# A Quantitative Analysis of Multi-Winner Rules

**Martin Lackner**[1] , **Piotr Skowron**[2]

[1]TU Wien, Vienna, Austria
[2]University of Warsaw, Warsaw, Poland
lackner@dbai.tuwien.ac.at, p.skowron@mimuw.edu.pl

## Abstract

To choose a suitable multi-winner voting rule is a hard and ambiguous task. Depending on the context, it varies widely what constitutes the choice of an "optimal" subset. In this paper, we offer a new perspective on measuring the quality of such subsets and—consequently—of multi-winner rules. We provide a quantitative analysis using methods from the theory of approximation algorithms and estimate how well multi-winner rules approximate two extreme objectives: diversity as captured by the Approval Chamberlin–Courant rule and individual excellence as captured by Multi-winner Approval Voting. With both theoretical and experimental methods we classify multi-winner rules in terms of their quantitative alignment with these two opposing objectives.

## 1 Introduction

A multi-winner rule is a voting method for selecting a fixed-size subset of alternatives, a so-called *committee*. More formally, it is a function that given a set of objects, preferences of a population of voters over these objects, and an integer $k$, returns a subset of exactly $k$ objects. Ideally, a multi-winner rule should select the "best" committee, but the suitability of a chosen committee strongly depends on the specific context. For instance, if voters are experts (e.g., judges in a sport competition) whose preferences reflect their estimates of the objective qualities of candidates, then the goal is typically to pick $k$ individually best candidates, e.g., those candidates who receive the highest scores from judges. Intuitively and somewhat simplified, in this and similar scenarios the quality of candidates can be assessed separately, and a suitable multi-winner rule should pick the $k$ best-rated ones. On the contrary, if the voters are citizens and the goal is to choose locations for $k$ public facilities (say, hospitals), then our goal is very different: assessing the candidates separately can result in building all the facilities in one densely populated area; yet, it is preferable to spread them in order to ensure that as many citizens as possible have access to *some* facility in their vicinity.

These two examples illustrate two very different goals of multi-winner rules, which can be informally described as fol-lows [Faliszewski *et al.*, 2017]: *Diversity* requires that a rule should select a committee which represents as many voters as possible; this translates to choosing a hospital distribution that covers as many citizens as possible. *Individual excellence* suggests picking those candidates that individually receive the highest total support from the voters; this translates to selecting a group of best contestants in the previous example. However, many real-life scenarios do not fall clearly into one of the two categories. For example, rankings provided by a search engine should list the most relevant websites but also provide every user at least one helpful link. In such cases, a mechanism designer would be interested in choosing a rule that guarantees some degree of diversity and individual excellence at the same time, putting more emphasis on either of them depending on the particular context.

Consequently, to properly match rules with specific applications, it is essential to understand to which degree committees chosen by established multi-winner rules are diverse or individually excellent. In this paper, we (1) develop *a set of tools* that allows a mechanism designer to better understand the nature of multi-winner rules and to assess the trade-offs with respect to diversity and individual excellence, and (2) provide *a classification* that clarifies the behavior of well-known rules with respect to the two criteria. We focus on the case where voters express their preferences by providing subsets of approved candidates (the approval-based model), yet our approach is applicable to other preference models as well.

### 1.1 Methodology and Contribution

In our approach, we choose two approval-based multi-winner rules, Chamberlin–Courant (CC) and Multi-winner Approval Voting (AV), as distinctive representatives of the principles of diversity and individual excellence, respectively. We measure how close certain rules are to AV and CC—we measure this distance by using the concept of the worst-case approximation. Thus, by investigating how well certain rules approximate AV or CC, we provide guarantees of how individually excellent or diverse these rules are; such guarantees can be viewed as *quantitative properties* of voting rules. This approach is fundamentally different from the traditional axiomatic approach, which is qualitative: a voting rule can either satisfy a property (axiom) or not. Our approach provides much more fine-grained information and allows us to estimate the degree to which a certain property is satisfied. With

these methods, we understand voting rules as a compromise between different (often contradictory) goals.

There are several reasons for our choice of AV and CC: First, both AV and CC are well-known, well understood rules. Second, due to their different nature they are a good choice for a two-dimensional evaluation. Third, they can be characterized by properties that capture individual excellence and diversity [Lackner and Skowron, 2018a]. Finally, they embody the ideas of utility-maximization and egalitarian maximization (as in the classic works on collective utility functions), which in turn capture the spirit of excellence and diversity.

Our main contribution lies in developing a new method for evaluating multi-winner rules. Specifically, we provide two types of analyses for a selection of prominent multi-winner rules. In Section 3, we derive theoretical upper bounds on how much an outcome of the considered multi-winner rules can differ from the outcomes of CC and AV. We call these bounds CC-guarantee and AV-guarantee. These can be interpreted as worst-case (over all possible preference profiles) guarantees for diversity and individual excellence. Our guarantees are given as functions of the committee size $k$ and return values between 0 and 1. Intuitively, a higher CC-guarantee (resp. AV-guarantee) indicates a better performance in terms of diversity (resp. individual excellence), where 1 denotes that the rule performs as good as CC (resp., AV). Table 1 summarizes our results.

In Section 4, we complement the worst-case analysis from Section 3 with an experimental study yielding approximation ratios for actual data sets. In extensive experiments we estimate how on average the outcomes of the considered rules differ from the outcomes of CC and AV.

