Optimization-Based Voting Rule Design: The Closer to Utopia the Better

Piotr Faliszewski, Stanislaw Szufa, Nimrod Talmon

Abstract

In certain situations, such as elections in the Euclidean domain, it is possible to specify clear requirements for the operation of a multiwinner voting rule, for it to provide committees that correspond to some desirable intuitive notions (such as individual excellence of committee members or their diversity). We formally describe several such requirements, which we refer to as "utopias". Supplied with such utopias, we develop an optimization-based mechanism for constructing committee scoring rules that provide results as close to these utopias as possible; we test our mechanism on weakly separable and OWA-based rules. Using our method we recovered some believed connections between known multiwinner voting rules and certain applications and got other interesting insights.

1 Introduction

Multiwinner voting is a formalism for selecting a set of items (a committee), based on the preferences of a group of agents (the voters) [26, 16, 25]. For example, a group of judges may need to select a set of finalists of a competition, a hiring committee may need to select a set of people to invite for on-site interviews, and an Internet store may need to decide which items to present on its homepage (depending on how the preferences of its customers are perceived). In each of these examples, we need committees with different properties; the judges should select individually best candidates, the Internet store should select a diverse set of items that covers interests of as many of its customers as possible, and the hiring committee should balance these two requirements (we should invite as good candidates as possible, but we also should maintain some diversity among the profiles of the interviewees).

More generally, following the recent overview of Faliszewski et al. [25], multiwinner elections might be categorized into three classes, based on what their goals are: *Individual Excellence*, for selecting individually best candidates; *Proportional Representation*, for accurately and proportionally representing the electorate views; and *Diversity*, for reflecting the wide spectrum of voters' views.

So far, to address these varied goals and needs researchers typically analyzed existing multiwinner voting rules, studied their computational complexity [33, 29, 7, 4, 12, 36], analyzed their axiomatic properties [16, 38, 2, 34, 22], evaluated them experimentally [14, 15, 24, 39, 9], and—based on this evidence—argued which rules are best for which application (e.g., the k-Borda rule [13] is seen as appropriate for choosing individually excellent candidates, whereas the Chamberlin–Courant rule [10] is appropriate for identifying diverse committees that cover a wide spectrum of opinions).[1] In other words, typically, researchers analyzed existing rules and checked which ones behave appropriately for a given setting. There are also cases where researchers hand-designed multiwinner rules to achieve their goals (e.g., for situations such as the hiring committee above, Faliszewski et al. [24] designed a spectrum of rules achieving various levels of compromise between the goals of excellence and diversity; Elkind et al. [18], Aziz et al. [3] and Sekar et al. [36] proposed multiwinner variants of the Condorcet rule).

[1] The references above are meant to present the wide range of results obtained, and are certainly not complete. We point readers interested in more systematic treatment to the survey of Faliszewski et al. [25].

In this paper we take a radically different approach from the previous ones: Given a specification of the kind of committees one is interested in, we use an optimization algorithm to automatically design—in a principled way—rules that match this specification. (Related to our approach, we mention the position paper of Xia [41], which takes a normative approach and suggests the use of machine learning to automatically design voting rules.)

Our work is driven by two main motivations. The first one, suggested above, is that we wish to develop a methodology for designing voting rules which would satisfy certain desired properties. There are many multiwinner rules (such as k-Borda [13], Bloc, Chamberlin–Courant [10], Proportional Approval Voting [40], Monroe [30], and many others) that seem to have good properties for *some* idealized goals (recall the three types of multiwinner elections discussed above), but here we wish to develop a general mechanism that, when supplied with *any* arbitrary goal (specified in an appropriate way), can output a multiwinner rule appropriate for this goal. Indeed, designing voting rules which satisfy certain properties is in the heart of social choice. As a proof of concept, in this paper we focus on (two subclasses of) the class of committee scoring rules [16] and design an algorithm that searches for appropriate rules among them. We focus on the rules from these classes because they are parameterized through sets of numeric parameters that we can tweak to manipulate their properties; this aspect is important for our optimization-based approach.

The idea of designing voting rules tailored to have specific properties is not new. For example, it has already motivated the view of voting rules as maximum likelihood estimators (see, e.g., the work of Conitzer et al. [11]), where the rules are meant to recover ground truth from noisy data, or the distance-rationalization approach (see, e.g., the work of Elkind et al. [17]), where the rules are viewed as seeking consensus. Our work can also be seen as providing means of specifying voting rules by non-experts (and, in this sense, it is related to the work of Cailloux and Endriss [8]).

Our second motivation relates to the richness of the class of committee scoring rules. So far, researchers have analyzed the general structure of committee scoring rules [38], considered a few of their subclasses [16, 23, 22], and studied several specific rules [16, 37, 2] or spectra of rules [24]. However, as there are so many committee scoring rules, it might be that, in spite of the effort outlined above, some important rules might have been missed. Our mechanism of designing rules tailored for particular goals explores specified subclasses of committee scoring rules and, for each setting, either finds one of the already-known rules (thus confirming that it is highly appropriate for a given setting) or discovers a new rule.

2 Preliminaries

An election $E = (C, V)$ consists of a set of candidates $C = \{c_1, \dots, c_m\}$ and a collection of voters $V = (v_1, \dots, v_n)$, where each voter v_i has a linear order $\succ_{v_i}$, ranking the candidates from the one that v_i appreciates most to the one that v_i appreciates least. We refer to $\succ_{v_i}$ as the preference order of voter v_i (and, sometimes, as the vote of v_i). For a voter v and a candidate c, we write $\mathrm{pos}_v(c)$ to denote the position of c in v's preference order (the top-ranked candidate has position 1, the next one has position 2, and so on). A multiwinner voting rule is a function $\mathcal{R}$ that, given an election $E = (C, V)$ and an integer k, $1 \le k \le |C|$, outputs a family of size-k subsets of C (i.e., a family of committees) that win this election.

For each integer t, we write $[t]$ to denote the set $\{1, \dots, t\}$. In particular, if m is the number of candidates, we often interpret $[m]$ as the set of positions that candidates may take in a preference order. A single-winner scoring function (for an election with m candidates) is a non-increasing function $\gamma_m \colon [m] \to \mathbb{R}$ that associates each position in a vote with a score value. We define the γ_m-score of a candidate c in an election $E = (C, V)$ to be $\gamma\text{-score}_E(c) = \sum_{v \in V} \gamma_m\big(\mathrm{pos}_v(c)\big)$. We use normalized scoring functions, so that $\gamma_m(1) = 1$

and $\gamma_m(0) = 0$. For example, the Borda scoring function is defined as $\beta_m(i) = {}^{m-i}/_{m-1}$, and the t-Approval scoring function (denoted α_t, where $t \in [m]$ is a parameter) is a function that associates score 1 with the first t positions, and score 0 with the remaining ones.

