

On the Complexity of Winner Verification and Candidate Winner for Multiwinner Voting Rules

Chinmay Sonar¹, Palash Dey² and Neeldhara Misra³

¹University of California, Santa Barbara

²Indian Institute of Technology Kharagpur

³Indian Institute of Technology Gandhinagar

csonar@cs.ucsb.edu, palash.dey@cse.iitkgp.ac.in, neeldhara.m@iitgn.ac.in,

Abstract

The Chamberlin-Courant and Monroe rules are fundamental and well-studied rules in the literature of multi-winner elections. The problem of determining if there exists a committee of size k that has a Chamberlin-Courant (respectively, Monroe) dissatisfaction score of at most r is known to be NP-complete. We consider the following natural problems in this setting: a) given a committee S of size k as input, is it an optimal k -sized committee, and b) given a candidate c and a committee size k , does there exist an optimal k -sized committee that contains c ? In this work, we resolve the complexity of both problems for the Chamberlin-Courant and Monroe voting rules in the settings of rankings as well as approval ballots. We show that verifying if a given committee is optimal is coNP-complete whilst the latter problem is complete for Θ_2^P . Our contribution fills an essential gap in the literature for these important multi-winner rules.

1 Introduction

We study preference aggregation in the multiwinner setting. Here, we have a set of voters who express preferences over a collection of candidates, and we are faced with the task of shortlisting a small number of candidates in a manner that is as satisfactory as possible for all the agents involved. This abstraction captures several application scenarios such as choosing a governing body of any institution, deciding which advertisements to show on TV during some program, recommending movies [Lu and Boutilier, 2011], selecting group of products for promotion [Skowron *et al.*, 2016], shortlisting candidates for a limited fellowship, etc.

A fundamental property that one often wants a multiwinner rule to satisfy is *proportional representation*. Intuitively speaking, proportional representation means that the fraction of seats that a party receives in the winning committee should be proportional to the fraction of votes it receives. Indeed, two of the most popular multiwinner rules, the Chamberlin-Courant rule [Chamberlin and Courant, 1983] and the Monroe rule [Monroe, 1995], are designed to achieve proportional

representation as best as one can hope for. In the Chamberlin-Courant (abbreviated CC) rule, we seek winners that can be hypothetically “assigned to” voters in such a way that every voter is reasonably satisfied with their alternative. In the Monroe rule, we have the additional requirement that each winning candidate takes on the responsibility of representing roughly the same number of voters. Depending on how one formalizes misrepresentation and how the preferences of the voters are modeled (popularly one of approval ballots or rankings), natural variants of these rules are used in practice.

Bartholdi *et al.* [1989] showed that determining winners for many, otherwise excellent, voting rules are NP-hard. Prominent examples of such single winner ($k=1$) rules include Kemeny’s voting rule [Kemeny, 1959], and Lewis Carroll’s rule (Dodgson Rule, [Dodgson, 1876]). Moreover, some of these single winner rules seem to be substantially harder than any NP-complete problem — they are complete for the complexity class P_{\parallel}^{NP} [Hemaspaandra *et al.*, 2005]. Papadimitriou and Zachos [1982] were the first to introduce the class P_{\parallel}^{NP} . Any language in this class can be decided in polynomial time using a polynomial number of *parallel* access to an NP oracle. Notice that, *parallel* access forbids adaptive queries and only allows ‘batch’ queries to an NP oracle.

We recall briefly that the preferences of voters in an election instance are typically solicited as either *rankings* (total orders over candidates) or *approval ballots* (subsets of “approved” candidates). The problem of finding a committee whose misrepresentation is bounded by a given threshold is known to be NP-complete for Chamberlin-Courant and Monroe [Lu and Boutilier, 2011; Procaccia *et al.*, 2008] in the setting of rankings as well as approval ballots. In a recent development ([Bredereck *et al.*, 2017a, Theorem 10], improving upon [Bredereck *et al.*, 2017b, Corollary 3]), it was shown that it is Θ_2^P -hard to determine whether a given candidate belongs to an optimal CC committee in the setting of rankings for the utilitarian method of aggregating misrepresentation scores. Following up on this, the main contribution of our work is to completely settle the complexity of two natural versions of the winner determination question in the context of the two fundamental multiwinner rules — Chamberlin-Courant and Monroe. We address these problems in the settings of both rankings and approval ballots, and for both the utilitarian and egalitarian methods of aggregating scores.

We consider the following problems for both the Chamberlin-Courant and Monroe rules, in the setting of approval ballots and rankings. In the WINNER VERIFICATION problem, we want to know if a proposed committee is optimal, and in the CANDIDATE WINNER problem, we are given a candidate c and a committee size k , and the question is if there exists an optimal k -sized committee containing c .