Our most important findings can be summarized as follows. Among the studied voting rules, Proportional Approval Voting (PAV) achieves the best compromise between AV and CC; this can be observed both from theoretical and experimental results. The sequential rules seq-PAV and Phragmén's rule, however, achieve almost the same quality while being polynomial-time computable (in contrast to PAV, which is computationally intractable [Skowron *et al.*, 2016; Aziz *et al.*, 2015]). Also the 2-Geometric rule achieves a very good compromise, but is slightly leaning towards diversity. More generally, we show that the $p$-Geometric rule spans the whole spectrum from AV to CC, controlled through the parameter $p$. Hence, by adjusting the parameter $p$, one can obtain any desired compromise between AV and CC.

Finally, we show that while proportional rules tend to achieve a good compromise between individual excellence (AV) and diversity (CC), proportionality does not yield an optimal compromise: we find a non-proportional rule that outperforms any proportional rule with respect to both their AV- and CC-guarantee.

## 1.2 Related Work

Our work is based on the idea of approximation algorithms, where computationally hard problems are solved by polynomial-time algorithms that can guarantee a certain (imperfect) solution quality. In our paper, we study how well popular multi-winner rules can approximate other, archetyp-

ical rules. The work of Brânzei *et al.* [2013] on the dynamic price of anarchy can be viewed in such a perspective: to which degree can the outcome of voting rules based on sincere preferences be approximated by the same voting rules with insincere preference (obtained via "selfish" best-response dynamics)? In a similar vein, Oren and Lucier [2014] study the performance of online social choice procedures in comparison to optimal (offline) procedures. Anshelevich *et al.* [2018] approximate an optimal social choice in a metric model with voting rules using rankings as input, i.e., using limited information.

The normative study of multi-winner election rules typically focuses on axiomatic analysis. For approval-based rules a number of axioms describing proportionality have been recently identified and explored, in particular in the context of the rules that we study in this paper [Aziz *et al.*, 2017a; Sánchez-Fernández *et al.*, 2017; Brill *et al.*, 2017; Skowron *et al.*, 2017; Aziz *et al.*, 2018; Lackner and Skowron, 2018a]. For a survey on properties of multi-winner rules, with the focus on the ideas of individual excellence, diversity, and proportionality, we refer the reader to a survey by Faliszewski et al. [2017].

Another approach to understanding the nature of different multi-winner rules is to analyze how these rules behave on certain subdomains of preferences, where their behavior is much easier to interpret, e.g., on two-dimensional geometric preferences [Elkind *et al.*, 2017], on party-list profiles [Brill *et al.*, 2018], or on single-peaked and single-crossing domains [Aziz *et al.*, 2017b]. Other approaches include quantifying regret and distortion in utilitarian models [Caragiannis *et al.*, 2017], assessing their robustness [Bredereck *et al.*, 2017], and evaluating them based on data collected from surveys [Rapoport *et al.*, 1988; Van der Straeten *et al.*, 2018].

## 2 Preliminaries

For each $t \in \mathbb{N}$, we let $[t] = \{1, \ldots, t\}$. For a set $X$, we write $\mathcal{S}(X)$ to denote the powerset of $X$, i.e., the set of all subsets of $X$. By $\mathcal{S}_k(X)$ we denote the set of all $k$-element subsets of $X$. Let $C = \{c_1, \ldots, c_m\}$ and $N = \{1, \ldots, n\}$ be sets of $m$ *candidates* and $n$ *voters*, respectively. Voters reveal their preferences by indicating which candidates they like: by $A(i) \subseteq C$ we denote the *approval set* of voter $i$ (that is, the set of candidates that $i$ approves of). For a candidate $c \in C$, by $N(c) \subseteq N$ we denote the set of voters who approve $c$. Given a set of candidates $X \subseteq C$, we write $N(X)$ to denote the set of voters who approve at least one candidate in $X$, that is $N(X) = \{i \in N : X \cap A(i) \neq \emptyset\}$. We call the collection of approval sets $A = (A(1), A(2), \ldots, A(n))$, one per each voter, an *approval profile*. We use the symbol $\mathcal{A}$ to represent the set of all possible approval profiles.

We call the elements of $\mathcal{S}_k(C)$ size-$k$ *committees*. Throughout the paper, we use the symbol $k$ to represent the desired size of the committee to be elected. An *approval-based committee rule* (in short, an ABC rule) is a function $\mathcal{R}$ that takes as an input an approval profile and an integer $k \in \mathbb{N}$ (the required committee size), and returns a set of size-$k$ committees. Below, we recall the definitions of ABC rules which are the objects of our study.