Committee scoring functions are defined analogously to the single-winner ones, but for a generalized notion of a position. Let us fix committee size k. Then, given a committee S and a vote v, we define the position of S in v, denoted $\mathrm{pos}_v(S)$, to be the sequence of positions of the members of S in v, sorted in the increasing order (i.e., we obtain $\mathrm{pos}_v(S)$ by sorting the set $\{\mathrm{pos}_v(s) \mid s \in S\}$ in the increasing order). We write $[m]_k$ to denote the set of all length-k increasing sequences of elements from $[m]$ (and we interpret elements of $[m]_k$ as committee positions). We say that committee position $I = (i_1, \dots, i_k)$ weakly dominates committee position $J = (j_1, \dots, j_k)$, denoted $I \succeq J$, if for each $t \in [k]$ it holds that $i_t \leq j_t$. A committee scoring function (for m candidates and committee size k) is a function $f_{m,k} \colon [m]_k \to \mathbb{R}$, such that for each two committee positions $I, J \in [m]_k$, if $I \succeq J$ then $f(I) \geq f(J)$. The $f_{m,k}$-score of committee S in election $E = (C, V)$ is defined as $\sum_{v \in V} f_{m,k}\big(\mathrm{pos}_v(S)\big)$. For a family $f = (f_{m,k})_{k \leq m}$ of committee scoring functions (one for each number of candidates and committee size), we define the committee scoring rule $\mathcal{R}_f$ as follows: Given an election $E = (C, V)$ with m candidates and committee size k, it outputs all size-k committees S with the highest $f_{m,k}$-score.

Example 2.1. Let us fix an election E with m candidates and committee size k. The SNTV rule is defined by committee scoring functions of the form $f^{\mathrm{sntv}}_{m,k}(i_1, \dots, i_k) = \alpha_1(i_1)$. This means that the rule selects a committee of k candidates that are ranked first most frequently (or several such committees, in case of ties). The Bloc rule uses functions of the form $f^{\mathrm{bloc}}_{m,k}(i_1, \dots, i_k) = \alpha_k(i_1) + \cdots + \alpha_k(i_k)$, which can be interpreted as saying that each voter gives one point to each of his or her k most favorite candidates, and the k candidates with the highest score form the winning committee. The k-Borda rule chooses k candidates with the highest Borda scores and is defined through the functions $f^{\mathrm{kb}}_{m,k}(i_1, \dots, i_k) = \beta_m(i_1) + \cdots + \beta_m(i_k)$. The Chamberlin–Courant rule (the CC rule) uses scoring functions of the form $f^{\mathrm{cc}}_{m,k}(i_1, \dots, i_k) = \beta_m(i_1)$. This means that given a committee S, each voter associates it with the Borda score of this member of S that he or she ranks highest (this candidate is called the representative of the voter). Finally, the Harmonic-Borda rule [24] (the HB rule) uses the scoring function $f^{\mathrm{cc}}_{m,k}(i_1, \dots, i_k) = \beta_m(i_1) + {}^1/_2\beta_m(i_2) + \cdots + {}^1/_k\beta_m(i_k)$.

Consider a setting with m candidates, where the desired committee size is k, and where $\mathcal{R}$ is a committee scoring rule:

1. We say that $\mathcal{R}$ is *weakly separable* if its committee scoring function is of the form $f(i_1, \dots i_k) = \gamma(i_1) + \cdots + \gamma(i_k)$, where γ is a single-winner scoring function.
2. We say that $\mathcal{R}$ is *OWA-based* if its committee scoring function is of the form $f(i_1, \dots, i_k) = \lambda_1\gamma(i_1) + \cdots + \lambda_k\gamma(i_k)$, where $\Lambda = (\lambda_1, \dots, \lambda_k)$ is a sequence of non-negative real numbers and γ is a single-winner scoring function (we refer to the vector Λ as the OWA vector).

Note that every weakly separable rule is OWA-based (with the all 1s OWA vector). For the purpose of this paper, we normalize OWA vectors, so that λ_1 is always 1. We say that a committee scoring rule is *OWA/Borda-based* if it is OWA-based, uses the Borda scoring function and a non-increasing OWA-vector. All the rules from Example 2.1 are OWA-based.

3 Methodology

In this section we describe our technique of designing voting rules. The technique is based on minimizing the distance between election results, computed for 1-dimensional Euclidean elections, and certain prespecified distributions (which we call *utopic distributions*).

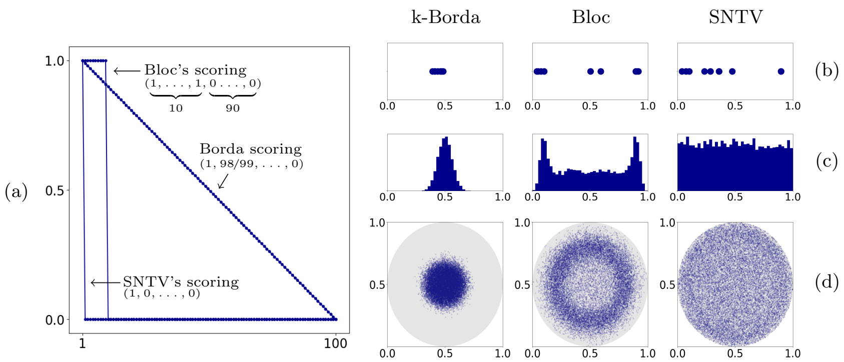


Figure 1: Visualization of example weakly separable rules (k-Borda, Bloc, and SNTV). Plot (a) shows the scoring functions, plots (b) show example election results on a 1D interval, plots (c) show histograms for the interval election model, and plots (d) show scatter plots for the disc model.

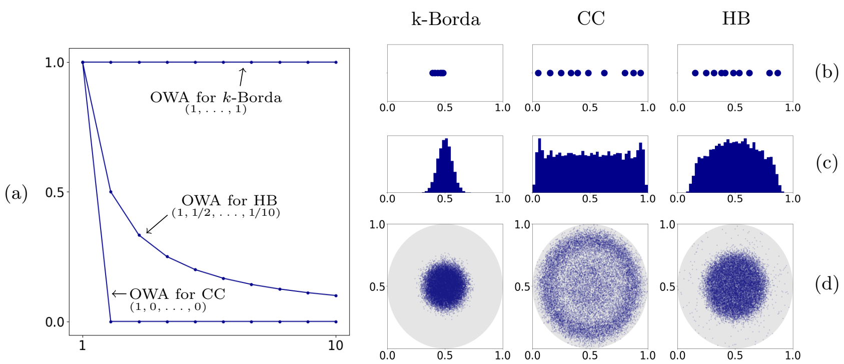


Figure 2: Visualization of example OWA/Borda-based rules (k-Borda, CC, and HB). Plot (a) shows the OWA vectors used, whereas plots (b)–(d) have the same meaning as in Figure 1.