Winner Verification. Our first set of contributions is for the WINNER VERIFICATION problem; we show that it is complete for the complexity class coNP. In this case, the membership is easy to establish. For a given committee, observe that it is easy to compute its score with respect to the Chamberlin-Courant rule (and also the Monroe rule, although this is less straightforward). Thus, our coNP certificate is simply a “rival” committee with a better score. We remark, as an aside, that this is in contrast with rules such as Dodgson for which computing the Dodgson score of a given candidate is intractable. To show hardness for coNP, we reduce from the complement of the HITTING SET problem in different ways depending on the setting. For showing the hardness of the *Monroe Rule* we employ a variant where the elements enjoy uniform occurrences among the sets. Apart from settling the complexity of fundamental question of winner verification, our contribution identifies a natural coNP-complete problem, in particular, one that is *not* merely the complement of a natural NP-complete problem.

Candidate Winner. This problem was shown to be Θ_2^P -complete for the Chamberlin-Courant rule, for the utilitarian aggregation mechanism, in the setting of rankings [Bredereck *et al.*, 2017a]¹. We show the analogous result for the egalitarian version of the rule, and also for both aggregation mechanisms in the context of the Monroe rule. Although these reductions are executed in a similar spirit, the different settings do require non-trivial techniques in the constructions. Our main contributions are summarized below. We refer the reader to the next section for the relevant terminology.

Theorem 1. WINNER VERIFICATION for Chamberlin-Courant and Monroe is coNP-complete in the setting of approval ballots and rankings. In the latter setting, the result holds for the ℓ_1 and ℓ_∞ -Borda misrepresentation functions.

Theorem 2. CANDIDATE WINNER for Chamberlin-Courant and Monroe is complete for Θ_2^P in the setting of approval ballots and rankings. In the latter setting, the result holds for the ℓ_1 and ℓ_∞ -Borda misrepresentation functions.

Each of the statements above addresses six distinct scenarios. Due to lack of space, we give complete proofs for three of these settings, which we believe to be representative of the overall flavor of the arguments. We defer the detailed proofs to the full version of this work [Sonar *et al.*, 2020]. In the full version of the paper, we also show that the CANDIDATE WINNER problem for Chamberlin-Courant can be solved in polynomial time on single-peaked voting profiles.

¹This result was independently discovered by the authors of the present work, while following up on a weaker version of the from [Bredereck *et al.*, 2017b]. We omit our version of the proof in this setting because of the similarity of techniques involved.

2 Preliminaries

For a positive integer ℓ , we denote the set $\{1, \dots, \ell\}$ by $[\ell]$. We first define some general notions relating to voting rules. Let V be a set of n voters and C be a set of m candidates. If not mentioned otherwise, we denote the set of candidates, the set of voters, the number of candidates, and the number of voters by C, V, m and n , respectively. Every voter v has a *preference* \succ_v which is typically a complete order over the set C of candidates (rankings) or a subset of approved candidates (approval ballots). An instance of an election consists of the set of candidates C and the preferences of the voters V , usually denoted as $E = (C, V)$ with the understanding that the voters in V are identified by their preferences.

We now recall some definitions in the context of rankings. We say voter v prefers a candidate $x \in C$ over another candidate $y \in C$ if $x \succ_v y$. For a ranking \succ , $\text{pos}_\succ(c)$ is given by one plus the number of candidates ranked above c in \succ . In particular, if there are m candidates and c is the top-ranked (respectively, bottom-ranked) candidate in the ranking \succ , then $\text{pos}_\succ(c)$ is one (respectively, m). We denote the set of all preferences over C by $\mathcal{L}(C)$. The n -tuple $(\succ_v)_{v \in V} \in \mathcal{L}(C)^n$ of the preferences of all the voters is called a *profile*. We note that a profile, in general, is a multiset of linear orders. For a subset $M \subseteq V$, we call $(\succ_v)_{v \in M}$ a sub-profile of $(\succ_v)_{v \in V}$. For a subset of candidates $D \subseteq C$, we use $\mathcal{P}|_D$ to denote the projection of the profile \mathcal{P} on the candidates in D alone. The definitions of profiles, sub-profiles, and projections are analogous for approval ballots.

Chamberlin-Courant for Rankings. The Chamberlin-Courant voting rule is based on the notion of a *dissatisfaction* or a *misrepresentation* function. This function specifies, for each $i \in [m]$, a voter’s dissatisfaction $\alpha^m(i)$ from being represented by the candidate she ranks in position i . A popular dissatisfaction function is Borda, given by $\alpha^m(i) = i - 1$.

We now turn to the notion of an assignment function. Let $k \leq m$ be a positive integer. A k -CC-assignment function for an election $E = (C, V)$ is a mapping $\Phi: V \rightarrow C$ such that $|\Phi(V)| = k$, where $\Phi(V)$ denotes the image of Φ . For a given assignment function Φ , we say that voter $v \in V$ is *represented* by candidate $\Phi(v)$ in the chosen committee. There are several ways to measure the quality of an assignment function Φ with respect to a dissatisfaction function $\alpha: [m] \rightarrow \mathbb{R}$; and we will use the following:

1. $\ell_1(\Phi, \alpha) = \sum_{v \in V} \alpha(\text{pos}_{\succ_v}(\Phi(v)))$, and
2. $\ell_\infty(\Phi, \alpha) = \max_{v \in V} \alpha(\text{pos}_{\succ_v}(\Phi(v)))$.

