What Do Multiwinner Voting Rules Do? An Experiment Over the Two-Dimensional Euclidean Domain

Edith Elkind
University of Oxford
Oxford, UK

Piotr Faliszewski
AGH University
Krakow, Poland

Jean-François Laslier
Paris School of Economics
Paris, France

Piotr Skowron
University of Oxford
Oxford, UK

Arkadii Slinko
University of Auckland
Auckland, New Zealand

Nimrod Talmon
Weizmann Institute of Science
Rehovot, Israel

Abstract

We visualize aggregate outputs of popular multiwinner voting rules—SNTV, STV, Bloc, k-Borda, Monroe, Chamberlin–Courant, and PAV—for elections generated according to the two-dimensional Euclidean model. We consider three applications of multiwinner voting, namely, parliamentary elections, portfolio/movie selection, and shortlisting, and use our results to understand which of our rules seem to be best suited for each application. In particular, we show that STV (one of the few nontrivial rules used in real high-stake elections) exhibits excellent performance, whereas the Bloc rule (also often used in practice) performs poorly.

Introduction

The goal of this paper is to develop a better understanding of a number of well-known multiwinner voting rules, by analyzing their behavior in elections where voters' preferences are generated according to a two-dimensional spatial model. By focusing on this preference domain, we can visualize the election results and check if they agree with the intuition and motivation behind these rules. Our study can be seen as an experimental counterpart of the work of of Elkind et al. (2014; 2017), who analyze multiwinner rules axiomatically.

In a multiwinner election, the goal is to select a size-k committee (i.e., a set of k candidates, where $k \in \mathbb{N}$ is part of the input) based on the voters' preferences. Usually, voters can express their preferences by listing the candidates from best to worst or by indicating which candidates they approve; we focus on the former setting, as it fits the spatial preference model better.

Applications of multiwinner voting range from choosing a parliament through preparing a portfolio of company's products (Lu and Boutilier 2011; 2015) or choosing movies to offer to passengers on a long flight (Elkind et al. 2014; 2017; Skowron, Faliszewski, and Lang 2016) to shortlisting runners-up for an award (Barberà and Coelho 2008; Elkind et al. 2014; 2017). As a consequence, there is also quite a variety of different multiwinner voting rules. For instance, for parliamentary elections an important desideratum is proportional representation of the voters, and there are voting rules such as STV or the Monroe rule (we define all rules considered in this paper in the next section) that have been designed with this idea in mind. On the other hand, in the context of portfolio or movie selection we primarily care about the diversity of the selected committee, and it has been argued that the Chamberlin–Courant rule is good for this purpose (Lu and Boutilier 2011; Skowron, Faliszewski, and Lang 2016). For shortlisting, our primary concern is fairness: if there are two similar candidates, we want to select both or neither, and increasing the target committee size should not result in any of the selected candidates being dropped; these requirements are satisfied by k-Borda. Naturally, there are other scenarios which require other normative properties.

The examples above indicate that choosing a good multiwinner rule is not a trivial task. It is therefore natural to ask how we can facilitate the decision-making process of a user who is facing this choice. There are several good answers to this question. First, some rules are specifically designed for certain tasks. For example, STV and the Monroe rule have explicit built-in mechanisms ensuring that every sufficiently large group of like-minded voters is represented. Second, we can analyze axiomatic properties of the rules. This line of work, was extensively pursued for single-winner rules; for the case of multiple winners in was initiated by Felsenthal and Maoz (1992) and Debord (1992), with recent contributions including the work of Elkind et al. (2014; 2017) and Aziz et al. (2015a; 2017). Finally, one can use empirical analysis to compare different rules under particular conditions. For example, Diss and Doghmi (2016) consider a few multiwinner voting rules and experimentally investigate how frequently they pick Condorcet committees.[1] All these approaches are useful, and the choice of a voting rule should take all of them into account.

Nonetheless, a non-expert user may still feel ill at ease when deciding which rule to choose for his or her particular application. In this case, a picture may be worth a thousand words: a simple graph that clearly explains differences between rules can be very informative. The contribution of this paper is to propose a novel approach to selecting a suitable mutiwinner rule, which is based on graphical informa-

[1] In a Condorcet committee, every committee member is preferred to every non-member by a majority of the voters.

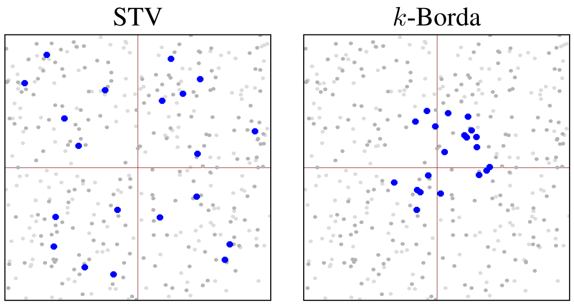

Figure 1: Results of an election (generated using the 2D Euclidean model) according to STV (left) and k-Borda (right). Voters are depicted as dark gray dots, candidates as light gray dots, and the winners as larger blue dots.

tion. That is, we provide images that we expect to be helpful in discussions of multiwinner voting rules. Naturally, reality is too complicated for a single picture to constitute a definite argument, but we believe that, on the one hand, our results provide good illustrations confirming intuitions regarding various multiwinner rules and, on the other hand, they highlight some faults of the rules that otherwise would not be easily visible.

Our Methodology. The outcome of an election depends both on the voting rule and on the set of candidates. In this work, we focus on the former aspect and ask what multiwinner rules do when choosing from a set of candidates that is representative of the electorate, i.e., under what one may call the *representative candidacy* assumption. We evaluate a number of multiwinner voting rules (SNTV, STV, Bloc, k-Borda, Chamberlin–Courant, Monroe, and PAV) on elections generated using the two-dimensional Euclidean model of preferences. In this model each candidate and each voter is represented by a point on a plane, and voters form their preference orders by ranking the candidates that are closer to them above the ones that are further away.

This model is very appealing and extensively studied (Enelow and Hinich 1984; 1990) because of its natural interpretations: A point representing a candidate or a voter simply specifies his or her position regarding two given issues. In the world of politics, these two issues could be, for example, the preferred levels of taxation and immigration, or the extent to which the individual believes in personal and economic freedom. While in some settings more dimensions may be necessary, the popularity of the Nolan Chart, which is used to represent the spectrum of political opinions, indicates that two dimensions are often sufficient to provide a good approximation of voters' preferences.

In Figure 1 we show a sample election (the points for candidates and voters are generated using uniform distribution over a square) and the committees selected by STV (left) and k-Borda (right). It is quite evident that the committee on the left would form a far more representative parliament than the one on the right, whereas the one on the right would probably be a better choice for the set of candidates that are shortlisted for a position, since they are similar to each other and receive broad support among the voters (in particular, no voter ranks them close to the bottom of their list).

Our main contributions are as follows:

1. For each of our voting rules and four distributions of candidates and voters (Gaussian, uniform on a disc, uniform on a square, and a mix of four Gaussians), we have generated $10\,000$ elections and built histograms (Figure 3) indicating how likely it is that a candidate from a given position will be selected.