3.1 Euclidean Elections

In the t-dimensional Euclidean model of elections, each individual u (i.e., each candidate and each voter) is represented by a point $p(u) \in \mathbb{R}^t$ in the t-dimensional space. Intuitively, the coordinates of this point may correspond to u's position regarding some t issues [19, 20]. Each voter forms his or her preference order by sorting the candidates in increasing order of the distances of the candidates' ideal points from the voter's ideal point (i.e., the closer a candidate is to a voter, the higher the voter ranks the candidate).

In our computations, we use either 1-dimensional Euclidean elections, where we generate the ideal points of candidates and voters by drawing them uniformly at random from the $[0, 1]$ interval, or 2-dimensional elections, where we draw the ideal points uniformly at random from a disc centered at point $(0.5, 0.5)$ with radius 0.5. We refer to the former as the *interval model* and to the latter as the *disc model*. We always generate elections with 100 candidates and 100 voters, and we seek committees of size 10. We chose these parameters to ensure that our results are comparable to those already present in the literature [15, 24, 21]. We use the interval model throughout the whole process of designing voting rules, and we use the disc model to check whether the rules that we produce maintain their features after changing (and, in a sense, generalizing) the setting.

Following Elkind et al. [15], we present the results of our elections visually. For a given voting rule $\mathcal{R}$ and a given election model (interval or disc), we generate a number of elections according to the model (1000 elections for the interval model and 2000 elections for the disc model), compute the $\mathcal{R}$ winning committee for each election (if there are ties, we break them arbitrarily), and—depending on the model—present them as follows:

1. For the interval model, we partition the $[0, 1]$ interval into 40 subintervals, count how many times a candidate from a given subinterval was in a winning committee, and present these numbers as a histogram. We do not normalize the histograms; different ones have different scales as their point is to show the "shape" of the election results.

2. For the disc model, we show a scatter plot, where each member of a winning committee is indicated as a blue dot (thus, as opposed to the work of Elkind et al. [15], our plots for the disc model are *not* histograms[2]). In addition to the blue dots, we also show the gray disc from which the candidates' and voters' points are drawn.

In Figures 1 and 2 we show visualizations of the results for the rules from Example 2.1. For weakly separable rules, we show plots of their scoring functions (so on the x-axis we have the 100 possible positions in a vote), and for OWA/Borda-based rules, we show their OWA vectors (so on the x-axis we have 10 entries); k-Borda is shown in both plots as it belongs to both classes of rules.

To compute results of weakly separable rules, we use their direct polynomial-time algorithms. For OWA/Borda-based rules, we compute winning committees by solving integer linear programs (ILPs) provided for this task by Peters [32] (we use the CPLEX ILP solver). Peters showed that using his formulations gives a polynomial-time algorithm for the case of single-peaked elections; since elections generated for the interval model are single-peaked, we enjoy this guaranteed efficiency (however, this no longer holds for the disc model).

3.2 Utopic Distributions and Distance Measures

We use probability distributions (which we call *utopic distributions*) to represent how, ideally, we would like the winners of our interval elections to be distributed (or, roughly speaking, how we would like their 1D histograms to look like). For example, the utopic distribution that models the goal of individual excellence associates the whole probability mass with the center of the interval, whereas the distribution associated with covering the whole spectrum is, simply, the uniform distribution over the interval.

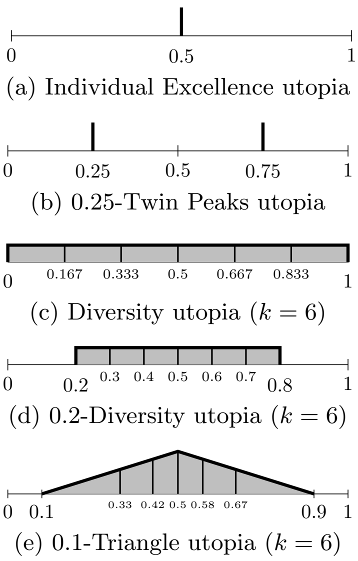
Let $\mathcal{U}$ be some utopic distribution. Given a committee $W = \{w_1, \ldots, w_k\}$ for some interval election, we define d_W, the distribution associated with W, so that for each $x \in [0, 1]$:

$$d_W(x) = \|\{w_i \mid p(w_i) = x\}\|/k.$$

To measure how closely W fits utopia $\mathcal{U}$, we use the intuitions underlying the Earth mover's distance [31]: We view the probability mass associated with each point (each interval) as the number of "grains of sand" that lie on this point (this interval). Moving a grain of sand from point x to point y costs $|x - y|$. The distance between two distributions is the lowest possible cost of moving the "grains of sand" needed to transform one of them into the other. While this intuition is discrete in its nature, our utopic distributions are sometimes continuous (in other words, sometimes we consider probability density functions). Instead of providing a general definition of our distance, below we describe the utopic distributions

[2] Our plotting tool draws the blue dots as "partially transparent," so areas with fewer winners appear in lighter shade of blue, whereas areas with high concentration of winners appear as dark blue. The reason to use scatter plot instead of 2D histograms is that, similarly to the histograms, it provides a good intuition on how the given rule behaves, but it requires far fewer election results.

that we consider and for each we derive the appropriate distance measure (they are defined so their values are comparable among each other, even though such comparisons are not necessary). In the descriptions below, we let k be the committee size and $W = \{w_1, \dots, w_k\}$ be a committee, whose members have ideal points $p(w_1) \dots, p(w_k) \in [0, 1]$. We assume that these points are sorted, i.e., $p(w_1) \le p(w_2) \le \dots \le p(w_k)$.



(a) Individual Excellence utopia

(b) 0.25-Twin Peaks utopia

(c) Diversity utopia ($k = 6$)

(d) 0.2-Diversity utopia ($k = 6$)

(e) 0.1-Triangle utopia ($k = 6$)

Figure 3: Utopic distributions.

Individual Excellence ($\mathcal{U}_{IE}$). The individual excellence utopic distribution, $\mathcal{U}_{IE}$, is defined as concentrating all the probability mass in the center of the interval, at point 0.5; see Figure 3a (this is inspired by the k-Borda rule, which is regarded as very good for the excellence goal, and which chooses candidates in the center). We define the distance between $\mathcal{U}_{IE}$ and d_W to be $\mathrm{EMD}(\mathcal{U}, d_W) = \sum_{i=1}^{k} {}^1\!/\!_k |p(w_i) - 0.5|$. That is, for each member of the committee we pay the cost of moving him or her to the center of the interval (we multiply each $|p(w_i) - 0.5|$ by ${}^1\!/\!_k$ as each member of the committee is associated with probability mass ${}^1\!/\!_k$).

