Kernelization Complexity of Possible Winner and Coalitional Manipulation Problems in Voting

Palash Dey[a], Neeldhara Misra[b], Y. Narahari[a]

[a] *Department of Computer Science and Automation*
Indian Institute of Science - Bangalore, India.
[b] *Department of Computer Science and Engineering*
Indian Institute of Technology - Gandhinagar, India.

Abstract

In the Possible winner problem in computational social choice theory, we are given a set of partial preferences and the question is whether a distinguished candidate could be made winner by extending the partial preferences to linear preferences. Previous work has provided, for many common voting rules, fixed parameter tractable algorithms for the Possible winner problem, with number of candidates as the parameter. However, the corresponding kernelization question is still open and in fact, has been mentioned as a key research challenge [10]. In this paper, we settle this open question for many common voting rules.

We show that the Possible winner problem for maximin, Copeland, Bucklin, ranked pairs, and a class of scoring rules that includes the Borda voting rule do not admit a polynomial kernel with the number of candidates as the parameter. We show however that the Coalitional manipulation problem which is an important special case of the Possible winner problem does admit a polynomial kernel for maximin, Copeland, ranked pairs, and a class of scoring rules that includes the Borda voting rule, when the number of manipulators is polynomial in the number of candidates. A significant conclusion of our work is that the Possible winner problem is harder than the Coalitional manipulation problem since the Coalitional manipulation problem admits a polynomial kernel whereas the Possible winner problem does not admit a polynomial kernel.

Keywords: Computational social choice; possible winner; voting; kernelization; parameterized complexity.

1. Introduction

In many real life situations including multiagent systems, agents often need to aggregate their preferences and agree upon a common decision (candidate). Voting is an immediate natural tool in these situations. Common and classical applications of voting rules in artificial

Email addresses: palash@csa.iisc.ernet.in (Palash Dey), mail@neeldhara.com (Neeldhara Misra), hari@csa.iisc.ernet.in (Y. Narahari)

intelligence include collaborative filtering [33], planning among multiple automated agents [20], etc.

Usually, in a voting setting, it is assumed that the votes are complete orders over the candidates. However, due to many reasons, for example, lack of knowledge of voters about some candidates, a voter maybe indifferent between some pairs of candidates. Hence, it is both natural and important to consider scenarios where votes are partial orders over the candidates. When votes are only partial orders over the candidates, the winner cannot be determined with certainty since it depends on how these partial orders are extended to linear orders. This leads to a natural computational problem called the Possible winner [27] problem: given a set of partial votes P and a distinguished candidate c, is there a way to extend the partial votes to linear ones to make c win? The Possible winner problem has been studied extensively in the literature [29, 34, 35, 37, 8, 6, 14, 5, 2, 28, 23] following its definition in [27]. The Possible winner problem is known to be NP-complete for many common voting rules, for example, scoring rules, maximin, Copeland, Bucklin, and ranked pairs etc. [37]. Walsh [35] showed, for a constant number of candidates, that the Possible winner problem can be solved in polynomial time for all the voting rules mentioned above. An important special case of the Possible winner problem is the Coalitional manipulation problem [1] where only two kinds of partial votes are allowed - complete preference and empty preference. The set of empty votes is called the manipulators' vote and is denoted by M. The Coalitional manipulation problem is NP-complete for maximin, Copeland, and ranked pairs voting rules even when $|M| \geq 2$ [21, 22, 38]. The Coalitional manipulation problem is in P for the Bucklin voting rule [38]. We refer to [37, 35, 38] for detailed overviews.

1.1. *Our Methodology*

Preprocessing, as a strategy for coping with hard problems, is universally applied in practice. The main goal here is *instance compression* - the objective is to output a smaller instance while maintaining equivalence. In the classical setting, NP-hard problems are unlikely to have efficient compression algorithms (since repeated application would lead to a polynomial time algorithm for the problem, which would imply P = NP). However, the breakthrough notion of *kernelization* in parameterized complexity provides a mathematical framework for analyzing the quality of preprocessing strategies. In parameterized complexity, each problem instance (x, k) comes with a parameter k. The parameterized problem is said to admit a *kernel* if there is a polynomial time algorithm (where the degree of polynomial is independent of k), called a *kernelization algorithm*, that reduces the input instance to an instance with size bounded by a function of k, while preserving the answer. This has turned out to be an important and widely applied notion in theory, and has also proven very successful in practice [36, 30]. Quantitatively, running a kernelization algorithm before solving it using an algorithm that runs in time $f(|x|)$ brings down the running time to $f(k)+p(|x|)$, where $|x|$ is the size of the input instance and the running time of the kernelization algorithm is $p(|x|)$.

A problem with parameter k is called *fixed parameter tractable* (FPT) if it is solvable in time $f(k) \cdot p(|x|)$, where f is an arbitrary function of k and p is a polynomial in the input size $|x|$. The existence of a fixed parameter tractable algorithm implies existence of a kernel for that problem. However, the size of the kernel need not be polynomial in the parameter. A polynomial kernel is said to exist if there is a kernelization algorithm that can output an equivalent problem instance of size polynomial in the parameter. We refer to [19, 32] for an excellent overview on fixed parameter algorithms and kernelization.

1.2. Contributions

Discovering kernelization algorithms is currently an active and interesting area of research in computational social choice theory [5, 3, 7, 12, 25, 11, 4, 18]. Betzler et al. [8] showed that the Possible winner problem admits fixed parameter tractable algorithm when parameterized by the total number of candidates for scoring rules, maximin, Copeland, Bucklin, and ranked pairs voting rules. Yang et al. [40, 39] provides efficient fixed parameter tractable algorithms for the Coalitional manipulation problem for the Borda, maximin, and Copeland voting rules. A natural and practical follow-up question is whether the Possible winner and Coalitional manipulation problems admit a polynomial kernel when parameterized by the number of candidates. This question has been open ever since the work of Betzler et al. and in fact, has been mentioned as a key research challenge in parameterized algorithms for computational social choice theory [10]. Betzler et al. showed non-existence of polynomial kernel for the Possible winner problem for the k-approval voting rule when parameterized by (t, k), where t is the number of partial votes [3]. The NP-complete reductions for the Possible winner problem for scoring rules, maximin, Copeland, Bucklin, and ranked pairs voting rules given by Xia et al. [38] are from the Exact 3 set cover problem. Their results do not throw any light on the existence of a polynomial kernel since Exact 3 set cover has a trivial $O(m^3)$ kernel where m is the size of the universe. In our work in this paper, we show that there is no polynomial kernel (unless $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$) for the Possible winner problem, when parameterized by the total number of candidates, with respect to maximin, Copeland, Bucklin, and ranked pairs voting rules, and a class of scoring rules that includes the Borda voting rule. These hardness results are shown by a parameter-preserving many-to-one reduction from the Small universe set cover problem for which there does not exist any polynomial kernel parameterized by universe size unless $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$ [17].

On the other hand, we show that the Coalitional manipulation problem admits a polynomial kernel for maximin, Copeland, and ranked pairs voting rules, and a class of scoring rules that includes the Borda voting rule when we have $poly(m)$ number of manipulators – specifically, we exhibit an $O(m^2|M|)$ kernel for maximin and Copeland voting rules, and an $O(m^4|M|)$ kernel for the ranked pairs voting rule, where m is the number of candidates and M is the set of manipulators. The Coalitional manipulation problem for the Bucklin voting rule is in P [38] and thus the kernelization question does not arise.

A significant conclusion of our work is that, although the Possible winner and Coalitional manipulation problems are both NP-complete for many voting rules, the Possible winner problem is harder than the Coalitional manipulation problem since the Coalitional manipulation problem admits a polynomial kernel whereas the Possible winner problem does not admit a polynomial kernel.

The rest of the paper is organized as follows. We first establish the setup and general notions in Section 2. Next, we discuss non-existence of polynomial kernels for the Possible winner problem in Section 3 followed by existence of polynomial kernels for the Coalitional manipulation problem in Section 4. A preliminary version of this work appeared at [16].

