Compiling the Votes of a Subelectorate for Multi-Winner Voting Rules

Neel Karia *[1][0000−0003−2751−9526] and Jérôme Lang[2][0000−0001−6789−9573]

[1] Columbia University, New York, USA
nmk2154@columbia.edu
[2] CNRS, LAMSADE, PSL, Université Paris Dauphine, Paris, France
lang@lamsade.dauphine.fr

Abstract. Compiling the votes of a subelectorate is a well-known problem in computational social choice. The goal is to store the information contained in the votes cast by a subelectorate in a space-efficient way, such that when the rest of the votes become available, the winners can be accurately ascertained. This problem has been studied for single-winner voting rules. We provide a comprehensive compilation complexity landscape for several ordinal and approval-based multi-winner voting rules.

Keywords: Voting · Compilation Complexity · Social Choice

1 Introduction

In a usual voting setting, it is assumed that the votes cast by the agents arrive simultaneously. A voting rule is then applied to elect a winner (or a set of co-winners). However, in most real-life scenarios, the votes are not obtained at the same time or at the same place. It is then beneficial to compile the votes that are already available, that is, to store the information contained in these votes, using as little space as possible, in such a way that when the rest of the votes are known, the winner(s) can be determined. The compilation complexity of a voting rule is the worst-case size of the most succinct compilation. Compilation has several advantages: first, the votes of the subelectorate can be stored succinctly. second, the compilation provides a synthetic, understandable view of the votes of the electorate, which makes verification easier. Lastly, it may help us develop dynamic programming algorithms for elections under uncertainty, by storing partial results (refer to [13]).

Compilation of single-winner rules has been studied in [4] and [19]. The compilation complexity of voting rules is related to their communication complexity (initiated in [6]), with a major difference: communication complexity allows several communication rounds, while compilation allows only one round (and is thus related to one-round communication complexity). It is also related to the possible and necessary winners problems (initiated in [15]). More recently, there has been related work on sample complexity for approximate winner predictions

*Corresponding author

[7] and algorithms for the approximate compilation of information from vote streams for some multi-winner voting rules [8]; a major difference of our work from [8] being that ours considers an offline and exact version of compilation.

We study the compilation complexity of multi-winner voting rules. The interpretation of compilation can be either temporal or spatial. We illustrate both with an example. Assume we are considering multi-winner approval voting (MAV) – voter ballots are subsets of candidates, and the winners are the candidates with the highest $k = 2$ scores (using tie-breaking if necessary). We have four candidates a, b, c, d, out of which we want to select two winners.

For the temporal interpretation, assume that votes come in two stages: say today's votes are two votes (abc)[‡], one vote (abd), one (bc), and one (cd); we will have more votes tomorrow (but we do not know how many). Which information should we keep from these 5 votes received today? The answer is intuitive enough (and proven formally in the paper): it is sufficient and necessary to store the information that both b and c have one more approval than a, which in turn has one more approval than d. Once we know the late votes, it will then be easy to determine the winners: for instance, if we get two votes (abd) and one vote (a), then we know that the winners for MAV are a and b.

For the spatial interpretation, assume that votes are collected at two different polling stations (say, two towns). At station 1 we have two votes (abc), one vote (abd), one (bc), and one (cd); at station 2 we have two votes (abd) and one vote (a). We need to publish local results, for several reasons: the people would like to know how many votes the candidates got in their town, and the local observers want to check that the results of ballot counting in their polling station (which they assisted in) correspond to the official figures published the day after the election. In most cases, it is undesirable to publish anonymized approval ballots: this could reveal too much information and would take up too much space if there are many candidates and voters. Instead, for each polling station we can simply publish the list all candidates ranked by non-increasing scores, together with their relative differences of the approval scores. This information is all we need to compute the winners (and it is also necessary, as we will later show).

For both interpretations, it is useful to have an order of magnitude estimate of the space needed to store the information coming for a subelectorate.

In Section 2 we give the necessary background on compilation functions, compilation complexity, and multi-winner voting rules. In Section 3 we extend the compilation framework from single-winner to multi-winner rules. In Section 4 we show results on optimal compilation equivalence conditions and compilation complexity for various multi-winner rules with ordinal input. In Section 5 we do the same for multi-winner rules with approval-based inputs.

[‡]For approval-based rules, by (abc) we indicate that a, b, c are approved in the vote. For ordinal rules we sometimes use (abc) to indicate the vote $a \succ b \succ c$.

2 Background

Let A be a set of candidates, with $|A| = m$. Let $[m] = \{1, \dots, m\}$. Let $\mathcal{P}_A$ be the space of *votes*, which, depending on the rule used, is either the set of linear orders over A (ranked/ordinal ballots) or the set of subsets of A (approval ballots). $\mathcal{P}_A^n$ represents the space of n votes (either ordinal or approval). A *preference profile* $P \in \mathcal{P}_A^n$ is a collection $(V_1, \dots, V_n)$ of n votes. For any given n it belongs to $\mathcal{P}_A^* = \bigcup_{n \geq 1} P_A^n$ (the set of profiles with an arbitrary number of votes).

An *irresolute* multi-winner voting rule is a function g that maps any profile P and any $k \in [m]$ to a nonempty subset of k-committees $g(P, k) \in \mathcal{S}_k(A)$. The restriction of an irresolute multi-winner rule to $k = 1$ is an irresolute single-winner rule. Note that all our results use irresolute rules when unspecified.

2.1 Compilation Complexity and Compilation Functions for Single-Winner Voting Rules

For single-winner voting rules, compilation complexity was introduced in [4], and further studied in [3,19]. We recall the setting for single-winner rules.

Let f be a single-winner voting rule defined for a profile of rankings or approvals over candidates. Consider any two profiles $P, Q \in \mathcal{P}_A^n$, that contain the votes of subelectorates composed of n voters. We say that P and Q are f-*equivalent* [4], which we denote by $P \sim_f Q$, if for each $t \geq 0$ and each profile $T \in \mathcal{P}_A^t$, we have $f(P \cup T) = f(Q \cup T)$. Clearly, $\sim_f$ is an equivalence relation. This notion of f-equivalence is further generalised in [19], where the number t of new votes is known beforehand; we will not use this notion in this paper.

A function $\sigma : \mathcal{P}_A^n \to \{0, 1\}^*$ is a *compilation function for* f if there exists a function $\rho : \{0, 1\}^* \times \mathcal{P}_A^* \to A$ such that for all $P \in \mathcal{P}_A^n$, $t \geq 0$ and $T \in \mathcal{P}_A^t$, we have $\rho(\sigma(P), T) = f(P \cup T)$. By $size(\sigma)$, we denote the number of bits needed to represent $\sigma(P)$. The *compilation complexity* of f, denoted $C(f)$, is the minimum value of $size(\sigma)$ over all compilation functions for f. It is known that $C(f)$ is the upper integer part of the logarithm (in base 2) of the number of equivalence classes for $\sim_f$ (*Proposition 1* of [4]).