2. We consider three applications of multiwinner voting, and, for each application, we identify the voting rules in our collection that are most appropriate for it. We make these recommendations based on our histograms and certain statistical properties of the elected committees. E.g., we confirm that STV is an excellent rule for parliamentary elections, even superior to the Monroe rule; PAV can also be seen as an interesting rule that chooses fairly representative committees, ignoring candidates with extreme opinions. We also provide evidence that Bloc should be treated very carefully since it may not perform as well as one might expect (this is particularly important because Bloc is among the most popular multiwinner rules).

Due to space restrictions, we omit some of our results (in particular, the analysis of approximation algorithms for the Monroe and Chamberlin–Courant rules); these results are available in the appendix.

Preliminaries

For every positive integer n, we write $[n]$ to denote the set $\{1, \dots, n\}$.

Elections. An *election* $E = (C, V)$ consists of a set $C = \{c_1, \dots, c_m\}$ of *candidates* and a list $V = (v_1, \dots, v_n)$ of *voters*. Each voter v_i has a *preference order* $\succ_i$, i.e., a ranking of the candidates from the most to the least favored one (according to this voter). 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 (where the top-ranked candidate has position 1). A *committee* is a subset of C.

A *multiwinner voting rule* is a function $\mathcal{R}$ that, given an election $E = (C, V)$ and a target committee size k ($1 \leq k \leq |C|$), outputs a nonempty set of size-k committees; these committees are said to *tie as election winners*. In practice, one has to use some tie-breaking mechanism. For our experiments, whenever we need to break a tie (possibly at an intermediate stage in the execution of the rule), we make a random choice with a uniform distribution over all possibilities.

(Single-Winner) Scoring Functions. For an election with m candidates, a *scoring function* γ_m associates each position j, $j \in [m]$, with a *score* $\gamma_m(j)$. The *γ_m-score that candidate c receives from voter v* is $\gamma_m(\mathrm{pos}_v(c))$. The *$\gamma_m$-score of candidate c in election E* is the sum of the γ_m-scores that c receives from the voters in E. We consider the following two prominent families of scoring functions:

1. The Borda scoring function, β_m, is defined as $\beta_m(j) = m - j$.

2. For each $t \in [m]$, the t-Approval scoring function, α_t, is defined as $\alpha_t(j) = 1$ if $j \leq t$ and $\alpha_t(j) = 0$ otherwise. The candidate's 1-Approval score is known as her Plurality score.

Multiwinner Rules. We focus on the following multiwinner rules (in the description below we consider an election $E = (C, V)$ and committee size k):

SNTV. The Single Nontransferable Vote rule (SNTV) outputs k candidates with the highest Plurality scores.

STV. The Single Transferable Vote rule (STV) executes a series of iterations, until it finds k winners. A single iteration operates as follows: If there is at least one candidate with Plurality score at least $q = \lfloor \frac{n}{k+1} \rfloor + 1$, then a candidate with the highest Plurality score is added to the committee; then q voters that rank him or her first are removed from the election (our randomized tie-breaking plays an important role here), and the selected candidate is removed from all voters' preference orders. If there is no such candidate, then a candidate with the lowest Plurality score is removed from the election (again, ties are broken uniformly at random). The Plurality scores are then recomputed.

Bloc. Under the Bloc rule we output k candidates with the highest k-Approval scores (intuitively, each voter is asked to name his or her k favorite committee members, and those mentioned most frequently are elected).

k-Borda. Under the k-Borda rule we output k candidates with the highest Borda score.

β-CC. The (classical) Chamberlin–Courant rule (β-CC) is defined as follows (Chamberlin and Courant 1983). A *k-CC-assignment function* is a function $\Phi\colon V \rightarrow C$ such that $|\Phi(V)| \leq k$ (i.e., Φ associates each voter with a candidate in a set $W \subseteq C$, $|W| \leq k$; for a voter v, candidate $\Phi(v)$ is referred to as *v's representative*). The *β-CC score* of an assignment Φ is defined as $\beta(\Phi) = \sum_{v \in V} \beta_m(\mathrm{pos}_v(\Phi(v)))$ (i.e., it is the sum of the Borda scores of voters' representatives). β-CC finds a k-CC-assignment Φ that maximizes $\beta(\Phi)$ and outputs the committee $\Phi(V)$ (if it happens that $|\Phi(V)| < k$—a situation that occurs, e.g., when all the voters have identical preference orders—then β-CC supplements $\Phi(V)$ with $k - |\Phi(V)|$ candidates selected at random).

β-Monroe. The (classical) Monroe rule (Monroe 1995) is similar to β-CC, except that it is restricted to k-Monroe-assignments. A *k-Monroe-assignment* is a k-CC-assignment that satisfies the following constraints: (a) $|\Phi(V)| = k$, and (b) for each candidate c such that $\Phi^{-1}(c) \neq \emptyset$ (i.e., for each selected representative) it holds that $\lfloor \frac{n}{k} \rfloor \leq |\Phi^{-1}(c)| \leq \lceil \frac{n}{k} \rceil$. Intuitively, under the Monroe rule each selected candidate represents, roughly, the same number of voters.

α_k-PAV. Consider a scoring function γ. For a voter v and a committee W such that v ranks the members of W on positions $p_1 < \cdots < p_k$, the *γ-PAV score* that v assigns to W is $\gamma\text{-PAV}(W, v) = \sum_{t=1}^{k} \frac{1}{t}\gamma(p_t)$. For an election $E = (C, V)$, the *γ-PAV score* of a committee W is defined as $\gamma\text{-PAV}(W, V) = \sum_{v \in V} \gamma\text{-PAV}(W, v)$. The rule outputs a committee with the highest γ-PAV score. In this paper, we consider α_k-PAV (originally the rule was defined for approval ballots, see e.g., the overview of Kilgour (2010); as we work with preference orders, we modify the definition accordingly).

With our tie-breaking, STV, SNTV, Bloc, and k-Borda are computable in polynomial time using straightforward algorithms. Unfortunately, the Chamberlin–Courant and Monroe rules are NP-hard to compute (Procaccia et al. (2008) show this for variants of these rules that use t-Approval scores α_t instead of β; for the Borda-based variants defined here, the results for the Chamberlin–Courant rule and the Monroe rule are due to Lu and Boutilier (2011) and Betzler et al. (2013), respectively). We compute these rules by solving their integer linear programming (ILP) formulations (suggested by Lu and Boutilier (2011) for the case of Chamberlin–Courant, and by Skowron et al. (2015) for the case of Monroe). PAV is also NP-hard to compute (Aziz et al. 2015b; Skowron, Faliszewski, and Lang 2016)[2], and we use a simplified version of the ILP formulation proposed by Skowron et al. (2016); see the appendix.

Euclidean Preferences. Given two points on the plane, $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$, we write $d(p_1, p_2)$ to denote the distance $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ between them.