2. Preliminaries

Let $\mathcal{V} = \{v_1, \dots, v_n\}$ be the set of all *voters* and $\mathcal{C} = \{c_1, \dots, c_m\}$ the set of all *candidates*. From here on, n denotes the number of voters and m denotes the number of candidates unless mentioned otherwise. We denote the set $\{1, 2, \dots\}$ by $\mathbb{N}^+$ and the set $\{1, \cdots, k\}$ by $[k]$ for any positive integer k. The *vote* of voter v_i for each $i \in [n]$, is a *preference* $\succ_i$ over the candidates which is a linear order over $\mathcal{C}$. For example, for two candidates a and b, $a \succ_i b$ means that the voter v_i prefers a to b. We will use $a >_i b$ to denote the fact that $a \succ_i b$ and $a \neq b$. We denote the set of all linear orders over $\mathcal{C}$ by $\mathcal{L}(\mathcal{C})$. The set of all preference profiles $(\succ_1, \dots, \succ_n)$ of the n voters is denoted by $\mathcal{L}(\mathcal{C})^n$. Let $\uplus$ denote the disjoint union of sets. A map $r : \uplus_{n,|\mathcal{C}|\in\mathbb{N}^+}\mathcal{L}(\mathcal{C})^n \longrightarrow 2^{\mathcal{C}} \setminus \{\emptyset\}$ is called a *voting rule*. For a voting rule r and a preference profile $\succ = (\succ_1, \dots, \succ_n)$, we say a candidate x wins uniquely if $r(\succ) = \{x\}$. In this paper, we say that a candidate wins an election to mean that the candidate is the unique winner of the election unless mentioned otherwise. A more general setting is an *election* where the votes are only *partial orders* over candidates. A *partial order* is a relation that is *reflexive, antisymmetric*, and *transitive*. A partial vote can be extended to possibly more than one linear votes depending on how we fix the order for the unspecified pairs of candidates. For example, in an election with the set of candidates $\mathcal{C} = \{a, b, c\}$, a valid partial vote can be $a \succ b$. This partial vote can be extended to three linear votes namely, $a \succ b \succ c$, $a \succ c \succ b$, $c \succ a \succ b$. However, the voting rules always take as input a set of votes that are complete orders over the candidates. Given an election E, we can construct a weighted graph G_E called *weighted majority graph* from E. The set of vertices in G_E is the set of candidates in E. For any two candidates x and y, the weight on the edge (x, y) is $D_E(x, y) = N_E(x, y) - N_E(y, x)$, where $N_E(x, y)(N_E(y, x))$ is the number of voters who prefer x to y (y to x). Some examples of common voting rules are as follows.

– **Positional scoring rules:** A collection of m-dimensional vectors $\overrightarrow{s_m} = (\alpha_1, \alpha_2, \dots, \alpha_m) \in \mathbb{R}^m$ with $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_m$ and $\alpha_1 > \alpha_m$ for every $m \in \mathbb{N}$ naturally defines a voting rule — a candidate gets score α_i from a vote if it is placed at the i^{th} position, and the score of a candidate is the sum of the scores it receives from all the votes. The winners are the candidates with maximum score. Scoring rules remain unchanged if we multiply every α_i by any constant $\lambda > 0$ and/or add any constant μ. Hence, we assume without loss of generality that for any score vector $\overrightarrow{s_m}$, there exists a j such that $\alpha_j - \alpha_{j+1} = 1$ and $\alpha_k = 0$ for all $k > j$. We call such a $\overrightarrow{s_m}$ a normalized score vector. If α_i is 1 for $i \in [k]$ and 0 otherwise, then we get the k-approval voting rule. For the k-veto voting rule, α_i is 0 for $i \in [m - k]$ and -1 otherwise. 1-approval is called the plurality voting rule and 1-veto is called the veto voting rule.

– **Bucklin:** A candidate x's Bucklin score is the minimum number ℓ such that more than half of the voters rank x in their top ℓ positions. The winners are the candidates with lowest Bucklin score.

– **Maximin:** The maximin score of a candidate x in an election E is $\min_{y \neq x} D_E(x, y)$. The winners are the candidates with maximum maximin score.

– **Copeland:** The Copeland score of a candidate x in an election E is the number of candidates $y \neq x$ such that $D_E(x, y) > 0$. The winners are the candidates with maximum Copeland score.

– **Ranked pairs:** Given an election E, we pick a pair $(c_i, c_j) \in \mathcal{C} \times \mathcal{C}$ such that $D_E(c_i, c_j)$ is maximum. We fix the ordering between c_i and c_j to be $c_i \succ c_j$ unless it contradicts previously fixed orders. We continue this process until all pairwise elections are considered. At this point, we have a complete order $\succ$ over the candidates. Now the top candidate of $\succ$ is chosen as the winner.

We use the parallel-universes tie breaking [15, 13] to define the winning candidate for the ranked pairs voting rule. In this setting, a candidate c is a winner if and only if there exists a way to break ties in all of the steps such that c is the winner.

The Possible winner problem with respect to r is the following:

Possible Winner $[r]$ **Parameter:** m
Input: A set $\mathcal{C}$ of candidates and votes $\mathcal{V}$, where each vote is a partial order over $\mathcal{C}$, and a candidate $c \in \mathcal{C}$.
Question: Is there a linear extension of $\mathcal{V}$ for which c is the unique winner with respect to r?

An important special case of the Possible winner problem is the Coalitional manipulation problem where every partial vote is either an empty order or a complete order. We call the complete votes the votes of the non-manipulators and the empty votes the votes of the manipulators. Formally the Coalitional manipulation problem is defined as follows.

Coalitional Manipulation $[r]$ **Parameter:** m
Input: A set $\mathcal{C}$ of candidates, a set $\mathcal{V}$ of complete votes, an integer t corresponding to the number of manipulators, and a candidate $c \in \mathcal{C}$.
Question: Does there exist a set of votes $\mathcal{V}'$ of size t such that c is the unique winner with respect to r for the voting profile $(\mathcal{V}, \mathcal{V}')$?

When the voting rule r is clear from the context, we often refer to the problems as Possible winner and Coalitional manipulation without any further qualification.

We note that one might also consider the variant of the problem where the designated candidate c is only required to be a co-winner, instead of being the unique winner. All the results in this work can be easily adapted to this variant as well.

We now briefly describe the framework in which we analyze the computational complexity of Possible winner and Coalitional manipulation problems.

Parameterized Complexity. A parameterized problem Π is a subset of $\Gamma^* \times \mathbb{N}$, where Γ is a finite alphabet. An instance of a parameterized problem is a tuple (x, k), where k is the parameter. We refer the reader to the books [26, 19, 24] for a detailed introduction to this paradigm, and below we state only the definitions that are relevant to our work.

A *kernelization* algorithm is a set of preprocessing rules that runs in polynomial time and reduces the instance size with a guarantee on the output instance size. This notion is formalized below.

Definition 1. [Kernelization] *[32, 24] A kernelization algorithm for a parameterized problem $\Pi \subseteq \Gamma^* \times \mathbb{N}$ is an algorithm that, given $(x, k) \in \Gamma^* \times \mathbb{N}$, outputs, in time polynomial in $|x| + k$, a pair $(x', k') \in \Gamma^* \times \mathbb{N}$ such that (a) $(x, k) \in \Pi$ if and only if $(x', k') \in \Pi$ and (b) $|x'|, k' \leq g(k)$, where g is some computable function. The output instance x' is called the kernel, and the function g is referred to as the size of the kernel. If $g(k) = k^{O(1)}$, then we say that Π admits a polynomial kernel.*

For many parameterized problems, it is well established that the existence of a polynomial kernel would imply the collapse of the polynomial hierarchy to the third level (or more precisely, $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$). Therefore, it is considered unlikely that these problems would admit polynomial-sized kernels. For showing kernel lower bounds, we simply establish reductions from these problems.

Definition 2. [Polynomial Parameter Transformation] [9] *Let Γ_1 and Γ_2 be parameterized problems. We say that Γ_1 is polynomial time and parameter reducible to Γ_2, written $\Gamma_1 \leq_{Ptp} \Gamma_2$, if there exists a polynomial time computable function $f : \Sigma^* \times \mathbb{N} \to \Sigma^* \times \mathbb{N}$, and a polynomial $p : \mathbb{N} \to \mathbb{N}$, and for all $x \in \Sigma^*$ and $k \in \mathbb{N}$, if $f((x, k)) = (x', k')$, then $(x, k) \in \Gamma_1$ if and only if $(x', k') \in \Gamma_2$, and $k' \leq p(k)$. We call f a polynomial parameter transformation (or a PPT) from Γ_1 to Γ_2.*

This notion of a reduction is useful in showing kernel lower bounds because of the following theorem.

Theorem 1. [9, Theorem 3] *Let P and Q be parameterized problems whose derived classical problems are P^c, Q^c, respectively. Let P^c be NP-complete, and $Q^c \in \mathsf{NP}$. Suppose there exists a PPT from P to Q. Then, if Q has a polynomial kernel, then P also has a polynomial kernel.*

3. Kernelization Hardness of Possible Winner

In this section, we show non-existence of polynomial kernels for the Possible Winner problem for the maximin, Copeland, Bucklin, and ranked pairs voting rules, and a class of scoring rules that includes the Borda voting rule. We do this by demonstrating polynomial parameter transformations from the Small universe set cover problem, which is the classic Set Cover problem, but now parameterized by the size of the universe and the budget.

> Small Universe Set Cover **Parameter:** $m + k$
> **Input:** A set $\mathcal{U} = \{u_1, \ldots, u_m\}$ and a family $\mathcal{F} = \{S_1, \ldots, S_t\}$.
> **Question:** Is there a subfamily $\mathcal{H} \subseteq \mathcal{F}$ of size at most k such that every element of the universe belongs to at least one $H \in \mathcal{H}$?

It is well-known [17] that Red-blue dominating set parameterized by k and the number of non-terminals does not admit a polynomial kernel unless $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$. It follows, by the duality between dominating set and set cover, that Set cover when parameterized by the solution size and the size of the universe (in other words, the Small universe set cover problem defined above) does not admit a polynomial kernel unless $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$.