2.2 Multi-Winner Voting Rules

Ordinal Multi-Winner Rules We start with rules whose input is a profile consisting of ranked ballots (*ordinal* rules), for which a good overview is in [11]. All of them but one (sequential plurality) are defined via *scores*: given a profile $P = (V_1, \dots, V_n)$, a score $sc(S, P)$ is associated with each committee $S \in \mathcal{S}_k(A)$, and the winning committee W is the one that maximizes $sc(S, P)$. If P contains a single vote V then we use $sc(S, V)$ instead of $sc(S, P)$. The multi-winner irresolute rule that selects all committees maximizing sc is denoted by g^{sc}.

For the first four rules below, the score of a k-committee $S \in \mathcal{S}_k(A)$ is the sum of the scores $sc(S, V_i)$ that it gets from all votes $V_1, \dots, V_n$.

- **Single Non Transferable Vote (SNTV)**: $sc(S, V_i) = 1$ if S contains the top candidate of V_i; else $sc(S, V_i) = 0$.

- **Bloc**: $sc(S, V_i)$ is the number of candidates in S ranked in the first k positions of V_i.
- k**-Borda**: $sc(S, V_i)$ is the sum of the Borda scores of the candidates in S; the Borda score of a candidate ranked in position j in a vote is $m - j$.
- **Chamberlin-Courant (**β**-CC)**: $sc(S, V_i)$ is the Borda score w.r.t. V_i of the best candidate in S according to V_i.
- **Gehrlein Stable Rules** [12]: We say that S is a weak Condorcet set if for every candidates c in S and d in $A \backslash S$, at least half of the voters prefer c to d. A multi-winner rule is Gehrlein stable if it outputs a weak Condorcet set whenever there exists one. Two Gehrlein-Stable rules are NED (*number of external defeats*) and SEO (*size of external opposition*) [5], which can be seen as the respective multi-winner counterparts of the Copeland and maximin single-winner rules. The NED rule outputs committees S that maximize the number of pairs $(x, y) \in S \times A \setminus S$ such that x majority-beats y. The SEO rule outputs committees S that maximise $\min_{x \in S, y \in A \setminus S} |\{i : x \succ_i y\}|$.
- **Theta-Winning Sets (**θ**)** [10] – also called LSE-maximin [1]. A k-committee S is a θ-winning set if, for every candidate d not in S, more than a θ-fraction of the voters prefer some member of S to d. A winning committee W is a k-committee being is a θ-winning set for the largest value of θ.
- **Sequential Plurality (SeqPl)** [2]: We proceed in rounds. Initially, $W = \emptyset$. The candidate ranked first by the largest number of votes is added to W, removed from the profile, and the procedure is repeated k times (breaking ties if resolute, or considering all possibilities if irresolute). W is the output.

SNTV, Bloc, k-Borda and β-CC belong to the larger family of *committee scoring rules* [9,18], for which (1) $sc(S, P) = \sum_{V_i \in P} sc(S, V_i)$, and (2) $sc(S, V_i)$ is a function of the vector containing the ranks of the elements of S in V_i. Moreover, SNTV, Bloc and k-Borda are *decomposable* committee scoring rules: each candidate x has a score function $sc(x, V)$, such that $sc(S, V) = \sum_{x \in S} sc(x, V)$.

An important ordinal rule that we do not consider is Single Transferable Vote (STV), as it seems quite challenging to characterise its compilation complexity.

Approval-based Multi-Winner Voting Rules For these three rules, the input is a profile consisting of approval ballots. Refer to [16] for a detailed survey.

- **Multiwinner Approval Voting (MAV)**: The winning committee consists of the k candidates that are approved most frequently.
- **Approval-based Chamberlin-Courant (**α**-CC)**: $sc(S, V_i) = 1$ if V_i votes for at least one candidate of S; else $sc(S, V_i) = 0$.
- **Proportional Approval Voting (PAV)**: $sc(S, V_i) = \sum_{x=1}^{j} \frac{1}{x}$ if V_i approves j candidates of S.

For the three approval-based rules above, $sc(S, P) = \sum_{V_i \in P} sc(S, V_i)$. These three rules belong to a family of voting rules called the *Thiele's Optimization Method* [14], where $sc(S, V_i) = h(|S \cap V_i|)$ where h is a function from $\mathbb{N}_0$ to $\mathbb{R}$ and $h(1)$ is normalised to 1. We focus on a subset of these rules for which

$h(|S \cap V_i|) \leq |S \cap V_i|$ (which include all the three rules above for $h(r) = r, h(r) = 1$, and $h(r) = \sum_{i=1}^{r} \frac{1}{r}$ respectively).

Two important approval-based rules that we are not considering are Phragmén [14] and the Method of Equal Shares [17] because we do not have characterisations of their compilation complexity. This is left for future work.

3 Compilation for Multi-Winner Voting Rules

The concepts of equivalence classes, compilation complexity and compilation functions used for single-winner rules can easily be extended to multi-winner rules (ordinal or approval-based) as defined below.

Definition 1. *Let g be a multi-winner voting rule. Let $k \in [m]$ be fixed. Consider any two profiles $P, Q \in \mathcal{P}_A^n$, that contain the votes of subelectorates composed of n voters. We say that P and Q are* g-equivalent*, which we denote by $P \sim_g Q$, if for each $t \geq 0$ and each profile $T \in \mathcal{P}_A^t$, we have $g(P \cup T, k) = g(Q \cup T, k)$. Clearly, $\sim_g$ is an equivalence relation.*

Definition 2. *A function $\sigma : \mathcal{P}_A^n \times [m] \to \{0,1\}^*$ is a* compilation function for *g if there exists a function $\rho : \{0,1\}^* \times \mathcal{P}_A^* \times [m] \to A$ such that for all $P \in \mathcal{P}_A^n$, $t \geq 0$ and $T \in \mathcal{P}_A^t$, we have $\rho(\sigma(P), T, k) = g(P \cup T, k)$. By size$(\sigma)$ we denote the number of bits needed to represent $\sigma(P)$.*

Definition 3. *The* compilation complexity *of g, denoted $C(g)$, is the minimum value of size(σ) over all compilation functions for g.*

Proposition 1 of [4] for single-winner voting rules continues to hold for multi-winner rules as well, implying that $C(g)$ is the upper integer part of the logarithm (in base two) of the number of equivalence classes for $\sim_g$.

4 Results for Ordinal Rules

Let $k \in [m]$ be fixed. It is easy to derive a sufficient condition for all the committee scoring rules:

Lemma 1. *Let g_C^{sc} be a committee scoring rule with score function sc. If for all committees $S \in \mathcal{S}_k(A)$ we have $sc(S, P) = sc(S, Q)$, then $P \sim_{g_C^{sc}} Q$.*

Proof. Let $sc(S, P) = sc(S, Q)$ for all $S \in \mathcal{S}_k(A)$. For any profile T, $sc(S, P \cup T) = sc(P) + sc(T)$ and $sc(S, Q \cup T) = sc(Q) + sc(T)$. Hence $sc(S, P \cup T) = sc(S, Q \cup T)$, which shows that $P \sim_{g_c^{sc}} Q$. ■

A weaker sufficient condition holds for the decomposable committee scoring rules. For $P = (V_1, \ldots, V_n)$, let $sc(x, P) = \sum_i sc(x, V_i)$.

