

Normative Comparison of Multiwinner Election Rules

Research Outline

Piotr Skowron

Abstract

A multiwinner election rule is a formal process for selecting a subset of candidates based on the preferences of voters. A number of different scenarios can be expressed in the framework of multiwinner elections; the examples include: shortlisting candidates for an interview, electing a group of representatives for a society (e.g., a country's parliament or a university senate), choosing locations for public facilities in a city, or designing a targeted advertising campaign. While all these scenarios can be described by the same abstract framework, they might call for fundamentally different election rules. It is important to be able to match these peculiar applications with rules that would be most appropriate for them. This motivates our study of tools that could allow the decision-maker to compare different election rules in a principled way. We propose several strategies that can be used to obtain a better understanding of the nature of various multiwinner rules.

1 Introduction, Motivation, and Impact

Multiwinner elections form an abstract model for selecting a subset of candidates (referred to as a committee), based on the preferences of a population of voters. We use the term “candidates” in the most general meaning; candidates can correspond to humans, facilities, objects, items, etc. A number of real-life scenarios fit this abstract model [30]. Naturally, multiwinner elections describe processes of selecting parliaments, university senates, or other groups of representatives for various societies. However, they also have many other applications. For example, an enterprise considering which products to advertise or which varieties of a certain product to push to production also faces a multiwinner election problem (with the potential customers being the voters, and the products/designs or prototypes being the candidates) [45, 46]. Other examples of using multiwinner elections include deciding which responses to a query should be displayed by a search engine [19, 64], selection operators in genetic algorithms [26], or shortlisting processes. Let us describe in more detail two examples, which will be useful in our further discussion:

Selecting finalists of a competition. Consider the problem of selecting a group of finalists in a competition. It is natural to view this problem as a multiwinner election: the contestants correspond to the candidates, the judges/experts/reviewers to the voters, and the selected finalists to the elected committee. When choosing finalists among contestants, one usually aims at selecting objectively best candidates. In particular, if there are two very similar candidates then, except for the borderline cases, either both should be selected as members of the winning committee or both should be rejected.

Facility location. Consider the problem of finding locations for a set of facilities in a city (e.g., hospitals, emergency stations, playgrounds, kindergartens, etc.). This is also an example of a multiwinner elections: possible locations correspond to candidates, and citizens to voters. However this problem is very different from the example of shortlisting the finalists of a competition. Indeed, for most types of facilities, the citizens prefer them to be built close to their places of living, and thus the location in the city center is objectively the best at the individual level of cognition (such a location usually minimizes the total travel distance of all the citizens). Yet, it is rather a bad choice to locate all the hospitals or all the playgrounds there. Since each citizen is usually perfectly satisfied when having one facility type in their vicinity, an election rule should rather choose a set of locations which are spread uniformly in the city. Thus we should use a completely different strategy than for shortlisting finalists among contestants.

With a wide range of applications, it is crucial to understand the nature of different multiwinner election rules. Otherwise, it would be very difficult to match specific applications with appropriate rules. Indeed, the examples of selecting finalists in a contest and of locating facilities in a city illustrate that even though all the aforementioned applications fit the same general framework, they are fundamentally different and may require election rules following different high-level principles [30]. We identified three such principles—*individual excellence*, *diversity*, and *proportionality*—which describe some ideal goals that election rules can aim at. For example, when selecting a group of finalists in a contest, the election rule should follow the principle of *individual excellence*. When selecting the locations of facilities, one should rather use a rule that follows the principle of *diversity*. Yet another different rule should be used for choosing a country’s parliament—here we would like minorities to be represented in the committee, but the numbers of representatives that different minorities get should match the sizes of these minorities. In this case we would expect the election rule to follow the principle of *proportionality*.

Let us describe these three principles in more detail through the following example. Consider a set of eight candidates $C = \{c_1, \dots, c_8\}$, and the following approval preferences of the voters. Assume that 73% of the voters approve candidates c_1, c_2, c_3, c_4 , and disapprove the remaining four, 23% of voters approve c_5 and c_6 , 2% approve only c_7 , and the remaining 2% approve c_8 . Further, let us assume that the goal is to select a committee of size $k = 4$. If the rule follows the principle of individual excellence, we feel that the committee of candidates $\{c_1, \dots, c_4\}$ should be selected—indeed each member of such a committee is supported by roughly $3/4$ of the voters, while any other candidate receives support that is at least three times lower. A proportional rule, on the other hand, should choose $\{c_1, c_2, c_3, c_5\}$ —for such a committee, a coherent group of $3/4$ of the voters approves three out of four elected candidates. At the same time, a group of $1/4$ of all voters gets proportionally fewer representatives, i.e., a single one, and the groups of 2% of voters are too small to get any representative. Finally, a rule following the principle of diversity would select $\{c_1, c_5, c_7, c_8\}$ —for this committee every voter approves one of the elected candidates and each of the remaining candidates can be viewed as a “clone” of some member of the selected committee¹.

The above example is very simple as the approval sets of any two voters are either the same or disjoint. With more complex preferences of the voters it is often not clear how a proportional, a diverse or an excellence-oriented rule should behave. Additionally, individual excellence, diversity, and proportionality only describe some ideal goals, and one often requires a rule which follows to some extent each of the three principles. In other words, these principles describe some extreme points in a wide spectrum of possible goals [27, 29], and there exist applications which call for more “intermediate rules”.

The goal of this project is to better understand the nature of different multiwinner rules. We propose several approaches that allow one to compare multiwinner rules in general, and in particular to compare them based on their suitability for the three idealized goals and their mixtures.

Apparently, choosing the right rule is often hard, and thus having a good understanding of election rules is extremely important. Successful research on election rules can have significant practical impact as demonstrated by the systematic works of Brams [8, 9] on single-winner Approval Voting. Indeed, Approval Voting has become very popular for low-stake elections and, for example, is used in the Doodle system for scheduling meetings. In recent years Laslier and Van der Straeten et al. [44] conducted a nation-wide experiment in France showing the feasibility of using Approval Voting for electing the president.

