Drawing a Map of Elections in the Space of Statistical Cultures

Stanisław Szufa
Jagiellonian University
Kraków, Poland
stanislaw.szufa@doctoral.uj.edu.pl

Piotr Faliszewski
AGH University
Kraków, Poland
faliszew@agh.edu.pl

Piotr Skowron
University of Warsaw
Warsaw, Poland
p.skowron@mimuw.edu.pl

Arkadii Slinko
University of Auckland
Auckland, New Zealand
a.slinko@auckland.ac.nz

Nimrod Talmon
Ben-Gurion University
Be'er Sheva, Israel
talmonn@bgu.ac.il

ABSTRACT

We consider the problem of forming a testbed of elections to be used for numerical experiments (such as testing algorithms or estimating the frequency of a given phenomenon). We seek elections that come from well-known statistical distributions and are as diverse as possible. To this end, we define a (pseudo)metric over elections, generate a set of election instances, and measure distances between them, to assess how diverse they are. Finally, we show how to use these elections to test election-related algorithms.

KEYWORDS

elections; distances; generative models; experimental social choice

ACM Reference Format:
Stanisław Szufa, Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. 2020. Drawing a Map of Elections in the Space of Statistical Cultures. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 9 pages.

1 INTRODUCTION

Alongside theoretical research, experimental studies lie in the very heart of *computational social choice* [9]. The computational aspects of elections, such as the problems of winner determination [3, 5, 35, 53], finding various forms of manipulation [4, 15, 34], control [6, 36], or bribery [18, 24], or measuring performance of candidates (e.g., via the possible/necessary winner notions [39, 61] or via the margin of victory notions [11, 51, 60]), are nowadays often investigated through experiments. For example, researchers evaluate running times of algorithms [30, 57, 59], or test what approximation ratios appear in practice [37, 53]. It is also common to test non-computational properties of elections—for instance to evaluate how frequently a given voting rule is manipulable [22, 32, 57], or how frequently particular candidates win [17] (naturally, the papers cited here are just a few examples). Yet, designing convincing experiments is not easy and, in particular, it is not clear what election data to use. Our goal is to propose a framework that can help in choosing synthetically generated elections for such tasks.

Motivating Example Let us say that we are interested in the Harmonic Borda [28] (HB) multiwinner voting rule. Under this rule we are given a set of candidates and a set of voters, where each voter ranks the candidates from the most to the least desirable one. The rule chooses a committee of k candidates that minimizes the sum of dissatisfaction scores assigned by the voters (see Section 5 for a detailed definition; the rule is a variant of the classic proportional approval voting rule, PAV [38, 55]). Finding a winning committee under this rule is NP-hard [52], but such a committee can be computed, e.g., using ILP solvers or using approximation algorithms [25, 28]. We want to assess how quickly an ILP solver can compute the winning committees and how good are the approximation algorithms.

Ideally, we should try all elections of a given size (for example, elections with about 100 candidates and 100 voters are common in the multiwinner literature [12, 17, 25]), but, of course, this is infeasible. Instead, a natural approach is to generate elections according to several standard distributions, referred to as *statistical cultures*, and test the algorithms on them.[1] Indeed, many of the above-cited papers focus on some subset of the following four models (see Section 2 for detailed descriptions of the distributions):

(1) The Impartial Culture (IC) model, where all votes are generated uniformly at random and independently [22, 32, 37, 53, 57, 59]; sometimes the Impartial Anonymous Culture model, IAC, is also used, where votes are very mildly dependent.
(2) The Polya-Eggenberger urn model, which introduces specific correlations (contagion) to the IC model [22, 30, 37, 53, 57]; the level of contagion is a parameter.
(3) The Mallows model (and its mixtures), which captures settings with ground truth [32, 53]; the dispersion of the generated votes is specified as a parameter.
(4) The Euclidean model [20, 21], which views the space of possible ideological positions in terms of Euclidean geometry [17, 30]; the way of generating the ideological positions of the candidates and voters is the parameter of the model.

However, which of these models should we use and how should we set their parameters? Perhaps we should use some other models as well, possibly generating elections that are single-peaked [8], single-crossing [46, 50], or are structured in some other way? Intuitively, we would like to have a set of elections that would be as varied as possible, so that, on the one hand, we would not spend too much time on very similar elections—for which we expect nearly

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

[1] It would also be natural to consider real-life elections (e.g., from PrefLib [44]). However, such elections typically involve relatively few candidates (often just three or four [43, 56]) and this would not suffice for our experiments (yet, such data is useful in other cases; see, e.g., the works of Brandt et al. [10] and Ayadi et al. [2]).

identical results of the experiments—and, on the other hand, we would not miss interesting families of elections—for which the results would be hard to predict. In other words, we would like to have a comprehensive testbed of elections that we could use in our experiments, and which could also be useful in other contexts.

Contributions Our main contribution is building a testbed of elections that satisfies the conditions specified above and that, as we show, is useful in practice. As a consequence of the assumed methodology, we also make a number of observations regarding various statistical cultures and relations between them. Below we describe our reasoning and resulting contributions in more detail.

To verify that our elections are not too similar to each other, a notion of a distance between elections is needed. Furthermore, this distance shall be neutral and anonymous—i.e., independent of the names of the candidates and the voters—because we will compare elections generated from statistical cultures in which this information is random. In particular, we cannot rely on the distance rationalizability framework [19, 45, 47] as its distances are not neutral. Thus we provide a new distance, called *positionwise*, that satisfies our desiderata and appears to give meaningful results.

We form our testbed as follows. First, we generate as many elections from as many different statistical cultures, with as many different parameter choices, as possible. Second, we compute the distances between each pair of generated elections. Third, we compute a mapping of the elections to the points in a 2D space in such a way that the Euclidean distances between these points reflect the positionwise distances between the respective elections (such an embedding cannot be perfect, but still helps in understanding the data). We present this mapping in Figure 2 and refer to it as our *map of elections*. This map—together with the positionwise distance—shows that the set of elections we generated is quite diverse, although, perhaps, some types of elections are overrepresented (such as, e.g., some Mallows elections). Further, along with the original distances, it allows us to draw a number of conclusions regarding various statistical cultures.

Finally, we show that both our testbed and the map are useful in practice, by performing experiments regarding the Harmonic Borda rule. For example, we measure the time needed for an ILP solver to find winning committees and superimpose the results on the map. We then see that the closer an election is to those coming from the IC model, the more time the ILP solver needs. We also obtain results regarding the quality of the committees computed by two approximation algorithms; the results are more varied here, but as the elections in the testbed are annotated with the statistical cultures from which they come, we make some interesting observations regarding the performance of our algorithms.

We believe that this paper is the first to provide a principled framework for forming diverse collections of synthetic elections. However, many other authors formed various other datasets, with different ideas in mind, often based on real-life data [43, 44, 56]. Our data and tools are available on GitHub.[2]

2 PRELIMINARIES

We write $\mathbb{R}_+$ to denote the set of non-negative real numbers. For an integer n, we write $[n]$ to denote the set $\{1, \dots, n\}$.