Unless specified otherwise, α will be the Borda dissatisfaction function described above. We are now ready to define the Chamberlin-Courant voting rule.

Definition 1 (Chamberlin-Courant). For $\ell \in \{\ell_1, \ell_\infty\}$, the ℓ -CC voting rule is a mapping that takes an election $E = (C, V)$ and a positive integer k with $k \leq |C|$ as its input, and returns the images of all the k -CC-assignment functions Φ for E that minimizes $\ell(\Phi, \alpha)$.

Chamberlin-Courant for Approval Ballots. Recall that an approval vote of a voter v on the set of candidates C is some subset S_v of C such that v approves all the candidates in S_v . We define the misrepresentation score of a k -sized committee W as the number of voters which do not have any of their approved candidates in W (i.e. $W \cap S_v = \emptyset$). Hence the optimal committees under approval Chamberlin-Courant are the committees which maximize the number of voters with at least one approved candidate in the winning committee [Lackner and Skowron, 2018].

We are now ready to describe the questions that we study in this paper. The first problem is CHAMBERLIN-COURANT WINNER VERIFICATION (CCWV). Here, the input is an election $E = (C, V)$ and a subset S of k candidates. The question is if S is a winning k -sized CC-committee for the election E , in other words, does S achieve the best Chamberlin-Courant score in the given election among all committees of size k ?

In the second problem, given an election $E = (C, V)$, a committee size k , and a candidate $c \in C$, we ask if c belongs to *some* optimal k -sized committee, in other words, if there exists $S \subseteq C$ such that $c \in S$, $|S| = k$, and S is a winning CC committee. We refer to this as the CHAMBERLIN-COURANT CANDIDATE WINNER problem (CCCW).

We now turn to the definition of the Monroe voting rule [Monroe, 1995]. Note that for $c \in C$, $\Phi^{-1}(c)$ denotes the set of voters *represented* by c .

Definition 2 (Monroe). For $\ell \in \{\ell_1, \ell_\infty\}$, the ℓ -Monroe voting rule is a mapping that takes an election $E = (C, V)$ and a positive integer k with $k \leq |C|$ as its input, and returns the image of any of the k -Monroe-assignment functions Φ such that $|\Phi^{-1}(c)|$ is either $\lfloor \frac{n}{k} \rfloor$ or $\lceil \frac{n}{k} \rceil$ where $c \in C$ for E that minimizes $\ell(\Phi, \alpha)$.

We note that MONROE WINNER VERIFICATION (MWV) and MONROE CANDIDATE WINNER (MCW) are defined in the natural way. We also recall the definitions of 3-HITTING SET and its *complement*. In the 3-HITTING SET problem, we are given a ground set \mathcal{U} , a family \mathcal{F} of three-sized subsets of \mathcal{U} , and an integer k , and the question is if there exists $S \subseteq \mathcal{U}$ of size at most k that intersects every set in \mathcal{F} , i.e. $\forall F \in \mathcal{F}, S \cap F \neq \phi$. In the c-3-HITTING SET problem, the input is the same, and is a YES-instance if and only if there is no hitting set of size k ; in other words, if for each $S \subseteq \mathcal{U}$ with $|S| \leq k$, there exists some $F_S \in \mathcal{F}$ such that $S \cap F_S = \phi$. We recall that 3-HITTING SET is a classic NP-complete problem, and c-3-HITTING SET is *co-NP* complete.

The Class $P_{\parallel}^{NP}(\Theta_2^P)$. The class P_{\parallel}^{NP} is the class of problems solvable using a P machine having parallel access to an NP oracle. The class Θ_2^P was introduced in [Papadimitriou and Zachos, 1982] and named in [Wagner, 1990]. The class Θ_2^P was shown to be equivalent to P_{\parallel}^{NP} by Hemachandra [1989]. The VERTEX COVER MEMBER problem is the following. Given a graph $G := (V, E)$ and a vertex $w \in V$, the question is if there exists a minimum sized vertex cover containing w . The problem was shown to be complete for P_{\parallel}^{NP} by [Hemaspaandra *et al.*, 2005].

3 Winner Verification Problems

In this section, we show the coNP-completeness of CHAMBERLIN-COURANT WINNER VERIFICATION in the setting of rankings for the ℓ_1 -Borda misrepresentation score. The argument for membership is, in brief, the following: a rival committee with a better misrepresentation score is a valid certificate for the NO instances of CCWV. This is an efficiently computable certificate since it is easy to compute the Chamberlin-Courant score of a given committee. We now turn to the proof of hardness.

Theorem 3. CHAMBERLIN-COURANT WINNER VERIFICATION is *coNP-hard* in the setting of rankings for the ℓ_1 -Borda misrepresentation score.