The problem of comparing rules in the multiwinner setting is even harder than in the single-winner one, since multiwinner rules need to take into account the relations between different committee members depending on objectives which can vary significantly (as we explained in our discussion on the principles of individual excellence, diversity, and proportionality). Further, our goal is challenging because, even though there are many multiwinner election rules, their definitions are not always self-explanatory and do not directly imply behavioral consequences. We will approach this goal with a new perspective, which we will summarize in Section 3.

¹In the approval model, we call candidate c a *clone* of candidate c' if both candidates are approved by exactly the same set of voters.

2 Model of Multiwinner Elections and Overview of Multiwinner Rules

A *multiwinner election* is a triple $E = (V, C, k)$, where $C = \{c_1, \dots, c_m\}$ and $V = \{v_1, \dots, v_n\}$ are the sets of m candidates and of n voters, respectively, and k is an integer, $k \in \{1, \dots, m\}$, specifying the desired committee size. The voters have preferences over the candidates; in the next two subsections we describe two common models of expressing these preferences. A *committee* is a k -element subset of the candidates, and a *multiwinner election rule* is a function that takes as input a multiwinner election and returns either a single winning committee or a set of tied winning committees.

2.1 Approval-Based Rules

In the *approval model*, each voter $v_i \in V$ provides a subset of candidates $A_i \subseteq C$ that she finds acceptable. A number of interesting approval-based rules can be defined as approval committee scoring rules: Each voter $v_i \in V$ assigns to each committee $S \subseteq C$ a score $sc(v_i, S)$. The total score of a committee is computed as the sum of the scores that it garners from all the voters. The committees with the highest scores are returned as winners. Particular approval committee scoring rules differ in a way in which the voters assign the scores to the committees. The examples of such rules include:

Multiwinner Approval Voting (MAV). MAV is perhaps the simplest rule: it selects k candidates that are approved by most voters. Thus, it is defined by $sc_{\text{MAV}}(v_i, S) = |S \cap A_i|$.

Satisfaction Approval Voting (SAV). Under SAV, each voter has one point and she distributes it equally among all the voters that she approves of. Formally, it is defined by the scoring function $sc_{\text{SAV}}(v_i, S) = \frac{|S \cap A_i|}{|A_i|}$.

Proportional Approval Voting (PAV). This rule, first introduced in the 19th century by the Danish polymath Thiele [65], is defined by the scoring function $sc_{\text{PAV}}(v_i, S) = \sum_{j=1}^{|A_i \cap S|} \frac{1}{j}$. That is, a voter examines the committee S and assigns a score of 1 to S for the first member of S that she approves, if she finds a second candidate that she approves, she adds $1/2$ to the score; for the third approved candidate she adds $1/3$, etc. The use of this particular scoring function guarantees certain appealing properties of the rule which make it suitable for tasks requiring proportionality [4, 12, 43].

Approval Chamberlin–Courant (CC). CC is defined by $sc_{\text{CC}} = \min(|S \cap A_i|, 1)$. Intuitively, according to CC each voter cares about having a representative that she finds acceptable in the committee—a representative of a voter is any elected candidate that she approves of. A voter assigns one point to a committee if she has an approved representative, and zero points otherwise.

For each of the aforementioned rules we can define its sequential variant, which operates as follows. It starts with an empty committee and performs k steps, where in each step it adds to the committee the candidate that would improve its score most. The reverse sequential variants of these rules are defined analogously: we start with the whole set of candidates and remove the candidate whose removal decreases the score least.

Another class of interesting rules can be attributed to the Swedish mathematician Phragmén [52, 53, 54, 38, 11]. The definitions of these rules are based on a different framework, where one assumes that candidates and voters are associated with units of load (for voters, the load can correspond, e.g., to the total amount of support to be distributed), and to select winning candidates one follows a certain resource allocation procedure.

For an overview of other approval-based rules we refer the reader to the book chapter by Kilgour [39].

2.2 Ordinal-Based Rules

In the *ordinal model*, a voter $v_i \in V$ expresses her preferences by providing a linear preference relation \succ_i , ranking the candidates from the most to the least preferred one (for instance, a ranking $c_2 \succ_i c_1 \succ_i c_3 \succ_i \dots$ indicates that voter v_i considers c_2 to be her most preferred candidate, c_1 to be her second most preferred candidate, etc.). By $\text{pos}_i(c)$ we denote the position of candidate c in the preference ranking of voter v_i ; e.g., the position of the voter's most preferred candidate in her ranking is one, of her second most preferred candidate is two, etc. A single-winner

positional scoring function is a mapping $s: \{1, \dots, m\} \rightarrow \mathbb{R}$ that assigns a score to each position. Examples of common positional scoring functions include the Borda count, $\beta(i) = m - i$, and the k -approval rule, $\alpha_k(i) = 1$ if $i \leq k$ and 0 otherwise.

Single-winner positional scoring functions can be used in the multiwinner setting to assign scores to whole committees. Indeed, as in the approval model, certain rules can be defined by providing appropriate functions describing how voters assign scores to committees. Let $sc(v_i, S)$ denote the score that v_i assigns to S . The winning committees are defined as those with the highest total score garnered from all the voters, i.e., the rule outputs $\operatorname{argmax}_{S \subseteq C: |S|=k} \sum_{v_i \in V} sc(v_i, S)$. The examples of rules which can be expressed in this way include:

Single Non-Transferable Vote (SNTV). SNTV uses the function $sc_{\text{SNTV}}(v_i, S) = \sum_{c \in S} \alpha_1(\text{pos}_i(c))$. In other words, SNTV picks the k candidates that are ranked first by most voters.

k -Borda. It picks k candidates with the highest Borda scores, i.e., $sc_{k\text{-Borda}}(v_i, S) = \sum_{c \in S} \beta(\text{pos}_i(c))$.