| | AV-guarantee | | CC-guarantee | |
|---|---|---|---|---|
| | lower | upper | lower | upper |
| AV | $1$ | $1$ | $\frac{1}{k}$ | $\frac{1}{k}$ |
| CC | $\frac{1}{k}$ | $\frac{1}{k}$ | $1$ | $1$ |
| seq-CC | $\frac{1}{k}$ | $\frac{1}{k}$ | $1 - {}^1/e$ | $1 - (1 - {}^1/k)^k$ |
| PAV | $\frac{1}{2+\sqrt{k}}$ | $\frac{2}{\lfloor\sqrt{k}\rfloor} - \frac{1}{k}$ | $\frac{1}{2}$ | $\frac{1}{2} + \frac{1}{4k-2}$ |
| seq-PAV | $\frac{1}{2\sqrt{k}}$ | $\frac{2}{\lfloor\sqrt{k}\rfloor} - \frac{1}{k}$ | $\frac{1}{\log(k)+2}$ | $\frac{1}{2} + \frac{1}{4k-2}$ |
| $p$-Geometric | $\frac{\mathcal{W}(k\log(p))}{k\log(p)+\mathcal{W}(k\log(p))}$ | $\frac{1}{k} + \frac{2\mathcal{W}(k\log(p))}{k\log(p)}$ | $\frac{p-1}{p}$ | $\frac{p}{p+\frac{k}{k+2}}$ |
| Monroe | $\frac{1}{k}$ | $\frac{1}{k}$ | $\frac{1}{2}$ | $\frac{1}{2} + \frac{1}{k-1}$ |
| seq-Phragmén | $\frac{1}{5\sqrt{k}+1}$ | $\frac{2}{\lfloor\sqrt{k}\rfloor} - \frac{1}{k}$ | $\frac{1}{2}$ | $\frac{1}{2} + \frac{1}{4k-2}$ |

Table 1: These worst-case guarantees are functions of the committee size $k$. A higher value means a better guarantee, with 1 denoting optimal performance. In most cases we could only find (accurate) estimates instead of the exact values of the guarantees: the "lower" and "upper" values in the table denote that the respective guarantee is between these two bounds.

**Multi-winner Approval Voting (AV).** This rule selects $k$ candidates which are approved by most voters. More formally, for a profile $A$ the AV-score of committee $W$ is defined as $\mathrm{sc_{av}}(A, W) = \sum_{c \in W} |N(c)|$, and AV selects committees $W$ that maximize $\mathrm{sc_{av}}(A, W)$.

**Approval Chamberlin–Courant (CC).** For a profile $A$ we define the CC-score of a committee $W$ as $\mathrm{sc_{cc}}(A, W) = \sum_{i \in N} \min\big(1, |A(i) \cap W|\big) = |N(W)|$; CC outputs committees $W$ that maximize $\mathrm{sc_{cc}}(A, W)$. In words, CC aims at finding a committee $W$ such that as many voters as possible have their representatives in $W$ (a representative of a voter is a candidate she approves of). The CC rule was first mentioned by Thiele [1895], and then introduced in a more general context by Chamberlin and Courant [1983].

**Proportional Approval Voting (PAV).** This rule [Thiele, 1895] selects committees with the highest PAV-scores, defined as $\mathrm{sc_{pav}}(A, W) = \sum_{i \in N} \mathrm{H}\big(|W \cap A(i)|\big)$, where $\mathrm{H}(t)$ is the $t$-th harmonic number, i.e., $\mathrm{H}(t) = \sum_{i=1}^{t} {}^1/i$. By using the harmonic function $\mathrm{H}(\cdot)$, voters who already have more representatives in the committee get less voting power than those with fewer representatives; this leads to strong proportionality guarantees [Aziz *et al.*, 2017a].

**$p$-Geometric.** This rule [Skowron *et al.*, 2016] is defined analogously to PAV but uses an exponentially decreasing function instead of $\mathrm{H}(\cdot)$. Formally, for a given parameter $p \geq 1$ the $p$-geometric rule assigns to each committee $W$ the score $\mathrm{sc}_{p\text{-geom}}(A, W) = \sum_{i \in N} \sum_{j=1}^{|A(i) \cap W|} \frac{1}{p^j}$, and picks the committees with the highest scores. Note that the 1-geometric rule is simply AV.

**Sequential CC/PAV/$p$-Geometric.** For each rule $\mathcal{R} \in \{\mathrm{CC}, \mathrm{AV}, \mathrm{PAV}, p\text{-geometric}\}$, we define its sequential variant, denoted as seq-$\mathcal{R}$, as follows. We start with an empty solution $W = \emptyset$ and in each of the $k$ consecutive steps we add to $W$ a candidate $c$ that maximizes $\mathrm{sc}_{\mathcal{R}}(A, W \cup \{c\})$,

i.e., the candidate that improves the committee's score most. We break ties lexicographically.

**Monroe.** Monroe's approval-based rule [1995], similarly to CC, aims at maximizing the number of voters who are represented in the elected committee. The difference is that Monroe additionally imposes that each committee member should represent roughly the same number of voters. Formally, a Monroe assignment of the voters to a committee $W$ is a function $\phi \colon N \to W$ such that each candidate $c \in W$ is assigned roughly the same number of voters, i.e., that $\lfloor {}^n/k \rfloor \leq |\phi^{-1}(c)| \leq \lceil {}^n/k \rceil$. Let $\Phi(W)$ be the set of all possible Monroe assignments to $W$. The Monroe-score of $W$ is defined as $\mathrm{sc_{Mon}}(A, W) = \max_{\phi \in \Phi(W)} \sum_{i \in N} |A(i) \cap \{\phi(i)\}|$; the rule returns all committees $W$ that maximize $\mathrm{sc_{Mon}}(A, W)$.