We now consider the Possible winner problem parameterized by the number of candidates for the maximin, Copeland, Bucklin, and ranked pairs voting rules, and a class of scoring rules that includes the Borda rule, and establish that they do not admit a polynomial kernel unless $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$, by polynomial parameter transformations from Small universe set cover.

3.1. Scoring Rules

We begin with proving the hardness of kernelization for the Possible winner problem for a class of scoring rules that includes the Borda voting rule. For that, we use the following lemma which has been used before [2].

Lemma 1. *Let $\mathcal{C} = \{c_1, \ldots, c_m\} \uplus D, (|D| > 0)$ be a set of candidates, and $\vec{\alpha}$ a normalized score vector of length $|\mathcal{C}|$. Then, for any given $\mathbf{X} = (X_1, \ldots, X_m) \in \mathbb{Z}^m$, there exists $\lambda \in \mathbb{R}$ and a voting profile such that the $\vec{\alpha}$-score of c_i is $\lambda + X_i$ for all $1 \leq i \leq m$, and the score of candidates $d \in D$ is less than λ. Moreover, the number of votes is $O(poly(|\mathcal{C}| \cdot \sum_{i=1}^{m} |X_i|))$.*

With the above lemma at hand, we now show hardness of polynomial kernel result for the class of strict scoring rules.

Theorem 2. *The* Possible winner *problem for any strict scoring rule, when parameterized by the number of candidates, does not admit a polynomial kernel unless* $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$.

Proof: Let $(\mathcal{U}, \mathcal{F}, k)$ be an instance of Small universe set cover, where $\mathcal{U} = \{u_1, \ldots, u_m\}$ and $\mathcal{F} = \{S_1, \ldots, S_t\}$. We use T_i to denote $\mathcal{U} \setminus S_i$ for $i \in [t]$. We let $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{2m+3})$ denote the score vector of length t, and let δ_i denote the difference $(\alpha_i - \alpha_{i+1})$ for $i \in [2m+2]$. Note that for a strict scoring rule, all the δ_i's will be strictly positive. We now construct an instance $(\mathcal{C}, V, c)$ of Possible winner as follows.

Candidates. $\mathcal{C} = \mathcal{U} \uplus \mathcal{V} \uplus \{w, c, d\}$, where $\mathcal{V} := \{v_1, \ldots, v_m\}$.

Partial Votes, P. The first part of the voting profile comprises t partial votes, and will be denoted by P. Let V_j denote the set $\{v_1, \ldots, v_j\}$. For each $i \in [t]$, we first consider a profile built on a total order η_i:

$$\eta_i := d \succ S_i \succ V_j \succ w \succ \text{others}, \text{where } j = m - |S_i|.$$

Now we obtain a partial order λ_i based on η_i for every $i \in [t]$ as follows:

$$\lambda_i := \eta_i \setminus (\{w\} \times (\{d\} \uplus S_i \uplus V_j))$$

That is, λ_i is the partial vote where the order between the candidates x and y for $x \neq w$ or $y \notin \{d\} \uplus S_i \uplus V_j$, is same as the order between x and y in η_i. Whereas, if $x = w$ and $y \in \{d\} \uplus S_i \uplus V_j$, then the order between x and y is unspecified. Let P' be the set of votes $\{\eta_i \mid i \in [t]\}$ and P be the set of votes $\{\lambda_i \mid i \in [t]\}$.

Complete Votes, Q. We now add complete votes, which we denote by Q, such that $s(c) = s(u_i), s(d) - s(c) = (k-1)\delta_1, s(c) - s(w) = k(\delta_2 + \delta_3 + \cdots + \delta_{m+1}) + \delta_1, s(c) > s(v_i) + 1$ for all $i \in [t]$, where $s(a)$ is the score of candidate a from the combined voting profile $P' \uplus Q$. From the proof of Lemma 1, we can see that such votes can always be constructed. In particular, also note that the voting profile Q consists of complete votes. Note that the number of candidates is $2m + 3$, which is polynomial in the size of the universe, as desired.

We now claim that the candidate c is a possible winner for the voting profile $P \uplus Q$ with respect to the strict scoring rule $\vec{\alpha}$ if and only if $(\mathcal{U}, \mathcal{F})$ is a YES instance of Set Cover.

In the forward direction, suppose, without loss of generality, that $S_1, \ldots, S_k$ form a set cover. Then we propose the following extension for the partial votes $\lambda_1, \ldots, \lambda_k$:

$$w > d > S_i > V_j > \text{others},$$

and the following extension for the partial votes $\lambda_{k+1}, \ldots, \lambda_t$:

$$d > S_i > V_j > w > \text{others}$$

For $i \in [k]$, the position of d in the extension of λ_i proposed above is one lower than its original position in η_i. Therefore, the score of d decreases by $k\delta_1$ making the final score of d less than the score of c. Similarly, since $S_1, \ldots, S_k$ form a set cover, the score of u_i decreases by at least $\min_{i=2}^{m}\{\delta_i\}$ for every $i \in [m]$, which is strictly positive as the scoring rule is strict. Finally, the score of w increase by at most $k(\delta_2 + \delta_3 + \cdots + \delta_{m+1})$, since there are at most k votes where the position of w in the extension of λ_i improved from it's original position in

η_i for $i \in [t]$. Therefore, the score of c is greater than any other candidate, implying that c is a possible winner.

For the reverse direction, notice that there must be at least k extensions where d is in the second position, since the score of d is $(k-1)\delta_1$ more than the score of c. In these extensions, observe that w will be at the first position. On the other hand, placing w in the first position causes its score to increase by $(\delta_2 + \delta_3 + \cdots + \delta_{m+1})$, therefore, if w is in the first position in ℓ extensions, its score increases by $\ell(\delta_2 + \delta_3 + \cdots + \delta_{m+1})$. Since the score difference between w and c is only $k(\delta_2 + \delta_3 + \cdots + \delta_{m+1}) + 1$, we can afford to have w in the first position in *at most* k votes. Therefore, apart from the extensions where d is in the second position, in all remaining extensions, w appears after V_j, and therefore the candidates from S_i continue to be in their original positions. Moreover, there must be exactly k votes where d is at the second position. We now claim that the sets corresponding to the k votes where d is at the second position form a set cover. Indeed, if not, suppose the element u_i is not covered. It is easily checked that the score of such a u_i remains unchanged in this extension, and therefore its score is equal to c, contradicting our assumption that we started with an extension for which c was a winner. □

The proof of Theorem 2 can be generalized to a wider class of scoring rules as stated in the following corollary.

Corollary 1. *Let r be a positional scoring rule such that there exists a polynomial function $f : \mathbb{N} \to \mathbb{N}$, such that for every $m \in \mathbb{N}$, there exists an index l in the $f(m)$ length score vector $\vec{\alpha}$ satisfying following,*

$$\alpha_i - \alpha_{i+1} > 0 \ \forall l \leq i \leq l + m$$

Then the Possible winner *problem for r, when parameterized by the number of candidates, does not admit a polynomial kernel unless* CoNP $\subseteq$ NP/Poly.

3.2. Maximin Voting Rule

We will need the following lemma in subsequent proofs. The lemma has been used before [31, 37].

Lemma 2. *Let $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{Z}$ be a function such that*

1. $\forall a, b \in \mathcal{C}, f(a, b) = -f(b, a)$.
2. $\forall a, b, c, d \in \mathcal{C}, f(a, b) + f(c, d)$ *is even.*

Then we can construct in time $O\left(|\mathcal{C}| \sum_{\{a,b\} \in \mathcal{C} \times \mathcal{C}} |f(a, b)|\right)$ an election E with n votes over the candidate set $\mathcal{C}$ such that for all $a, b \in \mathcal{C}$, $D_E(a, b) = f(a, b)$.

We now describe the reduction for the Possible winner problem for the maximin voting rule parameterized by the number of candidates.

Theorem 3. *The* Possible winner *problem for the maximin voting rule, when parameterized by the number of candidates, does not admit a polynomial kernel unless* CoNP $\subseteq$ NP/Poly.

Proof: Let $(\mathcal{U}, \mathcal{F}, k)$ be an instance of SMALL UNIVERSE SET COVER, where $\mathcal{U} = \{u_1, \ldots, u_m\}$ and $\mathcal{F} = \{S_1, \ldots, S_t\}$. We use T_i to denote $\mathcal{U} \setminus S_i$. We now construct an instance $(\mathcal{C}, V, c)$ of the POSSIBLE WINNER as follows.

Candidates. $C := \mathcal{U} \uplus W \uplus \{c, d, x\} \uplus L$, where $W := \{w_1, w_2, \ldots, w_m, w_x\}, L := \{l_1, l_2, l_3\}$.

