On the Complexity of Finding a Diverse and Representative Committee using a Monotone, Separable Positional Multiwinner Voting Rule

Kunal Relia*
New York University, USA
krelia@nyu.edu

Abstract
Fairness in multiwinner elections, a growing line of research in computational social choice, primarily concerns the use of constraints to ensure fairness. Recent work proposed a model to find a diverse and representative committee and studied the model’s computational aspects. However, the work gave complexity results under major assumptions on how the candidates and the voters are grouped. Here, we close this gap and classify the complexity of finding a diverse and representative committee using a monotone, separable positional multiwinner voting rule, conditioned only on the assumption that \( P \neq NP \).

1 Introduction
Fairness has recently received particular attention from the computer science research community. For context, the number of papers that contain the words “fair” or “fairness” in their titles and are published at top-tier computer science conferences like NeurIPS and AAAI grew at an average of 38% year on year since 2018. Moreover, the conference ACM FAT*, formerly known as FAT, was established in 2018 to “bring together researchers and practitioners interested in fairness, accountability, and transparency in socio-technical systems”. Similarly, there is a growing trend in the Computational Social Choice (COMSOC) community towards the use of “fairness” [Bredereck et al. 2018; Celis, Huang, and Vishnoi 2018; Cheng et al. 2019; Hershkowitz et al. 2021; Shrestha and Yang 2019].

However, the term “fairness” is used in varying contexts. For example, Bredereck et al. (2018) and Celis et al. (2018) call diversity of candidates in committee elections as fairness, while Cheng et al. (2019) call representation of voters in committee elections as fairness. Such context-specific use of the term narrates an incomplete story. Hence, Relia (2022) unified the framework using the DiRe Committee model that combines the use of diversity and representation constraints. In line with the conceptual difference, the use of constraints leads to setups of multiwinner elections that are technically different. For instance, diversity constraint is a property of candidates and representation constraint is a property of voters. The use of these constraints are mathematically as different as the regularity and uniformity properties of hypergraphs [2]. Hence, it is important to mathematically delineate the two notions.

A starting step towards understanding the difference between diversity and representation is having a classification of complexity of using diversity and representation constraints in multiwinner elections. The classification of complexity, while technically interesting and important in itself, enables a detailed understanding of the nuances that delineate the two notions. Our main contributions are the complexity results that are based on a singular assumption that \( P \neq NP \).

2 Related Work
2.1 COMSOC and Classification of Complexity
Computational Social Choice research has particularly focused on classifying the complexity of a known social choice problem. For instance, Konzcac and Lang (2005) introduced the problem of voting under partial information. This led to research of line that aimed to classify the complexity of the problem of possible winners (candidate is a winner in at least one completion) and necessary winners (candidate is a winner in all completions) over all pure positional scoring rules (Baumeister and Rothe 2012; Betzler and Dorn 2010; Chakraborty and Kolaitis 2021; Kenig 2019; Xia and Conitzer 2011).

2.2 Multiwinner Elections, Fairness, and Complexity
Our work primarily builds upon the literature on constrained multiwinner elections. Fairness from candidate’s perspective is discussed via the use of diversity constraints over multiple attributes and its use is known to be NP-hard, which has led to approximation algorithms and matching hardness of approximation results by Bredereck et al. (2018) and Celis et al. (2018). Additionally, goalbase score functions, which specify an arbitrary set of logic constraints and let the score capture the number of constraints satisfied, could be used to ensure diversity (Uckelman et al. 2009). On the other hand, the research on Fair Resource Allocation due to the specificity in the use of the word “fair” [1].

*This work was supported, in part, by Julia Stoyanovich’s NSF grants No. 1934464 and 1916505.

[1] We do not consider the research on Fair Resource Allocation due to the specificity in the use of the word “fair”.

[2] The mathematical differences between, say, the vertex cover problem on \( d \)-regular hypergraphs and \( k \)-uniform hypergraphs are well-known (Bansal and Khot 2014; Feige 2003).
hand, the study of fairness from voters’ perspective pertains the use of representation constraints \cite{Cheng19}. Finally, due to the hardeness of using diversity constraints over multiple attributes in approval-based multiwinner elections \cite{Brams90}, these have been formalized as integer linear programs (ILP) \cite{Potthoff90}.

Overall, our work is at the intersection of the interest COMSOC researchers have on classifying the complexity and fairness in multiwinner elections. Specifically, our work is the closest to the work by Bredereck et al. \cite{Bredereck2018}, Celis et al. \cite{Celis19} and Relia \cite{Relia22} but we differ as we: (i) provide a complete classification of complexity of using finding a diverse and representative committee using a monotone, separable positional multiwinner voting rule, (ii) our NP-hardness results hold for all integer values of attributes, and (iii) our NP-hardness results are conditioned only on the assumption that P ≠ NP.

3 Preliminaries and Notation

Multiwinner Elections. Let \( E = (C, V) \) be an election consisting of a candidate set \( C = \{c_1, \ldots, c_m\} \) and a voter set \( V = \{v_1, \ldots, v_n\} \), where each voter \( v \in V \) has a preference list \( >_v \) over \( m \) candidates, ranking all of the candidates from the most to the least desired. \( \text{pos}_v(c) \) denotes the position of candidate \( c \in C \) in the ranking of voter \( v \in V \), where the most preferred candidate has position 1 and the least preferred has position \( m \).