**Phragmén's Sequential Rule (seq-Phragmén).** This rule [Phragmén, 1894; Brill *et al.*, 2017] can be understood as a load distribution procedure. Each selected committee member $c$ is associated with one unit of load that needs to be distributed among those voters who approve $c$. The rule proceeds in $k$ steps. In each step it selects one candidate and distributes its load as follows: let $\ell_j(i-1)$ denote the total load assigned to voter $j$ just before the $i$-th step. In the $i$-th step the rule selects a candidate $c$ and finds a load distribution $\{x_j \colon j \in N\}$ that satisfies the following three conditions: (1) $x_j > 0$ implies that voter $j$ approves $c$, (2) $\sum_{j \in N} x_j = 1$, and (3) the maximum load assigned to a voter, $\max_{j \in N}(\ell_j(i-1) + x_j)$, is minimized. The new total load assigned to a voter $j \in N$ after the $i$-th step is $\ell_j(i) = \ell_j(i-1) + x_j$.

## 3 Worst-Case Guarantees

In this section we analyze the multi-winner rules from Section 2 with respect to how well they perform in terms of *diversity* and *individual excellence*. In our study we use the

established idea of approximation in a novel way: by estimating how well a given rule $\mathcal{R}$ approximates CC (resp., AV), we quantify how $\mathcal{R}$ performs with respect to diversity (resp., individual excellence). This differs from the typical use of the idea of approximation in the following aspects: (1) We do not seek new algorithms approximating a given objective function as well as possible, but rather analyze how well the existing known rules approximate given functions (even if it is apparent that better and simpler approximation algorithms exist, these algorithms might not share other important properties of the considered rules). (2) We are not approximating computationally hard rules with rules easier to compute. On the contrary, we will be even investigating how computationally hard rules (such as PAV, Monroe, etc.) approximate the algorithmically trivial AV rule.

**Definition 1.** *Recall that for a profile $A$, $\mathrm{sc}_{\mathrm{av}}(A, W)$ and $\mathrm{sc}_{\mathrm{cc}}(A, W)$ denote the AV-score and CC-score of committee $W$, respectively. The AV-guarantee of an ABC rule $\mathcal{R}$ is a function $\kappa_{\mathrm{av}} \colon \mathbb{N} \to [0, 1]$ that takes as input an integer $k$, representing the size of the committee, and is defined as:*

$$\kappa_{\mathrm{av}}(k) = \inf_{A \in \mathcal{A}} \frac{\min_{W \in \mathcal{R}(A,k)} \mathrm{sc}_{\mathrm{av}}(A, W)}{\max_{W \in \mathcal{S}_k(C)} \mathrm{sc}_{\mathrm{av}}(A, W)}.$$

*Analogously, the CC-guarantee of $\mathcal{R}$ is defined by*

$$\kappa_{\mathrm{cc}}(k) = \inf_{A \in \mathcal{A}} \frac{\min_{W \in \mathcal{R}(A,k)} \mathrm{sc}_{\mathrm{cc}}(A, W)}{\max_{W \in \mathcal{S}_k(C)} \mathrm{sc}_{\mathrm{cc}}(A, W)}.$$

The AV and CC-guarantees can be viewed as quantitative properties of multi-winner rules. In the remaining part of this section we evaluate the previously defined rules against their AV- and CC-guarantees. Clearly, the AV-guarantee of Approval Voting and the CC-guarantee of the Chamberlin–Courant rule are the constant-one function. We start by establishing the AV-guarantee of CC and vice versa. Due to space constraints, most proofs had to be omitted, but can be found in a longer version [Lackner and Skowron, 2018b].

**Proposition 1.** *The CC-guarantee of AV is $1/k$.*

**Proposition 2.** *The AV-guarantee of CC and seq-CC is $1/k$.*

Propositions 1 and 2 give a baseline for our further analysis. In particular, we would expect that "good" rules implementing a tradeoff between diversity and individual excellence should have AV and CC-guarantees better than $1/k$.

We note that the CC-guarantee of seq-CC is $1 - (1 - 1/k)^k$ (which approaches $1 - 1/e \approx 0.63$ for large $k$). This is the result of the fact that seq-CC is a $(1 - (1 - 1/k)^k)$-approximation algorithm for CC [Lu and Boutilier, 2011].

Let us turn our attention to the Monroe rule, which is often considered a proportional rule (e.g., it satisfies proportional justified representation [Sánchez-Fernández *et al.*, 2017]). Hence, one could expect that this rule offers a good compromise between AV and CC. Perhaps surprisingly, this is not the case: Monroe does not offer a better AV-guarantee than CC.

**Proposition 3.** *The AV-guarantee of Monroe is $1/k$, its CC-guarantee is between $\frac{1}{2}$ and $\frac{1}{2} + \frac{1}{k-1}$.*

Let us now move to multi-winner voting systems offering asymptotically better guarantees than Monroe. As we will see, the examination of such rules requires a more complex combinatorial analysis. We start with PAV:

**Theorem 1.** *The AV-guarantee of PAV is between $\frac{1}{2+\sqrt{k}}$ and $\frac{2}{\sqrt{k}}$; its CC-guarantee between $\frac{1}{2}$ and $\frac{1}{2} + \frac{1}{4k-2}$.*