β -Chamberlin–Courant (β -CC). β -CC is defined by $sc_{\beta\text{-CC}}(v_i, S) = \max_{c \in S} \beta(\text{pos}_i(c))$.

These multiwinner voting systems are called *committee scoring rules* by their analogy to the single-winner scoring rules. There is a broad variety of other committee scoring rules (see the survey by Faliszewski et al. [30]). Further, by providing different functions for aggregating the scores of the voters, we obtain other classes of rules—in particular *egalitarian committee scoring rules* are those that define the score of a committee S as the score that the least satisfied voter assigns to S . Thus, egalitarian committee scoring rules select $\operatorname{argmax}_{S \subseteq C: |S|=k} \min_{v_i \in V} sc(v_i, S)$ as winning committees.

Naturally, for each of the above rules there exists a sequential and a reverse-sequential counterpart. Single Transferable Vote (STV) is another example of an iterative rule taking ordinal preferences as input (this rule is currently used in a few contemporary democracies).

Other interesting multiwinner rules include those inspired by the Condorcet principle [5, 6, 15, 23, 33, 32, 35, 57, 60]. In the single-winner setting the Condorcet principle says that if there exists a candidate c who is preferred to any other candidate $c' \in C$ by a majority of voters (these might be different majorities for different candidates c'), then c should be considered as collectively the best candidate. There are various ways in which the Condorcet principle can be extended to the multiwinner setting—for an overview of them we refer the reader to the recent book chapter by Faliszewski et al. [30]—yet our understanding of such rules is still much less advanced compared to the scoring methods.

We also point the reader to the same book chapter of Faliszewski et al. [30] for an overview of other election methods. At this point, we want to emphasize that we do not plan to restrict our research to investigating the existing rules only; a normative analysis of existing rules can often lead to the discovery of new rules with appealing properties. This approach has already led us to identifying a number of new (classes of) rules [22, 27, 28, 29, 62, 63] (e.g., by looking for rules satisfying the multiwinner analogue of the simple majority property, we discovered the class of top- k -counting rules [28]).

3 Research Plan

In this section we describe concrete research ideas that we have for tackling the goal of the project.

3.1 Our New Approach versus Traditional Methods

The traditional approach to comparing election rules is to explore their axiomatic properties and to compare the rules against them (see, e.g., the classic book by Arrow [3] and the surveys by Chebotarev and Shamis [13] and by Merlin [49]). Such properties are binary indicators—the only type of information we can get from their analysis is that a given election rule either satisfies a certain property or that it violates one. However, especially in the multiwinner context, one can be interested in more fine-grained information: for instance we intuitively feel that some rules are “more proportional” than others, and that in fact there exists a whole spectrum of different proportionality levels. Thus, we would like to have mathematical tools which would allow us to assess to which extent a certain rule is

proportional. *We plan to address these types of questions by studying quantitative properties of multiwinner rules.* In contrast to traditional axiomatic properties, for the quantitative ones there might be more than two answers for the question “does a given rule have a given property?”. Formally, a quantitative property is a function mapping election rules to real numbers; in such a framework we could represent traditional (binary) axiomatic properties as functions mapping election rules to the elements of the binary set $\{0, 1\}$, with ‘1’ meaning that the rule satisfies the property and ‘0’ that it violates the property.

The second novel ingredient of this proposal is that we plan to analyze properties which are specific to multiwinner election rules, and which aim at capturing the principles of individual excellence, diversity, and proportionality. Several attempts have already been made to identify axioms formalizing the idea of proportionality; examples of such axioms in the approval model include justified representation [4], extended justified representation [4], proportional justified representation [58] and perfect representation [58]; for the ordinal model examples include Dummett’s proportionality [18], solid coalitions [22], consensus committee [22], and local stability [5]. While we believe that the axiomatic approach to formalizing proportionality is well-advanced, this is not the case for the individual excellence nor for the diversity principle. In particular, we hope to discover weak variants of the axioms describing individual excellence, diversity, and proportionality, such that there exist rules satisfying both the respective axiom for individual excellence and for diversity/proportionality. As one of the consequences, we aim at discovering and understanding different spectra of rules which implement certain tradeoffs between individual excellence, diversity, and proportionality.

The third new idea which we plan to pursue as a part of this proposal is to understand what multiwinner election rules do on certain simpler subdomains, where their behavior can be intuitively interpreted. One such appealing subdomain studied in the political and social choice literature is defined by Euclidean (geometric) preferences [16, 21, 24, 25, 48, 50, 55, 59]. For this subdomain, it is assumed that the voters and the candidates can be identified with points in a certain low-dimensional “issue space” (usually, one assumes that this is a 2-dimensional Euclidean space \mathbb{R}^2). Intuitively, the value of each coordinate of a point corresponding to a voter/candidate might reflect the position of such a voter/candidate on a certain issue such as the preferred level of taxation or immigration, or their position in the left-right spectrum of economic freedom². The voters derive their preferences over the candidates from the distances between respective points—formally, a voter v_i prefers candidate c to candidate c' if the distance between the points corresponding to v_i and c is smaller than the distance between the points corresponding to v_i and c' , $d(v_i, c) < d(v_i, c')$. This approach also relates to our first idea of quantitative properties since the geometric model allows for quantifying and for visualizing certain properties of election rules.

3.2 Methodology of Research

Our research will be mostly theoretical, and to achieve the aforementioned goals we will be using tools from combinatorics, discrete mathematics, and from theoretical computer science (e.g., the tools used in the analysis of approximation algorithms). However, we would like to complement our theoretical analysis with experiments using preference data from real-life elections; such data is publicly available, e.g., in the PrefLib repository [47]. We would like to devote about a quarter of the time within the project to tasks involving experiments. We believe that such experiments can potentially have a great impact, since they demonstrate the usage of new methods originating from computer science to answer fundamental problems from economics.