In a two-dimensional Euclidean election $E = (C, V)$, each entity e (i.e., either a candidate or a voter) is associated with a point $p(e) = (x(e), y(e))$. Given a pair of candidates $c_i, c_j \in C$, a voter $v \in V$ prefers c_i to c_j if $d(p(v), p(c_i)) < d(p(v), p(c_j))$. Note that this condition does not constrain voter's preferences over two equidistant candidates. In our case, since we draw our elections at random, such situations are unlikely to happen. When they do, we break the tie arbitrarily.

Euclidean preferences are very useful to realistically model political preferences and, in many cases, to model preferences in shortlisting tasks. Unfortunately, they are not nearly as useful for modeling preferences over movies. The reason is that people often do not have a single most favorite type of a movie, but rather like various genres for different reasons. Nonetheless, investigating rules meant for the movie selection application (i.e., for selecting diverse committees) in our framework is still important. On the one hand, movie selection is not the only application where diverse committees are needed, and, on the other hand, if a rule behaves badly on the Euclidean domain, then it is unlikely that it would behave well for richer preference models.

Main Results and Analysis

Experimental Setup. We assume that both the candidates and the voters have ideal positions in a two-dimensional Euclidean issue space that are drawn from the same distributions. For each voting rule and each distribution, we generated 10 000 elections, each with $m = 200$ candidates and

[2]the hardness proofs for this rule are in the approval model, but can be easily adapted to the preference-order based one.

$n = 200$ voters, and for each of them we computed a winning committee of size $k = 20$.

We consider four distributions of the ideal positions:

Gaussian. Ideal points are generated using symmetric Gaussian distribution with mean $(0, 0)$ and standard deviation 1.

Uniform Square. Ideal points are distributed uniformly on the square $[-3, 3] \times [-3, 3]$.

Uniform Disc. Ideal points are distributed uniformly on the disc with center $(0, 0)$ and radius 3.

4-Gaussian. Ideal points are generated using four symmetric Gaussian distributions with standard deviation 0.5, but different mean values, namely, $(-1, 0)$, $(1, 0)$, $(0, -1)$ and $(0, 1)$; each mean is used to generate 25% of the points.

We use the Gaussian distribution to model a society with one dominant idea (e.g., where being moderate is the most popular position, or where a single dominant party exists). Since the boundary plays a significant role in the case of uniform distributions (we will discuss this effect below), we have chosen the Gaussian distribution, as its density vanishes close to the boundary.

The 4-Gaussian distribution models a structured society, with four well-established positions (for the movie selection scenario, these might correspond to, e.g., a combination of two genres and two typical budget values; in the world of politics, these could be four political parties).

We also use the uniform distributions, on a square and on a disc, as intermediate cases, and in order to study specific behavior of voting rules at the border and, in case of the square, at the corners of the support of the distribution.

Raw Results. For each rule and each distribution, we have computed a histogram, showing how frequently winners from a given location were selected. These histograms, together with examples of elections and their winning committees, are presented in Figure 3 (the first row presents the distributions themselves).

The histograms were generated as follows. For each rule and distribution, all the winners were always within the $[-3, 3] \times [-3, 3]$ square. We have partitioned this square into 120×120 cells (each cell is a 0.05×0.05 square), and—for each given distribution and rule—counted how many times a member of the winning committee fell into a given cell (we refer to this value as the *frequency* of this cell). Then we have transformed the frequencies into color intensities (the more winners fall into a particular cell, the darker it is in Figure 3). Since there are big differences among frequencies of cells across various rules and distributions (e.g., the highest frequency of a cell for k-Borda with the Gaussian distribution is over 27 times larger than the highest frequency of a cell for SNTV under the uniform square distribution), we took the following approach. Given a cell of frequency x, we compute its color intensity y ($0 \leq y \leq 1$; the closer is y to 1 the darker is the cell) using the following formula:

$$y = \frac{1}{\pi/2} \arctan\left(\frac{x}{\varepsilon T}\right), \qquad (1)$$

where T is the sum of the frequencies of all the cells (so in our case $T = 20 \cdot 10000$) and ε is a parameter. We used $\varepsilon = 0.0004$, so for the highest frequency of a cell in all our experiments (found for k-Borda with the Gaussian distribution) we have $x/(\varepsilon T) = 10.9$; for most other rules and distributions this value is below 1.5 and thus falls into the part where our function behaves fairly linearly (see Figure 2). To present the distributions themselves, we computed histograms of the ideal points generated using our distributions (on the technical side, to generate these histograms, we used candidate positions from $10\,000$ generated elections for each distribution; since formula (1) is normalized, the pictures in the first row of Figure 3 are comparable to those in the other rows).

Analysis. We now consider the three applications of multiwinner rules that we mentioned in the introduction and analyze which of our rules are most suitable for each application.

Parliamentary Elections. We start with the case of parliamentary elections. Intuitively, in this application we value proportional representation, which requires that the distribution of the winners (as seen through the histograms) should be as close as possible to the underlying distribution of the voters. Thus, at first sight, among our rules SNTV would be the champion in this category. In addition, SNTV satisfies a number of axioms studied by Elkind et al. (2014; 2017), especially those geared towards proportional representation. However, at the same time, it is intuitively clear that SNTV is not a very good rule because it only takes the voters' top choices into account, thus ignoring most of the information in voters' preferences. A look at the sample elections for SNTV (Figure 3) shows that this intuition is correct: The reason why SNTV has such an appealing histogram is that it selects committee members in areas that, by random chance, have above-average density of voters and below-average density of candidates. Over all $10\,000$ elections such areas are distributed evenly, similarly to the distribution of the candidates and voters.

This means that, in addition to considering the histograms, we also need to check if results of individual elections are close to what the histograms show. To this end we have used an indirect approach that, nonetheless, turned out to be very effective. Let us fix some rule $\mathcal{R}$ and one of our

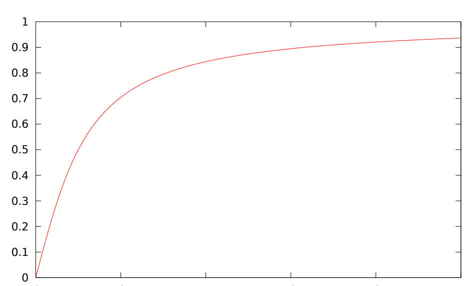

Figure 2: Plot of the function $y = \frac{1}{\pi/2} \arctan\left(\frac{x}{\varepsilon T}\right)$ that we use for converting cell frequencies to color intensities.