*Proof.* To give a flavor of the proof techniques used in this paper, we show that the AV-guarantee of PAV is at least equal to $\frac{1}{2+\sqrt{k}}$. Consider an approval profile $A$ and a PAV-winning committee $W_{\mathrm{pav}}$; let $\bar{n} = |N(W_{\mathrm{pav}})|$ denote the number of voters who approve some member of $W_{\mathrm{pav}}$. For each $i \in N$ we set $w_i = |A(i) \cap W_{\mathrm{pav}}|$. Let $W_{\mathrm{av}}$ be a committee with the highest AV-score. W.l.o.g., we can assume that $W_{\mathrm{av}} \neq W_{\mathrm{pav}}$. Now, consider a candidate $c \in W_{\mathrm{av}} \setminus W_{\mathrm{pav}}$ with the highest AV-score, and let $n_c = |N(c)|$ denote the number of voters who approve $c$. If we replace a candidate $c' \in W_{\mathrm{pav}}$ with $c$, the PAV-score of $W_{\mathrm{pav}}$ will change by:

$$
\begin{aligned}
\Delta(c, c') &= \sum_{i \colon c \in A(i) \wedge c' \notin A(i)} \frac{1}{w_i + 1} - \sum_{i \colon c' \in A(i) \wedge c \notin A(i)} \frac{1}{w_i} \\
&= \sum_{i \colon c \in A(i)} \frac{1}{w_i + 1} - \sum_{i \colon c' \in A(i)} \frac{1}{w_i} \\
&\quad + \sum_{i \colon \{c, c'\} \subseteq A(i)} \frac{1}{w_i} - \frac{1}{w_i + 1} \\
&\geq \sum_{i \in N(c)} \frac{1}{w_i + 1} - \sum_{i \in N(c')} \frac{1}{w_i}.
\end{aligned}
$$

Let us now compute the sum $\sum_{c' \in W_{\mathrm{pav}}} \Delta(c, c') =$

$$
\begin{aligned}
&= \sum_{c' \in W_{\mathrm{pav}}} \sum_{i \in N(c)} \frac{1}{w_i + 1} - \sum_{c' \in W_{\mathrm{pav}}} \sum_{i \in N(c')} \frac{1}{w_i} \\
&= k \sum_{i \in N(c)} \frac{1}{w_i + 1} - \sum_{i \in N} \sum_{c' \in W_{\mathrm{pav}} \cap A(i)} \frac{1}{w_i} \\
&= k \sum_{i \in N(c)} \frac{1}{w_i + 1} - \bar{n}
\end{aligned}
$$

We know that for each $c' \in W$ we have $\Delta(c, c') \leq 0$, thus $k \sum_{i \in N(c)} \frac{1}{w_i+1} - \bar{n} \leq 0$ and $\sum_{i \in N(c)} \frac{1}{w_i+1} \leq \frac{\bar{n}}{k}$. We now use the inequality between the harmonic and arithmetic mean to obtain:

$$\frac{\bar{n}}{k} \geq \sum_{i \in N(c)} \frac{1}{w_i + 1} \geq \frac{n_c^2}{\sum_{i \in N(c)}(w_i + 1)}.$$

This can be transformed to:

$$kn_c \leq \frac{\bar{n}\left(\sum_{i \in N(c)} w_i + n_c\right)}{n_c} = \frac{\bar{n} \sum_{i \in N(c)} w_i}{n_c} + \bar{n}.$$

Now, let us consider two cases. If $\bar{n} \leq n_c\sqrt{k}$, then we observe that:

$$\frac{\mathrm{sc}_{\mathrm{av}}(A, W_{\mathrm{av}})}{\mathrm{sc}_{\mathrm{av}}(A, W_{\mathrm{pav}})} \leq \frac{\sum_{i \in N} w_i + kn_c}{\sum_{i \in N} w_i} = 1 + \frac{kn_c}{\sum_{i \in N} w_i}$$

$$\leq 1 + \frac{\frac{\bar{n} \sum_{i \in N(c)} w_i}{n_c} + \bar{n}}{\sum_{i \in N} w_i} \leq 2 + \frac{\bar{n}}{n_c} \leq \sqrt{k} + 2.$$
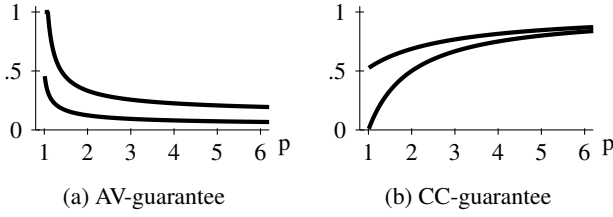
(a) AV-guarantee      (b) CC-guarantee

Figure 1: Guarantees for the $p$-Geometric rule for $k = 20$ and varying $p$. In both figures, the upper and the lower line depict the upper and lower bound from Theorem 3.

On the other hand, if $\bar{n} \geq n_c \sqrt{k}$, then:

$$\frac{\mathrm{sc}_{\mathrm{av}}(A, W_{\mathrm{av}})}{\mathrm{sc}_{\mathrm{av}}(A, W_{\mathrm{pav}})} \leq \frac{\sum_{i \in N} w_i + k n_c}{\sum_{i \in N} w_i} = 1 + \frac{k n_c}{\sum_{i \in N} w_i}$$
$$\leq 1 + {k n_c}/{\bar{n}} \leq 1 + \sqrt{k}.$$

In either case we have that $\frac{\mathrm{sc}_{\mathrm{av}}(A, W_{\mathrm{pav}})}{\mathrm{sc}_{\mathrm{av}}(A, W_{\mathrm{av}})} \geq \frac{1}{2 + \sqrt{k}}$. $\qquad\square$

For sequential PAV we can prove an AV-guarantee qualitatively similar to the one for PAV. Concerning its CC-guarantee, however, the gap between the lower and upper bounds is large; finding a more accurate estimate remains an interesting open question.