The experimental methods will require developing new algorithms for computing (approximate) outcomes of election rules which are NP-hard to compute. While the computational complexity of the standard ordinal-based rules is a well-explored area, and in most cases we can compute the outcomes of such rules efficiently (or at least we can find outcomes that are very close to the optimal ones), this is not the case for the approval-based rules, nor for the egalitarian ordinal ones. In our study we plan to analyze approximation algorithms, heuristics, and exact algorithms which can be computed efficiently when certain natural parameters are small (thus, we will be using tools of the parameterized complexity analysis [14, 17, 34, 51]).

²Indeed, even though one might think of many important issues that drive voters’ preferences over candidates, it is argued that two dimensions are often sufficient to provide an accurate estimation of voters’ preferences [59]

Objective 1. *Design effective algorithms for computing outcomes according to approval-based and egalitarian ordinal-based rules. In particular, design algorithms specifically suitable for Euclidean preferences.*

Further, a formulation of a new property always brings new questions about the computational complexity of testing it. Such questions are of fundamental importance, since we need to be able to compute whether a given outcome of a multiwinner rule violates a certain property in order to perform statistical analysis of properties using datasets of preferences (for instance, to check how often, and for what types of voters' preferences a given property is violated). This is often reflected in recent research, where the analysis of behavioral consequences of satisfying/violating certain properties receives as much attention as the analysis of the computational complexity of natural problems referring to these properties [4, 10]. Consequently, a part of this project will be devoted to developing a set of algorithmic tools that, in particular, would allow for an analysis of the studied properties through computer experiments.

We already prepared a preliminary open-source library that implements a selection of multiwinner rules and provides tools that can be used for a 2D-visualization of the outcomes of these rules³ (this library covers a representative selection of ordinal-based committee scoring rules, but it does not yet include approval rules nor rules based on the Condorcet principle). We plan to develop, enhance and improve this library as a part of this project. We hope that as an outcome of this project we will also obtain diagrams that illustrate the behavior of certain rules and that are understandable for a broader audience—in such case we plan to popularize our research by extending appropriate Wikipedia pages.

3.3 Quantitative Properties of Multiwinner Rules

The first type of quantitative properties of multiwinner rules that we plan to investigate within this project is inspired by the idea of approximation from the theory of computational complexity. Recent advancements in our understanding of multiwinner elections allow us to identify certain rules as good representatives for the classes of diversity- (respectively, excellence-) oriented rules. For instance, in the approval model the Chamberlin–Courant rule can be viewed as an archetypical example of a rule aiming at diversity, whereas Multiwinner Approval Voting is often considered a prime example of an excellence-oriented rule. Both rules are defined as optimization functions, maximizing certain measures of the satisfaction of the voters, so the concept of approximation is well defined for them. We plan to analyze how well other rules approximate the approval Chamberlin–Courant rule and the Multiwinner Approval Voting rule—the approximation ratios obtained for these two objective functions can be viewed as indicators on how diverse (respectively, excellence-oriented) a certain rule is. The better the rule approximates MAV (respectively, CC), the more excellence-oriented (respectively, diverse) it is⁴. Thus, we do not use the idea of approximation in the most typical way to approximate computationally hard rules by rules which are easier to compute. In particular, we ask how well we can approximate certain rules irrespective of whether they are easy to compute, which is the case for MAV, or computationally hard, which is the case for the Chamberlin–Courant rule. We also do not look for the best possible algorithms to approximate certain rules, but rather we investigate the existing rules and by viewing them as approximation algorithms, we aim at advancing our understanding of their nature. Our preliminary results show that the Phragmén's rule and PAV approximate CC very well, so they perform well from the point of view of diversity, and that these two rules provide worse (but still, much better than other rules) guarantees for MAV, so they are preferred from the perspective of individual excellence to other diverse or proportional rules, such as the Monroe rule.

Objective 2. *Analyze various approval rules as approximation algorithms for CC, MAV, and PAV. Analyze various ordinal-based rules as approximation algorithms for β -CC and k -Borda. Analyze the approximation guarantees of the studied rules in the Euclidean model.*

Another example of a quantitative property that we plan to study is *average justified satisfaction*. Consider the approval model. We call a group of voters $V' \subseteq V$ ℓ -cohesive if this group is large enough, $|V'| \geq \ell \cdot \frac{n}{k}$, and

³<https://github.com/elektronaj/MW2D>

⁴There are a few similarities between our approach and works of Kamler [40, 41, 42] and Eckert et al. [20], who, mostly through computer experiments, compared how similar are certain single-winner voting rules.

if there exist at least ℓ candidates who are approved by all voters from V' , that is, $|\bigcap_{v_i \in V'} A_i| \geq \ell$. Intuitively, such a group should be eligible to have ℓ representatives in the elected committee. We say that a rule \mathcal{R} has average justified satisfaction of $\alpha \in \mathbb{R}$ if for each ℓ -cohesive group of voters V' , the number of committee members approved, on average, by the voters from V' is at least equal to $\alpha\ell$. Average justified satisfaction can be viewed as a quantitative adaptation of the recently introduced family of binary properties that are based on the concept of justified representation [4, 58]. For instance, extended justified representation says that in each ℓ -cohesive group there should be a voter who approves at least ℓ committee members. A binary property such as EJR does not always explain a rule well. Intuitively, we feel that measuring the satisfaction of a large group of voters by taking the satisfaction of a single most-satisfied voter from this group does not tell much about how ℓ -cohesive groups are treated by the election rule in general. On the other hand, requirements on the satisfaction of *each* voter within a group would be too restrictive and there is no voting rule that would satisfy such a strong property. Average justified satisfaction is a tool for quantifying the extent to which ℓ -cohesive groups are satisfied.

