-dimensional ones), Escoffier, Spanjaard, and Tydričova (2022) have shown that already deciding if there is a Kemeny ranking with a given score is NP-complete, even if the embedding function is given. One of the reasons why this NP-hardness is possible is that there is no guarantee that the Kemeny ranking belongs to the given 2-Euclidean domain (which stands in contrast to Condorcet domains). Indeed, if we seek a ranking that minimizes the Kemeny score and belongs to the domain, then Hamm, Lackner, and Rapberger (2021) gave a polynomial-time algorithm for this problem. Briefly put, the size of each d -Euclidean domain is at most $O(m^{2d})$, where m is the number of candidates, so one can use brute-force search (the approach of Hamm, Lackner, and Rapberger (2021) is faster, though). We can also perform such a brute-force search for k -Kemeny scores.

Definition 3.2. *In the d -EMBEDDABLE k -KEMENY problem we are given an election E over some d -Euclidean domain, an embedding function x for this domain, and an integer q . We ask if there is a set $R = \{r_1, \dots, r_k\}$ of k rankings from the domain, such that k -Kemeny $_E(R) \leq q$.*

Corollary 3.4. *For each fixed d and k , d -EMBEDDABLE k -KEMENY is polynomial-time solvable.*

So, as opposed to k -KEMENY for SP, GS/bal and GS/cat domains, d -EMBEDDABLE k -KEMENY is tractable for $k = 2$. Yet, if k is part of the input then d -EMBEDDABLE k -KEMENY is NP-complete, even for 2-Euclidean elections.

Theorem 3.5. *d -EMBEDDABLE k -KEMENY is NP-complete for $d \geq 2$.*

3.2 Algorithms for Condorcet Domains

Faliszewski et al. (2023) have shown that for every $\varepsilon > 0$ there is an FPT approximation algorithm for k -KEMENY with $1 + \varepsilon$ approximation ratio, parameterized by the number n of the voters and running in time $O^*(n^n)$. Using dynamic programming, for Condorcet domains we improve this to an exact FPT algorithm, running in time $O^*(3^n)$. For the general domain such a result seems impossible, as 1-Kemeny is NP-hard even for $n = 4$ voters (Dwork et al. 2001; Biedl, Brandenburg, and Deng 2009).

Theorem 3.6. *There is an FPT algorithm parameterized by the number n of the votes that given an instance of k -KEMENY—where the votes come from a Condorcet domain—solves it in time $O^*(3^n)$.*

For single-crossing elections we even get a polynomial-time algorithm. The idea is to model k -KEMENY as a Chamberlin–Courant (CC) multiwinner election (Chamberlin and Courant 1983), where each vote from the original election is both a candidate and a voter in the CC one, ranking itself and the others with respect to the swap distance from itself. This produces single-peaked election, for which CC is tractable (Betzler, Slinko, and Uhlmann 2013; Sornat, Vassilevska Williams, and Xu 2022).

Theorem 3.7. *Single-crossing instances of k -KEMENY can be solved in time $O(nm + n^2)$.*

4 Diversity of Structured Domains

Next we move on to the diversity analysis of our domains and the elections that one can sample from them. Given an election E and an integer k , we write $\kappa_E(k)$ to denote its k -Kemeny score, i.e., the k -Kemeny score of its k -Kemeny set. In this section we often view domains as elections that include a single copy of each possible vote.