**Theorem 2.** *The AV-guarantee of sequential PAV is between* $\frac{1}{2\sqrt{k}}$ *and* $\frac{2}{\lfloor\sqrt{k}\rfloor} - \frac{1}{k}$; *its CC-guarantee is between* $\frac{1}{\log(k)+2}$ *and* $\frac{1}{2} + \frac{1}{4k-2}$.

The following theorem states guarantees for the $p$-geometric rule. Let $\mathcal{W}(\cdot)$ denote the Lambert W function, a function that increases asymptotically slower than $\log$.

**Theorem 3.** *The AV-guarantee of the $p$-geometric rule is between*

$$\frac{\mathcal{W}(k \log(p))}{k \log(p) + \mathcal{W}(k \log(p))} \quad and \quad \frac{2\mathcal{W}(k \log(p))}{k \log(p)} + \frac{1}{k},$$

*its CC-guarantee is between* $\frac{p-1}{p}$ *and* $\frac{p}{p + \frac{k}{k+2}}$.

The guarantees of Theorem 3 are visualized in Figure 1. We can see that $p$-geometric rules, for $p \in [1, \infty)$, form a spectrum connecting AV and CC (with $p \to 1$ we approach AV and with $p \to \infty$ we approach CC): by adjusting the parameter $p$ one can control the tradeoff between the diversity and individual excellence.

Finally, we consider seq-Phragmén, also a rule aimed at achieving proportional representation. It achieves guarantees similar to PAV.

**Theorem 4.** *The AV-guarantee of seq-Phragmén is between* $\frac{1}{5\sqrt{k}+1}$ *and* $\frac{2}{\lfloor\sqrt{k}\rfloor} - \frac{1}{k}$; *its CC-guarantee is between* $\frac{1}{2}$ *and* $\frac{1}{2} + \frac{1}{4k-2}$.

Let us conclude Section 3 with an investigation on how proportionality relates to diversity and individual excellence. In an initial stage of this study we conjectured that proportionality can be characterized as a certain compromise between diversity and individual excellence. However, as we

will argue below, proportionality should be rather viewed as a third, independent objective. Indeed, we will construct an ABC rule that is strictly better than any proportional rules with regard to both its AV- and CC-guarantee.

For each $\alpha \in [0, 1]$ we define $\alpha$-CC-AV as a linear combination of CC and AV. For an approval-based profile $A$, $\alpha$-CC-AV first computes a size-$\lceil\alpha k\rceil$ committee $W_1$ that maximizes $\mathrm{sc}_{\mathrm{cc}}(A, W_1)$, and a size-$\lfloor(1-\alpha)k\rfloor$ committee $W_2$ that maximizes $\mathrm{sc}_{\mathrm{av}}(A, W_2)$; then, the rule returns committee $W = W_1 \cup W_2$.

**Proposition 4.** *The CC-guarantee of $\alpha$-CC-AV is at least $\alpha$; its AV-guarantee is at least $1 - \alpha - {1}/{k}$.*

Let us now examine what AV- and CC-guarantees a proportional rule can achieve. We consider a very weak definition of proportionality, called lower quota [Balinski and Young, 1982]. Lower quota applies only to *party-list profiles*: approval profiles $A$ in which for all voters $i, j \in N$ it holds that either $A(i) = A(j)$ or $A(i) \cap A(j) = emptyset$; identical voters belong to "same party". Lower quota guarantees each group $N' \subseteq N$ to receive at least $\min\left(\left|\bigcap_{i \in N'} A(i)\right|, \left\lfloor\frac{k|N'|}{|N|}\right\rfloor\right)$ representatives in any winning committee, i.e., they receive a fair proportion of the committee seats (rounded down) as long as they have sufficient joint candidates. Lower quota is strictly weaker than proportionality axioms typically used for ABC rules (such as extended and proportional justified representation [Sánchez-Fernández *et al.*, 2017; Aziz *et al.*, 2017a]).

**Proposition 5.** *The AV-guarantee of a rule that satisfies lower quota is at most $\frac{2}{\lfloor\sqrt{k}\rfloor} - \frac{1}{k}$, its CC-guarantee is at most $\frac{3}{4} + \frac{3}{8k-4}$.*

These bounds yield that $\alpha$-CC-AV for $\alpha = 0.76$ achieves a better AV- and CC-guarantee for all $k \geq 81$. In particular, the AV-guarantee of $\alpha$-CC-AV is superior: it guarantees a constant fraction of the optimal AV score. We conclude that proportional rules can achieve a desirable compromise between diversity and individual excellence (e.g., PAV), but this compromise is not optimal (as we have just seen) and not all proportional rules achieve good AV-guarantees (Monroe performs no better than CC).