Twin Peaks ($\mathcal{U}_{TP}^{\epsilon}$). The Bloc rule motivates the study of the twin peaks utopic distributions (see Figure 1). An ϵ-twin peaks utopic distribution for parameter ϵ, denoted $\mathcal{U}_{TP}^{\epsilon}$, places half of the probability mass on point ϵ and half on point $1-\epsilon$; see Figure 3b. We let the distance between $\mathcal{U}_{TP}^{\epsilon}$ and d_W be:

$$\mathrm{EMD}(\mathcal{U}_{TP}^{\epsilon}, d_W) = \sum_{i=1}^{k/2} {}^1\!/\!_k |p(w_1) - \epsilon| + \sum_{i=k/2+1}^{k} {}^1\!/\!_k |p(w_1) - (1-\epsilon)|$$

(we assume that k is even); we we assign the left half of committee members to the left peak, and the right half to the right peak (recall that they are sorted).

Diversity ($\mathcal{U}_D$). The diversity (or, coverage) utopic distribution, denoted as $\mathcal{U}_D$ and defined to be the uniform distribution over $[0, 1]$, models the idea that a diverse committee should cover the whole interval as uniformly as possible. Our reasoning for the distance $\mathrm{EMD}(\mathcal{U}_D, d_W)$ is that the committee members are supposed to be distributed evenly along the interval $[0, 1]$ and, so, each of them is responsible for covering a ${}^1\!/\!_k$-length subinterval. We assign the subintervals to committee members so that w_1 is assigned to $[0, \frac{1}{k}]$, w_2 is assigned to $[\frac{1}{k}, \frac{2}{k}]$ and so on; see Figure 3c.

Let $\ell = {}^1\!/\!_k$ be the length of the subintervals. For each w_i, we define the cost of "spreading" his or her probability mass from d_W over the assigned subinterval $[\ell(i-1), \ell i]$ so:

1. If the committee member is to the left of his or her interval (i.e., $p(w_i) < \ell(i-1)$), then we need to pay the cost $(\ell(i-1) - p(w_i))\ell$ for moving his or her probability mass (which also is equal to ℓ) to the point $\ell(i-1)$, and then the cost ${}^1\!/\!_2 \ell^2 = \int_{\ell(i-1)}^{\ell} \big(1 - \frac{x - \ell(i-1)}{\ell}\big) dx$ for "spreading" his or her weight over the interval (note that this latter cost equals to the area of a triangle). If the committee member is to the right of his or her interval, we proceed analogously.

2. If the committee member is in his or her interval (i.e., $\ell(i-1) \le p(w_i) \le \ell i$), then it suffices to "spread" the $\frac{p(w_i) - \ell(i-1)}{\ell}$ fraction of his or her probability mass to the part of the interval left of him, at cost $\frac{1}{2}(p(w_i) - \ell(i-1))^2$ (analogously to the previous

cases, this can be expressed as the area of a right triangle, with two sides of length $p(w_i) - \ell(i-1))$, and the remaining mass, to the part of the interval to the right of him or her, at cost $\frac{1}{2}(\ell i - p(w_i))^2$.

Overall, the cost associated with w_i is:

$$\mathrm{cost}(w_i) = \begin{cases} (\ell(i-1) - p(w_i))\ell + \frac{\ell^2}{2} & \text{for } p(w_i) \leq \ell(i-1) \\ \frac{(p(w_i)-\ell(i-1))^2 + (\ell i - p(w_i))^2}{2} & \text{for } \ell(i-1) < p(w_i) < \ell i \\ (p(w_i) - \ell i)\ell + \frac{\ell^2}{2} & \text{for } \ell i \leq p(w_i) \end{cases}$$

and we define $\mathrm{EMD}(\mathcal{U}_D, d_W)$ to be $\sum_{i=1}^{k} \mathrm{cost}(w_i)$.

Diversity/Excellence Compromises ($\mathcal{U}_D^\epsilon$ and $\mathcal{U}_T^\epsilon$). We also consider two families of utopic distributions that achieve a certain level of compromise between the ideals of individual excellence and diversity. Let ϵ be a number in $[0, 0.5]$. We define the ϵ-diversity utopic distribution, denoted by $\mathcal{U}_D^\epsilon$, to be the uniform distribution over the interval $[\epsilon, 1-\epsilon]$.; see Figure 3d. Each committee member w_i is responsible for covering interval $I_i = [\epsilon + \ell(i-1), \epsilon + \ell i]$ of length $\ell = 1/k(1-2\epsilon)$.

Our second way of capturing a compromise between individual excellence and diversity is via a distribution whose probability density function, for a given $\epsilon \in [0, 0.5]$, is a triangle with a peak at 0.5, set over the interval $[\epsilon, 1-\epsilon]$ (the area of the interval is always one). We call it the ϵ-triangle utopic distribution and denote it by $\mathcal{U}_T^\epsilon$. We derive the values $\mathrm{EMD}(\mathcal{U}_D^\epsilon, d_W)$ and $\mathrm{EMD}(\mathcal{U}_T^\epsilon, d_W)$ following the same logic as in the case of $\mathrm{EMD}(\mathcal{U}_D, d_W)$ (omitted due to space restriction).

3.3 Search Algorithms

The final component of our method is an algorithm that, given a utopic distribution $\mathcal{U}$ and one of our two families of committee scoring rules, finds a rule $\mathcal{R}$ as close to $\mathcal{U}$ as possible.

For m candidates, a weakly separable rule is defined via a non-increasing vector $X = (x_1, \ldots, x_m)$, such that $x_1 = 1$ and $x_m = 0$; the vector X specifies the values of the underlying single-winner scoring function γ for the possible positions in a preference order. Given a committee size k, an OWA/Borda-based rule is defined by its non-increasing OWA vector $(\lambda_1, \ldots, \lambda_k)$, where $\lambda_1 = 1$. Correspondingly, given a vector Y of appropriate size, we write $\mathcal{R}_Y$ to denote the rule defined by this vector (when we consider weakly separable rules, Y gives the score values; when we consider OWA/Borda-based rules, it is the OWA vector). Given a vector Y', by *normalizing* it we mean sorting it, setting its first coordinate to 1, replacing all > 1 values with 1s and all < 0 values with 0, and—for weakly separable rules—setting its last coordinate to 0 (so that the vector describes a legal rule from the relevant class).

Let us fix the class of rules and the utopic distribution $\mathcal{U}$. Our goal is to find a vector Y so that the winning committees under $\mathcal{R}_Y$ follow $\mathcal{U}$ as closely as possible. To make this notion precise, our algorithm first computes a given number N of interval elections $E_1, \ldots, E_N$ (these are fixed throughout the whole optimization process). To evaluate the rule $\mathcal{R}_Y$, for each election E_i we compute the winning committee W_i (if there are ties, then we break them arbitrarily). Then we compute the average distance of these committees from the utopia, $\mathrm{EMD}(\mathcal{U}, \mathcal{R}_Y) = 1/N \sum_{i=1}^{N} \mathrm{EMD}(\mathcal{U}, d_{W_i})$; this value is referred to as the *score* of the rule (the lower, the better).