Lemma 2. *Let g_{DC}^{sc} be a decomposable committee scoring rule with score function sc. If for each candidate $x \in A$ we have $sc(x, P) = sc(x, Q)$, then $P \sim_{g_{DC}^{sc}} Q$.*

Proof. Decomposability implies $sc(S,V) = \sum_{x \in S} sc(x,V)$. Hence, $sc(S,P) = \sum_{x \in S} sc(x,P)$. Now, for all $a \in A$, let $sc(a,P) = sc(a,Q)$. This implies that $sc(S,P) = sc(S,Q)$, and $P \sim_{g^{sc}_{DC}} Q$ using Lemma 1, as a decomposable committee scoring rule is a committee scoring rule too. ■

The condition in Lemma 1 is also necessary for anonymous, neutral, and non-constant committee scoring rules (including SNTV, k-Borda, Bloc and Chamberlin-Courant). By non-constant, we mean that for a given vote V there exist $S, S' \in \mathcal{S}_k(A)$ such $sc(S',V) \neq sc(S,V)$.

Theorem 1. *Let g^{sc}_{AC} be an anonymous, neutral, and nonconstant committee scoring rule. $P \sim_{AC} Q$ holds if and only if for all $S, S' \in \mathcal{S}_k(A)$, $sc(S,P) = sc(S,Q)$.*

Proof. ($\Longleftarrow$) follows from Lemma 1. For ($\Longrightarrow$), the proof is as follows. First we show that if $P \sim_{AC} Q$, then for any S and S' in $\mathcal{S}_k(A)$, we have

$$sc(S,P) - sc(S,Q) = sc(S',P) - sc(S',Q) \quad (1)$$

Assume *Equation 1* fails for some S and S'. Let $sc(S,P) = \alpha$, $sc(S,Q) = \beta$, $sc(S',P) = \gamma$, and $sc(S',Q) = \delta$. Without loss of generality, $\alpha - \beta > \gamma - \delta$.

Let $S_1 = S \cap S'$, $S_2 = S \setminus S'$, and $S_3 = S' \setminus S$. Let $\mathcal{N}$ be an arbitrary bijective mapping from S_2 to S_3. We build a set a collection of votes T_1 obtained from P by swapping the positions of each candidate a in S_2 and $\mathcal{N}(a)$ in S_3. By neutrality, $sc(S,T_1) = \gamma$ and $sc(S',T_1) = \alpha$. Hence, $sc(S, P \cup T_1) = sc(S', P \cup T_1) = \alpha + \gamma$. However, $sc(S, Q \cup T_1) = \beta + \gamma < \alpha + \delta = sc(S', Q \cup T_1)$. In words, S' beats S in profile $Q \cup T_1$, but S' ties with S in profile $P \cup T_1$.

We define a voter set U_1 where the votes consist of a set of all permutations of S_2 followed by all permutations of S_1 followed by all permutations of S_3 followed by all permutations of the rest of the candidates (in cross multiplication, so that makes $|S_1|!|S_2|!|S_3|!(|A \setminus (S_1 \cup S_2)|!$ votes). We take u copies of each vote in U_1, for a large enough u, and call it T_2. Likewise, we define a voter set U_2 where the votes consist of a set of all permutations of S_3 followed by all permutations of S_1 followed by all permutations of S_2 followed by all permutations of the rest of the candidates. We take u copies of U_2 and call it T_3.

We consider a voter set W_1 where the votes consist of a set of all permutations of S_1 followed by all permutations of S_2 followed by all permutations of S_3 followed by all permutations of the rest of the candidates (in cross multiplication). We consider w copies of the above (for a large enough w) and call it as T_4. Likewise, we consider a voter set W_2 where the votes consist of a set of all permutations of S_1 followed by all permutations of S_3 followed by all permutations of S_2 followed by all permutations of the rest of the candidates (in cross multiplication). We consider w copies of W_2 and call it as T_5. Note that if S_1 is empty, then we do not need W_1 and W_2 because they are identical to U_1 and U_2 respectively.

For large enough u and w, adding votes $T_2 \cup T_3 \cup T_4 \cup T_5$ ensures that S and S' beat the rest of the committees in $\mathcal{S}_k(A)$, but increase their own scores by equal amounts (this uses non-constantness). Let $T = T_1 \cup T_2 \cup T_3 \cup T_4 \cup T_5$. For $P \cup T$, the winners are S and S', but for $Q \cup T$, the winner is solely S'. This implies (1).

Then we show that $\sum_{S \in \mathcal{S}_k(A)} sc(S, P)$ is a constant irrespective of the choice of $P \in \mathcal{P}_A^n$ (2). This must hold since for any vote V, whatever preference order they have, the committees cover all possible sets of size k. Hence, by neutrality, we see that $\sum_{S \in \mathcal{S}_k(A)} sc(S, V)$ is a constant, say ρ. $\sum_{S \in \mathcal{S}_k(A)} sc(S, P)$ is $n\rho$, using anonymity, and therefore a constant.

From *Equation 1*, if we had $sc(S, P) - sc(S, Q) = sc(S', P) - sc(S', Q) = c > 0$ (without loss of generality), then summing any side over $S \in \mathcal{S}_k(A)$ would contradict (2). Hence, we must have $sc(S, P) = sc(S, Q)$. ■

As an immediate corollary, we see that the condition in Lemma 2 is also necessary for any anonymous decomposable committee (ADC) scoring rule (which includes SNTV, k-Borda, and Bloc)

Corollary 1. *$P \sim_{g^{sc}_{ADC}} Q$ holds if and only if for every candidate $x \in A$ we have $sc(x, P) = sc(x, Q)$.*

Proof. ($\Longleftarrow$) follows from Lemma 1.

For ($\Longrightarrow$), take any two k-committees S_b and S_a such that $S_b \setminus S_a = \{b\}$ and $S_a \setminus S_b = \{a\}$. Applying Theorem 1 to S_a and S_b and then subtracting, we get $sc(S_a, P) - sc(S_b, P) = sc(S_a, Q) - sc(S_b, Q)$. Clearly, $sc(S_a, P) = \sum_{c \in S_a} sc(c, P)$, and likewise for the other terms. After simplification, we get $sc(a, P) - sc(b, P) = sc(a, Q) - sc(b, Q)$. Using neutrality and anonymity, we get ($\Longrightarrow$), as $\sum_{x \in A} sc(x, V)$ is a constant for any vote V, and hence $\sum_{x \in A} sc(x, P)$ is a constant. ■

4.1 Single Non Transferable Vote (SNTV)

As SNTV is an ADC scoring rule, Corollary 1 tells us that two profiles P and Q are SNTV-equivalent if and only if for each candidate x, the number of votes with x on top is the same in P and Q. For instance, consider a profile P with ranked ballots {(abc), (abc), (bca)} and a profile Q with ranked ballots {(abc), (bca), (bca)}. Let $k = 2$. $P \not\sim_{SNTV} Q$, as $sc(a, P) = 2$ but $sc(a, Q) = 1$. If we add $T = \{(cab), (cab)\}$ to P and Q, the winning committee will be (a, c) for $P \cup T$, but it will be (b, c) for $Q \cup T$.