## 4 Average Guarantees: An Experimental Analysis

To complement the theoretical analysis of Section 3, we have run experiments that aim at assessing *AV-ratios* and *CC-ratios* achieved by several voting rules. These two ratios are per-instance analogues of AV- and CC-guarantee and are defined as follows: Given an ABC rule $\mathcal{R}$ and a profile $A$, the AV-ratio for $\mathcal{R}(A, k)$ is defined as:

$$\frac{\min_{W \in \mathcal{R}(A,k)} \mathrm{sc}_{\mathrm{av}}(A, W)}{\max_{W \in \mathcal{S}_k(C)} \mathrm{sc}_{\mathrm{av}}(A, W)};$$

the CC-ratio is defined analogously. In the following experiments, we have calculated the AV- and CC-ratios for real-world and randomly generated profiles and compared them for different voting rules.
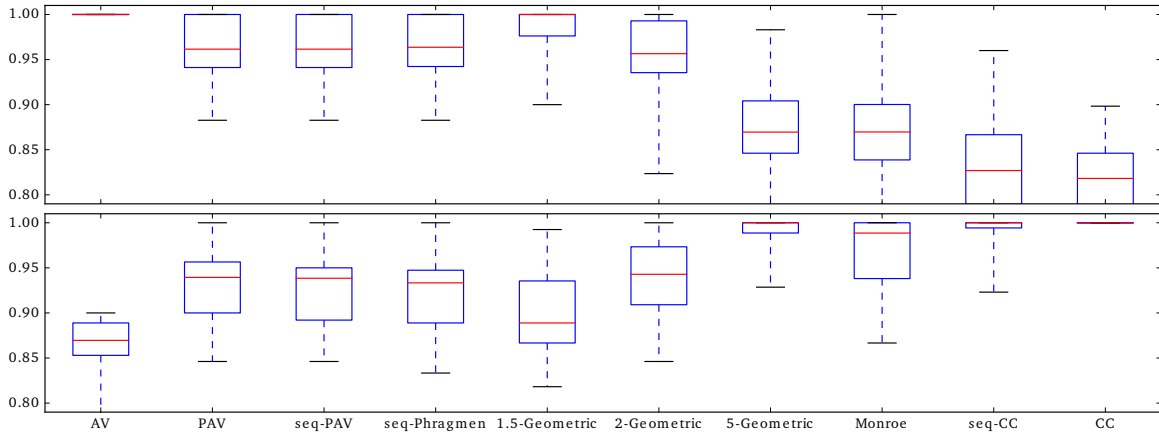
Figure 2: Results for the preflib dataset (upper boxplot shows AV-ratios, the lower CC-ratios).

We have used two data sets: profiles obtained from preflib. org [Mattei and Walsh, 2013] and profiles generated via a uniform distribution. We restricted our attention to profiles where both the AV-ratio of CC and the CC-ratio of AV is at most $0.9$. This excludes profiles where an (almost) perfect compromise between AV and CC exists. The uniform dataset consists of 500 profiles with 20 candidates and 50 voters, each. Voters' approval sets are of size 2–5 (chosen uniformly at random); the approval sets of a given size are also chosen uniformly at random. Experiments for the uniform dataset use a committee size of $k = 5$.

As preflib.org does not offer approval-based datasets, we extracted approval profiles from ranked ballots as follows: for each ranked profile and $i \in \{1, \ldots, k-1\}$, we generated an approval profile assuming that voters approve all candidates that are ranked in the top $i$ positions. As before, we excluded profiles that allowed an almost perfect compromise between AV and CC. For the preflib dataset we considered $k \in \{3, \ldots, 7\}$, obtaining a total number of 243 instances.

We considered the following ABC rules: AV, CC, seq-CC, PAV, seq-PAV, seq-Phragmén, Monroe, and the 1.5-, 2-, and 5-Geometric rule. Our results are displayed as boxplots in Figure 2 for the preflib dataset; results for the random data set are largely similar. The top and bottom of boxes represent first and third quantiles, the middle red bar shows the median. The dashed intervals (whiskers) show the range of all values, i.e., the minimum and maximum AV- /CC-ratios.

The main conclusion from the experiments is that the classification obtained from worst-case analytical bounds also holds in our (average-case) experiments. PAV, seq-PAV, and seq-Phragmén perform very well with respect to the AV-ratio, beaten only by 1.5-Geometric and AV itself. This is mirrored by our theoretical results as only PAV, seq-PAV, and seq-Phragmén achieve a $\Theta(1/\sqrt{k})$ AV-guarantee. Also the 2-Geometric rule achieves comparable AV-ratios. Even better AV-ratios are achieved only by 1.5-Geometric and AV.

Considering the CC-ratio, we see almost optimal performance of seq-CC, Monroe, and 5-Geometric, and good performance of PAV, seq-PAV, seq-Phragmén, and 2-Geometric. Minor variations within these groups seem to depend on the

chosen dataset. We also observe that 5-Geometric is better than Monroe's rule and seq-CC according to both criteria.

When looking at the three Geometric rules considered here, we see the transition from AV to CC as our theoretical findings predict (cf. Figure 1): 1.5-Geometric is close to AV, whereas 5-Geometric resembles CC; 2-Geometric performs very similarly to PAV, slightly favoring diversity over individual excellence. To sum up, our experimental findings indicate that PAV is the best compromise between AV and CC among the considered rules. Yet, seq-PAV, seq-Phragmén, and 2-Geometric achieve comparable ratios, and the former two are much cheaper to compute.

## 5 Conclusion and Future Work

This work presents a new tool for assessing the level of diversity and individual excellence that multi-winner rules provide. Our results can help to understand the landscape of multi-winner rules, specifically how they behave with respect to two contradictory goals.