Proof: We show a reduction from c-3-HITTING SET to the CC-WINNER problem. Let $\langle U, \mathcal{F}; k \rangle$ be an instance of c-3-HITTING SET with n elements in the universe U and m sets of size three in the family \mathcal{F} . We construct a profile \mathcal{P} over alternatives \mathcal{A} as follows. First, we introduce one candidate corresponding to each element of the universe U , k “dummy” candidates, and a large number of “filler” candidates, that is:

$$\mathcal{A} := \underbrace{\{c_u \mid u \in U\}}_C \cup \underbrace{\{d_1, \dots, d_k\}}_D \cup \underbrace{\{z_1, \dots, z_t\}}_Z,$$

where $t = 3(mk)^2$. Also, for every $1 \leq i \leq k$, and for every $X \in \mathcal{F}$, introduce a vote $v(i, X)$ that places the candidates corresponding to the elements in X in the top three positions, followed by d_i , followed by $3mk$ candidates from Z . We ensure that we use distinct candidates from Z in the top $3mk + 4$ positions of all the voters, in other words, no candidate from Z appears twice in the top $3mk + 4$ positions. Note that t is chosen to be large enough to make this possible. This is followed by the candidates in $U \setminus X$ ranked in an arbitrary order followed by the remaining filler candidates, also ranked in an arbitrary order.

In this instance, note that a committee corresponding to a hitting set has a score of at most $2mk$, while the score of the committee D is $3mk$. In the constructed instance, we now ask if the committee D consisting of k dummy candidates is a winning committee. This completes the construction of the instance. We now turn to the equivalence of two instances.

In the forward direction, suppose we have a YES instance of c-3-HITTING SET. This implies that there does not exist any hitting set of size at most k . Recall that the misrepresentation score for a committee consisting of a hitting set is at most $2mk$, while noting that any such committee must have size greater than k . Now, we show that for all other committees of size at most k , the misrepresentation score is greater than $3mk$.

Lemma 1. Consider an instance $\langle \mathcal{A}, V, \mathcal{D} \rangle$ of CC-winner Verification based on a YES-instance of c-3-Hitting Set $\langle U, \mathcal{F}; k \rangle$. For any feasible committee $C' \subseteq \mathcal{A}$ of size k different from \mathcal{D} , the ℓ_1 -Borda misrepresentation score of C' is greater than $3mk$.

Proof: (of Lemma 1.) Let U' , D' and Z' denote, respectively, the candidate subsets $C' \cap C$, $C' \cap D$ and $C' \cap Z$. Since C' is

different from \mathcal{D} , there is at least one candidate from \mathcal{D} that does not belong to C' (the only other possibility is that C' is a superset of \mathcal{D} , but this is not possible since $|C'| = |\mathcal{D}| = k$). Without loss of generality, suppose $d_1 \notin C'$. Now consider the votes given by $V' := \{v(1, X) \mid X \in \mathcal{F}\}$. We claim that there are at least $|Z'| + 1$ voters in V' whose misrepresentation score for the committee C' is strictly greater than three. Indeed, if not, then it is straightforward to verify that U' combined with an arbitrarily chosen element from each set not hit by U' comprises a subset of size at most $|U'| + |Z'| \leq k$ which intersects every set in \mathcal{F} , contradicting our assumption that \mathcal{F} has no hitting set of size at most k . To see this, observe that every vote in V' that has a misrepresentation score of three or less is necessarily represented by a candidate from U' , since $d_1 \notin C'$, and therefore, the sets corresponding to all of these votes are hit by U' , and the remaining sets can be hit “trivially” since there are at most $|Z'|$ of them. Now consider the voters who have a “high” misrepresentation score: $V'' := \{v(1, X) \mid X \in \mathcal{F} \text{ and } \tau(v(1, X), C') > 3\}$, where $\tau(v(1, X), C')$ is the Borda score of the highest-ranked candidate of C' according to $v(1, X)$, with respect to the ranking of $v(1, X)$. By the argument in the previous paragraph, we have that $|V''| > |Z'|$. Recalling that every vote has distinct filler candidates in the top $3mk$ positions after d_i , by the pigeon-hole principle, we conclude that there is at least one vote $v(1, X)$ in V'' such that $Z_X \cap Z' = \emptyset$, where Z_X denotes the filler candidates that appear in the top $3mk + 4$ positions of the vote $v(1, X)$. Since the candidates occupying the top four positions of this vote do not belong to C' either, it follows that the misrepresentation score of $v(1, X)$ for C' is greater than $3mk$, and this concludes our argument. ■

The committee \mathcal{D} has a misrepresentation score of $3mk$. Using Lemma 1, since \mathcal{F} has no hitting set of size at most k , we have that \mathcal{D} is a winning committee among all feasible committees, as desired.