[2] https://github.com/szufix/mapel

An election is a pair $E = (C, V)$, where $C = \{c_1, \dots, c_m\}$ is a set of candidates and $V = (v_1, \dots, v_n)$ is a collection of voters. Each voter v_i has a preference order, i.e., a ranking of the candidates from the most to the least desirable one. To simplify notation, v_i refers both to the voter and to his or her preference order; the exact meaning will always be clear from the context. We write $v\colon a \succ b$ to indicate that voter v ranks candidate a ahead of candidate b, and we write $\mathrm{pos}_v(c)$ to denote the position of candidate c in v's preference order (the top-ranked candidate has position 1, the next one has position 2, and so on). For an election $E = (C, V)$ and two candidates $a, b \in C$, we write $M_E(a, b)$ to denote the fraction of voters in V who prefer a to b. By $d_{\mathrm{swap}}(v, u)$ we denote the swap distance between votes u and v (over the same candidate set C), i.e., the minimal number of swaps of adjacent candidates needed to turn vote u into vote v. By $d_{\mathrm{Spear}}(v, u)$ we denote the Spearman's distance between v and u. It is defined as $d_{\mathrm{Spear}}(v, u) = \sum_{c \in C} |\mathrm{pos}_v(c) - \mathrm{pos}_u(c)|$. Let C_1 and C_2 be two equal-sized candidate sets. If v is a preference order over C_1 and $\delta\colon C_1 \to C_2$ is a bijection between C_1 and C_2, then by $\delta(v)$ we mean a preference order obtained from v by replacing each candidate $c \in C_1$ with candidate $\delta(c) \in C_2$.

2.1 Structured Preferences

We often consider elections where the voters' preferences have some particular structure, e.g., are *single-peaked* (or, *single-peaked on a circle (SPOC)*), *single-crossing*, or come from some *Euclidean* domain. Such elections are studied in the literature, e.g., to model various features observed in real-life scenarios (see the overview by Elkind [16] for more references on structured preferences).

Single-peaked preferences, introduced by Black [8], capture settings where it is possible to order the candidates in such a way that as we move along this order, then each voter's appreciation for the candidates first increases and then decreases (one typical example of such an order is the classic left-to-right spectrum of political opinions). Recently, Peters and Lackner [48] introduced the notion of preferences single-peaked on a circle, where instead of ordering the candidates in a line, we arrange them cyclically (such preferences may appear, e.g., when choosing a video-conference time and the voters are in different time zones) .

Definition 2.1 (Black [8], Peters and Lackner [48]). Let C be a set of candidates and let $c_1 \lhd c_2 \lhd \cdots \lhd c_m$ be a strict, total order over C, referred to as the *societal axis*. Let v be a preference order over C. We say that v is *single-peaked* with respect to $\lhd$ if for each $\ell \in [m]$ the set of ℓ top ranked candidates according to v forms an interval within $\lhd$. We say that v is *single-peaked on a circle* if for each $\ell \in [m]$, the set of ℓ top ranked candidates either forms an interval within $\lhd$ or a complement of an interval. An election is single-peaked (is single-peaked on a circle) if there is an axis such that each voter's preference order is single-peaked (single-peaked on a circle) with respect to this axis.

We note that if an election is single-peaked then each voter ranks one of the two extreme candidates from the societal axis last. Elections single-peaked on a circle are not restricted in this way.

Example 2.2. Consider candidate set $\{a, b, c, d\}$ and the axis $a \lhd b \lhd c \lhd d$. Vote $b \succ c \succ a \succ d$ is single-peaked with respect to $\lhd$ (sets $\{b\}$, $\{b, c\}$, $\{a, b, c\}$, and $\{a, b, c, d\}$ form intervals within $\lhd$), whereas $b \succ a \succ d \succ c$ is single-peaked on a circle with respect to $\lhd$, but is not single-peaked ($\{a, b, d\}$ is not an interval). Vote $a \succ c \succ b \succ d$ is not even single-peaked on a circle with respect to $\lhd$.

Euclidean preferences, discussed in detail by Enelow and Hinich [20, 21], are based on a similar idea as the single-peaked ones, but are defined geometrically: Each candidate and each voter corresponds to a point in a Euclidean space and voters form their preferences by ranking the candidates with respect to their distance.

Definition 2.3. Let t be a positive integer. An election $E = (C, V)$ is t-Euclidean if it is possible to associate each candidate and each voter with his or her ideal point in the t-dimensional Euclidean space $\mathbb{R}^t$ in such a way that the following holds: For each voter v and each two candidates $a, b \in C$, v prefers a to b if and only if v's point is closer to the point of a than to the point of b.

Naturally, 1-dimensional Euclidean elections are single-peaked. We also note that in a 2-dimensional election where the ideal points are arranged on a circle, the voters have SPOC preferences.

Finally, we consider single-crossing elections, introduced by Mirrlees [46] and Roberts [50] in the context of taxation.

Definition 2.4 (Mirrlees [46], Roberts [50]). An election $E = (C, V)$ is single crossing if it is possible to order the voters in such a way that for each pair of candidates $a, b \in C$, the set of voters that prefer a to b either forms a prefix or a suffix of this order.

We say that a set of preference orders $\mathcal{D}$ is a *single-crossing domain* if every election where each voter has a preference order from $\mathcal{D}$ is single-crossing. For a recent discussion of single-crossing domains, see the work of Puppe and Slinko [49].

2.2 Statistical Cultures

Below we describe a number of ways of generating random elections (i.e., statistical cultures). For each of the models we either describe explicitly how an election with m candidates and n voters is generated, or we describe the process of generating a single vote (and then it is implicit that this process is repeated n times).

Impartial Culture and Related Models Under the Impartial Culture model (IC), every preference order appears with the same probability. That is, to generate a vote we choose a preference order uniformly at random. Under Impartial Anonymous Culture (IAC), we require that each *voting situation* appears with the same probability (a voting situation specifies how many votes with a given preference order are present in a profile; thus IAC generates anonymized preference profiles uniformly at random). The Imapartial Anonymous Neutral Culture (IANC) model additionally abstracts away from the names of the candidates [23]. For elections with 100 candidates, these three models are nearly the same and, so, we focus on IC.

Polya-Eggenberger Urn Model The Polya-Eggenberger urn model [7] is parametrized with a nonnegative number α, the level of contagion, and proceeds as follows: Initially, we have an urn with one copy of each of the $m!$ possible preference orders. To generate a vote, we draw a preference order from the urn uniformly at random (this is the generated vote), and we return it to the urn together with additional $\alpha m!$ copies. For larger αs the generated votes are more correlated. For $\alpha = 0$ the model is equivalent to IC, for $\alpha = 1/m!$ it is equivalent to IAC, and for $\alpha = \infty$ it produces unanimous elections.

Mallows Model The Mallows model [42] is parameterized by a value $\phi \in [0, 1]$ and a center preference order v (we choose it uniformly at random and then use for all generated votes). We generate each vote independently at random, where the probability of generating vote u is proportional to $\phi^{d_{\mathrm{swap}}(v,u)}$. For $\phi = 1$, the model is equivalent to IC, whereas for $\phi = 0$ all generated votes are identical to the center v. See the work of Lu and Boutilier for an effective algorithm for sampling from the Mallows model [41].

Single-Peaked Models We consider two ways of generating single-peaked elections, one studied by Walsh [58] and one studied by Conitzer [14]; thus we refer to them as the *Walsh model* and the *Conitzer model*. Under both models, we first choose the axis (uniformly at random). To generate a vote, we proceed as follows:

(1) Under Walsh's model, we choose a single-peaked preference order (under the given axis) uniformly at random. Walsh [58] provided a sampling algorithm for this task.
(2) To generate a vote under the Conitzer model for the axis $c_1 \lhd c_2 \lhd \cdots \lhd c_m$, we first choose some candidate c_i (uniformly at random) to be ranked on top (so, at this point, c_i is the only ranked candidate). Then, we perform $m - 1$ steps as follows: Let $\{c_j, c_{j+1}, \ldots, c_k\}$ be the set of the currently ranked candidates. We choose the next-ranked candidate from the set $\{c_{j-1}, c_{k+1}\}$ uniformly at random.

To generate a single-peaked on a circle vote, we use the Conitzer model, except that we take into account that the axis is cyclical.