Given an election \( E = (C, V) \) and a positive integer \( k \in [m] \) (for \( k \in \mathbb{N}, \{k\} = \{1, \ldots, k\} \)), a multiwinner election selects a \( k \)-sized subset of candidates (or a committee) \( W \) using a multiwinner voting rule \( \mu \) (discussed later) such that the score of the committee \( \mu(W) \) is the highest. We assume ties are broken using a pre-decided priority order.

Candidate Groups. The candidates have \( \mu \) attributes, \( A_1, \ldots, A_\mu \), such that \( \mu \in \mathbb{Z} \) and \( \mu \geq 0 \). Each attribute \( A_i \), \( \forall i \in [\mu] \), partitions the candidates into \( g_i \in [m] \) groups, \( A_{i}(1), \ldots, A_{i}(g_i) \subseteq C \). Formally, \( A_{i}(j) \cap A_{i}(j') = \emptyset \), \( \forall j, j' \in [g_i], j \neq j' \). For example, candidates may have race and gender attribute (\( \mu = 2 \)) with disjoint groups per attribute, male and female \( (g_1 = 2) \) and Caucasian and African-American \( (g_2 = 2) \). Overall, the set \( G \) of all such arbitrary and potentially non-disjoint groups is \( A_{(1,1)}, \ldots, A_{(\mu,g_\mu)} \subseteq C \).

Voter Populations. The voters have \( \pi \) attributes, \( A_1', \ldots, A_\pi' \), such that \( \pi \in \mathbb{Z} \) and \( \pi \geq 0 \). The voter attributes may be different from the candidate attributes. Each attribute \( A_i' \), \( \forall i \in [\pi] \), partitions the voters into \( p_i \in [n] \) populations, \( P_{i}(1), \ldots, P_{i}(p_i) \subseteq V \). Formally, \( P_{i}(j) \cap P_{i}(j') = \emptyset \), \( \forall j, j' \in [p_i], j \neq j' \). For example, voters may have state attribute (\( \pi = 1 \)), which has populations California and Illinois \( (p_1 = 2) \). Overall, the set \( P \) of all such predefined and potentially non-disjoint populations will be \( P_{(1,1)}, \ldots, P_{(\pi,p_\pi)} \subseteq V \).

Additionally, we are given \( W_P \), the winning committee \( \forall P \in P \). We limit the scope of \( W_P \) to be a committee instead of a ranking of \( k \) candidates because when a committee selection rule such as CC rule is used to determine each population’s winning committee \( W_P \), then a complete ranking of each population’s collective preferences is not possible.

Multiwinner Voting Rules. There are multiple types of multiwinner voting rules, also called committee selection rules. In this paper, we focus on committee selection rules \( \mu \) that are based on single-winner positional voting rules, and are monotone and submodular \( (\forall A \subseteq B, \mu(A) \leq \mu(B) \text{ and } f(B) \leq f(A) + f(B \setminus A)) \), and specifically separable \cite{Bredereck2018, Celis19, Vishnoi18}. Specifically, a special case of submodular functions are separable functions: score of a committee \( W \) is the sum of the scores of individual candidates in the committee. Formally, \( \mu \) is separable if it is submodular and \( \mu(W) = \sum_{c \in W} \mu(c) \) \cite{Bredereck2018}. Monotone and separable selection rules are natural and are considered good when the goal of an election is to shortlist a set of individually excellent candidates:

Definition 1. \( k \)-Borda rule The \( k \)-Borda rule outputs committees of \( k \) candidates with the highest Borda scores.

4 Classification of Complexity

We now study the computational complexity of DRCWD due to the presence of candidate attributes and voter attributes. Specifically, we establish the NP-hardness of DRCWD for various settings of \( \mu, \pi, \mu, \pi \) via reductions from a single well known NP-hard problem, namely, the vertex cover problem on 3-regular \( 2 \)-uniform hypergraphs \cite{Garey1979, Alimonti97}.

Note that the reductions are designed to conform to the real-world stipulations of candidate attributes such that (i) each candidate attribute \( A_i, \forall i \in [\mu] \), partitions all the \( m \) candidates into two or more groups and (ii) either no two attributes partition the candidates in the same way or if they do, the lower bounds across groups of the two attributes are not the same. For stipulation (ii), note that if two attributes partition the candidates in the same way and if the lower bounds across groups of the two attributes are also the same, then mathematically they are identical attributes that can be combined into one attribute. We use the same stipulations for voter attributes.

Finally, all our results are based only on the assumption that P ≠ NP.

Theorem 1. If \( \mu = 1, \forall \pi \in \mathbb{Z}, \pi \geq 1, \text{ and } \mu \text{ is the monotone, separable function arising from an arbitrary single-winner positional scoring rule, then DRCWD is NP-hard.}

Proof Sketch. The proof extends from the reduction used in the proof of Theorem 5 in Relia \cite{Relia22} (full-version). Specifically, we have \( m + (n \cdot m) + 1 \) candidates such that \( m + (n \cdot m) \) from the previous reduction are in one group and we have a dummy candidate \( a \) in the second group. The diversity constraints are set to 1 for both the groups. The

---

3A 3-regular graph stipulates that each vertex is connected to exactly three other vertices, each one with an edge, i.e., each vertex has a degree of 3. The VC problem on 3-regular graphs is NP-hard.