In the reverse direction, we start with the assumption that \mathcal{D} is a winning committee. Therefore, the optimal misrepresentation for the constructed election instance is $3mk$. Observe that if there exists a hitting set S of size at most k , then the committee C' formed using the corresponding candidates of hitting set will have a misrepresentation score of at most $2mk$, as discussed above. Thus, \mathcal{D} would not be a winning committee, a contradiction — and this implies that $\langle U, \mathcal{F}; k \rangle$ was indeed a YES-instance of c-3-HITTING SET. This completes the argument of equivalence. ■

For ℓ_∞ -CC, we again reduce from the complement of the 3-HITTING SET problem with a similar construction. We introduce votes corresponding to sets in the family, where the top three candidates are the candidates corresponding to the elements contained in the set, and the fourth candidate is a dummy candidate. Once we construct k blocks with distinct dummy candidates in the fourth position, the possible misrepresentations will play out in an analogous fashion. For obtaining these three results in the setting of Monroe, while the reduction is similar, we have to reduce from a variant of the Hitting Set problem with additional structure where the elements appear with an uniform frequency in the set system.

4 Candidate Winner Problems

In this section, we turn to the MONROE CANDIDATE WINNER problem. Recall that the input is $\langle C, V, c, k \rangle$, and the question is if there exists an optimal Monroe committee of size k containing c . We demonstrate that the problem is complete for Θ_2^P in the setting of rankings for both the ℓ_1 -Borda misrepresentation function. The argument for the case of approval ballots for this problem are in a similar spirit, and are deferred to a full version of this paper.

We first consider the case of the ℓ_1 -Borda misrepresentation. Our focus here will be on showing hardness, and we informally justify the claim for membership. We use oracle queries to a variant of MCW where we additionally demand for the committee to achieve a particular target misrepresentation score. Note that the worst possible misrepresentation score in an instance with m candidates and n voters is mn . Thus, by guessing this target score, we can find the score of the optimal Monroe committee that contains c and the score of the optimal Monroe committee, and comparing these answers the question of whether there exists an optimal Monroe committee of size k containing c . We now turn to the reduction to demonstrate hardness.

Theorem 4. MONROE CANDIDATE WINNER is Θ_2^P -hard for the ℓ_1 -Borda misrepresentation function.

Proof: We reduce from the Θ_2^P -complete problem VERTEX COVER MEMBER. Recall that we are given a graph $G := (V, E)$ (with n vertices and m edges), and a vertex $w \in V$, the question is if there exists a minimum sized vertex cover containing w . Given an instance $\langle G := (V, E), w \rangle$ of VERTEX COVER MEMBER we construct an instance of CC CANDIDATE WINNER as follows. Let the set of candidates be $C := C_v \cup D \cup D' \cup S$, where C_v denotes the set of n candidates corresponding to vertices of G and D and D' denote type I and type II dummy candidates respectively. Let Δ denote a set of n^4m type I dummy candidates, and Δ' denote a set of $2(n^4m)$ type II dummy candidates. We note that the subsets of dummy candidates specified explicitly in different votes are always chosen so that there are no repeated dummy candidates in the explicitly defined blocks, in other words, the chosen dummy candidates are always distinct. Also, D (D') is the union of all the Δ 's (Δ' 's) respectively specified in the profile, which is given by the following five blocks of voters:

- ▷ **Block 1:** We construct m votes corresponding to the edges in G . For an edge (u, v) we add:

$$c_u \succ c_v \succ \Delta' \succ C_v \setminus \{c_u, c_v\} \succ \text{rest}$$

where “rest” denotes the set of remaining candidates placed in an arbitrary order.

- ▷ **Block 2:** For the desired vertex w from the VERTEX COVER MEMBER instance, we pick an arbitrary edge incident on w (say (w, x)) in G , and add $m + 1$ copies of the following vote:

$$c_w \succ c_x \succ \Delta' \succ C_v \setminus \{c_w, c_x\} \succ \text{rest}$$

- ▷ **Block 3:** We add n votes of the form:

$$\Delta \succ c_w \succ C_v \setminus \{c_w\} \succ \text{rest}$$

▷ **Block 4:** For each candidate $d_j \in D_s$, we add the vote:

$$d_j \succ \Delta' \succ \text{rest.}$$

▷ **Block 5:** Let D_α be a subset of dummy candidates $d \in D$ such that d appears in the top position for one of the votes in Block 3. Note that $|D_\alpha| = n$. Further, let $\mathcal{N} = 2(m + n + 1)$. For each $v \in \{C_v \cup D''\}$ and $\ell \in [\mathcal{N}]$, we add the following vote:

$$v \succ D_s \succ \Delta' \succ \text{rest}$$

In the constructed MONROE CANDIDATE WINNER instance, we ask if there exists an optimal committee of size $2n+1$ containing c_w . This completes the construction for our reduction. Before showing the equivalence of the two instances, we establish the following lemma.

Lemma 2. *Let q be the size of an optimal vertex cover in G . Then, the following holds for any optimal committee C' of size $2n + 1$ in the constructed election instance:*

1. C' does not contain any $d' \in D'$.
2. $S \subset C'$.
3. C' contains exactly q candidates corresponding to an optimal vertex cover.