Euclidean Models To generate a t-Euclidean election, we choose the ideal points for the candidates and the voters, and then derive the voters' preferences as in Definition 2.3. Given $t \in \{1, 2, \ldots\}$, we consider the following two ways of generating the ideal points:

(1) In the t-dimensional hypercube model (tD-Hypercube), we choose all the ideal points uniformly at random from $[-1, 1]^t$.
(2) In the t-dimensional hypersphere model (tD-Hypersphere), we choose all the ideal points uniformly at random from the hypersphere centered at $(0, \ldots, 0)$, with radius 1.

For $t \in \{1, 2, 3\}$, we refer to tD-Hypercube models as 1D-Interval, 2D-Square, and 3D-Cube, respectively. Similarly, by 2D-Circle and 3D-Sphere we mean the tD-Hypersphere models for $t \in \{2, 3\}$.

Single-Crossing Models We would like to generate single-crossing elections uniformly at random, but we are not aware of an efficient sampling algorithm. Thus, to generate a single-crossing election, we first generate a single-crossing domain $\mathcal{D}$ and then draw n votes from it uniformly at random. To generate this domain for a candidate set $C = \{c_1, \ldots, c_m\}$, we use the following procedure:

(1) We let v be a preference order $c_1 \succ c_2 \succ \cdots \succ c_m$ and we output v as the first member of our domain.
(2) We repeat the following steps until we output $c_m \succ c_{m-1} \succ \cdots \succ c_1$: (a) We draw candidate c_j uniformly at random and we let c_i be the candidate ranked right ahead of c_i in v (if c_j is ranked on top, then we repeat); (b) If $i < j$ then we swap c_i and c_j in v and output the new preference order.
(3) We randomly permute the names of the candidates.

Our domains always have cardinality $(1/2)m(m-1) + 1$.

3 DISTANCES BETWEEN ELECTIONS

In this section we discuss four natural pseudometrics between elections and select one to use in the remainder of the paper.

Formally, for a given set X, a pseudometric over X is a function $d\colon X \times X \to \mathbb{R}_+$ such that for each $x, y, z \in X$ it holds that (1) $d(x, x) = 0$, (2) $d(x, y) = d(y, x)$, and (3) $d(x, z) \le d(x, y) + d(y, z)$. In our case, we take X to be the set of all elections with a given number of candidates and a given number of voters. As our goal is to compare elections generated from statistical cultures, where the names of the candidates or the voters are chosen randomly, we require distances that are invariant under permutations of the names of candidates and voters; we refer to such distances as neutral/anonymous (*neutrality* refers to invariance with respect to permuting candidate names and *anonymity* has the same meaning for the case of voters); this is also why we seek pseudometrics and not metrics. So far, neutral/anonymous distances did not receive much attention in the literature, but recently Faliszewski et al. [26] introduced the class of *isomorphic distances*.

3.1 $d_{\mathrm{swap}}/d_{\mathrm{Spear}}$-Isomorphic Distances

The main idea of isomorphic distances is that given two elections, we find mappings between their candidates and between their voters, and then we sum up the distances between the individual pairs of matched votes (using, e.g., the swap distance or the Spearman's distance); we seek mappings that give the smallest final distance.

Definition 3.1. Let $E_1 = (C_1, V_1)$ and $E_2 = (C_2, V_2)$ be two elections such that $|C_1| = |C_2|$ and $|V_1| = |V_2|$. Let $\delta\colon C_1 \to C_2$ and $\sigma : V_1 \to V_2$ be two bijections.[3] For $D \in \{d_{\mathrm{swap}}, d_{\mathrm{Spear}}\}$, we define $d_D^{\delta,\sigma}\text{-ID}(E_1, E_2) = \sum_{v_1 \in V_1} d_D(\delta(v_1), \sigma(v_1))$, and let $d_D\text{-ID}(E_1, E_2)$ be the minimum of the $d_D^{\delta,\sigma}\text{-ID}(E_1, E_2)$ values, taken over δ and σ.

Example 3.2. Consider two elections, E_1 and E_2, over candidate sets $C_1 = \{a, b, c\}$ and $C_2 = \{x, y, z\}$. Election E_1 contains voters v_1, v_2, v_3 and election E_2 contains voters u_1, u_2, u_3:

$$v_1\colon a \succ b \succ c, \qquad v_2\colon b \succ a \succ c, \qquad v_3\colon b \succ c \succ a,$$
$$u_1\colon x \succ y \succ z, \qquad u_2\colon z \succ x \succ y, \qquad u_3\colon y \succ x \succ z.$$

We define $\delta(a) = x$, $\delta(b) = y$, and $\delta(c) = z$. Further, we let $\sigma(v_1) = u_1$, $\sigma(v_2) = u_3$, and $\sigma(v_3) = u_2$. We note that $\delta(v_1) = x \succ y \succ z$, so $d_{\mathrm{Spear}}(\delta(v_1), \sigma(v_1)) = 0$. Similarly, $d_{\mathrm{Spear}}(\delta(v_2), \sigma(v_2)) = 0$. However, $d_{\mathrm{Spear}}(\delta(v_2), \sigma(v_2)) = 4$ because $\delta(v_3) = y \succ z \succ x$ and $\sigma(v_3) = z \succ x \succ y$. All in all, $d_{\mathrm{Spear}}^{\delta,\sigma}\text{-ID}(E_1, E_2) = 4$. An exhaustive search over δ and σ shows that, indeed, $d_{\mathrm{Spear}}\text{-ID}(E_1, E_2) = 4$.

The isomorphic distances, d_{swap}-ID and d_{Spear}-ID, are intuitively very appealing—they capture even the smallest differences in the structure between elections and have very natural interpretations—but they are impractical; according to Faliszewski et al. [26], currently there is no simple way of computing d_{Spear}-ID for elections with more than a handful of candidates and voters. (the problem is NP-hard, hard to approximate, and the known FPT algorithms are too slow). Computing the d_{swap}-ID distance is usually even harder. Thus, while we would have liked to use d_{swap}-ID or d_{Spear}-ID, we need a different distance.

[3] Note that σ is a bijection over *voter names* so even if some two voters in V_1 have the same preference orders, σ maps them to different voters in V_2.

3.2 Positionwise and Pairwise Distances

We now introduce two new neutral/anonymous distances. The first one is based on analyzing how frequently the candidates are ranked on particular positions, and we call it the *positionwise distance*.

For an election $E = (C, V)$, a candidate $c \in C$ and a position $i \in [|C|]$, we write $\psi_E(c, i)$ to denote the fraction of votes from V that rank c on the i-th position. We define the *candidate distribution vector* of a candidate c as $\Psi_E(c) = (\psi_E(c, 1), \psi_E(c, 2), \dots, \psi_E(c, n))$. Given two vectors $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, we write $\mathrm{EMD}(x, y)$ to denote the earth mover's distance between x and y. Intuitively, this is the minimal cost of turning x into y, where the cost of moving a value Δ from position i to position j in the vector is $\Delta \cdot |i - j|$. Our EMD distance can be computed using a well-known greedy polynomial-time algorithm.

Definition 3.3. Let $E_1 = (C_1, V_1)$ and $E_2 = (C_2, V_2)$ be two elections such that $|C_1| = |C_2|$. For a bijection $\delta\colon C_1 \to C_2$, we define $\delta\text{-POS}(E_1, E_2) = \sum_{c \in C_1} \mathrm{EMD}(\Psi_{E_1}(c), \Psi_{E_2}(\delta(c)))$. The *positionwise distance* between elections E_1 and E_2, $\mathrm{POS}(E_1, E_2)$, is the minimum of the $\delta\text{-POS}(E_1, E_2)$ values, taken over δ.