To find a good vector Y, we use a local search algorithm that is similar to simulated annealing, but that never accepts worse solutions (we found local minima to not pose problems for our search space, and our approach turned out to be more effective). We use the following parameters: (a) the number of iterations T, (b) the probability $\omega(i) \in [0,1]$ of changing a given vector's coordinate, depending on the iteration number i, (c) the range parameter $r(i) \in [0,1]$, specifying how much vector coordinates can change depending on the iteration number i. The algorithm works as follows:

1. Draw vector Y with coordinates from $[0,1]$ uniformly at random and normalize it.
2. Repeat the following steps T times:
   (a) Create a vector Y' using the following procedure. Set $Y' = Y$. Then, for each of its coordinates y'_i, compute y''_i by adding to y'_i a value drawn uniformly at random from $[-r(i), r(i)]$. With probability $\omega(i)$, replace the value of y'_i with y''_i. Normalize Y'.
   (b) Replace Y with Y' if $\mathrm{EMD}(\mathcal{U}, R_{Y'}) < \mathrm{EMD}(\mathcal{U}, \mathcal{R}_Y)$; otherwise, keep Y as is.
3. Output the rule $\mathcal{R}_Y$.

For weakly separable rules, we use $T = 3000$ iterations, $\omega(i) = \max(\frac{T-i}{2T}, 0.05)$, $r(i) = 0.5\omega(i)$, and $N = 400$ test elections. For OWA/Borda-based rules, we use $T = 300$ iterations, $\omega(i) = \max(\frac{T-i}{2T}, 0.1)$, $r(i) = 0.3 \cdot \max(\frac{T-i}{2T}, 0.05)$, $N = 40$ test elections. To speed up the algorithm for the case of OWA/Borda-based rules, we first run it for elections with 50 candidates, 50 voters, and committee size 10, and only then we re-run it for full-sized elections (with 100 candidates, 100 voters, and committee size 10), using the result of the first run as the input for the second one. The algorithm does not provide any guarantees regarding the quality of the results, but we compared its performance to a brute-force search for small elections and it achieved very good results.

4 Results

We used our search algorithm to find the best weakly separable and OWA/Borda-based rules for the individual excellence, ϵ-diversity (with $\epsilon \in \{0, 0.1, 0.2\}$), ϵ-twin peaks (with $\epsilon \in \{0.167, 0.25, 0.333\}$), and ϵ-triangle (with $\epsilon \in \{0, 0.1, 0.2\}$) utopic distributions.

The results are given in Table 1 (where we show EMD distances for the best rules we computed using our algorithm, and for the five rules from Example 2.1) and in Figures 4–9. Each figure shows results for four utopic distributions: the individual excellence utopia, which can be seen as a border case for each of the other distributions, and either ϵ-diversity, ϵ-twin peak, or ϵ-triangle distributions, for appropriate values of ϵ. The largest plot on the left of each figure, marked (a), shows vectors computed for the respective four utopias. Next to it, as Plot (b), we show graphical representation of the respective utopia (drawn as a gray area over the [0,1] interval) and a sample result of a single interval election (the blue dots). As Plot (c), we show the 1D histograms achieved by the computed rules (we remind the reader that different histograms have different y-axis scales, as they only show the "shape" of the result). Finally, as Plot (d), we show the scatter plots computed for disc elections according to our four rules. The vectors computed for the utopias are marked with a number (1–4) and the respective figures in (b)–(d) are marked with matching numbers.

EMD Results Versus Histograms. The EMD results in Table 1 show that generally our best OWA/Borda-based rules are much closer to respective utopias than our best weakly separable rules. While this is in agreement with intuition, it may sometimes be surprising that weakly separable rules achieve far more visually appealing histograms for some settings than the OWA/Borda-based ones (e.g., for the twin-peaks distributions), in spite of having worse EMD values. The reason for this discrepancy is that weakly separable rules achieve good histograms "in the aggregate" (averaged over many elections), whereas OWA/Borda rules perform well (but not great) for every election instance.

Weakly Separable Rules. Next we discuss specific results for weakly separable rules. For individual excellence (see Figure 4) we obtained a nearly linear vector, very close to the Borda scoring function (vector 4). On the other extreme, for the diversity utopia, we found a rule very close to SNTV (vector 1, which is 0 for most positions, then slowly increases, and jumps to value 1 for the first position; in fact, it is a bit closer to $\mathcal{U}_D$ than SNTV is). We view it as a negative result: our hope was to find a weakly separable rule that would robustly implement the diversity utopia, but apparently such rule does not exist (SNTV does not implement this utopia robustly as its results seem to be a statistical artifact: SNTV chooses candidates from areas with lower density of candidates and increased density of voters, which, statistically, happens equally often in each area of the interval).

Utopia	W. Sep.	OWA/Borda	SNTV	Bloc	Borda	CC	HB
$\mathcal{U}_D$	0.104	**0.042**	0.092	0.215	0.227	**0.044**	0.094
$\mathcal{U}_D^{0.1}$	0.098	**0.042**	0.100	0.205	0.178	0.059	**0.053**
$\mathcal{U}_D^{0.2}$	0.073	**0.044**	0.126	0.210	0.130	0.100	**0.044**
$\mathcal{U}_{TP}^{0.167}$	0.235	**0.135**	0.159	0.237	0.307	**0.140**	0.172
$\mathcal{U}_{TP}^{0.25}$	0.185	**0.060**	0.144	0.222	0.224	0.124	**0.116**
$\mathcal{U}_{TP}^{0.333}$	0.135	**0.081**	0.156	0.221	0.141	0.136	**0.093**
$\mathcal{U}_T$	0.139	**0.043**	0.113	0.213	0.150	0.082	**0.043**
$\mathcal{U}_T^{0.1}$	0.136	**0.045**	0.137	0.220	0.127	0.116	**0.051**
$\mathcal{U}_T^{0.2}$	0.122	**0.054**	0.174	0.238	0.105	0.160	**0.083**
$\mathcal{U}_{IE}$	**0.041**	0.049	0.248	0.264	**0.049**	0.244	0.168

Table 1: EMD values for the best weakly separable and OWA/Borda-based rules computed for our utopias, and for the rules from Example 2.1 for the same utopias.

Perhaps the most interesting results are those achieved for 0.1-diversity and 0.2-diversity (vectors 2 and 3), as they show rules that, if at all, are very rarely discussed in the literature. Both vectors 2 and 3 resemble functions of the form $\gamma(i) = (1-x)\alpha_1(i) + x\beta_m(i)$, where $x \in [0,1]$ is a parameter (and, in our case, is close to 0.2). In other words, these functions give score 1 to position 1, score ≈0.2 to position 2, and then decrease linearly to 0. One could say that their Borda-score component is too small to be relevant, but this is not so. In our 1D elections we have 100 voters, which means that there are only 100 points to be distributed for being ranked on the first place, while there are ≈1000 points to be distributed for being ranked on the following places. This way, the rules described by vectors 2 and 3 achieve a compromise between SNTV and Borda.