4A 2-uniform hypergraph stipulates that each edge connects exactly two vertices.
voters, voter populations, and rankings are the same as before except candidate $a$ is ranked first by all the voters. The representation constraints are set to 2 for all voter populations. The committee size is set to $k + 1$. It is easy to see that the proof of correctness follows the proof of correctness of Theorem 5 [Relia202](full-version) with an addition of always selecting the candidate $a$.

**Theorem 2.** If $\mu = 2$, $\forall \pi \in Z : \pi \geq 1$, and $f$ is the monotone, separable function arising from an arbitrary single-winner positional scoring rule, then DRCWD is NP-hard.

**Proof Sketch.** We now build upon the reduction used in the proof of Corollary [1]. The only change in the reduction is the addition of the second candidate attribute. In addition two groups already present under one attribute, we create two more groups under the second attribute such that the first group contains the $m$ candidates from $C$ and the second group contains $(n \cdot m) + 1$ candidates ($(n \cdot m)$ dummy candidates $D$ and 1 dummy candidate $a$). The diversity constraints are set to 1 for both the groups. The voters, voter populations, rankings, and the representation constraints are the same as before. The committee size is again set to $k + 1$. It is easy to see that the proof of correctness follows the proof of correctness of Corollary [1].

**Theorem 3.** If $\forall \mu \in Z : \mu \geq 3$ and $\mu$ is an odd number, $\forall \pi \in Z : \pi \geq 1$, and $f$ is the monotone, separable function arising from an arbitrary single-winner positional scoring rule, then DRCWD is NP-hard.

**Proof.** We reduce an instance of vertex cover (VC) problem to an instance of DRCWD.

**PART 1: Construction**

**Candidates:** We have one candidate $c_i$ for each vertex $x_i \in X$, and $2\mu^2m - 7\mu m + 2\mu mn + 2mn + 3m$ dummy candidates $d \in D$ where $m$ corresponds to the number of vertices in the graph $H$, $n$ corresponds to the number of edges in the graph $H$, and $\mu$ is a positive, odd integer (hint: the number of candidate attributes). Specifically, we divide the dummy candidates into two types of blocks:

- Block type $B_1$ consists of $m\mu - 3m$ blocks and each block consists of three sets of candidates:
  - Set $T_1$ consists of single dummy candidate, $d_{i,1}^{T_1} \in T_1$, $\forall i \in [1, m\mu - 3m]$.
  - Set $T_2$ consists of $\mu - 1$ dummy candidates, $d_{i,j}^{T_2} \in T_2$, $\forall i \in [1, m\mu - 3m], j \in [1, \mu - 1]$.
  - Set $T_3$ consists of $\mu - 1$ dummy candidates, $d_{i,j}^{T_3} \in T_3$, $\forall i \in [1, m\mu - 3m], j \in [1, \mu - 1]$.

- Block type $B_2$ consists of $2mn$ blocks and each block consists of one set of candidates:
  - Set $T_4$ consists of $\mu + 1$ dummy candidates, $d_{i,j}^{T_4} \in T_4$, $\forall i \in [1, 2mn], j \in [1, \mu + 1]$.

Hence, there are $(m\mu - 3m) \cdot (1 + \mu - 1 + \mu - 1)$ dummy candidates in blocks of type $B_1$ and $(2mn) \cdot (\mu + 1)$ dummy candidates in blocks of type $B_2$. This results in a total of $2\mu^2m - 7\mu m + 3m$ dummy candidates of type $B_1$ and $2\mu mn + 2mn$ dummy candidates of type $B_2$. Thus, $|D| = 2\mu^2m - 7\mu m + 2\mu mn + 2mn + 3m$. Note that the types of blocks and sets discussed above are different and independent from the candidate groups (discussed later) that are used to enforce diversity constraints.

Overall, we set $A = \{c_1, \ldots, c_m\}$ and the dummy candidate set $D = \{d_1, \ldots, d_{2\mu^2m-7\mu m+2\mu mn+2mn+3m}\}$. Hence, the candidate set $C = A \cup D$ is of size $|C| = 2\mu^2m - 7\mu m + 2\mu mn + 2mn + 4m$ candidates.

**Committee Size:** We set the target committee size to be $k + m\mu^2 + 2mny - 3m$. 

**Candidate Groups:** We now divide the candidates in $C$ into groups such that each candidate is part of $\mu$ groups as there are $\mu$ candidate attributes.

**Candidates in $A$:** Each edge $e \in E$ that connects vertices $x_i$ and $x_j$ corresponds to a candidate group $G \in \mathcal{G}$ that contains two candidates $c_i$ and $c_j$. As our reduction proceeds from a 3-regular graph, each vertex is connected to three edges. This corresponds to each candidate $c \in A$ having three attributes and thus, belonging to three groups.