Partial Votes, P. The first part of the voting profile comprises t partial votes, and will be denoted by P. For each $i \in [t]$, we first consider a profile built on a total order η_i. We denote the order $w_1 \succ \cdots \succ w_m \succ w_x$ by $\vec{W}$. From this point onwards, whenever we place a set of candidates in some position of a partial order, we mean that the candidates in the set can be ordered arbitrarily. For example, the candidates in S_i can be ordered arbitrarily among themselves in the total order η_i below for every $i \in [t]$.

$$\eta_i := L \succ \vec{W} \succ x \succ S_i \succ d \succ c \succ T_i$$

Now we obtain a partial order λ_i based on η_i for every $i \in [t]$ as follows:

$$\lambda_i := \eta_i \setminus (W \times (\{c, d, x\} \uplus \mathcal{U}))$$

The profile P consists of $\{\lambda_i \mid 1 \in [t]\}$.

Complete Votes, Q. We now describe the remaining votes in the profile, which are linear orders designed to achieve specific pairwise difference scores among the candidates. This profile, denoted by Q, is defined according to Lemma 2 to contain votes such that the pairwise score differences of $P \cup Q$ satisfy the following.

- $D(c, w_1) = -2k$.
- $D(c, l_1) = -t$.
- $D(d, w_1) = -2k - 2$.
- $D(x, w_x) = -2k - 2$.
- $D(w_i, u_i) = -2t \ \forall \ i \in [m]$.
- $D(a_i, l_1) = D(w_x, l_1) = -4t$.
- $D(l_1, l_2) = D(l_2, l_3) = D(l_3, l_1) = -4t$.
- $D(l, r) \leq 1$ for all other pairs $(l, r) \in C \times C$.

We note that the for all $c, c' \in \mathcal{C}$, the difference $|D(c, c') - D_P(c, c')|$ is always even, as long as t is even and the number of sets in $\mathcal{F}$ that contain any element $u \in \mathcal{U}$ is always even. Note that the latter can always be ensured without loss of generality: indeed, if $u \in \mathcal{U}$ occurs in an odd number of sets, then we can always add the set $\{u\}$ if it is missing and remove it if it is present, flipping the parity in the process. In case $\{u\}$ is the only set containing the element u, then we remove the set from both $\mathcal{F}$ and $\mathcal{U}$ and decrease k by one. The number of sets t can be assumed to be even by adding a dummy element and a dummy pair of sets that contains the said element. It is easy to see that these modifications always preserve the instance. Thus, the constructed instance of POSSIBLE WINNER is $(\mathcal{C}, V, c)$, where $V := P \cup Q$. We now turn to the proof of correctness.

In the forward direction, let $\mathcal{H} \subseteq \mathcal{F}$ be a set cover of size at most k. Without loss of generality, let $|\mathcal{H}| = k$ (since a smaller set cover can always be extended artificially) and let $\mathcal{H} = \{S_1, \ldots, S_k\}$ (by renaming).

If $i \leq k$, let:

$$\lambda_i^* := L \succ x \succ S_i \succ d \succ c \succ \vec{W} \succ T_i$$

If $k < i \leq t$, let:

$$\lambda_i^* := L \succ \vec{W} \succ x \succ S_i \succ d \succ c \succ T_i$$

Clearly λ_i^* extends λ_i for every $i \in [t]$. Let V^* denote the extended profile consisting of the votes $\{\lambda_i^* \mid i \in [t]\} \cup Q$. We now claim that c is the unique winner with respect to the maximin voting rule in V^*.

Since there are k votes in V^* where c is preferred over w_1 and $(t-k)$ votes where w_1 is preferred to c, we have:

$$\begin{aligned} D_{V^*}(c, w_1) &= D_V(c, w_1) + k - (t-k) \\ &= -2k + k - (t-k) = -t \end{aligned}$$

It is easy to check that maximin score of c is $-t$. Also, it is straightforward to verify the following table.

Candidate	maximin score
$w_i, \forall i \in \{1, 2, \ldots, m\}$	$< -t$
$u_i, \forall i \in \{1, 2, \ldots, m\}$	$\leq -4t$
w_x	$\leq -4t$
l_1, l_2, l_3	$\leq -4t$
x	$\leq -t-2$
d	$\leq -t-2$

Therefore, c is the unique winner for the profile V^*.

We now turn to the reverse direction. Let P^* be an extension of P such that $V^* := P^* \cup Q$ admits c as a unique winner with respect to the maximin voting rule. We first argue that

P^* must admit a certain structure, which will lead us to an almost self-evident set cover for $\mathcal{U}$.

Let us denote by P^*_C the set of votes in P^* which are consistent with $c \succ w_1$, and let P^*_W be the set of votes in P^* which are consistent with $w_1 \succ c$. We first argue that P^*_C has at most k votes.

Claim 1. *Let P^*_C be as defined above. Then $|P^*_C| \leq k$.*

Proof: Suppose, for the sake of contradiction, that more than k extensions are consistent with $c \succ w_1$. Then we have:

$$\begin{aligned} D_{V^*}(c, w_1) &\geq D_V(c, w_1) + k + 1 - (t - k - 1) \\ &= -2k + 2k - t + 2 = -t + 2 \end{aligned}$$

Since $D_{V^*}(c, l_1) = -t$, the maximin score of c is $-t$. On the other hand, we also have that the maximin score of d is given by $D_{V^*}(d, w_1)$, which is now at least $(-t)$:

$$\begin{aligned} D_{V^*}(d, w_1) &\geq D_V(d, w_1) + k + 1 - (t - k - 1) \\ &= -2k - 2 + 2k - t + 2 = -t \end{aligned}$$

Therefore, c is no longer the unique winner in V^* with respect to the maximin voting rule, which is the desired contradiction.

We next propose that a vote that is consistent with $w_1 \succ c$ must be consistent with $w_x \succ x$.

Claim 2. *Let P^*_W be as defined above. Then any vote in P^*_W must respect $w_x \succ x$.*

Proof: Suppose there are r votes in P^*_C, and suppose that in at least one vote in P^*_W where $x \succ w_x$. Notice that any vote in P^*_C is consistent with $x \succ w_x$. Now we have:

$$\begin{aligned} D_{V^*}(c, w_1) &= D_V(c, w_1) + r - (t - r) \\ &= -2k + 2r - t \\ &= -t - 2(k - r) \end{aligned}$$

And further:

$$\begin{aligned} D_{V^*}(x, w_x) &\geq D_V(x, w_x) + (r + 1) - (t - r - 1) \\ &= -2k - 2 + 2r - t + 2 \\ &= -t - 2(k - r) \end{aligned}$$

It is easy to check that the maximin score of c in V^* is at most $-t - 2(k - r)$, witnessed by $D_{V^*}(c, w_1)$, and the maximin score of x is at least $-t - 2(k - r)$, witnessed by $D_{V^*}(x, w_x)$. Therefore, c is no longer the unique winner in V^* with respect to the maximin voting rule, and we have a contradiction.

We are now ready to describe a set cover of size at most k for $\mathcal{U}$ based on V^*. Define $J \subseteq [t]$ as being the set of all indices i for which the extension of λ_i in V^* belongs to P_C^*. Consider:

$$\mathcal{H} := \{S_i \mid i \in J\}.$$

The set $\mathcal{H}$ is our proposed set cover. Clearly, $|\mathcal{H}| \leq k$. It remains to show that $\mathcal{H}$ is a set cover.

We assume, for the sake of contradiction, that there is an element $u_i \in \mathcal{U}$ that is not covered by $\mathcal{H}$. This means that we have $u_i \in T_i$ for all $i \in J$, and thus $w_i \succ u_i$ in the corresponding extensions of λ_i in V^*. Further, for all $i \notin J$, we have that the extension of λ_i in V^* is consistent with:

$$w_1 \succ \cdots \succ w_i \succ \cdots \succ w_x \succ x \succ S_i \succ c \succ T_i,$$

implying again that $w_i \succ u_i$ in these votes. Therefore, we have:

$$D_{V^*}(w_i, u_i) = D_V(w_i, u_i) + k + (t - k) = -2t + t = -t.$$

We know that the maximin score of c is less than or equal to $-t$, since $D_{V^*}(c, l_1) = -t$, and we now have that the maximin score of w_i is $-t$. This overrules c as the unique winner in V^*, contradicting our assumption to that effect. This completes the proof. □

3.3. Copeland Voting Rule

We now describe the result for the Possible winner problem for the Copeland voting rule parameterized by the number of candidates.

Theorem 4. *The* Possible winner *problem for the Copeland voting rule, when parameterized by the number of candidates, does not admit a polynomial kernel unless* CoNP $\subseteq$ NP/Poly.

Proof: Let $(\mathcal{U}, \mathcal{F}, k)$ be an instance of Small universe set cover, where $\mathcal{U} = \{u_1, \ldots, u_m\}$ and $\mathcal{F} = \{S_1, \ldots, S_t\}$. For the purpose of this proof, we assume (without loss of generality) that $m \geq 6$. We now construct an instance $(\mathcal{C}, V, c)$ of Possible winner as follows.

Candidates. $\mathcal{C} := \mathcal{U} \uplus \{z, c, d, w\}$.