Proof: First, we analyze the Monroe score of a committee \mathcal{C} which contains all $n + 1$ special candidates S , q candidates corresponding to an optimal vertex cover (S'), and the remaining $(n - q)$ candidates from D which appear in the top positions of $(n - q)$ votes in Block 3. Note that in any Monroe committee C' , each candidate represents exactly \mathcal{N} votes. We now describe the Monroe assignment for \mathcal{C} . Each vote in Block 1 and 2 is represented by one of the top two candidates such that the corresponding vertex (v) belongs to the vertex cover (S'). The misrepresentation for \mathcal{C} in Block 1 is at most m , and in Block 2 it is at most $m+1$. In Block 3, exactly $n - q$ votes are represented by their first choice. For those votes that are not represented by the top candidate already, the misrepresentation for \mathcal{C} is at most $(|\Delta| + n - 1)$ per vote since all the votes in Block 3 are represented by the candidate in C_v for that vote in the worst case. In Block 4, all votes are represented by their top choice yielding zero misrepresentation. Votes in Block 5 are represented as follows: For each candidate c_i corresponding to a vertex $u \in S'$, if c_u represents t votes from first 3 blocks, then c_u also represents $(\mathcal{N} - t)$ votes from Block 5 among the ones she appears at the first position. Similarly, for $d \in \{D \cap \mathcal{C}\}$, d represents $\mathcal{N} - 1$ votes among the ones she appears at the top position. Next, each special candidate $s \in S$ represent $(\mathcal{N} - 1)$ votes in Block 5, yielding misrepresentation score at most $(n + 1)(\mathcal{N} - 1)$ for each s . Hence, the total misrepresentation for \mathcal{C} is strictly less than $n^4m + 3mn^2 + 3n^3 < 2(n^3m^2 + n^4m)$ for large enough n .

Towards showing the first statement, consider a committee C^* which contains $d' \in D'$. In any Monroe assignment, d' has to represent \mathcal{N} votes. Observe that d' appears in first (n^4m) positions exactly once, hence, $\text{misrepresentation}(C^*) > \text{misrepresentation}(\mathcal{C})$. To show the second statement, consider a committee C^* which excludes a special candidate $s \in S$. It is easy to see that

$\text{misrepresentation}(C^*) > \text{misrepresentation}(\mathcal{C})$ even if we only consider misrepresentation from a single vote from Block 4 with s at the first position.

We now turn to statement 3. Now, let C^* be an optimal committee which does not contain q candidates corresponding to some optimal vertex cover. We use statements 1 and 2 to analyze the following two cases:

- ▷ $|C^* \cap C_v| > q$: In this case, C^* contains at most $(n - q - 1)$ candidates from D . Hence, the misrepresentation of C^* from Block 3 is at least $|\Delta| \times (q + 1)$ which is greater than the misrepresentation score for \mathcal{C} . This contradicts the optimality of C^* .
- ▷ $|C^* \cap C_v| \leq q$: Since the size of an optimal vertex cover is q , any committee C^* with at most q candidates from C_v does not include any candidates corresponding to the endpoint of at least one edge due to the case we are in (i.e., C^* does not contain candidates corresponding to an optimal vertex cover). Hence, C^* incurs a misrepresentation of at least $|\Delta'|$ from one of the votes in Block 1 which implies $\text{misrepresentation}(C^*) > \text{misrepresentation}(\mathcal{C})$.

This completes the proof for Lemma 2. ■

We now turn to the proof of equivalence. In the forward direction, given an optimal vertex cover of size q containing w , we construct an optimal committee C' by choosing q candidates corresponding to the vertex cover, $(n - q)$ candidates from the set D which appears in the top position of exactly $(n - q)$ votes from Block 3, all $n + 1$ special candidates S . We compute the Monroe assignment of C' is the same way we did for committee \mathcal{C}' in Lemma 2. By Lemma 2, we already know that any optimal committee must contain all candidates from S , and candidates corresponding to an optimal sized vertex cover. Therefore, it suffices to show that committees corresponding to an optimal vertex covers not containing c_w are not optimal. Indeed, this follows from the fact that in Block 2, c_w is the top candidate in exactly $(m + 1)$ votes, and in Block 3, c_w leads all other candidates from the set C_v . Hence, it is easy to verify that an optimal committee must contain c_w .

In the reverse direction, given an optimal committee C' containing c_w , we need to construct an optimal vertex cover for G which includes the vertex w . Since C' is optimal, using Lemma 2 we know $C' \cap C_v$ is an optimal vertex cover of G . Since we are given that $c_w \in C'$, we have that the vertex cover corresponding to C' is an optimal vertex cover containing c_w , as desired. ■

Now, we show that CCCW is hard for Θ_2^P in the setting of rankings for the ℓ_∞ -Borda misrepresentation function. We recall that an analogous result for the ℓ_1 -Borda misrepresentation function was shown in [Bredereck *et al.*, 2017b]. The argument for membership is similar to the previous case and is omitted for brevity.