The results for the ϵ-twin peaks utopic distribution (Figure 5) are quite spectacular. For each $\epsilon \in \{0.167, 0.25, 0.333\}$ we find a rule whose 1D histogram matches the respective utopic distribution very well (but only in the aggregate; see Table 1). The twin-peaked distributions were inspired by the results for the Bloc rule (recall Figure 1); indeed, we find vectors consisting of several 1s followed by 0s (with a very rapid transition). However, as opposed to Bloc, our vectors have many more 1s (Bloc would have k of them, i.e., 10 in our case, but our rules have between 20 and 40). Apparently, this is the reason why our rules match the twin-peaked utopias better than Bloc, which selects more candidates "between the peaks". A final remark regarding the twin peaks distributions regards individual excellence: Our results show that, as we put the peaks closer to each other, we obtain vectors of more 1s, followed by fewer 0s, but when the peaks finally coincide, we should obtain the linear function. Either there is some sort of phase transition between these two extremes, or we did not put the peaks close enough to each other to observe a smooth transition.

For the triangle utopias (see Figure 6), we seem to find rules whose scoring vectors resemble the shape of the harmonic sequence $1, 1/2, \ldots, 1/m$. These rules, indeed, seem to achieve a compromise between individual excellence (i.e., choosing winners from the center) and diversity (choosing winners from the whole interval/disc). Further, on the intuitive level, these rules are more appealing than the SNTV/Borda compromise obtained for ϵ-diversity.

Indeed, this suggests that studying a variant of k-Borda that uses harmonic numbers instead of linearly decreasing scores might be interesting.

OWA/Borda-Based Rules. For individual excellence (see, e.g., Figure 7), our algorithm very quickly finds OWA vector of all 1s; thus, we obtain the k-Borda rule. For diversity (Figure 7, vector 1), we find a vector close to that of Chamberlin–Courant (a single 1 followed by 0s). For 0.1-diversity and 0.2-diversity we find, respectively, vector 2 that resembles (but, admittedly, quite poorly) the harmonic sequence and linearly decreasing vector 3. These two results are intriguing. First, the linear vector is a very natural solution to finding a compromise between excellence and diversity (which, in this case, would mean finding a compromise between k-Borda and Chamberlin–Courant) that has not been considered in the literature yet (even though Faliszewski et al. [24] look for rules that achieve such a compromise, they do not study this rule). On the other hand, the harmonic vector has received extensive treatment, both for Harmonic Borda [24] and for the PAV rule [40, 2, 26, 27] (which is approval-based and has motivated the study of Harmonic Borda). In fact, the OWA vectors that we have obtained for the triangle distributions (see Figure 9) seem to be closer to the harmonic sequence. This confirms the intuition of Faliszewski et al. [24] that this sequence achieves a good excellence/diversity compromise.

Finally, we consider the results for the twin peaks distributions (Figure 8). We find OWA vectors that consist of 1s followed by 0s. This means that the rules that we found are, in essence, the t-Borda rules of Faliszewski et al. [24] (for a given t, the t-Borda rule uses Borda scoring function and OWA vector of t 1s followed by 0s). Faliszewski et al. studied these rules in their search for excellence/diversity compromises, but concluded that they do not seem to work well for this case. The fact that they implement the twin peaks utopic distribution supports this conclusion.

Results for Disc Elections. Generally, our rules behave similarly on interval and disc elections, with the only exception of the rule from Figure 8 (1), which creates three peeks in the interval elections, but which does not select winners from the center in disc elections.

5 Conclusions

We have developed a methodology for designing multiwinner voting rules whose winning committees have properties specified via distributions on a 1D interval. Testing our method on weakly separable and OWA/Borda-based committee scoring rules, we confirmed many intuitions about the applicability of certain rules for certain tasks and discovered new rules.

Our work is a proof of concept and shows that our approach is indeed feasible. In particular, we have focused entirely on generating our rules using 1D Euclidean elections where voters' and candidates' ideal points are selected uniformly at random. In certainly would be interesting to consider how the results change when we vary these distributions (for the goals of individual excellence and diversity we would not except much difference, but modeling, e.g., proportionality might require more involved distributions).

One of the main motivations for our work was to seek rules that find diverse committees. Promoting diversity is an important thread in recent research in computational social choice [5, 25], not only in voting theory [1, 6, 35, 28], and it may be useful to merge the ideas from various papers that arise in this context.

Acknowledgments. Piotr Faliszewski was supported by the National Science Centre, Poland, under project 2016/21/B/ST6/01509. Stanisław Szufa was supported by the National Science Centre, Poland, under project 2016/21/B/HS5/00437.

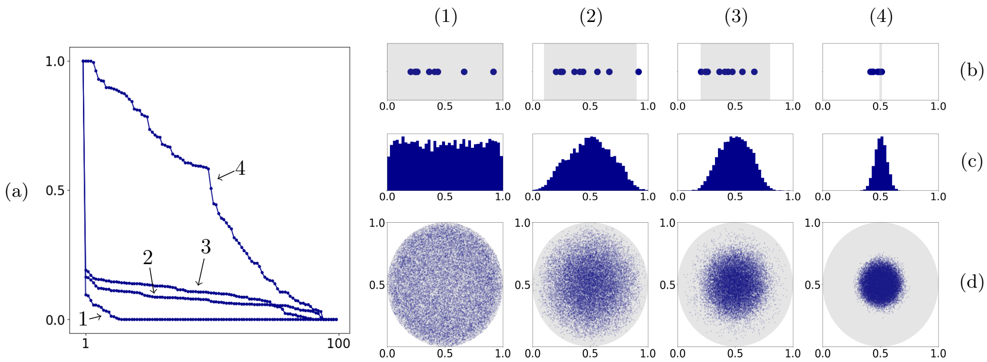


Figure 4: Results for weakly separable rules and utopic distributions (1) $\mathcal{U}_D$, (2) $\mathcal{U}_D^{0.1}$, (3) $\mathcal{U}_D^{0.2}$, and (4) $\mathcal{U}_{IE}$. Plots (a)–(d) have meaning as in Figure 1.