Corollary 2. $C(SNTV) = \Theta\left(m \log\left(1 + \frac{n}{m}\right) + n \log\left(1 + \frac{m}{n}\right)\right)$

The above result follows from the bounds of the plurality rule in Corollary 1 of [4] because we have the same equivalence classes for plurality and SNTV.

4.2 *k*-Borda

As k-Borda is an ADC scoring rule, Corollary 1 gives us its equivalence classes. For instance, consider a profile P with ranked ballots $\{(abc),(abc),(bac)\}$ and a profile Q with ranked ballots $\{(abc),(bca),(bac)\}$. Let $k=2$. $P \not\sim_{k-Borda} Q$, as $sc(a,P)=5$ but $sc(a,Q)=3$. If we add $T=\{(cba),(bca)\}$ to P and Q, the winning committee will be (a,b) for $P \cup T$, but it will be (b,c) for $Q \cup T$.

Corollary 3. $C(k - Borda) = \Theta(m \log nm)$

The above result follows from the bounds of Borda rule in Corollary 2 of [4] as we have the same equivalence classes for Borda and k-Borda rules.

4.3 Bloc

As Bloc is an ADC scoring rule, Corollary 1 gives us a characterisation of its equivalence classes. As an example, consider a profile P with ranked ballots $\{(abc),(abc),(abc)\}$ and a profile Q with ranked ballots $\{(abc),(abc),(acb)\}$. Let $k=2$. $P \not\sim_{Bloc} Q$, as $sc(b,P)=3$ but $sc(b,Q)=2$. If we add $T=\{(cba),(bca)\}$ to P and Q, the winning committee will be (a,b) for $P \cup T$, but it will be (b,c) for $Q \cup T$.

Corollary 4. $C(Bloc) = \Theta\left(m \log\left(1+\frac{n\hat{k}}{m}\right) + n\hat{k}\log\left(1+\frac{m}{n\hat{k}}\right)\right)$ *where* $\hat{k} = \min(k, m-k)$

The equivalence classes of Bloc rule are identical to those of l-Approval and hence the bounds are obtained from Theorem 1 of [19].

4.4 Chamberlin-Courant (*β*-CC)

As Chamberlin-Courant is an AC scoring rule, Theorem 1 characterises its equivalence classes. For instance, consider a profile P with ranked ballots $\{(abc),(abc),(bca)\}$ and a profile Q with ranked ballots $\{(abc),(bca),(bca)\}$. Let $k=2$. $P \not\sim_{\beta-CC} Q$, as $sc(bc,P)=4$ but $sc(bc,Q)=5$. If we add $T=\{(cab),(cab)\}$ to P and Q, the winning committee will be (a,c) for $P \cup T$, but it will be (b,c) for $Q \cup T$.

We obtain the following bounds by using some counting arguments.

Corollary 5. $C(\beta - CC) = O\left(\binom{m}{k}\log(n(m-k))\right)$

Proof. There are $\binom{m}{k}$ k-committees and the score received by each can be in the range $[n(k-1), n(m-1)]$, so there are at most $n(m-k)+1$ scores possible for each. This yields an upper bound of $(n(m-k)+1)^{\binom{m}{k}}$ on the number of equivalence classes. Hence, we get the upper bound on the compilation complexity. ■

Corollary 6. $C(\beta - CC) = \Omega(\log(n(m-k)))$

Proof. A single k-committee can take each of the possible scores in the range $[n(k-1), n(m-1)]$ and then there would possibly be constraints on the scores of some other committees based on these scores. This gives a lower bound of $n(m-k)+1$ on the number of equivalence classes. Hence, we get the lower bound on the compilation complexity. ■

Unlike the results we obtained for the decomposable committee scoring rules, our lower bound and upper bound results for β-CC are loose. This is because it is difficult to count the number of functions from $\mathcal{S}_k(A)$ to $\mathbb{N}$ corresponding to $sc(., P)$ for some profile P for each of these rules, because of the dependencies between the scores of intersecting committees. As an example, take $k = 2$. If $sc_B(., i)$ denotes the Borda score of a candidate for vote V_i, then $sc(bc, \{V_i\}) = \max(s_B(b,i), s_B(c,i)) \leq \max(sc(ab, \{V_i\}), sc(ac, \{V_i\}))$. Summing this inequality over all votes, we obtain that $sc(bc, P) \leq sc(ab, P) + sc(ac, P)$. More generally, if $X \subseteq X_1 \cup \ldots \cup X_q$ then $sc(X, P) \leq \sum_{j=1}^{q} sc(X_j, P)$.

4.5 Gehrlein-Stable Rules (GehrSta)

For any two candidates c and d, let $N_P(c,d)$ be the number of voters in P that prefer c to d. Let $\mathcal{M}_P$ represent the tournament over candidates in A with voters P. We get the following equivalence result.

Theorem 2. *$P \sim_{GehrSta} Q$ holds if and only if $\mathcal{M}_P = \mathcal{M}_Q$, where GehrSta is a Gehrlein Stable rule.*

Proof. ($\Longleftarrow$) is straightforward. For ($\Longrightarrow$), the proof is as follows.

Consider for contradiction that $P \sim_{GehrSta} Q$, but $\mathcal{M}_P \neq \mathcal{M}_Q$. Let there be $(a,b) \in A$ such that $N_P(a,b) \neq N_Q(a,b)$. Let $N_P(a,b) = N_Q(a,b) + l$ assuming without loss of generality that $l > 0$. Let T be a set of $m+1$ voters and choose any fixed set C of $k-1$ candidates other than a or b whom these voters place in the top $k-1$ positions of their preferences. In $m+1-N_P(a,b)$ of these votes let a be in the k^{th} place and b be in the $(k+1)^{th}$ place. In the rest of the $N_P(a,b)$ votes let b be in the k^{th} place and a be in the $(k+1)^{th}$ place. Similar to the idea mentioned in Lemma 4 of [4], it can be seen that for $P \cup T$, $C \cup \{a\}$ will form the winning committee, as it is a Condorcet set. But, for $Q \cup T$, $C \cup \{b\}$ will form the winning committee. Hence $P \not\sim_{GehrSta} Q$ here, which is a contradiction. ■

For instance, consider the rule NED, a profile P with ranked ballots $\{(abc), (abc), (abc)\}$ and a profile Q with ranked ballots $\{(bac), (bac), (bca)\}$. Let $k = 2$. $P \not\sim_{GehrSta} Q$, as $\mathcal{M}_P \neq \mathcal{M}_Q$ (a's out-degree is different in $\mathcal{M}_P$ and $\mathcal{M}_Q$). If we add $T = \{(cba), (cba)\}$ to P and Q, the winning committee will be (a, b) for $P \cup T$, but it will be (b, c) for $Q \cup T$.