Partial Votes, P. The first part of the voting profile comprises of m partial votes, and will be denoted by P. For each $i \in [t]$, we first consider a profile built on a total order:

$$\eta_i := \mathcal{U} \setminus S_i \succ z \succ c \succ d \succ S_i \succ w$$

Now we obtain a partial order λ_i based on η_i as follows for each $i \in [t]$:

$$\lambda_i := \eta_i \setminus (\{z, c\} \times (S_i \uplus \{d, w\}))$$

The profile P consists of $\{\lambda_i \mid i \in [t]\}$.

Complete Votes, Q. We now describe the remaining votes in the profile, which are linear orders designed to achieve specific pairwise difference scores among the candidates. This profile, denoted by Q, is defined according to Lemma 2 to contain votes such that the pairwise score differences of $P \cup Q$ satisfy the following.

– $D(c, d) = t - 2k + 1$

– $D(z, w) = t - 2k - 1$

– $D(c, u_i) = t - 1$

– $D(c, z) = t + 1$

– $D(c, w) = -t - 1$

– $D(u_i, d) = D(z, u_i) = t + 1 \; \forall \; i \in [m]$

– $D(z, d) = t + 1$

– $D(u_i, u_j) = t + 1 \; \forall \; j \in [i + 1 \pmod{m}, i + \lfloor m/2 \rfloor \pmod{m})]$

We note that the difference $|D(c, c') - D_P(c, c')|$ is always even for all $c, c' \in \mathcal{C}$, as long as t is odd and the number of sets in $\mathcal{F}$ that contain any element $a \in \mathcal{U}$ is always odd. Note that the latter can always be ensured without loss of generality: indeed, if $a \in \mathcal{U}$ occurs in an even number of sets, then we can always add the set $\{a\}$ if it is missing and remove it if it is present, flipping the parity in the process. In case $\{a\}$ is the only set containing the element a, then we remove the set from both $\mathcal{F}$ and $\mathcal{U}$ and decrease k by one. The number of sets t can be assumed to be odd by adding a dummy element in $\mathcal{U}$, adding a dummy set that contains the said element in $\mathcal{F}$, and incrementng k by one. It is easy to see that these modifications always preserve the instance.

Thus the constructed instance of Possible winner is $(\mathcal{C}, V, c)$, where $V := P \cup Q$. We now turn to the proof of correctness.

In the forward direction, let $\mathcal{H} \subseteq \mathcal{F}$ be a set cover of size at most k. Without loss of generality, let $|\mathcal{H}| = k$ (since a smaller set cover can always be extended artificially) and let $\mathcal{H} = \{S_1, \ldots, S_k\}$ (by renaming).

If $i \leq k$, let:

$$\lambda_i^* := \mathcal{U} \setminus S_i \succ z \succ c \succ d \succ S_i \succ w$$

If $k < i \leq t$, let:

$$\lambda_i^* := \mathcal{U} \setminus S_i \succ d \succ S_i \succ w \succ z \succ c$$

Clearly λ_i^* extends λ_i for every $i \in [t]$. Let V^* denote the extended profile consisting of the votes $\{\lambda_i^* \mid i \in [t]\} \cup Q$. We now claim that c is the unique winner with respect to the Copeland voting rule in V^*.

First, consider the candidate z. For every $i \in [m]$, between z and u_i, even if z loses to u_i in λ_j^*, for every $j \in [t]$, because $D(z, u_i) = t + 1$, z wins the pairwise election between z

and u_i. The same argument holds between z and d. Therefore, the Copeland score of z, no matter how the partial votes were extended, is at least $(m+1)$.

Further, note that all other candidates (apart from c) have a Copeland score of less than $(m+1)$, because they are guaranteed to lose to at least three candidates (assuming $m \geq 6$). In particular, observe that u_i loses to at least $\lfloor m/2 \rfloor$ candidates, and d loses to u_i (merely by its position in the extended votes), and w loses to u_i (because of way the scores were designed) for every $i \in [m]$. Therefore, the Copeland score of all candidates in $\mathcal{C} \setminus \{z, c\}$ is strictly less than the Copeland score of z, and therefore they cannot be possible (co-)winners.

Now we restrict our attention to the contest between z and c. First note that c beats u_i for every $i \in [m]$: since the sets of $\mathcal{H}$ form a set cover, u_i is placed in a position after c in some λ_j^* for $j \in [k]$. Since the difference of score between c and u_i was $(t-1)$, even if c suffered defeat in every other extension, we have the pairwise score of c and u_i being at least $t - 1 - (t-1) + 1 = 1$, which implies that c defeats every u_i in their pairwise election. Note that c also defeats d by getting ahead of d in k votes, making its final score $t - 2k + 1 + k - (t-k) = 1$. Finally, c is defeated by w, simply by the preset difference score. Therefore, the Copeland score of c is $(m+2)$.

Now all that remains to be done is to rule z out of the running. Note that z is defeated by w in their pairwise election: this is because z defeats w in k of the extended votes, and is defeated by w in the remaining. This implies that its final pairwise score with respect to w is at most $t - 2k - 1 + k - (t-k) = -1$. Also note that z loses to c because of its predefined difference score. Thus, the Copeland score of z in the extended vote is exactly $(m+1)$, and thus c is the unique winner of the extended vote.

We now turn to the reverse direction. Let P^* be an extension of P such that $V^* := P^* \cup Q$ admits c as a unique winner with respect to the Copeland voting rule. As with the argument for the maximin voting rule, we first argue that P^* must admit a certain structure, which will lead us to an almost self-evident set cover for $\mathcal{U}$.

Let us denote by P_C^* the set of votes in P^* which are consistent with $c \succ d$, and let P_W^* be the set of votes in P^* which are consistent with $w \succ z$. Note that the votes in P_C^* necessarily have the form:

$$\lambda_i^* := \mathcal{U} \setminus S_i \succ z \succ c \succ d \succ S_i \succ w$$

and those in P_W^* have the form:

$$\lambda_i^* := \mathcal{U} \setminus S_i \succ d \succ S_i \succ w \succ z \succ c$$

It is easy to check that this structure is directly imposed by the relative orderings that are fixed by the partial orders.

Before we argue the details of the scores, let us recall that in any extension of P, z loses to c and z wins over d and all candidates in $\mathcal{U}$. Thus the Copeland score of z is at least $(m+1)$. On the other hand, in any extension of P, c loses to w, and therefore the Copeland score of c is at most $(m+2)$. (These facts follow from the analysis in the forward direction.)

Thus, we have the following situation. If z wins over w, then c cannot be the unique winner in the extended vote, because the score of z goes up to $(m+2)$. Similarly, c cannot afford to lose to any of $\mathcal{U} \cup \{d\}$, because that will cause its score to drop below $(m+2)$, resulting in either a tie with z, or defeat. These facts will successively lead us to the correctness of the reverse direction.

Now let us return to the sets P^*_C and P^*_W. If P^*_C has more than k votes, then z wins over w: the final score of z is at least $t - 2k - 1 + (k + 1) - (t - k - 1) = 1$, and we have a contradiction. If P^*_C has fewer than k votes, then c loses to d, with a score of at most $t - 2k + 1 + (k - 1) - (t - k + 1) = -1$, and we have a contradiction. Hence, P^*_C must have exactly k votes.

Finally, suppose the sets corresponding to the votes of P^*_C do not form a set cover. Consider an element $u_i \in \mathcal{U}$ not covered by the union of these sets. Observe that c now loses the pairwise election between itself and u_i and is no longer in the running for being the unique winner in the extended vote. Therefore, the sets corresponding to the votes of P^*_C form a set cover of size exactly k, as desired. □

3.4. Bucklin Voting Rule

We now describe the result for the Possible winner problem for the Bucklin voting rule parameterized by the number of candidates.

Theorem 5. *The* Possible winner *problem for the Bucklin voting rule, when parameterized by the number of candidates, does not admit a polynomial kernel unless* CoNP $\subseteq$ NP/Poly.

Proof: Let $(\mathcal{U}, \mathcal{F}, k)$ be an instance of Small universe set cover, where $\mathcal{U} = \{u_1, \ldots, u_m\}$ and $\mathcal{F} = \{S_1, \ldots, S_t\}$. Without loss of generality, we assume that $t > k + 1$, and that every set in $\mathcal{F}$ has at least two elements. We now construct an instance $(\mathcal{C}, V, c)$ of Possible winner as follows.

Candidates. $\mathcal{C} := \mathcal{U} \uplus \{z, c, a\} \uplus W \uplus D_1 \uplus D_2 \uplus D_3$, where D_1, D_2, and D_3 are sets of *"dummy candidates"* such that $|D_1| = m$, $|D_2| = 2m$, and $|D_3| = 2m$. $W := \{w_1, w_2, \ldots, w_{2m}\}$.