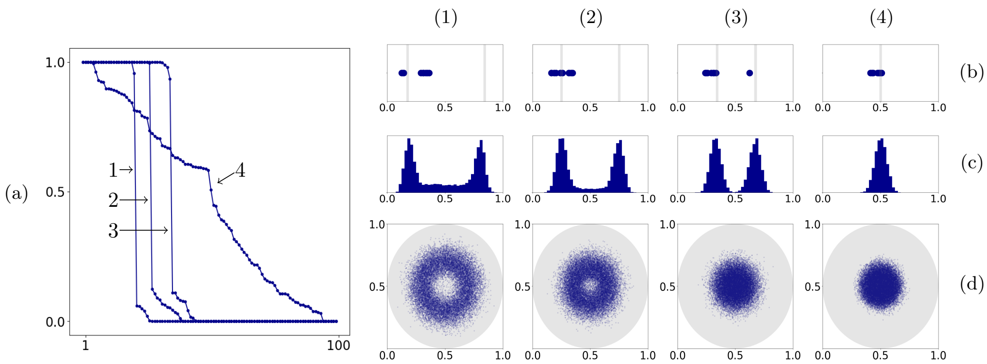


Figure 5: Results for weakly separable rules and utopic distributions (1) $\mathcal{U}_{TP}^{0.167}$, (2) $\mathcal{U}_{TP}^{0.25}$, (3) $\mathcal{U}_{TP}^{0.333}$, and (4) $\mathcal{U}_{IE}$. Plots (a)–(d) have meaning as in Figure 1.

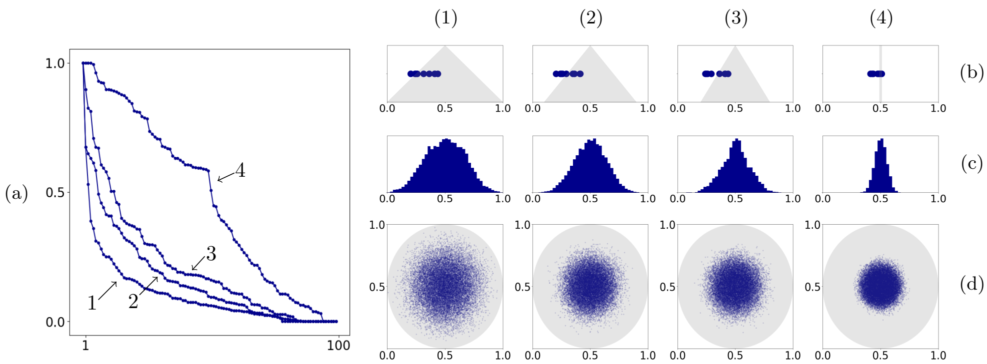


Figure 6: Results for weakly separable rules and utopic distributions (1) $\mathcal{U}_T^0$, (2) $\mathcal{U}_T^{0.1}$, (3) $\mathcal{U}_T^{0.2}$, and (4) $\mathcal{U}_{IE}$. Plots (a)–(d) have meaning as in Figure 1.

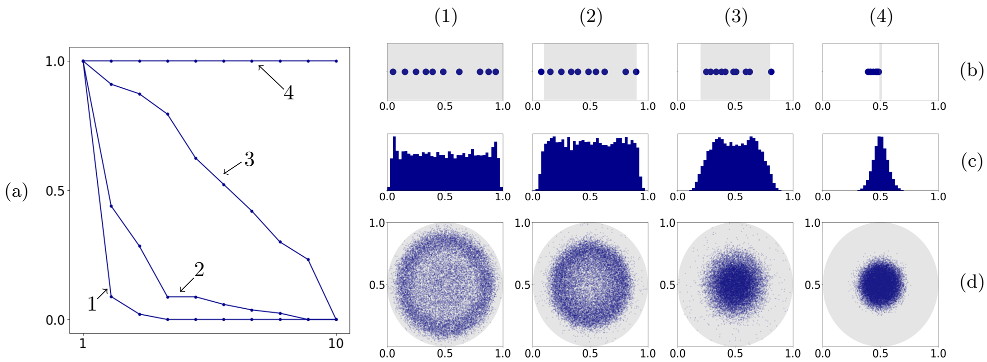


Figure 7: Results for OWA/Borda-based rules and utopic distributions (1) $\mathcal{U}_D$, (2) $\mathcal{U}_D^{0.1}$, (3) $\mathcal{U}_D^{0.2}$, and (4) $\mathcal{U}_{IE}$. Plots (a)–(d) have meaning as in Figure 2.

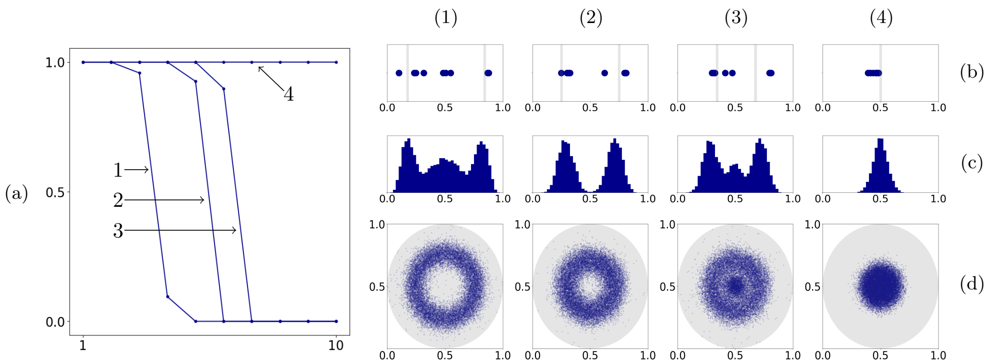


Figure 8: Results for OWA/Borda-based rules and utopic distributions (1) $\mathcal{U}_{TP}^{0.167}$, (2) $\mathcal{U}_{TP}^{0.25}$, (3) $\mathcal{U}_{TP}^{0.333}$, and (4) $\mathcal{U}_{IE}$. Plots (a)–(d) have meaning as in Figure 2.

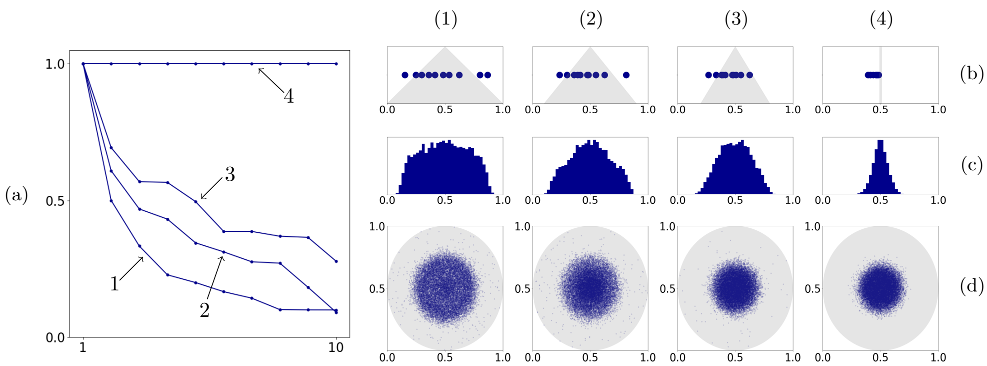


Figure 9: Results for OWA/Borda-based rules and utopic distributions (1) $\mathcal{U}_T^0$, (2) $\mathcal{U}_T^{0.1}$, (3) $\mathcal{U}_T^{0.2}$, and (4) $\mathcal{U}_{IE}$. Plots (a)–(d) have meaning as in Figure 2.