Additionally, each candidate $c \in A$ is part of $\mu - 3$ groups where each group is with the one candidate from Set $T_1$ of block type $B_1$. Specifically, candidate $c_i$ forms a group each with $d_{i,j}^{T_1} \in T_1 : j \in [1 + (i - 1)(\mu - 3), i(\mu - 3)]$. Hence, as each one of the $m$ candidates form $\mu - 3$ groups, we have a total of $m(\mu - 3) = m\mu - 3m$ blocks of type $B_1$ consisting of $m(\mu - 3)$ candidates in Set $T_1$. Overall, each candidate $c \in A$ has $\mu$ attributes and is part of $\mu$ groups.

**Candidates of Block type $B_2$:** Each candidate in the block type $B_1$ has $\mu$ attributes and are grouped as follows:

- Each dummy candidate $d_{i,j}^{T_1} \in T_1 : j \in [1 + (i - 1)(\mu - 3), i(\mu - 3)]$ is in the same group as candidate $c_i \in A$. Additionally, it is also in $\mu - 1$ groups, individually with each of $\mu - 1$ dummy candidates, $d_{i,j}^{T_2}, \forall i \in [1, \mu - 1]$, $\mu - 1$ attributes and is part of $\mu$ groups.

- For each dummy candidate $d_{i,j}^{T_2} \in T_2 : j \in [1 + (i - 1)(\mu - 3), i(\mu - 3)]$ and $\forall i \in [1, \mu - 1]$, it is in the same group as $d_{i,j}^{T_1}$ as described in the previous point. It is also in $\mu - 1$ groups, individually with each of $\mu - 1$ dummy candidates, $d_{i,j}^{T_3}, \forall i \in [1, \mu - 1]$. Thus, each dummy candidate $d_{i,j}^{T_3} \in T_3$ has $\mu$ attributes and is part of $\mu$ groups.

- For each dummy candidate $d_{i,j}^{T_3} \in T_3 : j \in [1 + (i - 1)(\mu - 3), i(\mu - 3)]$ and $\forall i \in [1, \mu - 1]$, it is in $\mu - 1$ groups, individually with each of $\mu - 1$ dummy candidates, $d_{i,j}^{T_2}, \forall i \in [1, \mu - 1]$, as described in the previous point. Next, note that when $\mu$ is an odd number, $\mu - 1$ is an even number, which means Set $T_3$ of each block has an even number of candidates. We randomly divide $\mu - 1$ candidates into two partitions. Then, we create $\frac{\mu - 1}{2}$ groups over one attribute where each group contains two candidates from Set $T_3$ such that one candidate is selected from each of the two partitions without replacement. Thus, each pair of groups is mutually disjoint. Thus, each dummy candidate $d_{i,j}^{T_3} \in T_3$
is part of exactly one group that is shared with exactly one another dummy candidate $d^T_{j,i} \in T_3$ where $j \neq j'$.

Overall, this construction results in one attribute and one group for each dummy candidate $d^T_{j,i} \in T_3$. Hence, each dummy candidate $d^T_{j,i} \in T_3$ has $\mu$ attributes and is part of $\mu$ groups.

Candidates of Block type $B_2$: Finally, we assign candidates from the Block type $B_2$ to groups. Each of the dummy candidate in Set $T_4$ of each of the $2mn$ blocks is grouped individually with each of the remaining $\mu$ dummy candidates in Set $T_4$ of a block. Formally, $\forall i \in [1, 2mn], \forall j \in [1, \mu + 1], d^T_{j,o} \in T_4$ and $\forall \alpha \in [1, \mu + 1], d^T_{j,\alpha} \in T_4$ such that $j \neq \alpha$, $d^T_{j,o}$ and $d^T_{j,\alpha}$ are in the same group. Hence, each block consists of $\mu + 1$ candidates and each candidate is grouped pairwise with each of the remaining $\mu$ candidates. This means that each dummy candidate $d^T_{j,i} \in T_4$ has $\mu$ attributes and is part of $\mu$ groups.

Diversity Constraints: We set the lower bound for each candidate group as follows:

- $i^D_1 = 1$ for all $G \in G : G \cap A \neq \phi$, which corresponds that each vertex in the vertex cover should be covered by some chosen edge.
- $i^D_2 = 1$ for all $G \in G$ such that at least one of the following holds:
  - $\forall i \in [1, \mu m - 3m], G \cap d^T_{i,1} \neq \phi$
  - $\forall i \in [1, \mu m - 3m], \forall j \in [1, \mu - 1], G \cap d^T_{i,j} \neq \phi$
  - $\forall i \in [1, \mu m - 3m], \forall j \in [1, \mu - 1], G \cap d^T_{i,j} \neq \phi$
- $i^D_3 = 2$ for all $G \in G$ such that $\forall i \in [1, 2mn], \forall j \in [1, \mu + 1], G \cap d^T_{i,j} \neq \phi$ and $\forall i \in [1, 2mn], G \cap d^T_{i,\mu+1} = \phi$.
- $i^D_4 = 1$ for all $G \in G$ such that $\forall i \in [1, 2mn], \forall j \in [1, \mu + 1], G \cap d^T_{i,j} \neq \phi$.

In summary, $\forall G \in G$.