Faliszewski et al. (2023, 2025a) proposed to measure the diversity of an election $E = (C, V)$ using function D defined as follows ($w_1 \geq w_2 \geq \dots$ are weights and $N(E)$ is a normalizing factor, depending on $|C|$ and $|V|$):

$$D(E) = N(E) \cdot (w_1 \kappa_E(1) + w_2 \kappa_E(2) + w_3 \kappa_E(3) + \dots);$$

the larger is the value $D(E)$, the more diverse is E . However, the choice of w_1, w_2, \dots and $N(E)$ is not obvious and the two above-cited papers make different ones. We follow their general approach, but focusing on qualitative comparisons between diversities of various elections/domains, by analyzing vectors $\kappa(E) = (\frac{1}{|V|} \kappa_E(1), \dots, \frac{1}{|V|} \kappa_E(|C|))$. We refer to values $\frac{1}{|V|} \kappa_E(k)$ as normalized k -Kemeny scores. We view an election $E' = (C, V')$ as more diverse than election $E'' = (C, V'')$ if $\kappa(E')$ dominates $\kappa(E'')$ —i.e., has greater-or-equal values on each coordinate—occasionally arguing how to resolve the situation when dominance does not hold either way.

4.1 Setup, Domains, and Statistical Cultures

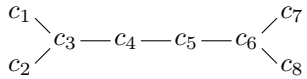
Throughout this section, we almost exclusively focus on the case of $m = 8$ candidates. This suffices for many realistic settings and is small enough for efficient computations and clear visualizations. Unless specified otherwise, all elections sampled from statistical cultures consist of 512 voters. All experiments calculating the k -Kemeny scores or domain sizes are averaged across 100 samples.

Let us fix $C = \{c_1, \dots, c_8\}$. We focus on the following nine domains, explained below:

1D-Int., 2D-Sq., 3D-Cb., SC, SP, SP/DF, SPOC, GS/bal, GS/cat.

SP, SPOC, GS/bal, GS/cat, were already introduced in Section 2, and are unique up to renaming the candidates.

SP/DF Domain. SP/double-forked domain, or SP/DF, is a domain of votes that are single-peaked on the following tree (it is also unique up to renaming the candidates):



Euclidean Domains. For a positive integer t , t D-HyperCube domain is a Euclidean domain where candidate points are selected uniformly at random from $[-1, 1]^t$ (so, in fact, it is a family of domains, one for each embedding function). Note that these domains include all preference orders that can be obtained by putting a voter point somewhere in \mathbb{R}^t , even outside of $[-1, 1]^t$. For $t \in \{1, 2, 3\}$, we refer to these domains as 1D-Interval, 2D-Square, and 3D-Cube, respectively (abbreviated as 1D-Int., 2D-Sq., and 3D-Cb.).

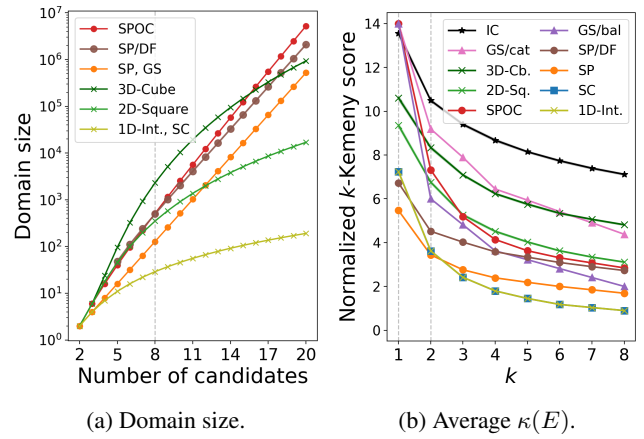


Figure 1: Sizes (a) and diversity (b) of considered domains.

Single Crossing Domains. To get a single-crossing domain (SC), we follow the approach of Szufa et al. (2025): We form a sequence $v_0, \dots, v_{\binom{m}{2}}$ of votes, where $v_0: c_1 \succ \dots \succ c_8$ and for each $i > 1$ we obtain v_i from v_{i-1} by swapping a pair of candidates c_p, c_q , $p < q$, that are ranked consecutively (we select such a pair uniformly at random).