Theorem 5. *CCCW is Θ_2^P -hard for the ℓ_∞ -Borda misrepresentation function.*

Proof: As before, we reduce from VERTEX COVER MEMBER. Given an instance $\langle G := (V, E), w \rangle$ of VERTEX

COVER MEMBER we construct an instance of CC CANDIDATE WINNER as follows. Let the set of candidates be $C := C_v \cup D \cup D'$, where C_v denotes the set of n candidates corresponding to vertices of G , and D and D' denote *type I* and *type II* dummy candidates respectively. Let Δ denote a set of $m + n + 1$ type I dummy candidates, and Δ' denote a set of 2 type II dummy candidates. We note that the subsets of dummy candidates specified explicitly in different votes are always chosen so that there are no repeated dummy candidates in the explicitly defined blocks, in other words, the chosen dummy candidates are always distinct. We construct the set of voters as the following three blocks:

- ▷ **Block 1:** We construct $(n + 2)$ copies of each of m votes corresponding to the edges in G . Specifically, for an edge (u, v) we add $(n + 2)$ copies of the vote:

$$c_u \succ c_v \succ \Delta \succ C_v \setminus \{c_u, c_v\} \succ \text{rest},$$

where *rest* denotes the set of remaining candidates in some arbitrary order.

- ▷ **Block 2:** We add the following n votes:

$$v_1 := \Delta' \succ d_{1,1} \succ d' \succ \Delta \succ C_v \succ \text{rest}$$

⋮

$$v_i := \Delta' \succ d_{i,1} \succ \dots \succ d_{i,i} \succ d' \succ \Delta \succ C_v \succ \text{rest}$$

⋮

$$v_n := \Delta' \succ d_{n,1} \succ \dots \succ d_{n,n} \succ d' \succ \Delta \succ C_v \succ \text{rest}$$

- ▷ **Block 3:** We also add $(n + 2)$ copies of the following vote to force d' in any optimal committee:

$$d' \succ \Delta \succ \text{rest}$$

In the constructed CCCW instance, we ask if there exists an optimal committee of size $n + 1$ containing c_w . This completes the construction for our reduction. We now state a lemma analogous to Lemma 2, whose proof is omitted due to space constraints.

Lemma 3 (\star). *Let q be the size of an optimal vertex cover in G such that $q \geq 2$. Then, the following holds for any committee C' of size $n + 1$ in the constructed election instance: (a) any optimal committee contains candidate d' , (b) if C' contains $d \in D$, then C' is not an optimal committee, and (c) if C' does not contain exactly q candidates corresponding to an optimal vertex cover, then C' is not optimal.*

Forward direction. Given an optimal vertex cover V' of size q containing w , we construct a committee \mathcal{C} as described in the paragraph one of the proof of Lemma 3. We claim that \mathcal{C} is an optimal committee. Using Lemma 3 we know that any optimal committee contains exactly q candidates corresponding to an optimal vertex cover and contains candidate d' . From the structure of Block 2, it is easy to observe that in order to minimize the overall misrepresentation from Block 2, the only way is to choose the candidates from position $n - q + 2$ for voters v_j for $j \in \{n - q + 1, n - q + 2, \dots, n\}$. Note that

first q voters from Block 2 are represented by d' while staying within the misrepresentation limit of $n - q + 1$.

Reverse direction. Let C' be an optimal committee containing c_w . Using Lemma 3, we know that C' contains candidates corresponding to an optimal vertex cover. Hence, we can recover a desired optimal vertex cover containing c_w . ■

For the analogous results for the CC and Monroe voting rules in the remaining settings, we refer the reader to the full version of this work [Sonar *et al.*, 2020].

5 Concluding Remarks

We have addressed the problems of WINNER VERIFICATION and CANDIDATE WINNER in the setting of multiwinner voting. We have resolved the complexity of both the problems for the Chamberlin-Courant and Monroe voting rules in various scenarios. In particular, for rankings, we have considered both the ℓ_1 and the ℓ_∞ notions of misrepresentation, and we have also studied variants of these rules in the context of approval ballots. We have shown that verifying if a given committee is optimal is coNP-complete whilst the second problem is complete for Θ_2^P in all the twelve cases. Overall, our results comprehensively settle the complexity of these two problems in the general setting. These outcomes primarily serve the purpose of deepening our understanding of where these problems lie in the complexity-theoretic landscape. Further, since the WINNER VERIFICATION family of problems are complete for Θ_2^P , our results also hint that natural heuristics for the question are unlikely to perform well in practice. Indeed, investigating the performance of heuristics (by possibly adapting greedy approaches for finding optimal committees and forcing the choice of a desired candidate) would be an interesting direction for complementing our theoretic considerations.