Voters and Preferences: We now introduce $2n^2$ voters, $2n$ voters for each edge $e \in E$. More specifically, an edge $e \in E$ connects vertices $x_i$ and $x_j$. Then, the corresponding $2n^2$ voters $v \in V$ rank the candidates as follows:

- **First 2 positions** - Set $U_1$ : The top two positions are occupied by candidates $c_i$ and $c_j$ that correspond to vertices $x_i$ and $x_j$. For voter $v^b_a$ where $a \in [2n]$ and $b \in [n]$, we denote the candidates $c_i$ and $c_j$ as $c_{ia}$ and $c_{ja}$. Note that we have two voters that correspond to each edge and hence, two voters $v^b_a$ and $v^b_{a-1}$ rank the same two candidates in the top two positions. Specifically, voter $v^b_a$ ranks the two candidates based on the non-increasing order of their indices. Voter $v^b_{a-1}$ ranks the two candidates based on the non-decreasing order of their indices.
- **Next $\mu \mu^2 - 3m \mu$ positions** - Set $U_2$ : All of the $\mu \mu^2 - 3m \mu$ dummy candidates from Set $T_1$ and all of the $2m \mu + 3m \mu$ dummy candidates from Set $T_3$ are ranked based on non-decreasing order of their indices. Note that $m \mu^2 - 3m \mu - 3m = m \mu - 3m$
- **Next $2m \mu$ positions** - Set $U_3$ : Out of the $2m \mu + 2mn$ dummy candidates from Set $T_4$ are ranked based on non-decreasing order of their indices. Specifically, a dummy candidate $d^T_{i,j}$ from Set $T_4$ is selected $\forall \alpha \in [1, 2mn] \text{ and } \forall j \in [1, \mu]$. Note that dummy candidate from Set $T_4$ of the kind $d^T_{i,\mu+1}$, for all $\forall \alpha \in [1, 2mn]$, is not ranked in these positions.
- **Next m positions** - Set $U_4$ : m out of $2mn$ unranked dummy candidates from Set $T_4$ of the kind $d^T_{i,o,\mu+1}$, for all $\forall \alpha \in [1, 2mn]$, are ranked in the next $m$ positions based on non-decreasing order of their indices. Specifically, the $m$ candidates that are ranked satisfy $(o \mod 2n) + 1 = a$. Note that for each type of voter of the kind $v_i$, these $m$ candidates are distinct as shown below. Hence, for all pairs of voters of the kind $v_i, v_j \in V$ : $v_i \neq v_j$, we know that $U_4^v \cap U_4^{v_j} = \phi$.

---

Table 1: An instance of preferences of voters created in the reduction for the proof of Theorem

| $c_{i_1} > c_{j_1}$ | $U_2 > U_3 > U_4^1 > U_5^2 > U_6^2 > U_7$ |
|------------------|-------------------------------------------|
| $c_{i_2} > c_{j_2}$ | $U_2 > U_3 > U_4^2 > U_5^3 > U_6^3 > U_7$ |
| $c_{i_3} > c_{j_3}$ | $U_2 > U_3 > U_4^3 > U_5^4 > U_6^4 > U_7$ |
| $\vdots$ | $\vdots$ |
| $c_{i_n} > c_{j_n}$ | $U_2 > U_3 > U_4^{2n-1} > U_5^{2n-1} > U_6^{2n-1} > U_7$ |

---

[^1]: This setup of Block type $B_2$ is analogous to creating $2mn$ $K_{\mu+1}$ graphs, i.e., a total of $2mn$ complete ($2$-uniform, $\mu$-regular) (hyper-)graphs, each with $\mu + 1$ vertices.
• **Next** \( m - 2 \) **positions** - Set \( U_5 \): The \( m - 2 \) candidates from \( A \), which are not ranked in the top two positions, are ranked based on non-decreasing order of their indices. Formally, \( U_5 = A \setminus \{c_{i,1}, c_{i,2}\} \).

• **Next** \( 2mn - m \) **positions** - Set \( U_6 \): The \( 2mn - m \) out of \( 2mn \) unranked dummy candidates from Set \( T_4 \) of the kind \( d_{o,a+1}^{T_4} \), for all \( o \in [1, 2mn] \), are ranked based on non-decreasing order of their indices. Specifically, the candidates that are ranked satisfy \( o \mod 2n) + 1 \neq a \).

• **Next** \( m\mu^2 - 4m\mu + 3m \) **positions** - Set \( U_7 \): All of the \( m\mu^2 - 4m\mu + 3m \) dummy candidates from Set \( T_2 \) are ranked based on non-decreasing order of their indices. These dummy candidates are part of a population \( \pi \) for all \( \pi \in \pi \).

More specifically, the voters rank the candidates as shown in Table II. The sets without a superscript (e.g., \( U_2 \)) denote the candidate rankings that are the same for all voters.

**Voter Populations:** We now divide the voters in \( V \) into populations such that each \( \pi \) voter is part of \( \pi \) populations as there are \( \pi \) voter attributes. Specifically, the voters are divided into disjoint population over one or more attributes whenever \( \pi \in \{1, 2, \ldots, \pi \} \). The voters are divided into populations as follows: \( \forall x \in [\pi], \forall y \in [2n], \forall z \in [n], \forall v \in P \) is a part of a population \( P \in \pi \) such that \( P \) contains all voters with the same \( (z \mod x) \) and \( y \). Each voter is part of \( \pi \) populations.

**Representation Constraints:** We set the lower bound for each voter population as follows: \( \forall P \in \pi, l_P = 1 + m\mu^2 - 3m\mu + 2m\mu \).