Sizes of the Domains. In Fig. 1a we plot the sizes of our domains—i.e., the numbers of distinct votes included in each—depending on the number of candidates (note that the y axis has a logarithmic scale). Interestingly, even though the sizes of SP (which is equal to the size of GS), SP/DF, and SPOC grow exponentially and the sizes of Euclidean domains grow polynomially, for $m = 8$, 2D-Square is larger than SP (GS) and almost as large as SPOC and SP/DF. 3D-Cube is outright larger and remains so until $m = 16$.

Statistical Cultures. Whenever we generate an election with votes from a given domain, we do so by sampling the required number of votes using a respective statistical culture. In particular, for each of our domains we consider its variant of impartial culture, i.e., sampling votes uniformly at random. For SP, this is known as the Walsh model (Walsh 2015). We also consider the Conitzer model of sampling SP votes (Conitzer 2009): First, we choose candidate c_i uniformly at random and place it on top of the vote. Next, we perform $|C| - 1$ iterations, so that in each we extend by one candidate, selected uniformly at random among at most two, so that the top-ranked candidates in the generated vote form an interval with respect to the axis. We assume $c_1 < \dots < c_8$ for both models. For t D-HyperCube domains, we use the r -Box model, $r \in \mathbb{R}$, which works as follows: Given a domain and its embedding function, to sample a vote we choose its point uniformly from $[-r, r]^t$ (recall that candidates are chosen from $[-1, 1]^t$). Large fraction of literature in computational social choice that uses the t D-HyperCube models uses the 1-Box model (or an isomorphic one) and not impartial culture over the given domain (Boehmer et al. 2024).

Computing k -Kemeny. Given an election E , an integer k , and a search space (i.e., a list of votes) we compute $\kappa_E(k)$ using a local search approximation algorithm of Faliszewski

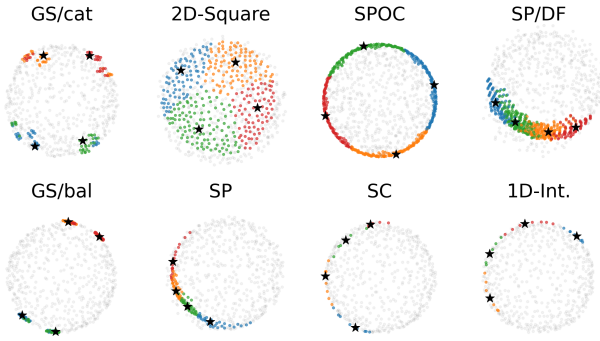


Figure 2: Microscopes of our domains with IC.

et al. (2023). We start with a set of k centers, i.e., rankings selected uniformly at random from the search space, and iteratively improve this set by replacing one of the centers with a better one, from the search space, until no improvement is found. For Condorcet domains, it suffices to search over the domain itself (Barbut 1980; Truchon 1998). For non-Condorcet domains, the search space consists of all the votes from the domain and of 512 votes sampled from IC (using more IC votes did not lead to notable improvements). We perform 10 random starts and select the best outcome.

4.2 Microscope View of Our Domains

Given an election $E = (C, V)$, with $V = (v_1, \dots, v_n)$, we visualize it using the approach of Faliszewski et al. (2023) as follows: We depict each vote $v \in V$ as a point on a plane, so that the Euclidean distance between points corresponding to votes $v_i, v_j \in V$ is as similar to $\text{swap}(v_i, v_j)$ as possible; to ensure such an embedding, we use the multidimensional scaling algorithm (Kruskal 1964; de Leeuw 2005). While not perfectly accurate, such visualizations—which we call *microscopes*—offer some intuition of the nature of the votes in the election. They follow the general ideas of the map framework of Szufa et al. (2025).

In Fig. 2, we present one microscope for each of our domains, except 3D-Cb. (for 1D-Int., 2D-Sq., and SC we show one representative example of a domain). The microscope shows the domain with 512 additional votes generated from IC (light gray dots). For each domain, we have computed an approximate 4-Kemeny set, we indicate the rankings from these sets with stars, and we colored the votes according to the closest one. As expected, votes of the same color cluster together, providing a sanity check. We refer to these plots in the following discussions.