We use EMD in Definition 3.3 because it captures the idea that being ranked on the top position is more similar to being ranked on the second position than to being ranked on the bottom one, but we could have used some other metric between vectors as well.

Example 3.4. Consider elections E_1 and E_2 from Example 3.2. The distribution vectors for our candidates are as follows (we omit the subscripts for Ψ to avoid clutter):

$$\Psi(a) = (1/3, 1/3, 1/3), \quad \Psi(b) = (2/3, 1/3, 0), \quad \Psi(c) = (0, 1/3, 2/3),$$
$$\Psi(x) = (1/3, 2/3, 0), \quad \Psi(y) = (1/3, 1/3, 1/3), \quad \Psi(z) = (1/3, 0, 2/3).$$

We see that $\mathrm{EMD}(\Psi(a), \Psi(y)) = 0$, $\mathrm{EMD}(\Psi(b), \Psi(x)) = 1/3$ because to transform $\Psi(b)$ into $\Psi(x)$, we need to move value $1/3$ from the first position to the second one (so we multiply $1/3$ by 1), and $\mathrm{EMD}(\Psi(c), \Psi(z)) = 1/3$. Thus for $\delta(a) = y$, $\delta(b) = x$, and $\delta(c) = z$ we have $\delta\text{-POS}(E_1, E_2) = 2/3$ and, in fact, $\mathrm{POS}(E_1, E_2) = 2/3$.

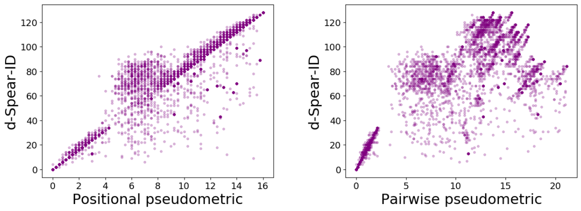
The positionwise distance is based on the idea that the most valuable information about a candidate can be extracted from the positions that he or she occupies in the voters' preference rankings. In this sense this distance is related to the family of positional scoring rules. Next we define the *pairwise distance*, which is inspired by a class of voting rules that rely on analyzing the results of head-to-head majority contests between the candidates.

Definition 3.5. Let $E_1 = (C_1, V_1)$ and $E_2 = (C_2, V_2)$ be two elections such that $|C_1| = |C_2|$. For a bijection $\delta\colon C_1 \to C_2$, we define $\delta\text{-PAIR}(E_1, E_2) = \sum_{(c,d) \in C_1 \times C_1} \left| M_{E_1}(c, d) - M_{E_2}(\delta(c), \delta(d)) \right|$. The *pairwise distance* between elections E_1 and E_2, $\mathrm{PAIR}(E_1, E_2)$, is the minimum value of the $\delta\text{-PAIR}(E_1, E_2)$ values, taken over δ.

Example 3.6. Let us again consider the two elections from Example 3.2. The matrices of head-to-head competitions look as follows:

$$M_{E_1} = \begin{array}{c} \\ a \\ b \\ c \end{array}\!\!\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 0 & 1/3 & 2/3 \\ 2/3 & 0 & 1 \\ 1/3 & 0 & 0 \end{bmatrix} \end{array} \qquad M_{E_2} = \begin{array}{c} \\ x \\ y \\ z \end{array}\!\!\begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \begin{bmatrix} 0 & 2/3 & 2/3 \\ 1/3 & 0 & 2/3 \\ 1/3 & 1/3 & 0 \end{bmatrix} \end{array}$$

For $\delta(a) = y$, $\delta(b) = x$, and $\delta(c) = z$, the $\delta\text{-PAIR}(E_1, E_2) = 2/3$, and this is also the value of $\mathrm{PAIR}(E_1, E_2)$.



(a) Positional distance vs d_{Spear}-ID; PCC = 0.937 **(b) Pairwise distance vs d_{Spear}-ID; PCC = 0.786**

Figure 1: Comparison of our distances. Each point on the plots corresponds to a pair of elections, its x coordinate gives either their positionwise or pairwise distance, and its y coordinate gives their d_{Spear}-ID distance.

Both the positionwise distance and the pairwise distance are pseudometrics defined to be neutral/anonymous. Yet, we can compute positionwise distances in polynomial-time, but the pairwise distance is intractable (indeed, it is similar to the NP-complete Approximate Graph Isomorphism problem [1, 33]). Nonetheless, we can compute the pairwise distance by formulating it as an integer linear program (we omit the details); in practice, this allows us to compute distances between elections of up to around 20 candidates.

Proposition 3.7. *There is a polynomial-time algorithm for computing the positionwise distance. The decision variant of the problem of computing the pairwise distance is* NP*-complete.*

3.3 Choosing the Distance

Our computational results suggest that the positionwise distance is our only option if we want to consider elections with 100 candidates and 100 voters. Yet, we want to get an idea as to how meaningful its results are, to find out if, perhaps, we should have relaxed our ambitions and considered smaller elections.

To this end, we have generated 100 elections with 8 candidates and 8 voters each,[4] and we have computed the d_{Spear}-ID, positionwise, and pairwise distances between each two (this gives 4950 pairs; we have looked at elections with only 8 voters and 8 candidates to be able to compute d_{Spear}-ID in reasonable time). In Figure 1 we plot the comparison of the results between d_{Spear}-ID and the positionwise distance, and between d_{Spear}-ID and the pairwise distance. In these plots, each point corresponds to a single pair of elections, the x coordinate is their positionwise distance (their pairwise distance, respectively), and the y-coordinate is their d_{Spear}-ID distance.

Already the visual inspection of Figure 1 indicates that d_{Spear}-ID is more strongly correlated with the positionwise distance than with the pairwise one. In order to obtain a more substantial evidence, we have ordered our pairs of elections (in an arbitrary way) and formed three 4950-dimensional vectors of their distances, according to d_{Spear}-ID, the positionwise, and the pairwise distance. Then we computed the Pearson Correlation Coefficient (PCC) between the vector for d_{Spear}-ID and the vectors for the positionawise and the pairwise distance, respectively. For two vectors $x = (x_1, \ldots, x_t)$ and $y = (y_1, \ldots, y_t)$, their PCC is defined as ($\overline{x}$ is the arithmetic average of the values from x; $\overline{y}$ is defined analogously):

$$\mathrm{PCC}(x, y) = \left(\sum_{i=1}^{t}(x_i-\overline{x})(y_i-\overline{y})\right) \Big/ \sqrt{\sum_{i=1}^{t}(x_i-\overline{x})^2 \sum_{i=1}^{t}(y_i-\overline{y})^2}.$$

PCC measures the level of linear correlation between two random variables and takes values between -1 and 1 (its absolute value gives the level of correlation and the sign indicates positive or negative correlation; in our case, the closer a value to 1, the better).

The PCC for d_{Spear}-ID and the positionewise distances turned out to be 0.937, and for d_{Spear}-ID and the pairwise distances, to be 0.786. Thus, for our application, where we compare elections sampled from statistical cultures, using positionwise distance is likely to be meaningful, at least to the extent to which we can draw conclusions based on small elections. (However, in general, the positionwise distance and d_{Spear} can differ arbitrarily as d_{Spear} is hard to approximate in polynomial time [26]).

4 CHARTING THE MAP OF ELECTIONS

We are ready to present our election testbed and its visualization. To build it, we proceeded as follows. First, we assembled a number of elections generated using the statistical cultures from Section 2.2; we list the exact distributions, their parameters, and numbers of generated elections in Table 1. All in all, we generated 800 elections (each with 100 candidates and 100 voters), some from very popular statistical cultures, and some from less typical ones, such as SPOC. Then, we computed the positionwise distance between each pair of them. We show statistics regarding (some of) these distances in Figure 3. For each set of elections listed there, we give their average distance to the elections from the other sets (or to the elections within the set, on the diagonal), divided by 10 and rounded.