This completes our construction for the reduction, which is a polynomial time reduction in the size of \( n \) and \( m \).

**PART II: Proof of Correctness**

**Claim 1.** We have a vertex cover \( S \) of size at most \( k \) that satisfies \( e \cap S \neq \emptyset \) for every \( e \in E \) if and only if we have a committee \( W \) of size at most \( k + m\mu^2 + 2m\mu - 3m\mu \) such that \( \forall G \in G \), \( |G \cap W| \geq l_G^D \) for all \( G \in G \), and \( f(W) = \max_{W \in \pi} f(W) \) where \( W \) is a set of committees that satisfy all constraints.

\( \Rightarrow \) If the instance of the VC problem is a yes instance, then the corresponding instance of DRCWD is a yes instance.

**Diversity constraints are satisfied:** Firstly, each and every candidate group will have at least one of their members in the winning committee \( W \), i.e., \( |G \cap W| \geq l_G^D \) for all \( G \in G \).

More specifically, for each of the \( m\mu - 3m \) blocks of type \( B_1 \) of candidates, we select:

• one dummy candidate from Set \( T_1 \)

• all \( \mu - 1 \) dummy candidates from Set \( T_3 \)

This helps to satisfy the condition \( l_G^D = 1 \) for all \( G \in G \) such that at least one of the following holds:

• \( \forall i \in [1, m\mu - 3m], G \cap d_{i,1}^{T_3} \neq \emptyset \).

• \( \forall i \in [1, m\mu - 3m], \forall j \in [1, \mu - 1], G \cap d_{i,j}^{T_3} \neq \emptyset \).

• \( \forall i \in [1, m\mu - 3m], \forall j \in [1, \mu - 1], G \cap d_{i,j}^{T_3} \neq \emptyset \).

Thus, we select \( m \) candidates from \( \mu - 3 \) blocks for each of the \( m \) candidates that correspond to vertices in the vertex cover. This results in \( (\mu \cdot (\mu - 3) \cdot m) = m\mu^2 - 3m\mu \) candidates in the committee.

Next, for each of the \( 2mn \) blocks of type \( B_2 \) of candidates, we select:

• \( m \) dummy candidates \( d_{i,j}^{T_4} \) from Set \( T_4 \) such that \( i \in [1, 2mn] \) and \( j \in [1, \mu] \).

This helps to satisfy the conditions: \( l_G^D = 2 \) for all \( G \in G \) such that \( \forall i \in [1, 2mn], \forall j \in [1, \mu + 1], G \cap d_{i,j}^{T_4} \neq \emptyset \) and \( \forall i \in [1, 2mn], G \cap d_{i,j+1}^{T_4} = \emptyset \) and \( l_G^D = 1 \) for all \( G \in G \) such that \( \forall i \in [1, 2mn], \forall j \in [1, \mu + 1], G \cap d_{i,j}^{T_3} \neq \emptyset \). Overall, we select \( m \) candidates from \( 2mn \) blocks. This results in additional \( 2m\mu \) candidates in the committee.

Finally, for groups that do not contain any dummy candidates, select \( k \) candidates \( c \in A \) that correspond to \( k \) vertices \( x \in X \) that form the vertex cover. These candidates satisfy the remainder of the constraints. Specifically, these \( k \) candidates satisfy \( |G \cap W| \geq 1 \) for all the candidate groups that do not contain any dummy candidates. Hence, we have a committee of size \( k + m\mu^2 - 2m\mu - 3m\mu \).

**Representation constraints are satisfied:** Next, if the instance of the VC problem is a yes instance, then we have a winning committee \( W \) of size \( k + m\mu^2 + 2m\mu - 3m\mu \) that consists of \( k \) candidates corresponding to the VC and \( m\mu^2 + 2m\mu - 3m\mu \) candidates from Sets \( U_2 \) and \( U_3 \). Also, each and every population’s winning committee, \( W_P \) for all \( P \in \pi \), will have at least \( 1 + m\mu^2 - 3m\mu + 2m\mu \) of their members in the winning committee \( W \) such that \( |W_P \cap W| \geq 1 + m\mu^2 - 3m\mu + 2m\mu \) for all \( P \in \pi \), because:

• as we have a yes instance of the VC problem, one of the two corresponding candidates occupying the first two positions of the ranking will be on the committee.

• each of the \( m\mu^2 - 3m\mu \) candidates from Set \( U_2 \) will be on the committee.

• each of the \( 2m\mu \) candidates from Set \( U_3 \) will be on the committee.

By construction, candidates in Set \( U_2 \) and candidates in Set \( U_3 \) will always be part of each population’s winning committee. Additionally, candidates in Set \( U_2 \) are the \( m \) candidates selected from each of the \( m\mu - 3m \) blocks of the type \( B_1 \) and these candidates are the same across all voter populations. Also, the candidates in Set \( U_3 \) are the \( m \) candidates selected from each of the \( 2mn \) blocks of the type \( B_2 \) and these candidates are the same across all voter populations. Thus, \( |W_P \cap W| \geq 1 + m\mu^2 - 3m\mu + 2m\mu \), for
all \( P \in \mathcal{P} \), and the same winning committee \( W \) satisfy the diversity constraints and the representation constraints.