4.3 Diversity of the Domains

For each of our domains—treated as an election \mathcal{D} —we have computed vector $\kappa(\mathcal{D})$ and present it on Fig. 1b. We repeated the computation 100 times and show averaged results, together with standard deviation in the shaded areas (nearly invisible in most cases; the randomness is over both the exact choice of the domain—for SC and the Euclidean ones—and the randomness within our k -Kemeny heuristic). The plot also includes IC for reference (in this case, for each

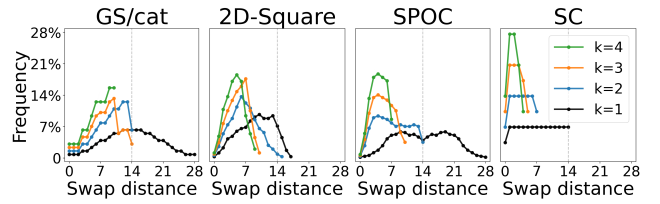


Figure 3: Histograms of the distances between the votes and their closest k -Kemeny ranking, for $k \in \{1, 2, 3, 4\}$. (For the remaining models, see the extended version.)

of the 100 repetitions we sampled new 512 votes from IC). As argued, for two domains \mathcal{D}' and \mathcal{D}'' where $\kappa(\mathcal{D}')$ dominates $\kappa(\mathcal{D}'')$, we say that \mathcal{D}' is (certainly) more diverse ($\mathcal{D}' \succ \mathcal{D}''$), leading to the following ranking (we also resolved some unclear cases and left others as ties):

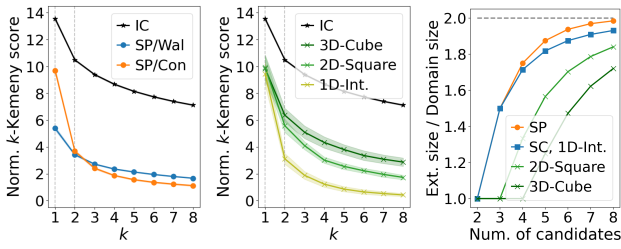
$$\begin{aligned} \text{GS/cat} &\succ 3\text{D-Cube} \succ \{2\text{D-Square, SPOC}\} \\ &\succ \{\text{SP/DF, GS/bal}\} \succ \text{SP} \succ \{\text{SC, 1D-Interval}\}. \end{aligned}$$

The three ties that we have left seem to be genuine—looking at the k -Kemeny scores of these pairs of domains we do not see strong arguments to view either of them as more diverse than the other. This is most clearly seen in case of SC and 1D-Interval, as their k -Kemeny scores are identical.

Remark 4.1. *It is well-known that every 1D-Interval election is both SP and SC, but there are elections that are SP and SC, which are not 1D-Interval (Chen, Pruhs, and Woeginger 2017; Elkind, Faliszewski, and Skowron 2020). Yet, the set of the maximal-sized 1D-Interval domains is a subset of the set of maximal-sized SC domains.*

Our diversity ranking is in agreement with many intuitions that one might get by looking at the microscopes in Fig. 2, but the fact that GS/cat is the most diverse among our domains, or that SPOC and GS/bal are ranked fairly highly, is surprising. Yet, this seems justified. Foremost, the microscopes only give approximate views of the domains and one of their features is that the points (i.e., the votes) that are presented on the outer part are typically much farther apart from each other than their Euclidean distance would suggest. Hence, e.g., the votes in the 2D-Square domain—which appear as a nice cluster in the inner part of the plot—are closer to each other than the SPOC votes, which are placed on the outer part. This is clearly seen in Fig. 3, where we show histograms of the distances of the votes from their closest k -Kemeny ranking, for $k \in \{1, 2, 3, 4\}$. In particular, for each 1-Kemeny ranking there are some GS/cat, SPOC, and GS/bal votes that are at the maximum possible distance from them (which is 28 for $m = 8$), meaning that these domains include reverses of their 1-Kemeny rankings (see also Section 4.5 below). Generally, GS/cat votes are typically farther from their k -Kemeny rankings than the 2D-Square ones.