Corollary 7. *$C(GehrSta) = O(m^2 \log n)$, and there exists a constant $q > 0$ such that $C(GehrSta) = \Omega\left(m^2 \log\left(\left\lfloor \frac{n}{qm} \right\rfloor - 2\right)\right)$*

The result follows from *Proposition 8* of [4] and *Proposition 5* of [19], as the equivalence classes are identical to those of a Condorcet Consistent WMG Rule.

4.6 θ-winning sets (θ)

We know that the winning committee W is the θ-winning set for the largest value of θ (after tie-breaking). Let $N_P(d, S)$ be the number of voters in P that prefer d to any candidate in S. Let the score of S for P be $sc(S, P) = n - \max_{d \in A \setminus S}(N_P(d, S))$. Then, the maximum θ for S is $\frac{sc(S,P)}{n}$.

Theorem 3. *$P \sim_\theta Q$ holds if and only if for all $d \in A \setminus S$ and for all $S \in \mathcal{S}_k(A)$, we have $N_P(d, S) = N_Q(d, S)$*

Proof. ($\Longleftarrow$) is straightforward from how the rule is defined. For ($\Longrightarrow$), the proof can be obtained as follows.

Let S_x^P be the x^{th} highest scoring k-committee for P according to sc. Choose the largest i such that $N_P(d, S_j^P) = N_Q(d, S_j^P)$ for all $d \in A \setminus S_j^P$ and for all $j \in [i-1]$. Initially, we consider $i \neq 1$ and denote the winner as W.

First consider the case where $sc(S_i^P, P) \neq sc(S_i^P, Q)$. Assume without loss of generality that $sc(S_i^P, P) > sc(S_i^P, Q)$. Here, we add voters T_1, each of whom contains every permutation of S_i one by one in the first k places until S_i^P becomes the winner for $P \cup T_1$ but not for $Q \cup T_1$. Hence it is not equivalent.

Next consider the case where $sc(S_i^P, P) = sc(S_i^P, Q)$ but there exists a $c \in A \setminus S_i^P$ such that $N_P(c, S_i^P) \neq N_Q(c, S_i)$. Assume without loss of generality that $N_P(c, S_i^P) < N_Q(c, S_i^P)$. Consider profile T_2 consisting of $n - sc(S_i^P, Q) + 1 - N_Q(c, S_i^P)$ voters where each voter votes in the order c followed by S_i in some order followed by the rest of the candidates in some order. This will increase S_i^P's score by $n - sc(S_i^P, Q) + 1 - N_Q(c, S_i^P)$ and $n - sc(S_i^P, Q) - N_Q(c, S_i^P)$ for $P \cup T_2$ and $Q \cup T_2$ respectively. Now since $sc(S_i^P, P) > s_Q(S_i^P, Q)$, T_1 can be obtained similarly as above to prove lack of equivalence.

Now consider the case where $i = 1$. Here, if $sc(W, P) \neq sc(W, Q)$, assume without loss of generality that $sc(W, P) < sc(W, Q)$. Take the highest scoring committee S_y^Q for Q for which $sc(S_y^Q, P) \geq sc(S_y^Q, Q)$ and add a set of voters T_3 each of whom keeps a permutation of S_y one by one at the top of their list. Continue this until S_y^Q wins for $P \cup T_3$ but not for $Q \cup T_3$.

If $sc(W, P) = sc(W, Q)$, that implies that there exists a c in $A \setminus W$ such that $N_P(c, W) \neq N_Q(c, W)$. Adding a T_4 to both P and Q which is similar to T_2, we reach a similar position as that required for adding T_3, which shows a lack of equivalence. Hence, ($\Longrightarrow$) must be true. ∎

For instance, consider a profile P with ranked ballots $\{(abc), (abc), (bca)\}$ and a profile Q with ranked ballots $\{(abc), (abc), (cab)\}$. Let $k = 2$. $P \not\sim_\theta Q$, as

$N_P(c, ab) = 0$ but $N_Q(c, ab) = 1$. If we add $T = \{(cab)\}$ to P and Q, the winning committee will be (a, b) for $P \cup T$, but it will be (a, b) or (a, c) for $Q \cup T$ depending on the tie-breaking rule.

Corollary 8. $C(\theta) = O\left(\binom{m}{k}(m-k)\log n\right)$

Proof. $N_P(d, S)$ must be in the range $[0, n]$ for each $d \in A \backslash S$ and for each $S \in \mathcal{S}_k(A)$. There are $\binom{m}{k}(m-k)$ such terms (where $\binom{m}{k}$ arises from choosing S and $m-k$ represents the choices of d), giving an upper bound of $(n+1)^{\binom{m}{k}(m-k)}$ on the number of equivalence classes. Hence $C(\theta) = O\left(\binom{m}{k}(m-k)\log n\right)$. ■

Corollary 9. $C(\theta) = \Omega((m-k)\log n)$

Proof. For any one particular k-committee say S_i^P, for all $d \in A \backslash S_i^P$, each $N_P(d, S_i^P)$ can take any value from $[0, n]$. However, this range is restricted for committees other than S_i so it is possible that they cannot take some values in $[0, n]$. Hence the lower bound for the number of equivalence classes is obtained by assigning any of $[0, n]$ independently to S_i^P, giving a lower bound of $(n+1)^{(m-k)}$. Hence $C(\theta) = \Omega((m-k)\log n)$. ■

Again we get a gap between the lower bound and upper bound results, for the same reason as in *Subsection 4.4*.

4.7 Sequential Plurality (SeqPl)

For each vote V, let V^j be the top-j truncation of V, where $j \in [k]$. For instance, if $V = (abcd)$ then $V^2 = (ab)$. Given a profile P, for each ordered sequence of j candidates λ^j, let $N(P, \lambda^j)$ be the number of votes V in P such that $V^j = \lambda^j$. For instance, if $P = \{(abcd), (abdc), (acbd), (dabc), (dabc)\}$ and $k = 2$, $N(P, ab) = N(P, da) = 2$, $N(P, ac) = 1$, and $N(P, \lambda^k) = 0$ for $\lambda^k \neq ab, ac, da$.

Theorem 4. *$P \sim_{SeqPl} Q$ if and only if $N(P, \lambda^k) = N(Q, \lambda^k)$ for each ordered sequence of k candidates λ.*

Proof. ($\Longleftarrow$) is straightforward as the voters are anonymous and to select the j^{th} winner for each $j \in [k]$, we need no more information than that in the first j places of each voter's ordering.

For ($\Longrightarrow$), the proof is as follows. It is a necessary condition for $k = 1$, as in that case, it reduces to the Plurality rule for which it is a necessary condition. From now on, assume that $k > 1$.

Suppose it was not a necessary condition for equivalence. There would be a $j \in [k]$ such that $N(P, \lambda^j) \neq N(Q, \lambda^j)$ for some j-sized ordered sequence λ^j. We will show that by taking the smallest such j satisfying the above condition, we can add voters T to both P and Q such that their winning committees will have different candidates at the j^{th} place and the same candidates in all other places.