With the concrete values of the positionwise distances in hand, we computed a mapping of the generated elections to a 2D space so that the Euclidean distances between the points in this mapping reflect the positionwise distances between the elections. To compute the embedding, we used the force-directed algorithm of Fruchterman and Reingold [31]. We present this visualization in Figure 2 and refer to it as our *map of elections*. We stress that the algorithm does not explicitly optimize the embedding so that the distances there are as proportional to the positionwise ones as possible. Instead, it seeks a compromise between the clarity of the presentation and the quality of the embedding. Yet, we have hand-verified for a number of elections that the map, indeed, roughly corresponds to the computed positionwise distances. Thus the map can be used as a source of intuitions, which, nonetheless, need to be verified.

4.1 Observations from the Map

Our map of elections (Figure 2) leads to a number of observations, both regarding how election-related algorithms and social choice phenomena should be tested, and regarding the statistical cultures themselves. We have verified that our conclusions are supported by the positionwise distances as well.

Impartial Culture and Urn Elections. Elections generated from the Impartial Culture model cover a relatively small area of our map. This confirms the well-accepted intuition that limiting one's

[4]We have generated 10 elections from each of the following statistical cultures: Impartial Culture, the Urn model with $\alpha \in 0.1, 0.2$, the Mallows model with $\phi = 0.01, 0.05, 0.1$, 1D-Interval, 2D-Square, Conitzer's single-peaked model, the single-crossing model.

Table 1: Elections generated to form our testbed.

Model	parameter	# of Elections
Impartial Culture	—	30
Urn model	$\alpha \in \{0.01, 0.02, 0.05\}$ $\alpha \in \{0.1, 0.2, 0.5\}$	30 for each α 30 for each α
Mallows model	$\phi \in \{0.999, 0.99\}$ $\phi \in \{0.95, 0.75, 0.5\}$ $\phi \in \{0.25, 0.1, 0.05\}$ $\phi \in \{0.01, 0.001\}$	20 for each ϕ 20 for each ϕ 20 for each ϕ 20 for each ϕ
Single-Peaked (Con.) Single-Peaked (Wal.) SPOC (Con.) Single-Crossing	— — — —	30 30 30 30
xD-Hypercube xD-Hypersphere	$x \in \{1, 2, 3, 5, 10, 20\}$ $x \in \{2, 3, 5\}$	30 for each x 30 for each x

experiments to this model is likely to produce biased results. On the other hand, if we consider the urn model (even for a few values of the parameter α, as in our case), then we cover quite a diverse set of elections. Indeed, in our experiment, the average distance between the urn elections was the largest among the considered statistical cultures. We believe that this is a strong argument to include urn elections in experimental research on elections. Nonetheless, the urn elections—as well as most other types of elections—have their own, particular structure. Indeed, even though the urn elections surround Mallows elections (for $\phi = 0.99$) or the xD-Hypercube elections, they almost never appear between them.

Mallows Model. The results concerning the Mallows model are intriguing. For values of ϕ that are not very close to 1, the Mallows model generates "islands" of elections that are similar to each other, but quite different from most other elections (except for single-peaked elections from the Walsh model; see below). In retrospect, this behavior is easy to explain: For each value of ϕ, the generated preference orders are, on the average, at some swap distance from the center order; the Mallows elections for different values of ϕ differ from each other because for them this average distance is different, and differ from the other models because this kind of correlation between the votes (which corresponds to the existence of ground truth) does not appear in these models.

If one wants to use the Mallows model and it is not clear what ϕ to use, then one might generate elections for random ϕ's. One natural approach, sometimes taken in the literature [32, 53], is to choose the ϕ values uniformly at random. As in our map we used values of ϕ that, roughly speaking, change exponentially, but our Mallows "islands" are, roughly speaking, equally spaced, we also suggest the following way of generating ϕ values: We first draw ϕ' from an exponential distribution, with probability density function $f(x) = (1/\beta)\exp(-x/\beta)$ and set $\phi = 1 - \phi'$ (repeat if $\phi < 0$).

We have generated 200 elections using the two approaches (for the latter, we used $\beta = 0.5$), as well as 10 Impartial Culture elections and a single unanimous election, for comparison; then we computed their visualizations—as for the full map—presented in Figure 4 (note that, due to our visualization technique, the maps are not aligned and, e.g., the positions of the Impartial Culture elections are not

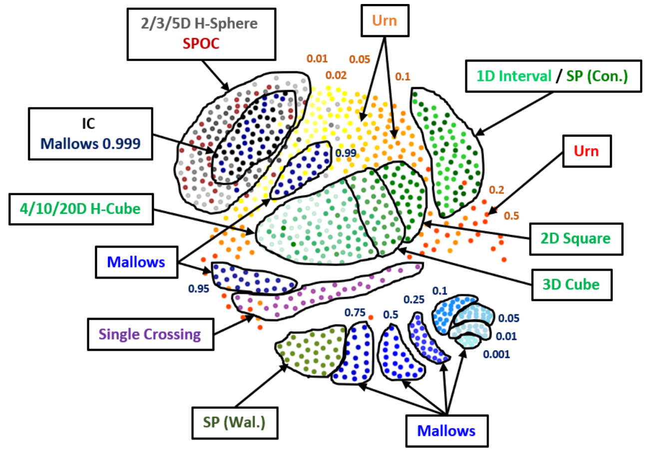


Figure 2: Visual representation of the election testbed. Each election is a dot whose colors give the statistical culture from which it was generated (the color of the statistical culture's label matches the color of its elections). For Urn and Mallows elections we also provide the value of their ϕ and α parameter.

	Impartial Culture	SPOC	5D H-Sphere	Urn model 0.05	2D Square	Conitzer SP	1D Interval	Urn model 0.5	Single Crossing	Mallows 0.95	Walsh SP	Mallows 0.5	Mallows 0.1
Impartial Culture	20	24	21	50	82	101	102	156	159	149	270	296	307
SPOC	24	24	24	56	86	104	106	161	163	152	273	299	310
5D H-Sphere	21	24	22	52	84	102	104	157	160	150	270	297	308
Urn model 0.05	50	56	52	45	68	92	92	125	126	120	235	262	273
2D Square	82	86	84	68	30	59	62	116	97	93	210	239	251
Conitzer SP	101	104	102	92	59	28	30	126	113	115	206	232	244
1D Interval	102	106	104	92	62	30	32	125	112	116	205	231	243
Urn model 0.5	156	161	157	125	116	126	125	91	90	100	147	164	172
Single Crossing	159	163	160	126	97	113	112	90	40	64	137	165	179
Mallows 0.95	149	152	150	120	93	115	116	100	64	22	133	165	178
Walsh SP	270	273	270	235	210	206	205	147	137	133	8	35	49
Mallows 0.5	296	299	297	262	239	232	231	164	165	165	35	4	14
Mallows 0.1	307	310	308	273	251	244	243	172	179	178	49	14	1

Figure 3: Average distances between elections generated using some of our statistical cultures. Each cell gives the average positionwise distance between elections generated from respective models, divided by 10 and rounded (diagonal gives the average distance between elections from the given distributions).

the same). Apparently, the latter approach gives more uniformly spaced elections and, thus, might be preferable in experiments.[5]

Single-Peaked and 1D-Interval Elections. First, we note that single-peaked elections generated using the Conitzer model are nearly indistinguishable from those generated using the 1D-Interval model (as far as the positionwise distances go). Indeed, if we draw

[5] To verify this, for both sets of elections we have computed each election's average distance to all the other ones in the set. Then, for these sets of values, we have computed the Gini inequality index (its values are between 0 and 1; the closer to 0, the more equal are the values). We expected that for a more uniform distribution we would get a value closer to 0 and, indeed, for the exponential distribution we got 0.218 and for the uniform one, we got 0.276. Thus the improvement is meaningful, but not huge.