Diversity of Conitzer and Walsh Elections. In Fig. 4a we show analogous plots as in Fig. 1b, but for elections sampled using the Walsh and Conitzer models. Intuitively, the Walsh model should produce more diverse elections as their votes cover SP more evenly, but this effect is not as strong as one



(a) SP elections. (b) Euc. elections. (c) Rev.-symmetry.

Figure 4: Average $\kappa(E)$ for SP elections sampled from Walsh and Conitzer models (a) and Euclidean elections sampled using 1-Box (b). Sizes of the reverse-symmetric extensions of the domains as a fraction of their original sizes (c).

might expect: 1-Kemeny score is higher for elections generated under the Conitzer model. However, if we disregard this one entry, then Walsh SP elections are indeed more diverse.

It is tempting to think that sampling votes from a domain uniformly at random maximizes the diversity of the resulting elections, but this is not always true. We present an example of a domain where this does not hold in the extended version.

4.4 Diversity of Euclidean Elections

Euclidean domains (particularly 1D and 2D ones) are among the most frequently used structured domains in experiments (Boehmer et al. 2024). While both voters’ and candidates’ ideal points can be drawn from various distributions, a common approach is to sample all points uniformly at random from the 1-Box models (or isomorphic ones).

In Fig. 4b we show an analogous plot as in Fig. 1b, but for elections sampled from 1D-Interval, 2D-Square, and 3D-Cube using the 1-Box model, with IC included as a reference point. As expected, the lower the dimension, the less diverse are the sampled elections. Notably, 1-Kemeny scores are nearly identical across all three Euclidean models.

Yet, we observe that Euclidean elections sampled from the 1-Box model are visibly less diverse than their respective domains. Let us explain this on the example of the 2D-Square domain. Here, each possible vote corresponds to a polygon, but given the candidates’ points, some polygons may lie outside of the $[-1, 1]^2$ square, or they may be so small that we never “hit” them during sampling. To measure how these two factors reduce the number of distinct votes in our samples, we conducted the following experiment. For each value of $r \in \{0.5, 0.75, \dots, 4\}$, we sampled 100 elections from the r -Box model.¹ Then, using the candidates’ positions in each election, we calculated both the maximum number of distinct votes possible within the given r -Box and the total number of possible votes overall. Results are presented in Fig. 5 (left). As expected, the maximum number of votes increases as we enlarge the box size. However, interestingly, the number of distinct sampled votes does not. Indeed, we find that the r -Box model produces most distinct sampled votes for $r = 1$ (we verified that this also gives the

¹For each dimension we sampled 10 times more votes than in the domain (starting from 290 for 1D, up to 222120 for 5D).

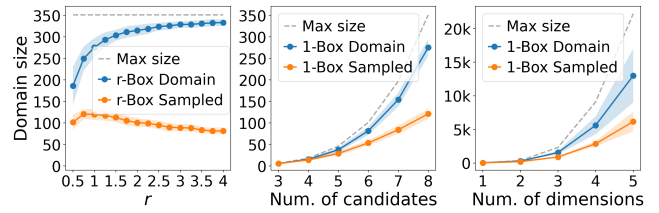


Figure 5: Changes in the sizes of Euclidean domains (the plots on the left and in the center plot regard the 2D-Square domains; for the plot on the right, the dimension of the considered domain is on the x axis).

highest diversity). Jointly with the fact that this model captures the scenario where candidates and voters come from the same population, we see it as an argument for using the 1-Box model in experiments, as already done. However, we also encourage the use of impartial culture over Euclidean domains, which so far does not seem to be done at all, as these cultures produce much more diverse elections.