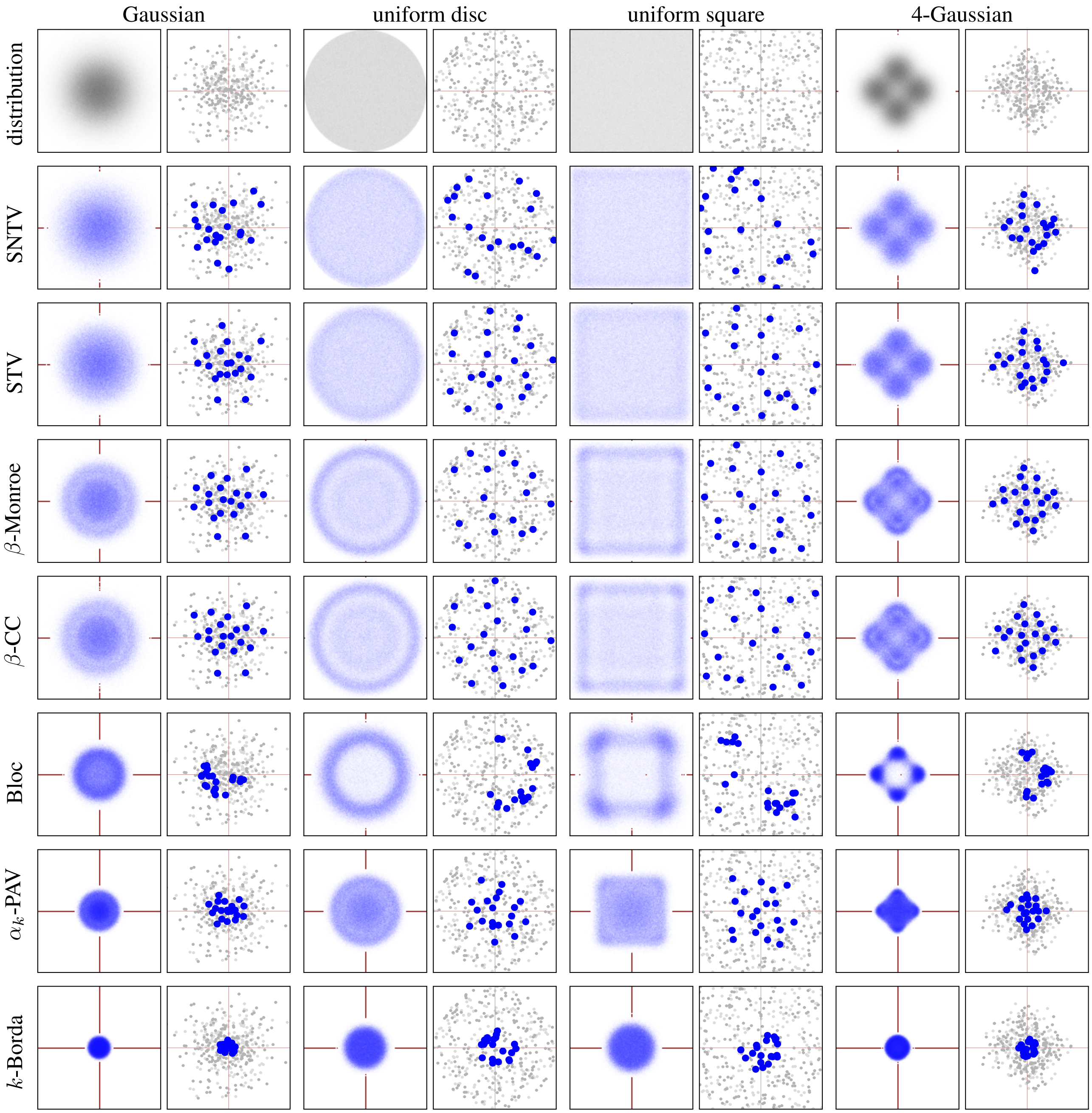

Figure 3: Histograms and sample elections for our rules and distributions. The first row shows the distributions only. For sample election, voters are depicted as dark gray dots, candidates as light gray dots, and the winners as larger blue dots.

distributions. For each generated election, we (1) count how many members of the winning committee are in each of the four quadrants $[0, \pm\infty) \times [0, \pm\infty)$, (2) collect these numbers in a sequence, and (3) compute the variance of this sequence; Table 1 shows the result of this computation, averaged over all instances. Since all our distributions are symmetric with respect to the x and y axes, for rules that represent voters proportionally *in individual instances* we expect this number to be small. Of course, the converse claim need not be true: Low variance does not guarantee proportional representation. That is, the variance-based approach can be used to eliminate 'bad' rules rather than to identify 'good' rules.

Table 1 clearly identifies a group of rules for which the variance of the number of winners per quadrant is close to or below 1.0, whereas for other rules the variance is significantly higher (in our experiments, typically close to or above 3.0). Thus, the performance of SNTV (close to 3.0) is a strong argument against it. On the other hand, the results

rule	square	disc	Gauss.	4 Gauss.
SNTV	**3.292**	**3.219**	**3.275**	**2.787**
STV	0.994	1.070	1.150	1.043
β-Monroe	0.738	0.797	0.864	0.765
β-CC	0.765	0.820	0.866	0.826
Bloc	**17.789**	**17.146**	**18.709**	**9.663**
α_k-PAV	1.323	1.391	1.463	1.289
***k*-Borda**	**4.605**	**4.653**	**4.736**	**3.653**

Table 1: Variance of the number of winners in each quadrant. Bold font indicates rules where this value suggests asymmetric placement of winners on the plane (for k-Borda, this turns out to be a false alarm).

for STV (both the shape of histograms and the variance) indicate that it is an exceedingly good rule for selecting parliaments. Indeed, this is the only rule with low variance that is computationally tractable. This is quite important, as STV is among just a few nontrivial voting rules used in practice, yet some researchers—including some of us, until recently—consider it unappealing. The axiomatic results of Elkind et al. (2014; 2017) and our experiments provide different arguments in favor of using STV for proportional representation.

The results for β-Monroe are slightly less appealing than those for STV. While the variance of the number of winners per quadrant is low, the histograms are farther from resembling the distributions of candidates and voters. They are very similar to those for β-CC, which should not be too surprising. In our experiments, the only difference between these rules is that β-Monroe is forced to assign exactly 10 voters to each selected committee member, whereas β-CC can choose an optimal assignment, where the number of voters assigned to each committee member may be arbitrary. Nonetheless, for each of the distributions, around 80% of the committee members selected by β-CC were assigned to between 7 and 13 voters each. In effect, the assignments computed by β-CC and β-Monroe were quite similar. Naturally, if the distributions of candidates and voters were not identical, the results would be different as well (we have run initial experiments to confirm this, available in the appendix). Below we discuss the intriguing patterns in the histograms for β-CC (a similar explanation applies to β-Monroe).

Portfolio/Movie Selection. Let us now consider the portfolio/movie selection scenario (Lu and Boutilier 2011; 2015; Elkind et al. 2014; 2017; Skowron, Faliszewski, and Lang 2016). Here we care mostly about the diversity of the committee and, intuitively, we would like to obtain histograms that cover a large chunk of the support of the distribution, but which—as compared to the parliamentary elections setting—are less responsive to the densities of the candidates and voters.

We first analyze the results for β-CC, a rule that seems to be designed exactly for this scenario. However, it does not quite fit the description above. As we will see, to some extent this is due to the nature of the rule, and to some extent this is because our initial expectations were not entirely reasonable. There are two main issues regarding β-CC.