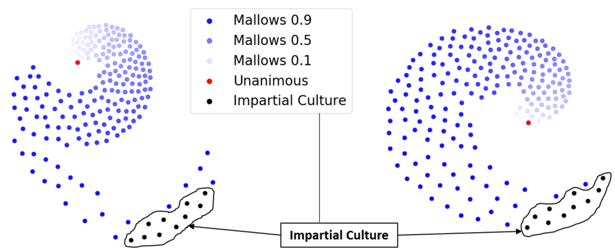


Figure 4: Two maps of Mallows elections for two distributions of the ϕ value (uniform distribution on the left, exponential distribution on the right). Each dot is an election; its color specifies the ϕ value used (the darker, the closer to 1).

the ideal points of the candidates and voters uniformly at random from an interval, then the process of forming the preference orders of the voters is similar to the one used in the Conitzer model, except that in the 1D-Interval model there are more correlations: A given voter v ranks the candidate with the closest ideal point on top, the next ranked candidate is either to the left or to the right of the first one, and this happens with probability close to $1/2$, and so on. The correlations occur because the decisions regarding which candidate should be ranked next are not made independently for each voter, but are derived from the positions of the ideal points; still, apparently, they are sufficiently small not to be easily detectable using the positionwise distances between elections.

Our second observation is that single-peaked elections generated using the Walsh model are very different from those generated using the Conitzer one and, in fact, are much closer to those from the Mallows model for $\phi = 0.75$ (and, even though the map does not show it clearly, are also quite close to Mallows elections for smaller values of ϕ). To understand why this is a natural result, let us consider a candidate set $C = \{l_m, \dots, l_1, c, r_1, \dots r_m\}$ and the corresponding societal axis $l_m \lhd \cdots \lhd l_1 \lhd c \lhd r_1 \lhd \cdots \lhd r_m$. Under the Conitzer model, the probability of generating a vote with a given candidate on top is $1/2m+1$ (by definition of the model), but in Walsh's model these probabilities differ drastically. The probability of a vote with l_m on top (or, with r_m on top) is $1/2^{2m}$ (because there are 2^{2m} different single-peaked votes for this axis, each of them is drawn uniformly at random, and only one of them starts with l_m), whereas the probability of generating a vote with c on top is $\Theta(1/\sqrt{m})$ (we omit the calculations). Generally, Walsh's model generates votes that are similar to $c \succ \{l_1, r_1\} \succ \{l_2, r_2\} \succ \cdots \succ \{l_m, r_m\}$; i.e., they are close to having a center order, as in the Mallows model.

Hypercube Elections. We observe that the x-dimensional hypercube elections are quite similar to each other, for a given x, and also fairly similar to hypercube elections generated for other dimensions. One exception is that the 1D-Interval elections are more different from the other hypercube one. This is understandable since 1D-Interval elections are single-peaked and single-crossing, whereas the other hypercube elections are not. Interestingly, the positionwise distances were sufficient to recognize this difference. Generally, hypercube elections form a large and diverse class, and we recommend using them in experiments (in particular, using *both* the 1D-Interval model and some model for a higher dimension).

Hypersphere and SPOC Elections. Similar observations as for the hypercube elections apply to hypersphere elections. Yet, it is quite interesting that hypersphere elections are generally much more similar to the IC ones than to the hypercube ones. This confirms that the distribution of the ideal points in the Euclidean models has a strong impact on the generated elections. Further studies are needed to recommend distributions that should be used in experiments (as some may lead to particularly appealing classes of elections, or to elections that are close to those appearing in reality).

The similarity between SPOC and hypersphere elections is reassuring. Indeed, 2D-Circle elections are, by their nature, a subset of SPOC ones, and for higher dimensions we would not expect big changes. Yet, the fact that SPOC/hypersphere elections are similar to the IC ones is intriguing, because the former ones have quite rigid structure, and the latter ones have none. The reason for this similarity is that both under the IC model and the SPOC/hypersphere models, if we look at a single candidate, then he or she appears at each position in a vote with roughly the same probability (though, in the SPOC/hypersphere models there are strong correlations between the positions of particular candidates). Thus elections generated from these models are similar with respect to the positionwise distance. While this may look worrisome, in Section 5 we show that this similarity is, in fact, meaningful.

Single-Crossing Elections. Elections generated using our single-crossing model are fairly close to the Mallows elections (for $\phi = 0.95$) and not too far away from urn elections (for $\alpha = 0.5$) and hypercube ones. Interestingly, 1D-Interval elections—which also are single-crossing—are farther away than higher-dimensional hypercube ones. We do not have a good explanation for these facts.

5 TESTING THE TESTBED

To demonstrate the usefulness of our testbed, we apply it to answer two questions regarding the *Harmonic Borda (HB)* voting rule [28]. Given an election and committee size k, the rule outputs a set of k candidates, referred to as the *winning committee*. It chooses this committee as follows: Consider a committee S, i.e., a subset of k candidates, a voter v, and denote by $p_1, \dots, p_k$ the positions of the members of S, sorted from the smallest (most preferred) to the largest (e.g., for a vote $v\colon c_2 \succ c_3 \succ c_1$ and committee $S = \{c_1, c_3\}$, we would have $p_1 = 2$, $p_2 = 3$). Then the *dissatisfaction* of v is $\sum_{i \in [k]} (p_i - 1)/i$. HB selects a committee S that minimizes the sum of the voters' dissatisfaction values.

Rules such as HB have received quite some attention from the research community (e.g., see the chapter of Faliszewski et al. [29]; HB is an OWA-based [40] committee scoring rule [27]). Unfortunately, identifying a winning committee under HB is NP-hard [28], but we can try to overcome this issue, e.g, by (1) formulating the problem as an integer linear program (ILP) and solving it with an off-the-shelf ILP solver, or by (2) designing polynomial-time approximation algorithms that find committees with close-to-optimal dissatisfaction values. We show how our testbed can be helpful in establishing how feasible is the ILP approach (i.e., how quickly can we compute winning committees) and which of two given approximation algorithms performs better.

Evaluating ILP Performance For each of the elections in our testbed, we computed a winning committee of size 10 using an

ILP solver (CPLEX; we used the ILP formulation for OWA-based rules of Skowron et al. [52], applied to the case of HB). We report the achieved running times in Figure 5, where the dots correspond to the elections—arranged as in Figure 2—and the colors give the running times (the darker a color, the longer the computation time; we also ran this experiment for the PAV rule [38, 55] and we found that the ILP solver needed very little time for each election).

Perhaps the most visible phenomenon is that the ILP solver needs most time on Impartial Culture elections, and the farther elections we consider (in terms of the positionwise distance) the less time is needed. Thus, even though we were surprised to see SPOC and hypersphere elections next to Impartial Culture ones, apparently their common features make them difficult for our task.

Our results say that if we wanted to analyze the pessimistic running time of the ILP solver in more detail (e.g., for different election sizes), then it would suffice to consider IC elections. Without the testbed, reaching and justifying this conclusion would be harder.

Evaluating Approximation Algorithms Two natural approximation algorithms for HB are *GreedyHB* and *RemovalHB*. GreedyHB starts with an empty committee and works in k iterations, where in each of them it adds to the committee a single candidate, so that the resulting committee has as small total dissatisfaction as possible. RemovalHB proceeds similarly, but it starts with a committee containing all candidates and works in $m-k$ iterations, in each of them removing a single candidate, so the resulting committee has as small total dissatisfaction as possible. Both algorithms are well-known in the literature and are used for various voting rules [25, 52, 54].