In Fig. 5, we also show the results of two similar experiments. In the first one, we fix $r = 1$ and vary the number of candidates in 2 dimensions (center plot). In the second, we fix 8 candidates, but vary the number of dimensions (right plot). In both cases, the number of distinct sampled votes is significantly smaller than the domain size.

4.5 Reverse-Symmetric Domains

We call a domain \mathcal{D} *reverse-symmetric* if for each vote $v \in \mathcal{D}$, its reversal is also in \mathcal{D} . Similarly, it is *reverse-free* if for each of its votes, its reversal is not in \mathcal{D} . Among the considered domains, only GS and SPOC are reverse-symmetric, while the SP/DF is reverse-free (in fact, all SP on a tree domains are reverse-free, except for the SP on path).

We can extend each domain by adding a reversal of each vote (unless it is already present). By definition, for a reverse-symmetric domain, the size of its extension equals the size of the original domain itself. For a reverse-free domain, the size of its extension is twice that of the original one. For domains that are neither reverse-symmetric nor reverse-free, we computed the ratio between the sizes of their extensions and the original ones. The results in Fig. 4c show that all curves appear to converge towards 2, with SP converging the fastest and 3D the slowest. Hence, we can classify our domains as either being reverse-symmetric or being (nearly) reverse-free. We encourage using domains of both types in experiments.

5 Conclusions

The most important take-home messages from our work are that caterpillar group-separable elections are much more diverse than one might think, and that sampling votes from Euclidean domains uniformly at random may lead to very different elections than using the natural approach based on sampling voter points. Consequently, we encourage the use of these means of sampling preference data in experiments.

Acknowledgements

Tomasz Wąs was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant EP/X038548/1. This project has received funding from the French government under the management of Agence Nationale de la Recherche as part of the France 2030 program, reference ANR-23-IACL-0008, and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101002854).