We also plan to investigate measures of statistical dispersion—such as the Gini index, the Palma ratio or the Hoover index—in the context of multiwinner elections. These indices are often used in economics to quantify the inequality in distribution of wealth. They can be adapted to the multiwinner setting by defining the notion of the wealth of a voter as her satisfaction from a committee. In the approval model, the satisfaction of a voter can be defined in a natural way, as the number of committee members that a given voter approves. In the ordinal model we can use positional scoring functions, such as the Borda rule, to map the positions of committee members in the preference ranking of a voter to real values quantifying her satisfaction with the committee.

Objective 3. *Analyze the average justified satisfaction for various rules in the approval model. Find rules which obtain the optimal average justified satisfaction and derive some lower bounds. Investigate properties analogous to the average justified satisfaction in the ordinal model. Analyze measures of statistical dispersion, such as income inequality indices, for a selection of multiwinner rules (with various interpretations of the notion of the satisfaction of a voter).*

As an outcome of our quantitative analysis, we would like to be able to represent rules as points in a three-dimensional space, with dimensions corresponding to the three principles. Such a visualization would allow us to compare different rules with respect to how proportional, diverse, or excellence-oriented they are, and to understand the tradeoffs between these principles that certain rules implement. Finally, we would like to characterize the Pareto frontier with optimality criteria corresponding to the previously defined measures of the quality for the three principles. As a result we will either identify that certain known rules belong to the Pareto frontier or we will design new Pareto-optimal rules.

Objective 4. *Analyze the Pareto frontier with optimality criteria corresponding to certain measures of quality for the three principles.*

One of the main challenges of our project is conceptual: we plan to come up with new binary and quantitative properties that would capture the ideas of diversity, individual excellence, and proportionality. We also observe that even though the axiomatic properties of single-winner rules have been extensively studied in the literature, this is not the case for the multiwinner setting. Most of the known results in this area are very recent and they usually concern scoring rules. In particular, a systematic study of axiomatic properties for Phragmén’s rules [52, 53, 54, 38, 11], and for the rules based on the Condorcet principle [5, 6, 15, 23, 33, 32, 35, 57, 60] is missing. Apparently, even though the Condorcet principle plays an instrumental role in the studies on single-winner elections, the multiwinner Condorcet-based rules are only very poorly understood.

Objective 5. *Find properties that should be satisfied by rules aiming at diversity and individual excellence. Analyze existing rules against these properties. Provide a systematic study of (existing and new) axiomatic properties of the multiwinner rules based on the Condorcet principle, sequential and reverse-sequential (approval) committee scoring rules and of Phragmén’s rules.*

3.4 Analysis of the Distortion for Multiwinner Rules

The objectives presented in this section are a part of the analysis of quantitative properties of multiwinner rules, but they are more specific since they refer to the model of Euclidean preferences of the voters.

Euclidean spaces provide more information than the approval or the ordinal models of voters' preferences. Indeed, a distance not only indicates which among the two candidates is more preferred, but also describes the fine-grained intensities of the preferences of the voters. It turns out that for such a richer and more complex model it is sometimes easier to define the winning committees than in the approval or in the ordinal framework. For instance, if we seek excellence-oriented committees, one can argue that the most preferred committees S would be the ones that minimize the total distance of the voters from the committee members, $\sum_{v \in V} \sum_{c \in S} d(v, c)$. If we are primarily concerned about the diversity, then our goal should rather be to minimize the distance of the voters to their representatives in the committee (i.e., to their most preferred candidates among the committee members), $\sum_{v \in V} \min_{c \in S} d(v, c)$. Such diversity-inspired measure of optimality, however, might not be appropriate to find "proportional committees", since the numbers of the voters that different committee members represent might vary significantly. Thus, it is more reasonable to define "proportional committees" as the ones that minimize $\min_{\Phi \in \text{Assign}(V, S)} \sum_{v \in V} d(v, \Phi(v))$, where $\text{Assign}(V, S)$ is the space of balanced assignments of the voters to the committee members, i.e., each $\Phi \in \text{Assign}(V, S)$ is a mapping $\Phi: V \rightarrow S$ between the voters and the committee members such that $\lfloor n/k \rfloor \leq |\Phi^{-1}(\{c\})| \leq \lceil n/k \rceil$ for each $c \in S$.

Unfortunately, the accurate information carried by the Euclidean model is almost never available. Typically, the voters cannot position themselves in the issue space and assigning (even inaccurate) cardinal (dis)utilities to the candidates would require a substantial cognitive effort from the voters. Nevertheless, the Euclidean model can be used as a benchmark for comparing various election rules which use preference rankings or approval sets. The concept of *distortion*, first introduced and studied in the context of single-winner rules [1, 2, 7, 31, 36, 37, 56, 61], quantifies the loss of voters' utility due to using certain election rules (and thus, indirectly, due to the limited information available for the rule). The definition of distortion can be naturally generalized to the multiwinner setting in a number of meaningful ways. We are specifically interested in the three ways which correspond to the three high-level principles—individual excellence, diversity, and proportionality:

$$\text{dist}_{\mathcal{P}}(\mathcal{R}) = \max_{E=(V, E, k) \in \text{Euclid}(n, m, k)} \frac{d_{\mathcal{P}}(V, \mathcal{R}(E))}{\min_{S \subseteq C} d_{\mathcal{P}}(V, S)}, \quad \mathcal{P} \in \{\text{Excel}, \text{Diver}, \text{Prop}\},$$

where $d_{\text{Excel}}(V, S)$, $d_{\text{Diver}}(V, S)$, and $d_{\text{Prop}}(V, S)$ are metrics designed to reflect, respectively, individual excellence, diversity, and proportionality (cf., our previous discussion in this subsection):

1. $d_{\text{Excel}}(V, S) = \sum_{v \in V} \sum_{c \in S} d(v, c)$ measures the sum of the distances between all the voters and all the committee members.
2. $d_{\text{Diver}}(V, S) = \sum_{v \in V} \min_{c \in S} d(v, c)$ measures the sum of the distances of each voter to her closest (most preferred) committee member.
3. $d_{\text{Prop}}(V, S) = \min_{\Phi \in \text{Assign}(V, S)} \sum_{v \in V} d(v, \Phi(v))$ is a measure where we assume that the voters are clustered to roughly equal-size groups, and for each group there is a single candidate who represents it. The distance between the voters and the committee is the sum of the distances of each voter to the representative of her cluster.