We ask which of the algorithms achieves better results. To this end, for each election in the testbed we (1) compute a committee of size 10 using both algorithms, and (2) for each of these committees we compute its *misrepresentation ratio*, i.e., we divide the dissatisfaction it provides by the dissatisfaction of the winning committee. The lower the misrepresentation ratio (i.e., the closer it is to 1), the better is the committee. We show our results in Figure 6, where the dots represent the elections—arranged as in Figure 2—their color specifies the algorithm which achieved lower misrepresentation ratio (blue for GreedyHB, red for RemovalHB, green for a tie; these are usually the cases where both algorithms computed the same committee), and the intensity specifies the achieved misrepresentation ratio (the darker the color, the higher it is).

The results are not as conclusive as before, but still useful. Foremost, we see that there is no clear winner among the algorithms and both perform similarly (technically, one cannot draw this conclusion from Figure 6, but we verified that in the areas where blue and red dots are interleaved, both algorithms perform very similarly). Yet, there are some surprises. For example, RemovalHB performs significantly better on 1D-Interval/Conitzer's single-peaked elections, whereas GreedyHB is significantly better on 2D-Square and 3D-Cube elections (for higher dimensions, this effect diminishes). Also, RemovalHB performs somewhat better for SPOC/hypersphere elections, but not for Impartial Culture ones (so, sometimes the structural differences between these elections are significant).

Generally, GreedyHB and RemovalHB are complementary and we should use them both. Indeed, in 90% of our elections the better of the algorithms achieved misrepresentation ratio below 1.05; in many cases the other algorithm gave a notably worse result.

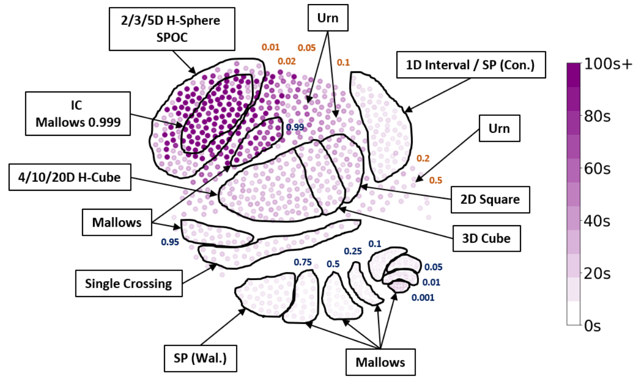


Figure 5: Map of elections, where the color corresponds to the running time of the ILP solver for Harmonic Borda on a given election (the darker, the longer the running time).

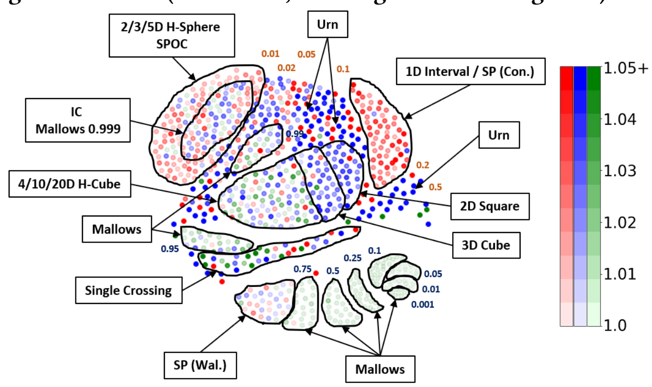


Figure 6: Map of elections, where the color corresponds to the algorithm achieving lower misrepresentation ratio (blue for GreedyHB, red for RemovalHB, green for a tie); the intensity of the color corresponds to the value of this algorithms' misrepresentation ratio (the darker, the higher).

6 FUTURE WORK

Our work leads to a number of open problems, such as: (1) How to evaluate the quality of neutral/anonymous distances? What further distances to consider? (Perhaps, the ideas regarding compiling voting rules [13] may be inspiring.) (2) How to generate single-crossing elections uniformly at random? How to generate single-peaked elections between those of Conitzer's and Walsh's models? (3) What real-life elections are similar to those from our testbed?

ACKNOWLEDGMENTS

Stanisław Szufa was supported by NCN project 2018/29/N/-ST6/01303. Piotr Faliszewski was supported by the funds of the Polish Ministry of Science and Higher Education assigned to AGH University. Piotr Skowron was supported by the Foundation for Polish Science (Homing programme). Arkadii Slinko was supported by Marsden Fund grant 3706352 of The Royal Society of New Zealand. Nimrod Talmon was supported by the Israel Science Foundation (grant No.630/19). We are very grateful to the AAMAS reviewers. Our ideas were first presented at Dagstuhl Seminar 19381.

REFERENCES

[1] V. Arvind, J. Köbler, S. Kuhnert, and Y. Vasudev. 2012. Approximate Graph Isomorphism. In *Proceedings of MFCS-2012*. 100–111.

[2] M. Ayadi, N. Ben Amor, J. Lang, and D. Peters. 2019. Single Transferable Vote: Incomplete Knowledge and Communication Issues. In *Proceedings of AAMAS-19*. 1288–1296.

[3] H. Aziz, S. Gaspers, J. Gudmundsson, S. Mackenzie, N. Mattei, and T. Walsh. 2015. Computational Aspects of Multi-Winner Approval Voting. In *Proceedings of AAMAS-2015*. 107–115.

[4] J. Bartholdi, III, C. Tovey, and M. Trick. 1989. The Computational Difficulty of Manipulating an Election. *Social Choice and Welfare* 6, 3 (1989), 227–241.

[5] J. Bartholdi, III, C. Tovey, and M. Trick. 1989. Voting Schemes for Which it Can Be Difficult to Tell Who Won The Election. *Social Choice and Welfare* 6, 2 (1989), 157–165.

[6] J. Bartholdi, III, C. Tovey, and M. Trick. 1992. How Hard is it to Control an Election? *Mathematical and Computer Modeling* 16, 8/9 (1992), 27–40.

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

[8] D. Black. 1958. *The Theory of Committees and Elections*. Cambridge University Press.

[9] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia (Eds.). 2016. *Handbook of Computational Social Choice*. Cambridge University Press.

[10] F. Brandt, C. Geist, and M. Strobel. 2016. Analyzing the Practical Relevance of Voting Paradoxes via Ehrhart Theory, Computer Simulations, and Empirical Data. In *Proceedings of AAMAS-16*. 385–393.

[11] R. Bredereck, P. Faliszewski, A. Kaczmarczyk, R. Niedermeier, P. Skowron, and N. Talmon. 2017. Robustness Among Multiwinner Voting Rules. In *Proceedings of SAGT-2017*. 80–92.

[12] L. Celis, L. Huang, and N. Vishnoi. 2018. Multiwinner Voting with Fairness Constraints. In *Proceedings of IJCAI-2018*. 144–151.

[13] Y. Chevaleyre, J. Lang, N. Maudet, and G. Ravilly-Abadie. 2009. Compiling the Votes of a Subelectorate. In *Proceedings of IJCAI-2009*. 97–102.

[14] V. Conitzer. 2009. Eliciting Single-Peaked Preferences Using Comparison Queries. *Journal of Artificial Intelligence Research* 35 (2009), 161–191.

[15] V. Conitzer, T. Sandholm, and J. Lang. 2007. When Are Elections with Few Candidates Hard to Manipulate? *J. ACM* 54, 3 (2007), Article 14.

[16] E. Elkind. 2018. Restricted Preference Domains in Social Choice: Two Perspectives. In *Proceedings of SAGT-2018*. 12–18.