References

- Ailon, N.; Charikar, M.; and Newman, A. 2008. Aggregating Inconsistent Information: Ranking and Clustering. *Journal of the ACM*, 55(5): 23:1–23:27.
- Alcalde-Unzu, J.; and Vorsatz, M. 2013. Measuring the Cohesiveness of Preferences: An Axiomatic Analysis. *Social Choice and Welfare*, 41(4): 965–988.
- Alcalde-Unzu, J.; and Vorsatz, M. 2016. Do We Agree? Measuring the Cohesiveness of Preferences. *Theory and Decision*, 80(2): 313–339.
- Ammann, M.; and Puppe, C. 2025. Preference Diversity. *Review of Economic Design*. <https://doi.org/10.1007/s10058-025-00386-0>.
- Arrighi, E.; Fernau, H.; Lokshantov, D.; de Oliveira Oliveira, M.; and Wolf, P. 2021. Diversity in Kemeny Rank Aggregation: A Parameterized Approach. In *Proceedings of IJCAI-2021*, 10–16.
- Baigent, N. 1987. Metric Rationalisation of Social Choice Functions according to Principles of Social Choice. *Mathematical Social Sciences*, 13(1): 59–65.
- Barbut, M. 1980. Médianes, Condorcet et Kendall. *Mathématiques et Sciences Humaines*, 69: 5–13.
- Bartholdi, J., III; Tovey, C.; and Trick, M. 1989. Voting Schemes for Which it Can Be Difficult to Tell Who Won The Election. *Social Choice and Welfare*, 6(2): 157–165.
- Betzler, N.; Bredereck, R.; and Niedermeier, R. 2014. Theoretical and Empirical Evaluation of Data for Exact Kemeny Rank Aggregation. *Autonomous Agents and Multiagent Systems*, 28(5): 721–748.
- Betzler, N.; Fellows, M.; Guo, J.; Niedermeier, R.; and Rosamond, F. 2009. Fixed-Parameter Algorithms for Kemeny Scores. *Theoretical Computer Science*, 410(45): 4554–4570.
- Betzler, N.; Slinko, A.; and Uhlmann, J. 2013. On the Computation of Fully Proportional Representation. *Journal of Artificial Intelligence Research*, 47: 475–519.
- Biedl, T.; Brandenburg, F. J.; and Deng, X. 2009. On the Complexity of Crossings in Permutations. *Discrete Mathematics*, 309(7): 1813–1823.
- 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.; Faliszewski, P.; Janeczko, Ł.; 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. In *Proceedings of IJCAI-2024*, 7962–7970.
- Boehmer, N.; Faliszewski, P.; Niedermeier, R.; Szufa, S.; and Wąs, T. 2022b. Understanding Distance Measures Among Elections. In *Proceedings of IJCAI-2022*, 102–108.
- Can, B.; Ozkes, A.; and Storcken, T. 2017. Generalized Measures of Polarization in Preferences. Technical report, HAL.
- Can, B.; Ozkes, A. I.; and Storcken, T. 2015. Measuring Polarization in Preferences. *Mathematical Social Sciences*, 78: 76–79.
- Chamberlin, B.; and Courant, P. 1983. Representative Deliberations and Representative Decisions: Proportional Representation and the Borda Rule. *American Political Science Review*, 77(3): 718–733.
- Chen, J.; Pruhs, K.; and Woeginger, G. 2017. The One-Dimensional Euclidean Domain: Finitely Many Obstructions Are Not Enough. *Social Choice and Welfare*, 48(2): 409–432.
- Conitzer, V. 2009. Eliciting Single-Peaked Preferences Using Comparison Queries. *Journal of Artificial Intelligence Research*, 35: 161–191.
- Conitzer, V.; Davenport, A.; and Kalagnanam, J. 2006. Improved Bounds for Computing Kemeny Rankings. In *Proceedings of AAAI-2006*, 620–626.
- Cornaz, D.; Galand, L.; and Spanjaard, O. 2012. Bounded Single-Peaked Width and Proportional Representation. In *Proceedings of ECAI-2012*, 270–275.
- Cornaz, D.; Galand, L.; and Spanjaard, O. 2013. Kemeny Elections with Bounded Single-Peaked or Single-Crossing Width. In *Proceedings of IJCAI-2013*, 76–82.
- De, K.; Mittal, H.; Dey, P.; and Misra, N. 2024. Parameterized Aspects of Distinct Kemeny Rank Aggregation. *Acta Informatica*, 61(4): 401–414.
- de Leeuw, J. 2005. Modern Multidimensional Scaling: Theory and Applications. *Journal of Statistical Software*, 14: 1–2.
- Demange, G. 1982. Single-Peaked Orders on a Tree. *Mathematical Social Sciences*, 3(4): 389–396.
- Dwork, C.; Kumar, R.; Naor, M.; and Sivakumar, D. 2001. Rank Aggregation Methods for the Web. In *Proceedings of WWW-2001*, 613–622.
- Elkind, E.; Faliszewski, P.; and Skowron, P. 2020. A Characterization of the Single-Peaked Single-Crossing Domain. *Social Choice and Welfare*, 54(1): 167–181.
- Elkind, E.; Faliszewski, P.; and Slinko, A. 2015. Distance Rationalization of Voting Rules. *Social Choice and Welfare*, 45(2): 345–377.