Here our objectives are the following:

Objective 6. *Analyze theoretically the distortion of various multiwinner election rules with respect to the above three definitions of distortion. Perform experiments to assess the average and certain positional statistics of the ratios used in the definition of distortion. In other words, use computer experiments to analyze the loss measured by distortion, but beyond the worst case.*

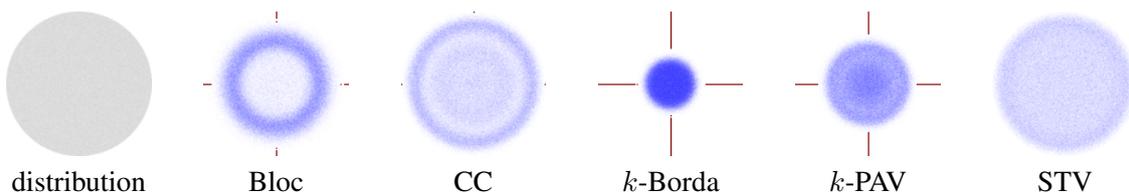


Figure 1: Visualization of the outcomes of the selected multiwinner rules for Euclidean preferences [21]. The voters and the candidates are drawn from the uniform distribution on a disc. In ten thousand experiments, each time a candidate corresponding to a given point was selected by a rule, we put a blue dot on the respective picture. Thus, the intensity of a color of each point reflects the probability that a candidate corresponding to this point is elected as a member of a winning committee, provided the candidates and the voters are drawn from the aforementioned distribution.

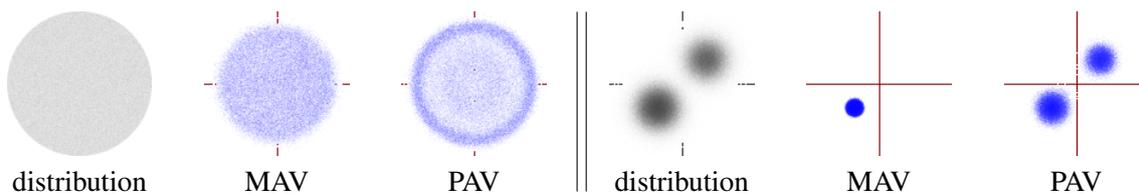


Figure 2: Visualization of the outcomes of Multiwinner Approval Voting (MAV) and Proportional Approval Voting (PAV) for two different distributions of the voters and candidates. In the three left plots the voters and the candidates were distributed uniformly on a disc. In the three right plots the voters and the candidates were divided into two groups. The larger group was concentrated in the lower-left quarter of the plot, and the smaller one—in the upper-right quarter.

3.5 Experiments Over the Two-Dimensional Euclidean Domain

In our recent paper we described a new method for visualizing the outcomes of multiwinner rules in the Euclidean model [21]. Figure 1 depicts the outcomes for five selected multiwinner election rules, for the case when the points describing voters and candidates are distributed uniformly at random on a disc (the distribution of the voters and candidates is depicted in the left plot in Figure 1). Each plot depicts a histogram describing 10,000 independent elections: the intensity of the color of each blue dot reflects the likelihood that a candidate corresponding to this point is a member of a winning committee. These types of experiments allowed us to make interesting claims about the behavior of various multiwinner rules. For instance, the plot for CC closely resembles the original distribution of the voters, which suggests that the committees returned by this rule can well represent the diversity of opinions in a society; k -Borda, on the other hand, identifies the center of the distribution, and so seems particularly well-suited for applications requiring excellence-oriented rules⁵. The outcomes returned by k -PAV are in-between, that is they represent voters more proportionally than the outcomes of k -Borda, yet k -PAV avoids selecting extremist candidates (in particular, we plan to explain this observed phenomenon theoretically, by measuring how well PAV approximates k -Borda and CC, cf., Objective 2).

We plan to further use this new method since the results that we obtained so far are very preliminary. In particular, we have run experiments only for four basic distributions of voters and candidates, and our plan is to extend these results.

Objective 7. *Visualize in histograms the behavior of various rules by considering more complex distributions of voters and candidates which model certain real political situations, and by considering scenarios where voters and candidates come from different distributions.*

This objective includes, but is not limited to, the investigation of scenarios with multiple groups of voters (minorities) concentrated in different parts of the issue space. Indeed, our preliminary results suggest that an analysis of

⁵This is because, on the one hand, k -Borda selects candidates which are similar to each other and, on the other hand, which individually minimize the total distance to the voters. However, the committees selected by k -Borda implement what is called the “rotten consensus”—most of the voters are not particularly happy about any member of the elected committee.

such scenarios can shed more light on the behavior of multiwinner election rules than the analysis of the basic distributions. In particular, Figure 2 depicts the comparison of Proportional Approval Voting and Multiwinner Approval Voting for two distributions. On the left, we see the results for the case where both candidates and voters are drawn from the uniform distribution on a disc. On the right, the voters and the candidates were divided into two groups: 60% of the voters and the candidates were drawn from the Gaussian distribution with the mean centered in point $(-1, -1)$ (lower-left quarter of the plot), and the remaining 40%—from the distribution with the mean centered in $(1, 1)$. While the simple distribution over the disc would suggest that the two rules behave rather similarly, the distribution with two groups of voters clearly shows that one of the rules selects only individually best candidates, while the other one is much more proportional. These preliminary results indicate that considering more complex distributions can highlight very different natures of election rules. Another set of our preliminary results indicates that also by considering different distributions for candidates than for voters, we can observe interesting differences between rules, which we do not observe when looking at the same distributions.