References

[1] F. Ahmed, J. Dickerson, and M. Fuge. Diverse weighted bipartite b-matching. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 35–41, 2017.

[2] H. Aziz, M. Brill, V. Conitzer, E. Elkind, R. Freeman, and T. Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485, 2017.

[3] H. Aziz, E. Elkind, P. Faliszewski, M. Lackner, and P. Skowron. The condorcet principle for multiwinner elections: From shortlisting to proportionality. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 84–90, 2017.

[4] D. Baumeister, S. Dennisen, and L. Rey. Winner determination and manipulation in minisum and minimax committee elections. In *Proceedings of the 4th International Conference on Algorithmic Decision Theory*, pages 469–485, 2015.

[5] D. Baumeister, P. Faliszewski, A. Laruelle, and T. Walsh. Voting: Beyond simple majorities and single-winner elections (Dagstuhl Seminar 17261). Technical Report Dagstuhl Reports 7(6), Leibniz-Zentrum für Informatik, 2017.

[6] N. Benabbou, M. Chakraborty, V. Xuan, J. Sliwinski, and Y. Zick. Diversity constraints in public housing allocation. Technical Report arXiv:1711.10241 [cs.AI], arXiv.org, 2017.

[7] N. Betzler, A. Slinko, and J. Uhlmann. On the computation of fully proportional representation. *Journal of Artificial Intelligence Research*, 47:475–519, 2013.

[8] O. Cailloux and U. Endriss. Arguing about voting rules. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pages 287–295, 2016.

[9] I. Caragiannis, S. Nath, A. D. Procaccia, and N. Shah. Subset selection via implicit utilitarian voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, pages 151–157, 2016.

[10] B. Chamberlin and P. Courant. Representative deliberations and representative decisions: Proportional representation and the Borda rule. *American Political Science Review*, 77(3):718–733, 1983.

[11] V. Conitzer, M. Rognlie, and L. Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, pages 109–115. AAAI Press, July 2009.

[12] M. Cygan, L. Kowalik, A. Socala, and K. Sornat. Approximation and parameterized complexity of minimax approval voting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 459–465, 2017.

[13] B. Debord. An axiomatic characterization of Borda's k-choice function. *Social Choice and Welfare*, 9(4):337–343, 1992.

[14] M. Diss and A. Doghmi. Multi-winner scoring election methods: Condorcet consistency and paradoxes. Technical Report WP 1613, GATE Lyon Saint-Étienne, March 2016.

[15] E. Elkind, P. Faliszewski, J. Laslier, P. Skowron, A. Slinko, and N. Talmon. What do multiwinner voting rules do? An experiment over the two-dimensional euclidean domain. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 494–501, 2017. Full version available at `http://home.agh.edu.pl/~faliszew/2d.pdf`.

[16] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. Properties of multiwinner voting rules. *Social Choice and Welfare*, 48(3):599–632, 2017.

[17] E. Elkind, P. Faliszewski, and A. Slinko. Distance rationalization of voting rules. *Social Choice and Welfare*, 2015. To appear.

[18] E. Elkind, J. Lang, and A. Saffidine. Condorcet winning sets. *Social Choice and Welfare*, 44(3):493–517, 2015.

[19] J. M. Enelow and M. J. Hinich. *The spatial theory of voting: An introduction.* CUP Archive, 1984.

[20] J. M. Enelow and M. J. Hinich. *Advances in the spatial theory of voting.* Cambridge University Press, 1990.

[21] P. Faliszewski, M. Lackner, D. Peters, and N. Talmon. Effective heuristics for committee scoring rules. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, February 2018. To appear.

[22] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Committee scoring rules: Axiomatic classification and hierarchy. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, pages 250–256, 2016.

[23] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Multiwinner analogues of the plurality rule: Axiomatic and algorithmic views. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 482–488, 2016.

[24] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Multiwinner rules on paths from k-Borda to Chamberlin–Courant. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 192–198, 2017.

[25] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Multiwinner voting: A new challenge for social choice theory. In U. Endriss, editor, *Trends in Computational Social Choice*. AI Access Foundation, 2017.

[26] M. Kilgour. Approval balloting for multi-winner elections. In *Handbook on Approval Voting*. Springer, 2010. Chapter 6.

[27] M. Lackner and P. Skowron. Consistent approval-based multi-winner rules. Technical Report arXiv:1704.02453 [cs.GT], April 2017.

[28] J. Lian, N. Mattei, R. Noble, and T. Walsh. The conference paper assignment problem: Using order weighted averages to assign indivisible goods. Technical Report arXiv:1705.06840 [cs.AI], arXiv.org, 2017.

[29] T. Lu and C. Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 280–286, 2011.

[30] B. Monroe. Fully proportional representation. *American Political Science Review*, 89(4):925–940, 1995.

[31] S. Peleg, M. Werman, and H. Rom. A unified approach to the change of resolution: Space and gray-level. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):739–742, 1989.

[32] D. Peters. Single-peakedness and total unimodularity: Efficiently solve voting problems without even trying. Technical Report arXiv:1609.03537 [cs.GT], September 2016.

[33] A. Procaccia, J. Rosenschein, and A. Zohar. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30(3):353–362, 2008.

[34] L. Sánchez-Fernández, E. Elkind, M. Lackner, N. Fernández, J. A. Fisteus, P. Basanta Val, and P. Skowron. Proportional justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 670–676, 2017.

[35] C. Schumann, S. Counts, J. Foster, and J. Dickerson. The diverse cohort selection problem: Multi-armed bandits with varied pulls. Technical Report arXiv:1709.03441 [cs.LG], arXiv.org, 2018.

[36] S. Sekar, S. Sikdar, and L. Xia. Condorcet consistent bundling with social choice. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems*, pages 33–41, 2017.

[37] P. Skowron, P. Faliszewski, and J. Lang. Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence*, 241:191–216, 2016.

[38] P. Skowron, P. Faliszewski, and A. Slinko. Axiomatic characterization of committee scoring rules. Technical Report arXiv:1604.01529 [cs.GT], arXiv.org, April 2016.

[39] P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. Proportional rankings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 409–415, 2017.

[40] T. Thiele. Om flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, pages 415–441. 1895.

[41] L. Xia. Designing social choice mechanisms using machine learning. In *Proceedings of AAMAS 2013*, pages 471–474, 2013.

Piotr Faliszewski
AGH University of Science and Technology
Krakow, Poland
Email: `faliszew@agh.edu.pl`

Stanislaw Szufa
Jagiellonian University
Krakow, Poland
Email: `s.szufa@gmail.com`

Nimrod Talmon
Ben-Gurion University of the Negev
Be'er Sheva, Israel
Email: `nimrodtalmon77@gmail.com`