The first one concerns what we call the *edge effect* and the *corner effect*. Let us consider the uniform square distribution. If a candidate is located far from the edges, then he or she is also surrounded by a relatively large number of other candidates with whom he or she needs to compete for a high position in voters' preference orders. On the other hand, if a candidate is located near an edge (or, better yet, near a corner) then the competition is less stiff. However, if a candidate is close to the edge/corner, the number of voters for whom he or she would be a representative also decreases. In effect, for the uniform square and uniform disc distributions, we see increased frequencies of winners near (but not exactly on) the edges and corners. The edge and corner effects are visible also for SNTV and STV (though to a lesser extent), and they are very prominent for Bloc (especially in conjunction with cases where an area near edge/corner has an above-average density of voters).

The second issue regarding β-CC is that when some candidate is included in the committee, other candidates that are very close to him or her are unlikely to be selected; indeed, this behavior is quite desirable when one wants to maintain diversity of the committee. This explains why for the uniform square and uniform disc distributions the near-edge area with increased frequencies is surrounded by an area with lower frequencies. This effect also explains the interesting pattern for the 4-Gaussian distribution. Since there are many voters in the centers of the four Gaussians, candidates from these locations are likely to be included in the committee. But this very fact strongly decreases the chances of the candidates that are located just a bit further away from the centers of the Gaussians.

Our visual inspection of the election results for β-CC shows that every single committee appears to be diverse and appealing for the portfolio/movie selection problem (this is also supported by the low value of the variance of the number of winners per quadrant). However, the histograms show that the rule also has an implicit, systematic bias against certain candidates (the nature of this bias depends on the distribution) that users of the rule should take into account.

α_k-PAV also appears to be a very interesting rule for the portfolio/movie selection task (and, perhaps, even for parliamentary elections). In our experiments, α_k-PAV chose committees distributed fairly uniformly in the central areas, ignoring candidates with extreme opinions.

Shortlisting. Here our guiding principle is that the committee should consist of similar candidates (i.e., located close to each other). For this criterion, k-Borda is our rule of choice. In all of the experiments it consistently chose candidates located in the center, close to each other. Table 1 indicates that k-Borda has high variance of the number of winners per quadrant. We believe that this is caused not by any faults of the rule itself, but by a fairly natural statistical property of our distributions. Since k-Borda selects 20 candidates from the center, due to random perturbations, sometimes the central candidates are not distributed over the quadrants in a perfectly balanced way, and our variance-based measure does not take into account the candidates' centrality.

The Strange Case of Bloc. In the situation where k candidates are to be selected (e.g., to a city council), it is quite common to ask the voters to come up with k names (ranked or non-ranked). Bloc, in particular, is quite a popular rule. Our histograms show that Bloc is very sensitive to the edge and corner effects (the pattern is similar to that for β-CC, but the effects are much stronger). Worse yet, Table 1 shows very high variance of the number of winners in each quarter and, indeed, the example elections for Bloc in Figure 3 show very asymmetric placements of the winners. These two arguments by themselves make Bloc a questionable voting rule.

Bloc is also the only rule in our collection that shows the following *inversion* effect: For the Gaussian distribution, the frequencies of the cells near the center (i.e., near the mean of the Gaussian distribution) are lower than the frequencies of the cells in the ring surrounding it. This is a very counterintuitive and unexpected phenomenon: The most popular views in the society are represented less frequently than the not-so-popular ones. We believe that the mechanism behind this effect is similar to that behind the edge/corner effect: Even though the center has the highest density of the voters, it also has the highest density of the candidates, who therefore "steal points away" from each other. As a consequence, the slightly less popular candidates in the ring get enough support (both from some of the voters in the center and from those on the ring and beyond) to be elected.[3]

Robustness of the Results

So far we have considered elections with $m = 200$ candidates, $m = 200$ voters, and committee size $k = 20$ only. Thus it is natural to wonder if our conclusions remain valid as we vary these parameters.

Except for STV and β-Monroe, all our rules belong to the class of committee scoring rules (Elkind et al. 2014; 2017; Faliszewski et al. 2016a), i.e., they define a per-voter score of each possible committee and select committees for which the sums of these scores are the highest. In consequence, the results for these rules should not change significantly with the number of voters (unless this number becomes very small). Since STV and β-Monroe are similar in spirit to committee scoring rules (indeed, STV is similar to SNTV and β-Monroe is very closely related to β-CC), the results for them should be similarly robust.

We also do not expect strong qualitative differences in our results for different numbers of candidates or different committee sizes (again, except for very small values). Nonetheless, we do observe quantitative differences.

In Figure 4 we present histograms for our rules with respect to the disc distribution, for committee sizes 10, 20, and 30 (the histogram for committee size 20 is the same as in Figure 3; we repeat it for the sake of comparison). We note that the results for SNTV and STV are nearly the same irrespective of the committee size.[4] The results for Bloc, α_k-PAV, and k-Borda also look very similar, and the differences are only in the radii of the discs/rings generated by these rules (this is especially natural for k-Borda; as we choose more and more of the centrally located candidates, they form a larger and larger disc). The results for β-CC and β-Monroe for different committee sizes also look similar, but for $k = 10$ (especially for the case of β-CC) the artifacts in the histograms become much more visible (e.g., for $k = 10$ and β-CC, there are two very clearly visible consecutive rings). This indicates that our observations about β-CC and β-Monroe do not necessarily carry over to the case of very small committees.

[3]Indeed, this can be seen as a type of approximate cloning (see the discussion in the papers of of Tideman (1987), Laffond et al. (1996), and Elkind et al. (2011)).

[4]For $k = 30$, the quota for STV is $q = \lfloor \frac{200}{31} \rfloor + 1 = 7$. Thus, in the first 28 stages we remove 196 voters, so the 29th candidate is chosen by 4 voters and the 30th candidate is selected randomly.

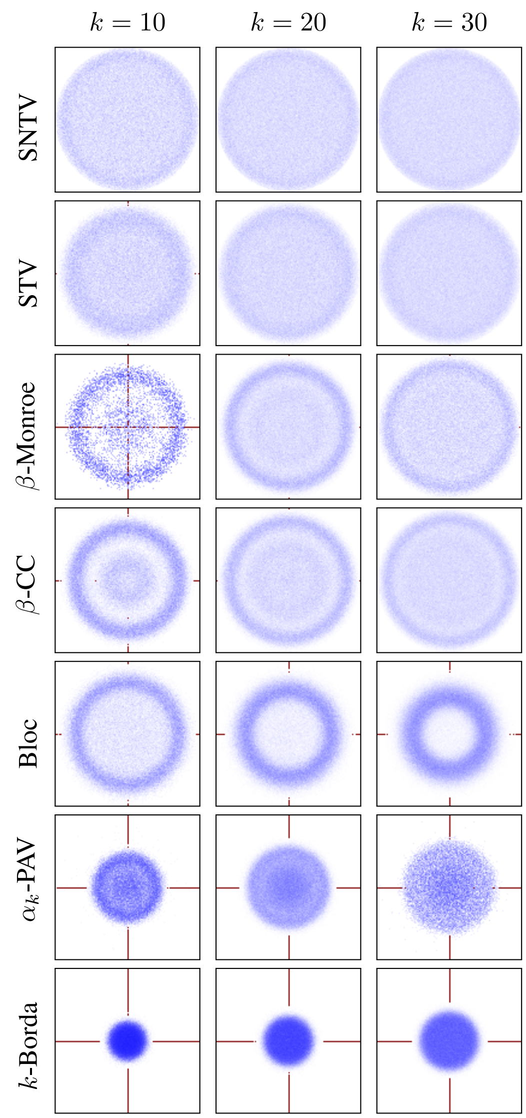

Figure 4: Histograms for our rules under the disc distribution, for committee sizes 10, 20, and 30. For α_k-PAV ($k \in \{10, 30\}$) and Monroe ($k = 30$) we computed only 5000 elections. Due to technical issues, for β-Monroe with $k = 10$ we computed only about 500 elections.