- Elkind, E.; Lackner, M.; and Peters, D. 2017. Structured Preferences. In Endriss, U., ed., *Trends in Computational Social Choice*, 187–207. AI Access Foundation.
- Enelow, J.; and Hinich, M. 1984. *The Spatial Theory of Voting: An Introduction*. Cambridge University Press.
- Enelow, J.; and Hinich, M. 1990. *Advances in the Spatial Theory of Voting*. Cambridge University Press.
- Escoffier, B.; Spanjaard, O.; and Tydrichová, M. 2022. Weighted Majority Tournaments and Kemeny Ranking with 2-Dimensional Euclidean Preferences. *Discrete Applied Mathematics*, 318: 6–12.
- 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.; Mertlová, J.; Nunn, P.; Szufa, S.; and Wąs, T. 2025a. Distances Between Top-Truncated Elections of Different Sizes. In *Proceedings of AAAI-2025*, 13823–13830.
- Faliszewski, P.; Skowron, P.; Slinko, A.; Sornat, K.; Szufa, S.; and Talmon, N. 2025b. How Similar Are Two Elections? *Journal of Computer and System Sciences*, 150: 103632.
- Feige, U. 2014. NP-Hardness of Hypercube 2-Segmentation. Technical Report arXiv:1411.0821 [cs.CC], arXiv.org.
- Hamm, T.; Lackner, M.; and Rapberger, A. 2021. Computing Kemeny Rankings from d -Euclidean Preferences. In *Proceedings of ADT-2021*, 147–161.
- Hashemi, V.; and Endriss, U. 2014. Measuring Diversity of Preferences in a Group. In *Proceedings of ECAI-2014*, 423–428.
- Hemaspaandra, E.; Spakowski, H.; and Vogel, J. 2005. The Complexity of Kemeny Elections. *Theoretical Computer Science*, 349(3): 382–391.
- 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.
- Karpov, A. 2019. On the Number of Group-Separable Preference Profiles. *Group Decision and Negotiation*, 28(3): 501–517.
- Karpov, A.; Markström, K.; Riis, S.; and Zhou, B. 2024. Local Diversity of Condorcet Domains. Technical Report arXiv:2401.11912 [econ.TH], arXiv.org.
- Kemeny, J. 1959. Mathematics Without Numbers. *Daedalus*, 88: 577–591.
- Kleinberg, J.; Papadimitriou, C.; and Raghavan, P. 1998. Segmentation Problems. In *Proceedings of STOC-1998*, 473–482.
- Kleinberg, J.; Papadimitriou, C.; and Raghavan, P. 2004. Segmentation Problems. *Journal of the ACM*, 51(2): 263–280.
- Kruskal, J. 1964. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. *Psychometrika*, 29(1): 1–27.
- Meskanen, T.; and Nurmi, H. 2008. Closeness Counts in Social Choice. In Braham, M.; and Steffen, F., eds., *Power, Freedom, and Voting*. Springer-Verlag.
- Mirrlees, J. 1971. An Exploration in the Theory of Optimal Income Taxation. *Review of Economic Studies*, 38: 175–208.
- Nehring, K.; and Puppe, C. 2002. A Theory of Diversity. *Econometrica*, 70(3): 1155–1198.
- Peters, D.; and Lackner, M. 2020. Preferences Single-Peaked on a Circle. *Journal of Artificial Intelligence Research*, 68: 463–502.
- Roberts, K. 1977. Voting Over Income Tax Schedules. *Journal of Public Economics*, 8(3): 329–340.
- Sornat, K.; Vassilevska Williams, V.; and Xu, Y. 2022. Near-Tight Algorithms for the Chamberlin-Courant and Thiele Voting Rules. In *Proceedings of IJCAI-2022*, 482–488.
- Szufa, S.; Boehmer, N.; Bredereck, R.; Faliszewski, P.; Niedermeier, R.; Skowron, P.; Slinko, A.; and Talmon, N. 2025. Drawing a Map of Elections. *Artificial Intelligence*, 343: 104332.
- Truchon, M. 1998. An Extension of the Condorcet Criterion and Kemeny Orders. Technical Report Cahier 98-15 du Centre de Recherche en Économie et Finance Appliquées, Université Laval, Québec, Canada.
- Walsh, T. 2015. Generating Single Peaked Votes. Technical Report arXiv:1503.02766 [cs.GT], arXiv.org.