Partial Votes, P. The first part of the voting profile comprises of t partial votes, and will be denoted by P. For each $i \in [t]$, we first consider a profile built on a total order:

$$\eta_i := \mathcal{U} \setminus S_i \succ S_i \succ w_{i \pmod m} \succ w_{i+1 \pmod m} \succ z \succ c \succ D_3 \succ \text{others}$$

Now we obtain a partial order λ_i based on η_i for every $i \in [t]$ as follows:

$$\lambda_i := \eta_i \setminus \left(\left(\{w_{i \pmod m}, w_{i+1 \pmod m}, z, c\} \uplus D_3 \right) \times S_i \right)$$

The profile P consists of $\{\lambda_i \mid i \in [t]\}$.

Complete Votes, Q.

$$t-k-1 \quad : \quad D_1 \succ z \succ c \succ \text{others}$$
$$1 \quad : \quad D_1 \succ c \succ a \succ z \succ \text{others}$$
$$k-1 \quad : \quad D_2 \succ \text{others}$$

We now show that $(\mathcal{U}, \mathcal{F}, k)$ is a YES instance if and only if $(\mathcal{C}, V, c)$ is a YES instance. Suppose $\{S_j : j \in J\}$ forms a set cover. Then consider the following extension of P:

$$(\mathcal{U} \setminus S_j) \succ w_{j \,(\text{mod } m)} \succ w_{j+1 \,(\text{mod } m)} \succ z \succ c \succ D_3 \succ S_j \succ \text{others}, \text{ for } j \in J$$

$$(\mathcal{U} \setminus S_j) \succ S_j \succ w_{j \,(\text{mod } m)} \succ w_{j+1 \,(\text{mod } m)} \succ z \succ c \succ D_3 \succ \text{others}, \text{ for } j \notin J$$

We claim that in this extension, c is the unique winner with Bucklin score $(m+2)$. First, let us establish the score of c. The candidate c is already within the top $(m+1)$ choices in $(t-k)$ of the complete votes. In all the sets that form the set cover, c is ranked within the first $(m+2)$ votes in the proposed extension of the corresponding vote (recall that every set has at least two elements). Therefore, there are a total of t votes where c is ranked within the top $(m+2)$ preferences. Further, consider a candidate $v \in \mathcal{U}$. Such a candidate is not within the top $(m+2)$ choices of any of the complete votes. Let S_i be the set that covers the element v. Note that in the extension of the vote λ_i, v is not ranked among the top $(m+2)$ spots, since there are at least m candidates from D_3 getting in the way. Therefore, v has strictly fewer than t votes where it is ranked among the top $(m+2)$ spots, and thus has a Bucklin score more than c.

Now the candidate z is within the top $(m+2)$ ranks of at most $(t-k-1)$ votes among the complete votes. In the votes corresponding to the sets *not* in the set cover, z is placed beyond the first $(m+2)$ spots. Therefore, the number of votes where z is among the top $(m+2)$ candidates is at most $(t-1)$, which makes its Bucklin score strictly more than $(m+2)$.

The candidates from W are within the top $(m+2)$ positions only in a constant number of votes. The candidates $D_1 \cup \{a\}$ have $(t-k)$ votes (among the complete ones) in which they are ranked among the top $(m+2)$ preferences, but in all extensions, these candidates have ranks below $(m+2)$. Finally, the candidates in D_3 do not feature in the top $(m+2)$ positions of any of the complete votes, and similarly, the candidates in D_2 do not feature in the top $(m+2)$ positions of any of the extended votes. Therefore, the Bucklin scores of all these candidates is easily seen to be strictly more than $(m+2)$, concluding the argument in the forward direction.

Now consider the reverse direction. Suppose $(\mathcal{C}, V, c)$ is a YES instance. For the same reasons described in the forward direction, observe that only the following candidates can win depending upon how the partial preferences get extended - either one of the candidates in $\mathcal{U}$, or one of z or c. Note that the Bucklin score of z in any extension is at most $(m+3)$. Therefore, the Bucklin score of c has to be $(m+2)$ or less. Among the complete votes Q, there are $(t-k)$ votes where the candidate c appears in the top $(m+2)$ positions. To get majority within top $(m+2)$ positions, c should be within top $(m+2)$ positions for at least

k of the extended votes in P. Let us call these set of votes P'. Now notice that whenever c comes within top $(m+2)$ positions in a valid extension of P, the candidate z also comes within top $(m+2)$ positions in the same vote. However, the candidate z is already ranked among the top $(m+2)$ candidates in $(t-k-1)$ complete votes. Therefore, z can appear within top $(m+2)$ positions in *at most* k extensions (since c is the unique winner), implying that $|P'| = k$. Further, note that the Bucklin score of c cannot be strictly smaller than $(m+2)$ in any extension. Indeed, candidate c features in only one of the complete votes within the top $(m+1)$ positions, and it would have to be within the top $(m+1)$ positions in at least $(t-1)$ extensions. However, as discussed earlier, this would give z exactly the same mileage, and therefore its Bucklin score would be $(m-1)$ or even less; contradicting our assumption that c is the unique winner.

Now we claim that the S_i's corresponding to the votes in P' form a set cover for $\mathcal{U}$. If not, there is an element $x \in \mathcal{U}$ that is uncovered. Observe that x appears within top m positions in all the extensions of the votes in P', by assumption. Further, in all the remaining extensions, since z is not present among the top $(m+2)$ positions, we only have room for two candidates from W. The remaining positions must be filled by all the candidates corresponding to elements of $\mathcal{U}$. Therefore, x appears within the top $(m+2)$ positions of all the extended votes. Since these constitute half the total number of votes, we have that x ties with c in this situation, a contradiction. □

3.5. Ranked Pairs Voting Rule

We now describe the reduction for Possible winner parameterized by the number of candidates, for the ranked pairs voting rule.

Theorem 6. *The* Possible winner *problem for the ranked pairs voting rule, when parameterized by the number of candidates, does not admit a polynomial kernel unless* CoNP $\subseteq$ NP/Poly.

Proof: Let $(\mathcal{U}, \mathcal{F}, k)$ be an instance of Small universe set cover, where $\mathcal{U} = \{u_1, \dots, u_m\}$ and $\mathcal{F} = \{S_1, \dots, S_t\}$. Without loss of generality, we assume that t is even. We now construct an instance $(\mathcal{C}, V, c)$ of Possible winner as follows.

Candidates. $\mathcal{C} := \mathcal{U} \uplus \{a, b, c, w\}$.

Partial Votes, P. The first part of the voting profile comprises of t partial votes, and will be denoted by P. For each $i \in [t]$, we first consider a profile built on a total order:

$$\eta_i := \mathcal{U} \setminus S_i \succ S_i \succ b \succ a \succ c \succ \text{others}$$

Now we obtain a partial order λ_i based on η_i for every $i \in [t]$ as follows:

$$\lambda_i := \eta_i \setminus (\{a, c\} \times (S_i \uplus \{b\}))$$

The profile P consists of $\{\lambda_i \mid i \in [t]\}$.

Complete Votes, Q. We add complete votes such that along with the already determined pairs from the partial votes P, we have following.

- $D(u_i, c) = 2 \ \forall \ i \in [m]$
- $D(c, b) = 4t$
- $D(c, w) = t + 2$
- $D(b, a) = 2k + 4$
- $D(w, a) = 4t$
- $D(a, c) = t + 2$
- $D(w, u_i) = 4t \ \forall \ i \in [m]$

We now show that $(\mathcal{U}, \mathcal{F}, k)$ is a YES instance if and only if $(\mathcal{C}, V, c)$ is a YES instance. Suppose $\{S_j : j \in J\}$ forms a set cover. Then consider the following extension of P :

$$\mathcal{U} \setminus S_j \succ a \succ c \succ S_j \succ b \succ \text{others} \ \forall j \in J$$

$$\mathcal{U} \setminus S_j \succ S_j \succ b \succ a \succ c \succ \text{others} \ \forall j \notin J$$

We claim that the candidate c is the unique winner in this extension. Note that the pairs $(w \succ a)$ and $(w \succ u_i)$ for every $i \in [t]$ get locked first (since these differences are clearly the highest and unchanged). The pair (c, b) gets locked next, with a difference score of $(3t+2k)$. Now since the votes in which $c \succ b$ are based on a set cover of size at most k, the pairwise difference between b and a becomes at least $2k + 4 - k + (t - k) = t + 4$. Therefore, the next pair to get locked is $b \succ a$. Finally, for every element $u_i \in \mathcal{U}$, the difference $D(u_i, c)$ is at most $2 + (t - 1) = t + 1$, since there is at least one vote where $c \succ u_i$ (given that we used a set cover in the extension). It is now easy to see that the next highest pairwise difference is between c and w, so the ordering $c \succ w$ gets locked, and at this point, by transitivity, c is superior to w, b, a and all u_i. It follows that c wins the election irrespective the sequence in which pairs are considered subsequently.