Conclusions

Our results lead to several interesting observations. Foremost, within the framework of our study STV stands out as an exceptionally good rule for parliamentary elections. On the other hand, the Monroe rule, which is also an appealing rule for this application, did not do quite as well. We also found that the Monroe and Chamberlin–Courant rules may have (somewhat surprising) implicit biases against some candidates. Further, we discovered that in our experiments α_k-PAV tends to ignore extremist candidates and fairly uniformly covers central areas (this seems quite related to the results of Aziz et al. (2015a; 2017) on justified representation). We confirmed that k-Borda has good properties as a shortlisting rule and provided strong arguments against the Bloc rule.

Acknowledgments. Edith Elkind and Piotr Skowron were supported by the ERC grant 639945 (ACCORD), Piotr Faliszewski was supported by the NCN grant 2016/21/B/ST6/01509, Arkadii Slinko was supported in part by the Marsden Fund 3706352 of The Royal Society of New Zealand, and Nimrod Talmon was supported by a postdoctoral fellowship from I-CORE ALGO. Jean-François Laslier thanks the ANR project ANR13-BSH1-0010 DynaMITE.

References

Aziz, H.; Brill, M.; Conitzer, V.; Elkind, E.; Freeman, R.; and Walsh, T. 2015a. Justified representation in approval-based committee voting. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 784–790.

Aziz, H.; Gaspers, S.; Gudmundsson, J.; Mackenzie, S.; Mattei, N.; and Walsh, T. 2015b. Computational aspects of multi-winner approval voting. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, 107–115.

Aziz, H.; Brill, M.; Conitzer, V.; Elkind, E.; Freeman, R.; and Walsh, T. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare*. To appear.

Barberà, S., and Coelho, D. 2008. How to choose a non-controversial list with k names. *Social Choice and Welfare* 31(1):79–96.

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

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

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

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

Elkind, E.; Faliszewski, P.; Skowron, P.; and Slinko, A. 2014. Properties of multiwinner voting rules. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, 53–60.

Elkind, E.; Faliszewski, P.; Skowron, P.; and Slinko, A. 2017. Properties of multiwinner voting rules. *Social Choice and Welfare*. To appear.

Elkind, E.; Faliszewski, P.; and Slinko, A. 2011. Cloning in elections: Finding the possible winners. *Journal of Artificial Intelligence Research* 42:529–573.

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

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

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

Faliszewski, P.; Slinko, A.; Stahl, K.; and Talmon, N. 2016b. Achieving fully proportional representation by clustering voters. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, 296–304.

Felsenthal, D., and Maoz, Z. 1992. Normative properties of four single-stage multi-winner electoral procedures. *Behavioral Science* 37:109–127.

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

Laffond, G.; Laine, J.; and Laslier, J. 1996. Composition consistent tournament solutions and social choice functions. *Social Choice and Welfare* 13(1):75–93.

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

Lu, T., and Boutilier, C. 2015. Value-directed compression of large-scale assignment problems. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 1182–1190.

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

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

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

Skowron, P.; Faliszewski, P.; and Slinko, A. 2015. Achieving fully proportional representation: Approximability result. *Artificial Intelligence* 222:67–103.

Tideman, T. 1987. Independence of clones as a criterion for voting rules. *Social Choice and Welfare* 4(3):185–206.

Appendix

Overview

In this appendix we present some of the results that we had to omit from the main part of the paper due to space restrictions. First, we present approximation algorithms for the Chamberlin–Courant and Monroe rules (including one that is due to this paper) and discuss the results for them. Then we show our preliminary results for a scenario where the distributions of candidates and voters are not similar. Finally, we show our Integer Linear Program (ILP) formulation for α_k-PAV.

Approximation Algorithms

Let us first consider approximation algorithms for β-CC. Recall that if $E = (C, V)$ is an election, k is a committee size, and Φ is a k-CC-assignment, then by $\beta(\Phi)$ we denote the sum of Borda scores that the voters assign to their representatives (with respect to Φ). Given a committee W and election E, by the *grab-your-best* assignment of W to the voters in E we mean the function $\mathrm{gyb}(W, E)$ which assigns to each voter the member of W which this voter ranks highest.

We consider the following three approximation algorithms for β-CC (we use the same notation as in the description of our multiwinner rules; $E = (C, V)$ is the election at hand and k is the committee size):

GreedyCC. The algorithm starts by setting the initial committee W to be empty, and then executes the following k iterations: In each iteration, it extends the committee W with a candidate c (previously not included in W) that maximizes $\beta(\mathrm{gyb}(W \cup \{c\}))$. (In particular, the algorithm always starts by including the candidate with the highest Borda score.) Finally, it outputs the computed committee W. GreedyCC is due to Lu and Boutilier (2011) and guarantees approximation ratio of at least $1 - \frac{1}{e} \approx 0.63$.

Algorithm P. This algorithm proceeds as follows. First, it computes a threshold value $x = \frac{|C|\mathrm{w}(k)}{k}$ (where $\mathrm{w}(\cdot)$ is Lambert's w function; $\mathrm{w}(k)$ is $o(\log k)$). Then it sets the initial committee W to be empty and executes k iterations as follows: In each iteration, it finds a candidate c that is ranked among the top x positions by the largest number of voters. Then it adds c to W and deletes all the voters that rank c among their top x positions. (Thus, the algorithm can be seen as an incarnation of a greedy SetCover algorithm, where voters are items to be covered and each candidate covers those voters that rank him or her among their top x positions). Finally, it outputs W. The algorithm is due to Skowron et al. (2015) and achieves approximation ratio of $1 - \frac{2\mathrm{w}(k)}{k}$. We mention that it is also a basis of a polynomial-time approximation scheme for β-CC.

RangingCC. This is an extension of Algorithm P introduced in this paper. RangingCC computes the committees using Algorithm P for threshold values x between 1 and $\frac{|C|\mathrm{w}(k)}{k}$ and outputs the one with the highest β-CC score.