**Highest scoring committee:** It remains to be shown that \( W \) is the highest scoring committee among all the committees that satisfy the given constraints.

Note that for a given population \( P \in \mathcal{P} \), \( \forall c \in C : \text{pos}_P(c) = 1 \) or \( \text{pos}_P(c) = 2 \), \( \forall d_{i,j} \in S, \forall v \in V \), \( \text{pos}_c(c) \geq \text{pos}_v(d_{i,j}) \). This holds based on the prerequisite condition that we are interested in committees that satisfy the constraints. Additionally, \( \forall S \cup \forall U \cup \forall S \cup \forall U \)

\( \forall d_{i,j} \in S \cup T_4 \), \( \forall d_{i,j} \in S \cup T_3 \), each holds because \( d_{i,j} \) is either in Set \( S_4 \) or Set \( S_3 \), and even after accounting for varying preferences of these two sets, Sets \( S_2 \) and \( S_3 \) are always ranked higher. Thus, the contribution of each candidate in Sets \( S_2 \) and \( S_3 \) to the total score will be greater than equal to a candidate \( d_{i,j} \). As noted earlier, the ties are broken based on ranking. Hence, \( W \) will be the highest scoring committee.

Overall, a yes instance of the VC problem implies a yes instance of the DRCWD problem such that the committee \( W \) s of size at most \( k + m \mu^2 + 2mn \mu - 3m\mu \), \( \forall \mu \geq 3 \) and \( \mu \) is an even number, \( \forall \pi \in \mathcal{P} : \pi \geq 1 \), and \( \pi \) is the monotone, separable function arising from an arbitrary single-winner positional scoring rule, then DRCWD is NP-hard.

**Proof Sketch.** The proof follows from the proof of Theorem 3. Hence, we only discuss the major changes in the construction of the proof.

**PART I: Construction**

**Candidates:** We have two candidates \( c_i \) and \( c_{m+i} \) for each \( i \in \mathcal{P} \), and \( 2 \cdot (2m^2 + 2mn + 2mn + 3m) \) dummy candidates \( d \) in \( D \) where \( m \) corresponds to the number of vertices in the graph \( H \), \( n \) corresponds to the number of edges in the graph \( H \), and \( \mu \) is a positive, even integer (hint: the number of candidate attributes). We divide the dummy candidates into two types of blocks in line with the proof of Theorem 3, but there twice the number of blocks and hence, by transitivity, twice the number of candidates.

**Candidate Groups:** We divide the candidates in \( C \) into groups such that each candidate is part of \( \mu \) groups as there are \( \mu \) candidate attributes. The division is in line with the proof of Theorem 3, but each set contains twice the number of candidates and there is one exception. For candidates in Block Type \( B_2 \):

- For each dummy candidate \( d_{i,j} \in T_3 \), it is in \( \mu - 1 \) groups, individually with each of \( \mu - 1 \) dummy candidates, \( d_{i,j} \in T_2 \), as described in the previous point. Next, note that when \( \mu \) is an even number, \( \mu - 1 \) is an odd number, which means Set \( T_3 \) has an odd number of candidates. We randomly divide \( \mu - 2 \) candidates into two partitions. Then, we create \( \frac{\mu - 2}{2} \) groups over one attribute where each group contains two candidates from Set \( T_3 \) such that one candidate is selected from each of the two partitions without replacement. Thus, each pair of groups is mutually disjoint. Hence, each dummy candidate \( d_{i,j} \in T_3 \) is part of exactly one group that is shared with exactly one another dummy candidate \( d_{i,j} \in T_3 \) where \( j \neq j' \). Overall, this construction results in one attribute and one group for all but one dummy candidate \( d_{i,j} \in T_3 \), which results in a total of \( \mu \) attributes and \( \mu \) groups for these \( \mu - 2 \) candidates. This is because \( \frac{\mu - 2}{2} \) groups can hold \( \mu - 2 \) candidates. Hence, one candidate still has \( \mu - 1 \) attributes and is part of \( \mu - 1 \) groups. If this block of dummy candidates is for candidate \( c_i \in A \), then another corresponding block of dummy candidates for candidate \( c_{m+i} \in A \) will also have one candidate \( d_{i,j} \in T_3 \) which will have \( \mu - 1 \) attributes and is part of \( \mu - 1 \) groups. We group these two candidates from separate blocks. Hence, now that one remaining candidate also has \( \mu \) attributes and is part of \( \mu \) groups. As there is always an even number of candidates in set \( A (|A| = 2m) \), representation constraints are set to zero (\( \forall P \in \mathcal{P}, l_P^P = 0 \)).
PART II: Proof of Correctness

Diversity Constraints: We set the lower bound for each candidate group in line with the proof of Theorem 3.

Voters and Preferences: We now introduce \(4n^2\) voters, \(4n\) voters for each edge \(e \in E\) to connect vertices \(x_i\) and \(x_j\). Then, the corresponding \(4n^2\) voters \(v \in V\) rank the candidates in line with the proof of Theorem 3 but each set now consists of twice the number of candidates. Specifically, the major change is in \(U_1\) as it now consists of \(c_i, c_j, c_{i+m}, \) and \(c_{j+m}\).

Voter Populations: We divide the voters into populations in line with the proof of Theorem 3.