The first $j-1$ winners (call them C_1) are the same for P and Q as $N(P, \lambda^i) = N(Q, \lambda^i)$, for each ordered sequence of i candidates for all $i \in [j-1]$. First, consider a group of voters V_a^P and V_a^Q for P and Q whose first $j-1$ positions contain the first $j-1$ winners in some fixed order (which could be different from the final ranking) and their j^{th} member is some c which is not the j^{th} member of the winning committee, which is d. Consider the first case where we assume without loss of generality that $|V_a^P| > |V_a^Q|$. Here we add voters T_1 who have their preference order until $j-1$ same as in V_a^P and with c in the j^{th} position and in the rest of the positions the $k-j$ candidates, call them C_2 (apart from c and d and the first $j-1$ candidates) that we want as the last $k-j$ winners, in the order needed. Continue this until c is the j^{th} winner for $P \cup T_1$ but it is still d for $Q \cup T_1$.

After this, if the last $k-j$ winners are not C_2 for $P \cup T_1$ and $Q \cup T_1$ both, add the profile T_2 which is constructed as follows. The first $j-1$ candidates for each added voter are C_1. Put the next $k-j+1$ preferences as follows: 2 voters having c followed by C_2, 2 voters having d followed by C_2 and 1 voter each having C_2 followed by c and C_2 followed by d respectively. Add the above collection of 6 voters one by one repeatedly until the last $k-j$ winners are C_2 for both $P \cup T_1 \cup T_2$ and $Q \cup T_1 \cup T_2$. Hence the winning committee is $C_1 \cup \{c\} \cup C_2$ for $P \cup T_1 \cup T_2$ and $C_1 \cup \{d\} \cup C_2$ for $Q \cup T_1 \cup T_2$.

Note that if c was the j^{th} winner for P, consider voters V_b^P and V_b^Q with the same $j-1$ candidates at the top as above and a d in the j^{th} position such that $|V_b^P| < |V_b^Q|$ and such a d must exist given the assumption on V_a above. We can make $C_1 \cup \{c\} \cup C_2$ as the winning committee for $P \cup T_1 \cup T_2$ and $C_1 \cup \{d\} \cup C_2$ as the winning committee for $Q \cup T_1 \cup T_2$, by choosing profiles similar to T_1 and T_2 using the above method.

Now, for the second case where there is a difference in $|V_c^P|$ and $|V_c^Q|$, where V_c^P and V_c^Q are the sets of voters whose first $j-1$ candidates C_1' are not the first $j-1$ members of the winning committee (in any order). We can make C_1' the first $j-1$ winners by adding a profile T_3 consisting of these C_1' in the first $j-1$ positions and all possible permutations for the following $k-j+1$ places. Add the above multiple times if needed until C_1' become the first $j-1$ winners for both $P \cup T_3$ and $Q \cup T_3$. Now it again reduces to one of the cases above profiles similar to T_1 and T_2 can be added to show the lack of equivalence.

Hence ($\Longrightarrow$) must be true. ■

For instance, consider a profile P with ranked ballots $\{(abc), (abc), (abc)\}$ and a profile Q with ranked ballots $\{(abc), (abc), (cba)\}$. Let $k = 2$. $P \not\sim_{SeqPl} Q$, as $N(P, ab) = 3$ but $N(Q, ab) = 2$. If we add $T = \{(cba), (cba)\}$ to P and Q, the winning committee will be (a, b) for $P \cup T$, but it will be (b, c) for $Q \cup T$.

Corollary 10. $C(SeqPl) = \Theta\left(\frac{m!}{(m-k)!}\log\left(1 + \frac{nm!}{(m-k)!}\right) + n\log\left(1 + \frac{m!}{n(m-k)!}\right)\right)$

Proof. For each of the $\frac{m!}{(m-k)!}$ possible k-orderings we can have a certain fixed number $[0, n]$ of voters. There are n voters in total, which yields the result. ■

5 Results for Approval-based Rules

We begin with a general result for the equivalence classes of Thiele's Optimization Method for voting rules for which score from a vote V, $sc(S, V) = h(|S \cap V|) \leq |S \cap V|$. We refer to this class of rules as $ThiOpt$. Let $sc(S, P)$ below be the $ThiOpt$ score received by S from P.

Theorem 5. *$P \sim_{ThiOpt} Q$ holds if and only if for all $S, S' \in \mathcal{S}_k(A)$, $sc(S', P) - sc(S, P) = sc(S', Q) - sc(S, Q)$.*

Proof. ($\Longleftarrow$) is straightforward because if the relative differences between scores of all pairs of committees are equal within P and Q, they will remain equal after adding new voters, hence the winning committee W will remain the same for both. For ($\Longrightarrow$), the idea of the proof is as follows.

First, we show that ($\Longrightarrow$) holds for Thiele rules with $h(r) < r$ for all $r > 1$ Assume that the right-hand side of the equation in the theorem is not satisfied for P and Q. Then, there exist $S_x, S_y \in \mathcal{S}_k(A)$ such that $sc(S_x, P) - sc(S_y, P) > sc(S_x, Q) - sc(S_y, Q)$ without losing generality. First, we add a set of voters T_1 with approval ballots such that $sc(S_x, Q \cup T_1) = sc(S_y, Q \cup T_1)$ (Note that this can be done by interchanging the candidates in S_x and S_y from the votes in Q, just like in the proof of *Theorem 1*). So, $sc(S_x, Q \cup T_1) - sc(S_y, Q \cup T_1) = 0$, whereas $sc(S_x, P \cup T_1) - sc(S_y, P \cup T_1) > 0$.

Then, add a profile T_2 consisting of a large number (t_2) of voters, each of them voting for every 'single' candidate in $S_x \cap S_y$ in a round-robin manner. Note that if the intersection is empty, then $T_2 = \phi$, and we skip this step. After this, S_x and S_y beat all committees for $P \cup T_1 \cup T_2$ and $Q \cup T_1 \cup T_2$ except possibly those which contain all candidates of $S_x \cap S_y$.

Consider a third set of voters T_3 with t_3 voters who vote for all pairs (c_1, c_2) for each $c_1 \in S_x \setminus S_y$ and for each $c_2 \in S_y \setminus S_x$ in a round-robin manner. Also, note that we must have $t_2 \gg t_3 \gg n$. After adding these voters, it can be shown that S_x and S_y also beat those committees that contain $S_x \cap S_y$ for both $P \cup T_1 \cup T_2 \cup T_3$ and $Q \cup T_1 \cup T_2 \cup T_3$, using the fact that $h(2) < 2$. As an example, if the committees S_x and S_y are (ab) and (cd), then in T_3 we add copies of votes for $(ac), (ad), (bc), (bd)$. This produces an increase in score of $2h(1) + h(2)$ to each one of $(ac), (ad), (bc), (bd)$. But (ab) and (cd) each get $4h(1)$. Since $h(2) < 2 = 2h(1)$, S_x and S_y will eventually beat the other committees.

After adding the profile $T = T_1 \cup T_2 \cup T_3$, we see that S_x continues to be the sole winner for $P \cup T$, whereas there is a tie between S_x and S_y for $Q \cup T$, demonstrating a lack of equivalence. Hence we must have $sc(S', P) - sc(S, P) = sc(S', Q) - sc(S, Q)$ for all $S, S' \in \mathcal{S}_k(A)$, implying that ($\Longrightarrow$) is true if $h(r) < r$ for all $r > 1$.