For β-Monroe, we consider the GreedyMonroe algorithm of Skowron et al. (2015);[5] again, we use the same notation as in the description of multiwinner rules above (so we seek k winners for election $E = (C, V)$ with m candidates and n voters):

GreedyMonroe. The algorithm starts by setting W to be the empty committee. Then it constructs a Monroe assignment iteratively as follows (for simplicity, let us assume that k divides n). At the beginning of each iteration, the algorithm finds a candidate c and $\frac{n}{k}$ voters, denoted by $V(c)$, that jointly maximize the Borda score of c in the election $(C, V(c))$. Then, the algorithm adds c to W, assigns c to each voter in $V(c)$, and removes the voters in $V(c)$ from further considerations. The algorithm guarantees an approximation ratio of $1 - \frac{k-1}{2(m-1)} - \frac{H_k}{k}$, where $H_k = 1 + \frac{1}{2} + \cdots + \frac{1}{k}$ is the k'th harmonic number.

Results for Approximation Algorithms

The histograms for the approximation algorithms are presented in Figure 5 (together with the repeated histograms for β-CC and β-Monroe) and their variances for the number of winners per quadrant are in Table 2.

Approximation Algorithms for β-CC. The results of the approximation algorithms for β-CC are rather varied, but even a quick glance shows that RangingCC seems to be the closest to the original β-CC rule. While we provide explanation as to why the other two algorithms are not doing well, the performance of RangingCC came as a surprise to us and we still do not have a very good explanation for its behavior.

To understand the behavior of GreedyCC, it suffices to recall that—by definition—in the first iteration the algorithm chooses the Borda winner. In our elections, the Borda winner is always located very close to the center, so the histograms for GreedyCC show a spike there. Then, due to the nature of the Chamberlin–Courant rule (as described in the main body of the paper), the algorithm selects candidates that are not too close to this first winner. This explains the patterns that we see for all our distributions. These patters are far more visible for GreedyCC than for β-CC, in particular, because the first iteration chooses a candidate from almost the same location irrespective of the actual distribution of the points. GreedyCC achieves good results for the variance for the number of winners per quadrant.

The behavior of Algorithm P can be explained similarly to that of GreedyCC, but by an analogy to the Bloc rule.

[5] In the paper of Skowron et al., it is denoted as Algorithm A.

rule	square	disc	Gauss.	4 Gauss.
GreedyCC	1.019	1.083	1.106	1.132
Algorithm P	2.551	2.453	2.418	2.381
RangingCC	0.907	0.944	1.015	0.959
GreedyMonroe	0.848	0.926	0.978	0.877

Table 2: Variance of the number of winners in each quadrant.

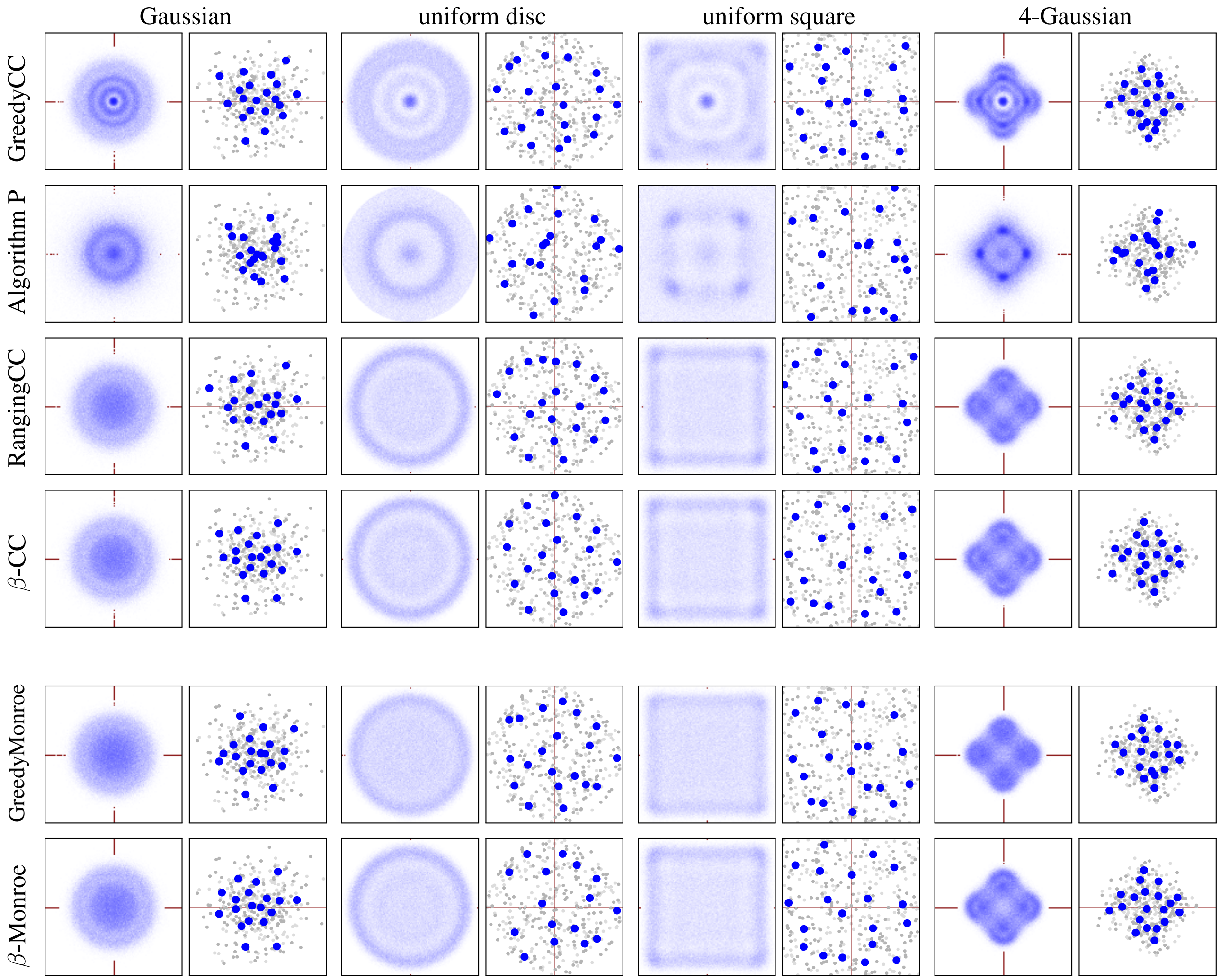

Figure 5: Results for approximation algorithms for β-CC and β-Monroe.

Algorithm P considers candidates ranked at the top x positions, where x is a prespecified threshold value (recall the description of the algorithm). In effect, the first iteration is almost the same as in Bloc, except that Bloc chooses a candidate ranked most frequently among the top k positions and Algorithm P considers the top x positions. In the second iteration, Algorithm P chooses a candidate that is ranked among the top-x positions by many voters who are far from the candidate chosen in the first iteration. Such a candidate is likely to also be included in the Bloc committee (again, taking into account the fact that both rules consider slightly different numbers of top candidates). We believe that similar effect lasts for a few iterations and is sufficient to create those patterns in the histograms of Algorithm P which resemble Bloc. However, in further iterations Algorithm P starts behaving differently than Bloc and, for example, chooses candidates from the center (especially for the Gaussian and uniform disc distributions). Unfortunately, Algorithm P has poor variance of the number of winners per quadrant (on the order of 2.4-2.5) and, indeed, visual inspection of its results shows that they are not satisfactory. Thus we believe that it should not be used (even though, in most settings, its guaranteed approximation ratio is better than that of GreedyCC).