We further plan to analyze certain properties of multiwinner rules in the Euclidean models and to visualize them using the aforementioned histograms. For instance, we plan to investigate what happens when: (i) certain groups of voters move away, (ii) when the distribution of the candidates changes and the distribution of the voters stays the same (which corresponds to strategic behavior of the candidates), (iii) when certain groups of candidates are cloned, etc.

Objective 8. *Analyze properties of multiwinner election rules for Euclidean preferences and visualize them using histograms.*

This research will not only allow us to better understand the nature of different multiwinner rules but also the nature of voters' preferences. In particular, it will allow us to explain what are the geometric characteristics of voters' preferences that cause certain rules to behave in noticeably different ways, and will allow us to relate the consequences of violating certain properties with the shapes of voters' preferences.

4 Tasks Assignment

There will be three people working full-time on the project—the principal investigator (Piotr Skowron) and two PhD students. The project can be naturally divided into three parts, described by Sections 3.3–3.5. The tasks referring to the analysis of binary and quantitative properties of multiwinner rules (these are the tasks described in Section 3.3) will be assigned to the principal investigator and to one of the PhD students. The tasks concerning the analysis of the distortion of various multiwinner rules, and the tasks involving experiments using the new method of visualization (these are the tasks described in Sections 3.4 and 3.5) will be carried out by the principal investigator and the other PhD student. Additionally we plan that the whole team (the principal investigator and both PhD students) will be assigned to Objective 1, that is to the study of algorithms for computing the (approximate) outcomes of multiwinner rules.

This project, however, allows for a certain degree of flexibility in the assignment of the tasks. For instance, if one of the PhD students will be particularly attracted to studies concerning the Euclidean model, he or she can be assigned to tasks concerning the part of Objective 2, which refers to the Euclidean model, and to tasks for Objectives 6, 7, and 8. Similarly, a student who is more interested in the theoretical analysis of multiwinner rules can be assigned to the appropriate parts of Objectives 2–6, and a student who is more interested in analysis involving experiments can be assigned to the other parts of these objectives, and to Objectives 7–8. Nevertheless, we would like each student to be at least partially involved in some tasks involving the theoretical analysis.

Some of the tasks will require a collaboration with external partners, which is described in a separate document attached to this proposal. In particular, when working on the tasks related to the analysis of properties of multiwinner rules, we plan to collaborate with partners who have a wide knowledge and experience in economics (e.g., Prof. Arkadii Slinko from the University of Auckland, and Prof. Jean-François Laslier from the Paris School of Economics). In our studies of the algorithms for multiwinner rules, we plan to closely collaborate with Prof. Rolf Niedermeier from TU Berlin, and in our studies on restricted domains—with Prof. Edith Elkind from the University of Oxford.

References

- [1] E. Anshelevich, O. Bhardwaj, and J. Postl. Approximating optimal social choice under metric preferences. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI-2015)*, pages 777–783, 2015.
- [2] E. Anshelevich and J. Postl. Randomized social choice functions under metric preferences. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-2016)*, pages 46–59, 2016.
- [3] K. J. Arrow. *Social Choice and Individual Values*. Wiley, New York, 2nd edition, 1963.
- [4] 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.
- [5] 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 (IJCAI-2017)*, pages 84–90, 2017.
- [6] S. Barberá and D. Coelho. How to choose a non-controversial list with k names. *Social Choice and Welfare*, 31(1):79–96, 2008.
- [7] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A. D. Procaccia, and O. Sheffet. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213, 2015.
- [8] S. J. Brams and P. C. Fishburn. *Approval Voting*. Birkhäuser, 1983.
- [9] S. J. Brams and D. R. Herschbach. The science of elections. *Science*, 292(5521):1449, 2001.
- [10] R. Bredereck, P. Faliszewski, A. Kaczmarczyk, R. Niedermeier, P. Skowron, and N. Talmon. Robustness among multiwinner voting rules. In *Proceedings of the 10th International Symposium on Algorithmic Game Theory (SAGT-2017)*, pages 80–92, 2017.
- [11] M. Brill, R. Freeman, S. Janson, and M. Lackner. Phragmén’s voting methods and justified representation. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*, pages 406–413, 2017.
- [12] M. Brill, J.-F. Laslier, and P. Skowron. Multiwinner approval rules as apportionment methods. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*, pages 414–420, 2017.
- [13] P. Chebotarev and E. Shamis. Characterizations of scoring methods for preference aggregation. *Annals of Operations Research*, 80:299–332, 1998.
- [14] M. Cygan, F. Fomin, Ł. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh. *Parameterized Algorithms*. Springer, 2015.
- [15] A. Darmann. How hard is it to tell which is a Condorcet committee? *Mathematical Social Sciences*, 66(3):282–292, 2013.
- [16] O. A. Davis and M. J. Hinich. A mathematical model of preference formation in a democratic society. In J. Bernd, editor, *Mathematical Applications in Political Science II*, pages 175–208. Southern Methodist University Press, 1966.
- [17] R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Springer Publishing Company, Incorporated, 2013.
- [18] M. Dummett. *Voting Procedures*. Oxford University Press, 1984.
- [19] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation methods for the web. In *Proceedings of the 10th International World Wide Web Conference*, pages 613–622. ACM Press, Mar. 2001.