Next, for ($\implies$), we consider the case where $h(r) = r$ for every $r \in [k]$. This corresponds to multiwinner approval voting. This needs an alternate analysis, where we consider an arbitrary tie-breaking rule over the committees (this does not conflict with the rule being irresoluteness, because tie-breaking is used to simplify the proof). There are $\kappa = \binom{m}{k}$ committees of size k. For a given preference profile P, arrange them in the decreasing order of the scores they receive from P from S_1 to S_κ. Given that $P \sim_{ThiOpt} Q$ for some Q, the winning committee W must be identical for P and Q.

Now, assume without loss of generality that for all $i \in [x-1]$, $sc(W,P) - sc(S_i,P) = sc(W,Q) - sc(S_i,Q)$, and $sc(W,P) - sc(S_x,P) < sc(W,Q) - sc(S_x,Q)$. From now on, we refer to S_x as S'. Let S'' be the next higher-ranked committee for P (here, S_{x-1}). Choose an arbitrary candidate $a \in S' \setminus S''$ and add just enough copies of votes (call the profile T_1) approving only a to both P and Q so that S' scores 1 higher than S'' when considering voter $P \cup T_1$, but S' scores equal to or less than S'' when considering votes $(Q \cup T_1)$. Iterate this process till we get $S'' = W$, and let the final added vote profile be T_τ for some $\tau \geq 1$. Let $T = \bigcup_{j=1}^{\tau} T_j$ Now, S' is the sole winner for the profile $P \cup T$, but at best the shared winner for $Q \cup T$, hence demonstrating lack of equivalence. ∎

5.1 Multiwinner Approval Voting (MAV)

Let $sc(a,P)$ be the approval score of $a \in A$ received from P, and let $sc(S,P)$ be the approval score of $S \in \mathcal{S}_k(A)$ from P. Then, the following result follows directly from Theorem 5 by following the idea from the proof of Corollary 1.

Corollary 11. *$P \sim_{MAV} Q$ holds if and only if for all $a, b \in A$, $sc(a,P) - sc(b,P) = sc(a,Q) - sc(b,Q)$*

Take the following example. Let there be a profile P with approval ballots $\{(a),(a),(b)\}$ and a profile Q with approval ballots $\{(a),(b),(b)\}$. Let $k = 2$. $P \not\sim_{AV} Q$, as $sc(a,P) = 2$ but $sc(a,Q) = 1$. If we add $T = \{(c),(c)\}$ to P and Q, the winning committee will be (a,c) for $P \cup T$ but it will be (b,c) for $Q \cup T$.

Corollary 12. $C(MAV) = \Theta(m \log n)$

Proof. From Corollary 11, two sets of approval scores are equivalent if and only if all their pairwise differences are equal. Defining the effective AV score of a candidate a as $sc_e(a,P) = sc(a,P) - \min_{i \in [m]}(sc(A_i,P))$, it is clear that $P \sim_{MAV} Q$ if and only if they produce same sc_e set. The sc_e of the minimum scoring candidate must be 0, hence we remove the scores which do not contain any $0's$, from the total possible MAV scores. Hence the total effective MAV score sets possible are $(n+1)^m - n^m$. Hence $C(MAV) = \Theta(m \log(n))$. ∎

5.2 Approval-based Chamberlin-Courant rule (α-CC)

Theorem 5 immediately characterises the equivalence classes for Approval-based Chamberlin-Courant. For example, consider a profile P with approval ballots $\{(a),(ab)\}$ and a profile Q with approval ballots $\{(b),(ab)\}$. Let $k=2$. $P \not\sim_{AV} Q$, as $sc(bc,P)=1$ but $sc(bc,Q)=2$. If we add $T=\{(c)\}$ to P and Q, the winning committee will be (a,c) for $P \cup T$, but it will be (b,c) for $Q \cup T$.

Corollary 13. $C(\alpha - CC) = O\left(\binom{m}{k}\log n\right)$

Proof. There are $\binom{m}{k}$ k-committees. Let the effective α-CC score be $sc_e(S,P) = sc(S,P) - \min_{S_i \in \mathcal{S}_k(A)} sc(S_i,P)$. The number of possible values of $sc_e(S,P)$ is upper bounded by $(n+1)^{\binom{m}{k}} - n^{\binom{m}{k}} = O\left(n^{\binom{m}{k}}\right)$ (after eliminating the set with no effective 0 scores). Hence proved that $C(\alpha - CC) = O\left(\binom{m}{k}\log n\right)$. ■

Corollary 14. $C(\alpha - CC) = \Omega\left(\left\lfloor \frac{m}{k} \right\rfloor \log n\right)$

Proof. There are $\left\lfloor \frac{m}{k} \right\rfloor$ possible committees with no candidates in common which can be assigned scores independently from $[0,n]$. The number of possible values of $sc_e(S,P)$ is lower bounded by $(n+1)^{\left\lfloor \frac{m}{k} \right\rfloor} - n^{\left\lfloor \frac{m}{k} \right\rfloor} = \Omega\left(n^{\left\lfloor \frac{m}{k} \right\rfloor}\right)$. Hence proved. ■

The lower bound and upper bound do not match, for the same reason as for β-CC discussed in *Subsection 4.4*.

5.3 Proportional Approval Voting (PAV)

We get the equivalence classes for PAV directly from Theorem 5. Consider the following example. There is a profile P with approval ballots $\{(a),(ab)\}$ and a profile Q with approval ballots $\{(b),(ab)\}$. Let $k=2$. $P \not\sim_{PAV} Q$, as $sc(ac,P) = 2$ but $N(ac,Q)=1$. If we add $T=\{(c)\}$ to P and Q, the winning committee will be (a,c) for $P \cup T$ but it will be (b,c) for $Q \cup T$.

Corollary 15. $C(PAV) = O\left(\binom{m}{k}(k \log k + \log(n \log n))\right)$

Proof. There are $\binom{m}{k}$ k-committees. For a given committee $S_x \in \mathcal{S}_k(A)$, its score can be represented as $\sum_{i=1}^{k} \frac{z_i}{i}$, where z_1 is an integer which is formed by adding the scores from the first approval, pairs of second approvals, triplets of third approvals and the rest of the fractions that can be combined to 1. This can take at most $O(n\log(n))$ values (because the harmonic series converges to $n\log(n)$). $z_2,\ldots,z_k$ (where each z_i represents the number of parts of $\frac{1}{i}$ remaining after combining the rest of the terms to integers) can take $1,\ldots,(k-1)$ values respectively. So the number of equivalence classes is $O\left(((n\log n)(k-1)!)^{\binom{m}{k}}\right)$. The bound then follows. ■

Corollary 16. $C(PAV) = \Omega\left(\lfloor\frac{m}{k}\rfloor(\log(n\log n))\right)$