Now suppose $(\mathcal{C}, V, c)$ is a YES instance. Notice that, irrespective of the extension of the votes in P, $c \succ b, w \succ a, w \succ u_i \ \forall i \in [m]$ are locked first. Now if $b \succ c$ in all the extended votes, then it is easy to check that $b \succ a$ gets locked next, with a difference score of $2k+4+t$; leaving us with $D(u_i, c) = t + 2 = D(c, w)$, where $u_i \succ c$ could be a potential lock-in. This implies the possibility of a u_i being a winner in some choice of tie-breaking, a contradiction to the assumption that c is the unique winner. Therefore, there are at least some votes in the extended profile where $c \succ b$. We now claim that there are at most k such votes. Indeed, if there are more, then $D(b, a) = 2k + 4 - (k + 1) + (t - k - 1) = t + 2$. Therefore, after the forced lock-ins above, we have $D(b, a) = D(c, w) = D(a, c) = t+2$. Here, again, it is possible for $a \succ c$ to be locked in before the other choices, and we again have a contradiction.

Finally, we have that $c \succ b$ in at most k many extensions in P. Call the set of indices of these extensions J. We claim that $\{S_j : j \in J\}$ forms a set cover. If not, then suppose an element $u_i \in \mathcal{U}$ is not covered by $\{S_j : j \in J\}$. Then the candidate u_i comes before c in all the extensions which makes $D(u_i, c)$ become $(t+2)$, which in turn ties with $D(c, w)$. This again contradicts the fact that c is the unique winner. Therefore, if there is an extension that makes c the unique winner, then we have the desired set cover. □

4. Polynomial Kernels for Coalitional Manipulation

We now describe a kernelization algorithm for every scoring rule which satisfies certain properties mentioned in Theorem 7 below. Note that the Borda voting rule satisfies these properties.

Theorem 7. *For $m \in \mathbb{N}$, let $(\alpha_1, \dots, \alpha_m)$ and $(\alpha'_1, \dots, \alpha'_{m+1})$ be the normalized score vectors for a scoring rule r for an election with m and $(m+1)$ candidates respectively. Let $\alpha'_1 = poly(m)$ and $\alpha_i = \alpha'_{i+1}$ for every $i \in [m]$. Then the* Coalitional manipulation *problem for r admits a polynomial kernel when the number of manipulators is $poly(m)$.*

Proof: Let c be the candidate whom the manipulators aim to make winner. Let M be the set of manipulators and $\mathcal{C}$ the set of candidates. Let $s_{NM}(x)$ be the score of candidate x from the votes of the non-manipulators. Without loss of generality, we assume that, all the manipulators place c at top position in their votes. Hence, the final score of c is $s_{NM}(c) + |M|\alpha_1$, which we denote by $s(c)$. Now if $s_{NM}(x) \geq s(c)$ for any $x \neq c$, then c cannot win and we output *no*. Hence, we assume that $s_{NM}(x) < s(c)$ for all $x \neq c$. Now let us define $s^*_{NM}(x)$ as follows.

$$s^*_{NM}(x) := \max\{s_{NM}(x), s_{NM}(c)\}$$

Also define $s^*_{NM}(c)$ as follows.

$$s^*_{NM}(c) := s_{NM}(c) - |M|(\alpha'_1 - \alpha_1)$$

We define a Coalitional manipulation instance with $(m+1)$ candidates as $(\mathcal{C}', NM, M, c)$, where $\mathcal{C}' = \mathcal{C} \uplus \{d\}$ is the set of candidates, M is the set of manipulators, c is the distinguished candidate, and NM is the non-manipulators' vote such that it generates score of $x \in \mathcal{C}$ to be $K + (s^*_{NM}(x) - s_{NM}(c))$, where $K \in \mathbb{N}$ is same for $x \in \mathcal{C}$, and the score of d is less than $K - \alpha'_1|M|$. The existence of such a voting profile NM of size $poly(m)$ is due to Lemma 1 and the fact that $\alpha'_1 = poly(m)$. Hence, once we show the equivalence of these two instances, we have a kernel whose size is polynomial in m. The equivalence of the two instances is due to the following facts: (1) The new instance has $(m+1)$ candidates and c is always placed at the top position without loss of generality. The candidate c recieves $|M|(\alpha'_1 - \alpha_1)$ score more than the initial instance and this is compensated in $s^*_{NM}(c)$. (2) The relative score difference from the final score of c to the current score of every $x \in \mathcal{C} \setminus \{c\}$ is same in both the instances. (3) In the new instance, we can assume without loss of generality that the candidate d will be placed in the second position in all the manipulators' votes. □

We now move on to the voting rules that are based on the weighted majority graph. The reduction rules modify the weighted majority graph maintaining the property that there exists a set of votes that can realize the modified weighted majority graph. In particular, the final weighted majority graph is realizable with a set of votes.

Theorem 8. *The* Coalitional manipulation *problem for the maximin voting rule admits a polynomial kernel when the number of manipulators is poly(m).*

Proof: Let c be the distinguished candidate of the manipulators. Let M be the set of all manipulators. We can assume that $|M| \geq 2$ since for $|M| = 1$, the problem is in P [1]. Define s to be $\min_{x \in C \setminus \{c\}} D_{(\mathcal{V} \setminus M)}(c, x)$. So, s is the maximin score of the candidate c from the votes except from M. Since the maximin voting rule is monotone, we can assume that the voters in M put the candidate c at top position of their preferences. Hence, c's final maximin score will be $s + |M|$. This provides the following reduction rule.

Reduction rule 1. *If $s + |M| \geq 0$, then output* Yes.

In the following, we will assume $s + |M|$ is negative. Now we propose the following reduction rules on the weighted majority graph.

Reduction rule 2. *If $D_{(\mathcal{V} \setminus M)}(c_i, c_j) < 0$ and $D_{(\mathcal{V} \setminus M)}(c_i, c_j) > 2|M| + s$, then make $D_{(\mathcal{V} \setminus M)}(c_i, c_j)$ either $2|M| + s + 1$ or $2|M| + s + 2$ whichever keeps the parity of $D_{(\mathcal{V} \setminus M)}(c_i, c_j)$ unchanged.*

If $D_{(\mathcal{V} \setminus M)}(c_i, c_j) > 2|M| + s$, then $D_{\mathcal{V}}(c_i, c_j) > |M| + s$ irrespective of the way the manipulators vote. Hence, given any votes of the manipulators, whether or not the maximin score of c_i and c_j will exceed the maximin score of c does not gets affected by this reduction rule. Hence, Reduction rule 1 is sound.

Reduction rule 3. *If $D_{(\mathcal{V} \setminus M)}(c_i, c_j) < s$, then make $D_{(\mathcal{V} \setminus M)}(c_i, c_j)$ either $s - 1$ or $s - 2$ whichever keeps the parity of $D_{(\mathcal{V} \setminus M)}(c_i, c_j)$ unchanged.*

The argument for the correctness of Reduction rule 3 is similar to the argument for Reduction rule 1. Here onward, we may assume that whenever $D_{(\mathcal{V} \setminus M)}(c_i, c_j) < 0$, $s - 2 \leq D_{(\mathcal{V} \setminus M)}(c_i, c_j) \leq 2|M| + s + 2$

Reduction rule 4. *If $s < -4|M|$ then subtract $s + 5|M|$ from $D_{(\mathcal{V} \setminus M)}(x, y)$ for every $x, y \in \mathcal{C}, x \neq y$.*

The correctness of Reduction rule 4 follows from the fact that it adds linear fixed offsets to all the edges of the weighted majority graph. Hence, if there a voting profile of the voters in M that makes the candidate c win in the original instance, the same voting profile will make c win the election in the reduced instance and vice versa.

Now we have a weighted majority graph with $O(|M|)$ weights for every edge. Also, all the weights have uniform parity and thus the result follows from Lemma 2. $\square$

We next present a polynomial kernel for the COALITIONAL MANIPULATION problem for the Copeland voting rule.

Theorem 9. *The* COALITIONAL MANIPULATION *problem for the Copeland voting rule admits a polynomial kernel when the number of manipulators is poly*(m)*.*

Proof: We apply the following reduction rule.

Reduction rule 1. *If* $D_{(\mathcal{V}\setminus M)}(x,y) > |M|$ *for* $x, y \in \mathcal{C}$*, then make* $D_{(\mathcal{V}\setminus M)}(x,y)$ *either* $|M|+1$ *or* $|M|+2$ *whichever keeps the parity of* $D_{(\mathcal{V}\setminus M)}(x,y)$ *unchanged.*

Given any votes of M, we have $D_{\mathcal{V}}(x,y) > 0$ in the original instance if and only if $D_{\mathcal{V}}(x,y) > 0$ in the reduced instance for every $x, y \in \mathcal{C}$. Hence, each candidate has the same Copeland score and thus the reduction rule is correct.

Now we have a weighted majority graph with $O(|M|)$ weights for every edges. Also, all the weights have uniform parity. From Lemma 2, we can realize the weighted majority graph using $O(m^2.|M|)$ votes. $\square$

Now we move on to the ranked pairs voting rule.

Theorem 10. *The* COALITIONAL MANIPULATION *problem for the ranked pairs voting rule admits a polynomial kernel when the number of manipulators is poly*(m)*.*

Proof: Consider all non-negative $D_{(\mathcal{V}\setminus M)}(c_i, c_j)$ and arrange them in non-decreasing order. Let the ordering be $x_1, x_2, \ldots, x_l$ where $l = \binom{m}{2}$. Now keep applying following reduction rule till possible. Define $x_0 = 0$.