Representation Constraints: We set the lower bound for each voter population as follows: \(\forall P \in \mathcal{P}, l_P^R = 2 \cdot (1 + n\mu^2 - 3n\mu + 2m\mu^2)\).

This completes our construction for the reduction, which is a polynomial time reduction in the size of \(n\) and \(m\) in line with the proof of Theorem 3.

PART II: Proof of Correctness

Claim 2. We have a vertex cover \(S\) of size at most \(k\) that satisfies \(e \cap S \neq \emptyset \forall e \in E\) if and only if we have a committee \(W\) of size at most \(2 \cdot (k + m\mu^2 + 2m\mu - 3m\mu)\) such that \(\forall G \in \mathcal{G}, |G \cap W| \geq l_P^R, \forall P \in \mathcal{P}, |W_P \cap W| \geq l_P^R, \) and \(f(W) = \max_{W' \in W} f(W')\) where \(W\) is a set of committees that satisfy all constraints.

The proof of correctness easily follows from the proof of correctness of Theorem 3.

In line with the previous theorem, our reduction holds for a weaker result such that the hardness holds even when either all diversity constraints are set to zero (\(\forall G \in \mathcal{G}, l_P^D = 0\)) or all representation constraints are set to zero (\(\forall P \in \mathcal{P}, l_P^R = 0\)).

5 Conclusion
We classified the complexity of finding a diverse and representative committee using a monotone, separable positional multiwinner voting rule. In doing so, we established the close association between DRCWD and the vertex problem. This can lead to interesting future work about having an understanding of the existence of a diverse and representative outcome and complexity of finding one. Specifically, we can study if DRCWD under certain realistic assumptions is PPAD-complete or not based on knowledge about the vertex cover problem and its possible PPAD-completeness. Additionally, this association can be used to find realistic settings where the model becomes tractable and validate it by giving examples of real-world datasets where such algorithms may work.

Acknowledgments
I am grateful to Julia Stoyanovich for her insights. I thank Markus Brill, Benny Kimelfeld, and Phokion G. Kolaitis for many interesting and thought-provoking questions.

References
Alimonti, P.; and Kann, V. 1997. Hardness of approximating problems on cubic graphs. In Italian Conference on Algorithms and Complexity, 288–298. Springer.

Bansal, N.; and Khot, S. 2010. Inapproximability of hypergraph vertex cover and applications to scheduling problems. In International Colloquium on Automata, Languages, and Programming, 250–261. Springer.

Baumeister, D.; and Rothe, J. 2012. Taking the final step in line with the proof of Theorem 3 but each set now consists of twice the number of candidates. Specifically, the major change is in \(U_1\) as it now consists of \(c_i, c_j, c_{i+m}, \) and \(c_{j+m}\).

Betzler, N.; and Dorn, B. 2010. Towards a dichotomy for the possible winner problem in elections based on scoring rules. Information Processing Letters, 112(5): 186–190.

Bredereck, R.; Faliszewski, P.; Igarashi, A.; Lackner, M.; and Skowron, P. 2018. Multiwinner elections with diversity constraints. In AAAI.

Celis, L. E.; Huang, L.; and Vishnoi, N. K. 2018. Multiwinner voting with fairness constraints. In IJCAI.

Chakraborty, V.; and Kolaitis, P. G. 2021. Classifying the Complexity of the Possible Winner Problem on Partial Chains. In Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems, 297–305.

Cheng, Y.; Jiang, Z.; Munagala, K.; and Wang, K. 2019. Group fairness in committee selection. In Proceedings of the 2019 ACM Conference on Economics and Computation, 263–279.

Feige, U. 2003. Vertex cover is hardest to approximate on regular graphs. Technical Report MCS03–15.

Garey, M. R.; and Johnson, D. S. 1979. Computers and intractability, volume 174. Freeman San Francisco.

Hershkowitz, D. E.; Kahng, A.; Peters, D.; and Procaccia, A. D. 2021. District-fair participatory budgeting. arXiv preprint arXiv:2102.06115.

Kenig, B. 2019. The complexity of the possible winner problem with partitioned preferences. In Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, 2051–2053.

Király, T.; and Pap, J. 2013. PPAD-completeness of polyhedral versions of Sperner’s Lemma. Discrete Mathematics, 313(15): 1594–1599.

Konczak, K.; and Lang, J. 2005. Voting procedures with incomplete preferences. In Proc. IJCAI-05 Multidisciplinary Workshop on Advances in Preference Handling, volume 20. Citeseer.

Potthoff, R. 1990. Use of linear programming for constrained approval voting. Interfaces, 20(5): 79–80.
Relia, K. 2022. DiRe Committee: Diversity and Representation Constraints in Multiwinner Elections. In IJCAI-22.

Shrestha, Y. R.; and Yang, Y. 2019. Fairness in algorithmic decision-making: Applications in multi-winner voting, machine learning, and recommender systems. Algorithms, 12(9): 199.

Uckelman, J.; Chevaleyre, Y.; Endriss, U.; and Lang, J. 2009. Representing utility functions via weighted goals. Mathematical Logic Quarterly, 55(4): 341–361.

Xia, L.; and Conitzer, V. 2011. Determining possible and necessary winners given partial orders. Journal of Artificial Intelligence Research, 41: 25–67.