Proof. There are $\left\lfloor\frac{m}{k}\right\rfloor$ possible committees with no candidates in common which can be assigned at least $\Omega(n\log n)$ scores each. This gives us $\Omega\left((n\log n)^{\lfloor\frac{m}{k}\rfloor}\right)$ equivalence classes. Taking the logarithm produces the result. ■

Note that our upper bound and lower bound results do not match for the same reason as for β-CC discussed in *Subsection 4.4*

6 Conclusion and Future Work

We characterise the equivalence classes and compilation complexity bounds of many common ordinal and approval-based multi-winner rules. This helps us understand how they fare against each other regarding the amount of memory needed to store the partial results. Although some results are inspired by ideas used in [4,19] for single-winner voting rules, several of our results, such as those for θ-Winning sets, Sequential Plurality and Thiele's Optimization Method are quite novel. Recall that the reason why the lower and upper bounds for α-CC, β-CC, θ-Winning Sets, and PAV do not match is that it is not easy to count the number of functions from $\mathcal{S}_k(A)$ to $\mathbb{N}$ that correspond to $sc(.,P)$ for some profile P for each of these rules. These gaps could be reduced in future. Another open problem is the extension of our results to the setting where the number of additional voters is already known. Finally, equivalence classes and compilation complexity bounds for Single Transferrable Vote, Phragmén, and the Method of Equal Shares, are still open. We present below a summary of our results.

Acknowledgements

We are grateful to the anonymous reviewers for their careful reading and helpful comments. This work was supported in part by project ANR-22-CE26-0019 “Citizens”, funded by the French Agence Nationale de la Recherche.

References

1. Aziz, H., Elkind, E., Faliszewski, P., Lackner, M., Skowron, P.: The Condorcet principle for multiwinner elections: From shortlisting to proportionality. In: Proceedings of IJCAI 2017. pp. 84–90 (2017)

† $\hat{k} = \min(k, m-k)$

‡ for some $q > 0$

Table 1. A summary of results of compilation complexity for multi-winner voting rules.

Rule	Equiv. Condition	Upper Bound	Lower Bound
SNTV	$sc(x, P) = sc(x, Q)$	$\Theta\left(m \lg\left(1 + \frac{n}{m}\right) + n \lg\left(1 + \frac{m}{n}\right)\right)$	
k-Borda		$\Theta(m \lg(nm))$	
Bloc		$\Theta\left(m \lg\left(1 + \frac{n\hat{k}^{\dagger}}{m}\right) + n\hat{k} \lg\left(1 + \frac{m}{n\hat{k}}\right)\right)$	
β-CC	$sc(S, P) = sc(S, Q)$	$O\left(\binom{m}{k} \lg(n(m-k))\right)$	$\Omega(\lg(n(m-k))$
GehrSta	$\mathcal{M}_P = \mathcal{M}_Q$	$O(m^2 \lg n)$	$\Omega\left(m^2 \lg\left(\left\lfloor \frac{n}{q^{\ddagger} m} \right\rfloor - 2\right)\right)$
θ	$N_P(d, S) = N_Q(d, S)$	$O\left(\binom{m}{k}(m-k) \lg n\right)$	$\Omega((m-k) \lg n)$
SeqPl	$N(P, \lambda^k) = N(Q, \lambda^k)$	$\Theta\left(\frac{m!}{(m-k)!} \lg\left(1 + \frac{nm!}{(m-k)!}\right) + n \lg\left(1 + \frac{m!}{n(m-k)!}\right)\right)$	
MAV	$sc(a, P) - sc(b, P)$ $= sc(a, Q) - sc(a, Q)$	$\Theta(m \lg n)$	
α-CC	$sc(W, P) - sc(S, P)$	$O\left(\binom{m}{k} \lg n\right)$	$\Omega\left(\left\lfloor \frac{m}{k} \right\rfloor \lg n\right)$
PAV	$= sc(W, Q) - sc(S, Q)$	$O\left(\binom{m}{k}(k \lg k + \log(n \lg n)\right)$	$\Omega\left(\left\lfloor \frac{m}{k} \right\rfloor (\lg(n \lg n))\right)$

2. Barberà, S., Coelho, D.: How to choose a non-controversial list with k names. Soc. Choice Welf. **31**(1), 79–96 (2008)
3. Chevaleyre, Y., Lang, J., Maudet, N., Monnot, J.: Compilation and communication protocols for voting rules with a dynamic set of candidates. In: Proceedings of TARK 2011. pp. 153–160 (2011)
4. Chevaleyre, Y., Lang, J., Maudet, N., Ravilly-Abadie, G.: Compiling the votes of a subelectorate. In: Proceedings of IJCAI 2009. pp. 97–102 (2009)
5. Coelho, D.: Understanding, Evaluating and Selecting Voting Rules Through Games and Axioms. Ph.D. thesis (2004)
6. Conitzer, V., Sandholm, T.: Vote elicitation: Complexity and strategy-proofness. In: Proceedings of AAAI-2001. pp. 392–397 (2002)
7. Dey, P., Bhattacharyya, A.: Sample complexity for winner prediction in elections. In: Proceedings of AAMAS 2015. pp. 1421–1430 (2015)
8. Dey, P., Talmon, N., Van Handel, O.: Proportional representation in vote streams. arXiv preprint arXiv:1702.08862 (2017)
9. Elkind, E., Faliszewski, P., Skowron, P., Slinko, A.: Properties of multiwinner voting rules. Soc. Choice Welf. **48**(3), 599–632 (2017)
10. Elkind, E., Lang, J., Saffidine, A.: Condorcet winning sets. Soc. Choice Welf. **44**(3), 493–517 (2015)
11. Faliszewski, P., Skowron, P., Slinko, A., Talmon, N.: Multiwinner voting: A new challenge for social choice theory. Trends in Computational Social Choice **74**, 27–47 (2017)
12. Gehrlein, W.: The condorcet criterion and committee selection. Mathematical Social Sciences **10**(3), 199–209 (1985)
13. Hazon, N., Aumann, Y., Kraus, S., Wooldridge, M.: On the evaluation of election outcomes under uncertainty. Artificial Intelligence **189**, 1–18 (2012)
14. Janson, S.: Phragmén's and thiele's election methods. Tech. rep., Technical report (2016)
15. Konczak, K., Lang, J.: Voting procedures with incomplete preferences. In: Proc. IJCAI-05 Multidisciplinary Workshop on Advances in Preference Handling. vol. 20. Citeseer (2005)

16. Lackner, M., Skowron, P.: Multi-Winner Voting with Approval Preferences - Artificial Intelligence, Multiagent Systems, and Cognitive Robotics. Springer Briefs in Intelligent Systems, Springer (2023)
17. Peters, D., Pierczyński, G., Skowron, P.: Proportional participatory budgeting with additive utilities. In: Proceedings of NeurIPS21. pp. 12726–12737 (2021)
18. Skowron, P., Faliszewski, P., Slinko, A.: Axiomatic characterization of committee scoring rules. J. Econ. Theory **180**, 244–273 (2019)
19. Xia, L., Conitzer, V.: Compilation complexity of common voting rules. In: Proceedings of AAAI 2010. pp. 915–920 (2010)