Finally, RangingCC achieves nearly the same histograms as β-CC and has very good results for the variances for the number of winners per quadrant (but still slightly higher than β-CC). Since RangingCC winners can be computed quite efficiently, it appears to be the best choice among the three algorithms we have tested (in practice, one might also try the clustering technique of (Faliszewski et al. 2016b). Nonetheless, we are quite baffled with the performance of RangingCC and do not really have convincing explanations for its superiority against its component algorithms (various incarnations of Algorithm P).

GreedyMonroe. It appears that GreedyMonroe is a very good approximation algorithm for β-Monroe. The histograms we obtained for it are very similar to those for β-Monroe and the variance for the number of winners per

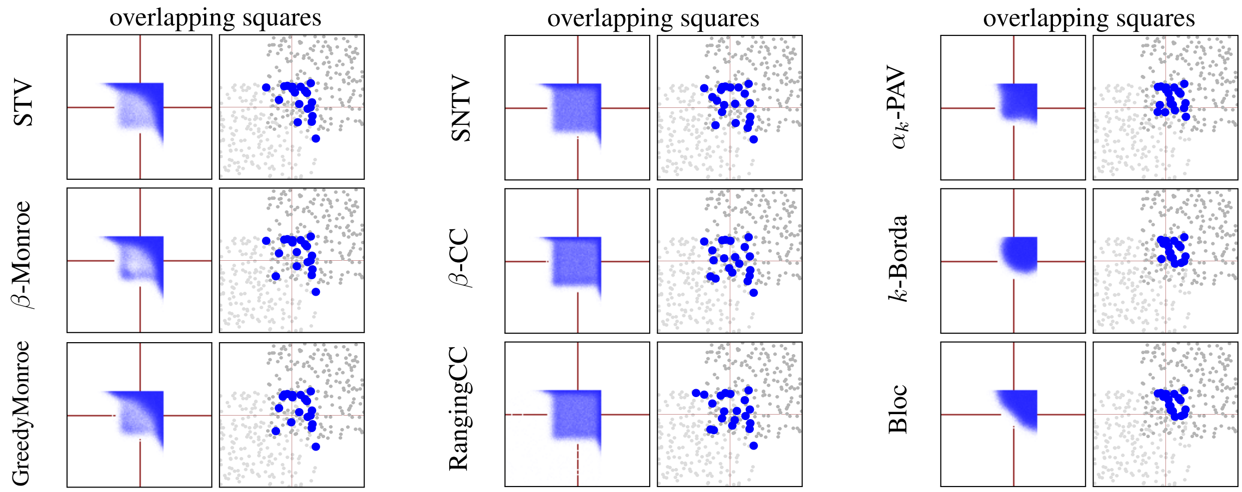

Figure 6: Histograms and sample elections for several voting rules, for the case where the candidates and voters are distributed uniformly on two overlapping squares.

quadrant is low (if a bit higher than for β-Monroe). In fact, GreedyMonroe's histograms appear to be a bit more similar to the underlying distributions of candidates and voters than those of β-Monroe, which by our criteria makes it a slightly better rule for parliamentary elections than the latter. Indeed, this also shows in the results of Elkind et al. (2014; 2017), who prove that GreedyMonroe satisfies the solid coalitions property—a property desirable for proportional representation[6]—and that β-Monroe does not. Interestingly, while Elkind at al. (2014; 2017) did not insist strongly on this property, our three rules with histograms most similar to the underlying distributions (STV, SNTV, and GreedyMonroe) do satisfy it.

Overlapping Squares Distribution

So far, our results for β-Monroe and β-CC were quite similar. To show that the rules are, indeed, different, we have performed a quick experiment for a setting where the distributions of the ideal points of candidates and voters are not the same:

Overlapping Squares. The ideal points of the candidates are distributed uniformly on the $[-3, 1] \times [-3, 1]$ square, whereas the ideal points of the voters are distributed uniformly on the $[-1, 3] \times [-1, 3]$ square.

Naturally, we should not expect any society to really follow such a distribution and we use it only as a test case.

It turns out that STV, SNTV, β-CC, β-Monroe, RangingCC, and GreedyMonroe can be partitioned into two groups. STV, β-Monroe and GreedyMonroe aim for proportional representation and, thus, their histograms put more emphasis on the candidates near the $(1, 1)$-corner. β-Monroe also puts some emphasis on the $(-1, -1)$ corner, while GreedyMonroe and STV do it only to a very minor extent. On the other hand, SNTV, β-CC, and RangingCC are more geared towards covering the intersection of the supports of the distributions of candidates and voters. We view this as further evidence that these rules (or, rather, only β-CC and RangingCC, since we already argued against SNTV) are well-suited for portfolio/movie selection tasks.

As to the three other rules, α_k-PAV, k-Borda, and Bloc, note that they concentrate on a support that is strictly smaller than the intersection of the two distributions, and tilted towards the center of the voters' distribution. This confirms the tendency of these rules to be detrimental to extreme candidates.

Integer Linear Program for α_k-PAV

In this section we describe the integer linear program that we have used for computing α_k-PAV.

Let $E = (C, V)$ be an input election with m candidates and n voters. We are interested in a winning committee S of size k. We define the following binary variables. For $j \in [m]$, we define x_j, with the intent that $x_j = 1$ if and only if $c_j \in S$. For $i \in [n], j \in [m], \ell \in [k]$, we define $y^{\ell}_{i,j}$, with the intent that $y^{\ell}_{i,j} = 1$ if and only if the jth-ranked candidate of voter v_i is chosen as her l-th best committee member. We have the following optimization goal:

$$\max \sum_{i\in[n]} \sum_{j\in[m]} \sum_{\ell\in[k]} \frac{1}{\ell} \cdot \alpha_k(j) \cdot y^{\ell}_{i,j};$$

and we include the following constraints:

1. The committee includes exactly k candidates:
$$\sum_{j\in[m]} x_j = k.$$
2. For a given voter and position j, the candidate on position j can be ℓ-th best committee member for this voter for at

[6] Formally, the property says the following: if we have an election with n voters, we want to choose a committee of size k, and there is a candidate c that is ranked on the first place by at least $\frac{n}{k}$ voters, then this candidate shall be included in the winning committee.

most one value of ℓ. Formally, for each $i \in [n]$ and for each $j \in [m]$, we have the constraint:

$$\sum_{\ell \in [k]} y_{i,j}^{\ell} \leq 1;$$

3. For a given voter, there is exactly one candidate that this voter ranks as ℓ-th best in the committee. Formally, for each $i \in [n]$ and for each $\ell \in [k]$, we have the constraint:

$$\sum_{j \in [m]} y_{i,j}^{\ell} = 1;$$

4. A candidate cannot be the ℓ-th best committee member for a given voter if this candidate is not even a committee member. Formally, for each $i \in [n]$, for each $j \in [m]$, and for each $\ell \in [k]$, we have the constraint:

$$y_{i,j}^{\ell} \leq x_t;$$

where c_t is the j-th ranked candidate of voter v_i.