[17] E. Elkind, P. Faliszewski, J. Laslier, P. Skowron, A. Slinko, and N. Talmon. 2017. What Do Multiwinner Voting Rules Do? An Experiment Over the Two-Dimensional Euclidean Domain. In *Proceedings of AAAI-2017*. 494–501.

[18] E. Elkind, P. Faliszewski, and A. Slinko. 2009. Swap Bribery. In *Proceedings of SAGT-2009*. 299–310.

[19] E. Elkind, P. Faliszewski, and A. Slinko. 2012. Rationalizations of Condorcet-consistent rules via distances of Hamming type. *Social Choice and Welfare* 39, 4 (2012), 891–905.

[20] J. Enelow and M. Hinich. 1984. *The Spatial Theory of Voting: An Introduction*. Cambridge University Press.

[21] J. Enelow and M. Hinich. 1990. *Advances in the Spatial Theory of Voting*. Cambridge University Press.

[22] G. Erdélyi, M. Fellows, J. Rothe, and L. Schend. 2015. Control Complexity in Bucklin and Fallback Voting: An Experimental Analysis. *J. Comput. System Sci.* 81, 4 (2015), 661–670.

[23] Ö. Eğecioğlu and A. Giritligil. 2013. The Impartial, Anonymous, and Neutral Culture Model: A Probability Model for Sampling Public Preference Structures. *Journal of Mathematical Sociology* 37, 4 (2013), 203–222.

[24] P. Faliszewski, E. Hemaspaandra, and L. Hemaspaandra. 2009. How Hard is Bribery in Elections? *Journal of Artificial Intelligence Research* 35 (2009), 485–532.

[25] P. Faliszewski, M. Lackner, D. Peters, and N. Talmon. 2018. Effective Heuristics for Committee Scoring Rules. In *Proceedings of AAAI-2018*. 1023–1030.

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

[27] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. 2016. Committee Scoring Rules: Axiomatic Classification and Hierarchy. In *Proceedings of IJCAI-2016*. 250–256.

[28] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. 2017. Multiwinner Rules on Paths From k-Borda to Chamberlin–Courant. In *Proceedings of IJCAI-2017*. 192–198.

[29] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. 2017. Multiwinner Voting: A New Challenge for Social Choice Theory. In *Trends in Computational Social Choice*, U. Endriss (Ed.). AI Access Foundation.

[30] P. Faliszewski, A. Slinko, K. Stahl, and N. Talmon. 2018. Achieving Fully Proportional Representation by Clustering Voters. *Journal of Heuristics* 24, 5 (2018), 725–756.

[31] T. Fruchterman and E. Reingold. 1991. Graph drawing by force-directed placement. *Software: Practice and Experience* 21, 11 (1991), 1129–1164.

[32] J. Goldsmith, J. Lang, N. Mattei, and P. Perny. 2014. Voting with Rank Dependent Scoring Rules. In *Proceedings of AAAI-14*. 698–704.

[33] M. Grohe, G. Rattan, and G. Woeginger. 2018. Graph Similarity and Approximate Isomorphism. In *Proceedings of MFCS-2018*. 20:1–20:16.

[34] E. Hemaspaandra and L. Hemaspaandra. 2007. Dichotomy for Voting Systems. *J. Comput. System Sci.* 73, 1 (2007), 73–83.

[35] E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. 1997. Exact Analysis of Dodgson Elections: Lewis Carroll's 1876 Voting System is Complete for Parallel Access to NP. *J. ACM* 44, 6 (1997), 806–825.

[36] E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. 2007. Anyone But Him: The Complexity of Precluding an Alternative. *Artificial Intelligence* 171, 5–6 (2007), 255–285.

[37] O. Keller, A. Hassidim, and N. Hazon. 2019. New Approximations for Coalitional Manipulation in Scoring Rules. *Journal of Artificial Intelligence Research* 64 (2019), 109–145.

[38] M. Kilgour. 2010. Approval Balloting for Multi-Winner Elections. In *Handbook on Approval Voting*. Springer, 105–124. Chapter 6.

[39] K. Konczak and J. Lang. 2005. Voting Procedures with Incomplete Preferences. In *Proceedings of M-PREF-2005*. 124–129.

[40] J. Lang and P. Skowron. 2018. Multi-Attribute Proportional Representation. *Artificial Intelligence* 263 (2018), 74–106.

[41] T. Lu and C. Boutilier. 2014. Effective Sampling and Learning for Mallows Models with Pairwise-Preference Data. *Journal of Machine Learning Research* 15, 1 (2014), 3783–3829.

[42] C. Mallows. 1957. Non-null ranking models. *Biometrica* 44 (1957), 114–130.

[43] N. Mattei, J. Forshee, and J. Goldsmith. 2012. An Empirical Study of Voting Rules and Manipulation with Large Datasets. In *Proceedings of COMSOC-2012*.

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

[45] T. Meskanen and H. Nurmi. 2008. Closeness Counts in Social Choice. In *Power, Freedom, and Voting*, M. Braham and F. Steffen (Eds.). Springer-Verlag.

[46] J. Mirrlees. 1971. An Exploration in the Theory of Optimal Income Taxation. *Review of Economic Studies* 38 (1971), 175–208.

[47] S. Nitzan. 1981. Some Measures of Closeness to Unanimity and Their Implications. *Theory and Decision* 13, 2 (1981), 129–138.

[48] D. Peters and M. Lackner. 2017. Preferences Single-Peaked on a Circle. In *Proceedings of AAAI-2017*. 649–655.

[49] C. Puppe and A. Slinko. 2019. Condorcet Domains, Median Graphs and the Single-Crossing Property. *Economic Theory* 67, 1 (2019), 285–318.

[50] K. Roberts. 1977. Voting Over Income Tax Schedules. *Journal of Public Economics* 8, 3 (1977), 329–340.

[51] D. Shiryaev, L. Yu, and E. Elkind. 2013. On Elections with Robust Winners. In *Proceedings of AAMAS-13*. 415–422.

[52] P. Skowron, P. Faliszewski, and J. Lang. 2016. Finding a Collective Set of Items: From Proportional Multirepresentation to Group Recommendation. *Artificial Intelligence* 241 (2016), 191–216.

[53] P. Skowron, P. Faliszewski, and A. Slinko. 2015. Achieving Fully Proportional Representation: Approximability Result. *Artificial Intelligence* 222 (2015), 67–103.

[54] P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. 2017. Proportional Rankings. In *Proceedings of IJCAI-2017*. 409–415.

[55] T. Thiele. 1895. Om Flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*. 415–441.

[56] T. Tideman and F. Plassmann. 2012. Modeling the Outcomes of Vote-Casting in Actual Elections. In *Electoral Systems: Paradoxes, Assumptions, and Procedures*, D. Felsenthal and M. Machover (Eds.). Springer, 217–251.

[57] T. Walsh. 2011. Where Are the Hard Manipulation Problems. *Journal of Artificial Intelligence Research* 42, 1 (2011), 1–29.

[58] T. Walsh. 2015. *Generating Single Peaked Votes*. Technical Report arXiv:1503.02766 [cs.GT]. arXiv.org.

[59] J. Wang, S. Sikdar, T. Shepherd, Z. Zhao, C. Jiang, and L. Xia. 2019. Practical Algorithms for Multi-Stage Voting Rules with Parallel Universes Tiebreaking. In *Proceedings of AAAI-2019*. 2189–2196.

[60] L. Xia. 2012. Computing the Margin of Victory for Various Voting Rules. In *Proceedings of EC-12*. 982–999.

[61] L. Xia and V. Conitzer. 2011. Determining Possible and Necessary Winners Given Partial Orders. *Journal of Artificial Intelligence Research* 41 (2011), 25–67.