Our work can be extended in several directions. First, we have focused on approval-based multi-winner rules—a natural next step is to perform a similar analysis for multi-winner rules that take rankings over candidates as input. Second, we have excluded some interesting voting rules from our analysis, in particular reverse-sequential PAV [Skowron et al., 2017] and Minimax Approval Voting [Brams et al., 2007]; it is unclear how they compare to rules considered in this paper. Finally, we have chosen AV and CC as extreme notions that represent diversity and individual excellence. Another natural approach would be to take a proportional rule (such as PAV) as a standard and see how well others rules approximate it.

# References

[Anshelevich *et al.*, 2018] E. Anshelevich, O. Bhardwaj, E. Elkind, J. Postl, and P. Skowron. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264:27–51, 2018.

[Aziz *et al.*, 2015] H. Aziz, S. Gaspers, J. Gudmundsson, S. Mackenzie, N. Mattei, and T. Walsh. Computational aspects of multi-winner approval voting. In *Proc. of AAMAS-2015*, pages 107–115, 2015.

[Aziz *et al.*, 2017a] 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.

[Aziz *et al.*, 2017b] H. Aziz, E. Elkind, P. Faliszewski, M. Lackner, and P. Skowron. The Condorcet principle for multiwinner elections: From shortlisting to proportionality. In *Proc. of IJCAI-2017*, pages 84–90, 2017.

[Aziz *et al.*, 2018] H. Aziz, E. Elkind, S. Huang, M. Lackner, L. Sánchez-Fernández, and P. Skowron. On the complexity of extended and proportional justified representation. In *Proc. of AAAI-2018*, 2018. To appear.

[Balinski and Young, 1982] M. Balinski and H. P. Young. *Fair Representation: Meeting the Ideal of One Man, One Vote*. Yale University Press, 1982. (2nd Edition [with identical pagination], Brookings Institution Press, 2001).

[Brams *et al.*, 2007] S. J. Brams, D. M. Kilgour, and M. R. Sanver. A minimax procedure for electing committees. *Public Choice*, 132(3–4):401–420, 2007.

[Brânzei *et al.*, 2013] S. Brânzei, I. Caragiannis, J. Morgenstern, and A.D. Procaccia. How bad is selfish voting? In *Proc. of AAAI-2013*, pages 138–144, 2013.

[Bredereck *et al.*, 2017] R. Bredereck, P. Faliszewski, A. Kaczmarczyk, R. Niedermeier, P. Skowron, and N. Talmon. Robustness among multiwinner voting rules. In *Proc. of SAGT-2017*, pages 80–92, 2017.

[Brill *et al.*, 2017] M. Brill, R. Freeman, S. Janson, and M. Lackner. Phragmén's voting methods and justified representation. In *Proc. of AAAI-2017*, pages 406–413, 2017.

[Brill *et al.*, 2018] M. Brill, J. Laslier, and P. Skowron. Multiwinner approval rules as apportionment methods. *Journal of Theoretical Politics*, 30(3):358–382, 2018.

[Caragiannis *et al.*, 2017] I. Caragiannis, S. Nath, A. D. Procaccia, and N. Shah. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152, 2017.

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

[Elkind *et al.*, 2017] 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 *Proc. of AAAI-2017*, 2017. 494–501.

[Faliszewski *et al.*, 2017] 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*. AAAI Press, 2017.

[Lackner and Skowron, 2018a] M. Lackner and P. Skowron. Consistent approval-based multi-winner rules. In *Proc. of ACM-EC-2018*, pages 47–48, 2018.

[Lackner and Skowron, 2018b] M. Lackner and P. Skowron. A quantitative analysis of multi-winner rules. *arXiv preprint arXiv:1801.01527*, 2018.

[Lu and Boutilier, 2011] T. Lu and C. Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proc. of IJCAI-2011*, pages 280–286, 2011.

[Mattei and Walsh, 2013] N. Mattei and T. Walsh. Preflib: A library for preferences http: //www.preflib.org. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT-2013)*, pages 259–270, 2013.

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

[Oren and Lucier, 2014] J. Oren and B. Lucier. Online (budgeted) social choice. In *Proc. of AAAI-2014*, pages 1456–1462, 2014.

[Phragmén, 1894] 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.

[Rapoport *et al.*, 1988] A. Rapoport, D. Felsenthal, and Z. Maoz. Proportional representation: An empirical evaluation of single-stage, non-ranked voting procedures. *Public Choice*, 59(2):151–165, 1988.

[Sánchez-Fernández *et al.*, 2017] L. Sánchez-Fernández, E. Elkind, M. Lackner, N. Fernández, J. A. Fisteus, P. Basanta Val, and P. Skowron. Proportional justified representation. In *Proc. of AAAI-2017*, pages 670–676, 2017.

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

[Skowron *et al.*, 2017] P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. Proportional rankings. In *Proc. of IJCAI-2017*, pages 409–415, 2017.

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

[Van der Straeten *et al.*, 2018] K. Van der Straeten, R. Lachat, and J.-F. Laslier. Strategic voting in multiwinner elections with approval balloting: An application to the 2011 regional government election in Zurich. In J. Aldrich, A. Blais, and L. Stephenson, editors, *The Many Faces of Strategic Voting*. CBS, 2018. forthcoming.