Another natural direction for further thought is the setting of restricted domains, which have received much attention for capturing structure in real-world data sets and for providing natural “islands of tractability” for several hard voting problems [Elkind *et al.*, 2017]. Indeed, although determining optimal committees for the Monroe rule remains intractable even in the setting of single-crossing profiles [Skowron *et al.*, 2015]), we can find an optimal Chamberlin-Courant committee efficiently if the input is single-peaked [Betzler *et al.*, 2013] or single-crossing [Skowron *et al.*, 2015]. Further, the rule is tractable also for structured approval ballots [Elkind and Lackner, 2015]. With this background, it would be interesting to explore the complexity of the problems we study in the setting of restricted domains. The WINNER VERIFICATION problems are tractable whenever the naturally associated WINNER DETERMINATION problem is tractable, but the CANDIDATE WINNER problem is less immediate to resolve. In the single-peaked setting, with the ℓ_1 -Borda misrepresentation score, the CANDIDATE WINNER problem can be resolved by adding several dummy voters who place the desired candidate at the top position, and comparing the optimal CC scores of the original and modified instances. The situation for other restricted domains remains open.

References

- [Bartholdi *et al.*, 1989] John Bartholdi, Craig A Tovey, and Michael A Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and welfare*, 6(2):157–165, 1989.
- [Betzler *et al.*, 2013] Nadja Betzler, Arkadii Slinko, and Johannes Uhlmann. On the computation of fully proportional representation. *Journal of Artificial Intelligence Research*, 47:475–519, 2013.
- [Bredereck *et al.*, 2017a] Robert Bredereck, Piotr Faliszewski, Andrzej Kaczmarczyk, Rolf Niedermeier, Piotr Skowron, and Nimrod Talmon. Robustness among multiwinner voting rules. *CoRR*, abs/1707.01417, 2017.
- [Bredereck *et al.*, 2017b] Robert Bredereck, Piotr Faliszewski, Andrzej Kaczmarczyk, Rolf Niedermeier, Piotr Skowron, and Nimrod Talmon. Robustness among multiwinner voting rules. In *10th International Symposium of Algorithmic Game Theory (SAGT)*, pages 80–92, 2017.
- [Chamberlin and Courant, 1983] John R Chamberlin and Paul N Courant. Representative deliberations and representative decisions: Proportional representation and the Borda rule. *American Political Science Review*, 77(3):718–733, 1983.
- [Dodgson, 1876] Charles L. Dodgson. A method of taking votes on more than two issues. In *Classics of Social Choice*, pages 288–297. University of Michigan Press (Reprinted in 1995), 1876.
- [Elkind and Lackner, 2015] Edith Elkind and Martin Lackner. Structure in dichotomous preferences. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, pages 2019–2025, 2015.
- [Elkind *et al.*, 2017] Edith Elkind, Martin Lackner, and Dominik Peters. Structured preferences. *Chapter 10, Trends in computational social choice (Ed. Ulle Endriss)*, pages 187–207, 2017.
- [Hemachandra, 1989] Lane A Hemachandra. The strong exponential hierarchy collapses. *Journal of Computer and System Sciences*, 39(3):299–322, 1989.
- [Hemaspaandra *et al.*, 2005] Edith Hemaspaandra, Holger Spakowski, and Jörg Vogel. The complexity of Kemeny elections. *Theoretical Computer Science*, 349(3):382–391, 2005.
- [Kemeny, 1959] John G Kemeny. Mathematics without numbers. *Daedalus*, 88(4):577–591, 1959.
- [Lackner and Skowron, 2018] Martin Lackner and Piotr Skowron. Consistent approval-based multi-winner rules. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 47–48. ACM, 2018.
- [Lu and Boutilier, 2011] Tyler Lu and Craig Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Twenty-Second International Joint Conference on Artificial Intelligence*, pages 280–286, 2011.
- [Monroe, 1995] Burt L Monroe. Fully proportional representation. *American Political Science Review*, 89(4):925–940, 1995.
- [Papadimitriou and Zachos, 1982] Christos H Papadimitriou and Stathis K Zachos. Two remarks on the power of counting. In *Theoretical Computer Science*, pages 269–275. 1982.
- [Procaccia *et al.*, 2008] Ariel D Procaccia, Jeffrey S Rosenschein, and Aviv Zohar. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30:353–362, 2008.
- [Skowron *et al.*, 2015] Piotr Skowron, Lan Yu, Piotr Faliszewski, and Edith Elkind. The complexity of fully proportional representation for single-crossing electorates. *Theoretical Computer Science*, 569:43–57, 2015.
- [Skowron *et al.*, 2016] Piotr Skowron, Piotr Faliszewski, and Jérôme Lang. Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence*, 241:191–216, 2016.
- [Sonar *et al.*, 2020] Chinmay Sonar, Palash Dey, and Neeldhara Misra. On the complexity of winner verification and candidate winner for multiwinner voting rules. *CoRR*, abs/2004.13933, 2020.
- [Wagner, 1990] Klaus W Wagner. Bounded query classes. *SIAM Journal on Computing*, 19(5):833–846, 1990.