- [20] D. Eckert, C. Klamler, J. Mitlöhner, and C. Schlötterer. A distance-based comparison of basic voting rules. *Central European Journal of Operations Research*, 14(4):377–386, 2006.
- [21] E. Elkind, P. Faliszewski, J. F. 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 Conference on Artificial Intelligence (AAAI-2017)*, 2017. 494–501.
- [22] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. Properties of multiwinner voting rules. *Social Choice and Welfare*, 48(3):599–632, 2017.
- [23] E. Elkind, J. Lang, and A. Saffidine. Condorcet winning sets. *Social Choice and Welfare*, 44(3):493–517, 2015.
- [24] J. M. Enelow and M. J. Hinich. *The Spatial Theory of Voting: An introduction*. CUP Archive, 1984.
- [25] J. M. Enelow and M. J. Hinich. *Advances in the Spatial Theory of Voting*. Cambridge University Press, 1990.
- [26] P. Faliszewski, J. Sawicki, R. Schaefer, and M. Smolka. Multiwinner voting in genetic algorithms for solving illposed global optimization problems. In *Proceedings of the 19th International Conference on the Applications of Evolutionary Computation*, pages 409–424, 2016.
- [27] 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 (IJCAI-2016)*, pages 250–256, 2016.
- [28] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Multiwinner analogues of the plurality rule: Axiomatic and algorithmic views. In *Proceedings of the 30th Conference on Artificial Intelligence (AAAI-2016)*, pages 482–488, 2016.
- [29] 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 (IJCAI-2017)*, pages 192–198, 2017.
- [30] 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, 2017. To appear.
- [31] M. Feldman, A. Fiat, and I. Golomb. On voting and facility location. In *Proceedings of the 17th ACM Conference on Economics and Computation*, pages 269–286, 2016.
- [32] P. C. Fishburn. An analysis of simple voting systems for electing committees. *SIAM Journal on Applied Mathematics*, 41(3):499–502, 1981.
- [33] P. C. Fishburn. Majority committees. *Journal of Economic Theory*, 25(2):255–268, 1981.
- [34] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer-Verlag, 2006.
- [35] W. Gehrlein. The Condorcet criterion and committee selection. *Mathematical Social Sciences*, 10(3):199–209, 1985.
- [36] A. Goel, A. K. Krishnaswamy, and K. Munagala. Metric distortion of social choice rules: Lower bounds and fairness properties. Technical Report arXiv:1612.02912, arXiv.org, 2016.
- [37] S. Gross, E. Anshelevich, and L. Xia. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*, pages 544–550, 2017.

- [38] S. Janson. Phragmén’s and Thiele’s election methods. Technical Report arXiv:1611.08826 [math.HO], arXiv.org, 2016.
- [39] D. Kilgour. Approval balloting for multi-winner elections. In J. Laslier and R. Sanver, editors, *Handbook on Approval Voting*, pages 105–124. Springer, 2010.
- [40] C. Klamler. The Dodgson ranking and its relation to Kemeny’s method and Slater’s rule. *Social Choice and Welfare*, 23(1):91–102, 2004.
- [41] C. Klamler. The Dodgson ranking and the Borda count: a binary comparison. *Mathematical Social Sciences*, 48(1):103–108, 2004.
- [42] C. Klamler. A distance measure for choice functions. *Social Choice and Welfare*, 30(3):419–425, 2008.
- [43] M. Lackner and P. Skowron. Consistent approval-based multi-winner rules. Technical Report arXiv:1704.02453v1 [cs.GT], arXiv.org, Apr. 2017.
- [44] J.-F. Laslier and K. Van der Straeten. A live experiment on approval voting. *Experimental Economics*, 11(1):97–105, 2008.
- [45] 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 (IJCAI-2011)*, pages 280–286, 2011.
- [46] T. Lu and C. Boutilier. Value directed compression of large-scale assignment problems. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI-2015)*, pages 1182–1190, 2015.
- [47] N. Mattei and T. Walsh. PrefLib: A library of preference data. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT-2013)*, pages 259–270, 2013.
- [48] R. D. McKelvey, P. C. Ordeshook, et al. A decade of experimental research on spatial models of elections and committees. *Advances in the Spatial Theory of Voting*, pages 99–144, 1990.
- [49] V. Merlin. The axiomatic characterization of majority voting and scoring rules. *Mathematical Social Sciences*, 41(161):87–109, 2003.
- [50] S. Merrill and B. Grofman. *A Unified Theory of Voting: Directional and Proximity Spatial Models*. Cambridge University Press, 1999.
- [51] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford University Press, 2006.
- [52] E. Phragmén. Sur une méthode nouvelle pour réaliser, dans les élections, la représentation proportionnelle des partis. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 51(3):133–137, 1894.
- [53] E. Phragmén. *Proportionella val. En valteknisk studie*. Svenska spörsmål 25. Lars Hökersbergs förlag, Stockholm, 1895.
- [54] E. Phragmén. Sur la théorie des élections multiples. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 53:181–191, 1896.
- [55] C. R. Plott. A notion of equilibrium and its possibility under majority rule. *The American Economic Review*, 57(4):787–806, 1967.
- [56] A. D. Procaccia and J. S. Rosenschein. The distortion of cardinal preferences in voting. In *Proceedings of 10th International Workshop on Cooperative Information Agents*, pages 317–331. Springer, 2006.
- [57] T. Ratliff. Some startling inconsistencies when electing committees. *Social Choice and Welfare*, 21(3):433–454, 2003.

- [58] 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 Conference on Artificial Intelligence (AAAI-2017)*, pages 670–676, 2017.
- [59] N. Schofield. *The Spatial Model of Politics*. Routledge, 2007.
- [60] 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 (AAMAS-2017)*, pages 33–41, 2017.
- [61] P. Skowron and E. Elkind. Social choice under metric preferences: Scoring rules and STV. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*, pages 706–712, 2017.
- [62] P. Skowron, P. Faliszewski, and J. Lang. Finding a collective set of items: From proportional multirepresentation to group recommendation. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI-2015)*, pages 2131–2137, 2015.
- [63] P. Skowron, P. Faliszewski, and A. Slinko. Axiomatic characterization of committee scoring rules. Technical Report arXiv:1604.01529 [cs.GT], arXiv.org, 2016.
- [64] P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. Proportional rankings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-2017)*, pages 409–415, 2017.
- [65] T. N. Thiele. Om flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, pages 415–441. 1895.