Reduction rule 1. *If there exist any* i *such that,* $x_i - x_{i-1} > |M| + 2$*, subtract an even offset to all* $x_i, x_{i+1}, \ldots, x_l$ *such that* x_i *becomes either* $(x_{i-1} + |M| + 1)$ *or* $(x_{i-1} + |M| + 2)$*.*

The reduction rule is correct since for any set of votes by M, for any four candidates $a, b, x, y \in \mathcal{C}$, $D(a,b) > D(x,y)$ in the original instance if and only if $D(a,b) > D(x,y)$ in the reduced instance. Now we have a weighted majority graph with $O(m^2.|M|)$ weights for every edges. Also, all the weights have uniform parity and hence can be realized with $O(m^4.|M|)$ votes Lemma 2. $\square$

5. Conclusion and Future Work

Here we showed that the Possible winner problem does not admit a polynomial kernel for many common voting rules under the complexity theoretic assumption that $\mathsf{CoNP} \subseteq \mathsf{NP/Poly}$ is not true. We also showed the existence of polynomial kernels for the Coalitional manipulation problem for many common voting rules. This shows that the Possible winner problem is a significantly harder problem than the Coalitional manipulation problem, although both the problems are NP-complete.

There are other interesting parameterizations of these problems for which fixed parameter tractable algorithms are known but the corresponding kernelization questions are still open. One such parameter is the total number of pairs s in all the votes for which an ordering has not been specified. With this parameter, a simple $O(2^s.\ \text{poly}(m, n))$ algorithm is known [8]. However, the corresponding kernelization question is still open.

Acknowledgement

Palash Dey wishes to gratefully acknowledge support from Google India for providing him with a special fellowship for carrying out his doctoral work. Neeldhara Misra is supported by the INSPIRE Faculty Scheme, DST India (project DSTO-1209).

[1] Bartholdi, J., Tovey, C., and Trick, M. (1989). The computational difficulty of manipulating an election. *Social Choice and Welfare (SCW)*, 6(3):227–241.

[2] Baumeister, D., Roos, M., and Rothe, J. (2011). Computational complexity of two variants of the possible winner problem. In *The 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 853–860. International Foundation for Autonomous Agents and Multiagent Systems.

[3] Betzler, N. (2010). On problem kernels for possible winner determination under the k-approval protocol. In *Mathematical Foundations of Computer Science (MFCS)*, pages 114–125. Springer.

[4] Betzler, N., Bredereck, R., Chen, J., and Niedermeier, R. (2012). The multivariate algorithmic revolution and beyond. chapter Studies in Computational Aspects of Voting: Open Problems of Downey and Fellows, pages 318–363. Springer-Verlag, Berlin, Heidelberg.

[5] Betzler, N., Bredereck, R., and Niedermeier, R. (2010). Partial kernelization for rank aggregation: theory and experiments. In *International Symposium on Parameterized and Exact Computation (IPEC)*, pages 26–37. Springer.

[6] Betzler, N. and Dorn, B. (2009). Towards a dichotomy of finding possible winners in elections based on scoring rules. In *Mathematical Foundations of Computer Science (MFCS)*, pages 124–136. Springer.

[7] Betzler, N., Fellows, M. R., Guo, J., Niedermeier, R., and Rosamond, F. A. (2009a). Fixed-parameter algorithms for kemeny rankings. *Theoretical Computer Science (TCS)*, 410(45):4554–4570.

[8] Betzler, N., Hemmann, S., and Niedermeier, R. (2009b). A Multivariate Complexity Analysis of Determining Possible Winners given Incomplete Votes. In *International Joint Conference on Artificial Intelligence (IJCAI)*, volume 9, pages 53–58.

[9] Bodlaender, H. L., Thomassé, S., and Yeo, A. (2009). Kernel Bounds for Disjoint Cycles and Disjoint Paths. In Fiat, A. and Sanders, P., editors, *Proceedings of the 17th Annual European Symposium,on Algorithms (ESA 2009), Copenhagen, Denmark, September 7-9, 2009.*, volume 5757 of *Lecture Notes in Computer Science*, pages 635–646. Springer.

[10] Bredereck, R., Chen, J., Faliszewski, P., Guo, J., Niedermeier, R., and Woeginger, G. J. (2014a). Parameterized algorithmics for computational social choice: Nine research challenges. In *Tsinghua Science and Technology*, volume 19, pages 358–373. IEEE.

[11] Bredereck, R., Chen, J., Faliszewski, P., Nichterlein, A., and Niedermeier, R. (2014b). Prices matter for the parameterized complexity of shift bribery.

[12] Bredereck, R., Chen, J., Hartung, S., Niedermeier, R., Suchỳ, O., and Kratsch, S. (2012). A multivariate complexity analysis of lobbying in multiple referenda. In *International Conference on Artificial Intelligence (AAAI)*.

[13] Brill, M. and Fischer, F. A. (2012). The price of neutrality for the ranked pairs method. In *International Conference on Artificial Intelligence (AAAI)*.

[14] Chevaleyre, Y., Lang, J., Maudet, N., and Monnot, J. (2010). Possible winners when new candidates are added: The case of scoring rules. In *International Conference on Artificial Intelligence (AAAI)*.

[15] Conitzer, V., Rognlie, M., and Xia, L. (2009). Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, volume 9, pages 109–115.

[16] Dey, P., Misra, N., and Narahari, Y. (2015). Kernelization complexity of possible winner and coalitional manipulation problems in voting. In *Proceeding of the 14th International Conference on Autonomous Systems and Multiagent Systems (AAMAS-15)*.

[17] Dom, M., Lokshtanov, D., and Saurabh, S. (2009). Incompressibility through colors and IDs. In *Automata, Languages and Programming (ICALP), 36th International Colloquium, Part I*, volume 5555 of *LNCS*, pages 378–389.

[18] Dorn, B. and Schlotter, I. (2012). Multivariate complexity analysis of swap bribery. *Algorithmica*, 64(1):126–151.

[19] Downey, R. G. and Fellows, M. R. (1999). *Parameterized Complexity*, volume 3. springer Heidelberg.

[20] Ephrati, E. and Rosenschein, J. (1991). The Clarke tax as a consensus mechanism among automated agents. In *Proceedings of the Ninth International Conference on Artificial Intelligence (AAAI)*, pages 173–178.

[21] Faliszewski, P., Hemaspaandra, E., and Schnoor, H. (2008). Copeland voting: Ties matter. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 983–990. International Foundation for Autonomous Agents and Multiagent Systems.

[22] Faliszewski, P., Hemaspaandra, E., and Schnoor, H. (2010). Manipulation of copeland elections. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 367–374. International Foundation for Autonomous Agents and Multiagent Systems.

[23] Faliszewski, P., Reisch, Y., Rothe, J., and Schend, L. (2014). Complexity of manipulation, bribery, and campaign management in bucklin and fallback voting. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1357–1358. International Foundation for Autonomous Agents and Multiagent Systems.

[24] Flum, J. and Grohe, M. (2006). *Parameterized Complexity Theory*, volume 3. Springer.

[25] Froese, V., Van Bevern, R., Niedermeier, R., and Sorge, M. (2013). A parameterized complexity analysis of combinatorial feature selection problems. In *Mathematical Foundations of Computer Science (MFCS)*, pages 445–456. Springer.

[26] Guo, J. and Niedermeier, R. (2007). Invitation to data reduction and problem kernelization. *ACM SIGACT News*, 38(1):31–45.

[27] Konczak, K. and Lang, J. (2005). Voting procedures with incomplete preferences. In *Proceedings of the International Joint Conference on Artificial Intelligence-05 Multidisciplinary Workshop on Advances in Preference Handling*, volume 20.

[28] Lang, J., Pini, M. S., Rossi, F., Salvagnin, D., Venable, K. B., and Walsh, T. (2012). Winner determination in voting trees with incomplete preferences and weighted votes. *Autonomous Agents and Multiagent Systems*, 25(1):130–157.

[29] Lang, J., Pini, M. S., Rossi, F., Venable, K. B., and Walsh, T. (2007). Winner determination in sequential majority voting. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*, volume 7, pages 1372–1377.

[30] Lokshtanov, D., Misra, N., and Saurabh, S. (2012). Kernelization–preprocessing with a guarantee. In *The Multivariate Algorithmic Revolution and Beyond*, pages 129–161. Springer.

[31] McGarvey, D. C. (1953). A theorem on the construction of voting paradoxes. *Econometrica: Journal of the Econometric Society*, pages 608–610.

[32] Niedermeier, R. (2002). Invitation to fixed-parameter algorithms. *Habilitationschrift, University of Tübingen*.

[33] Pennock, D. M., Horvitz, E., and Giles, C. L. (2000). Social choice theory and recommender systems: Analysis of the axiomatic foundations of collaborative filtering. In *Proceedings of the 17th International Conference on Artificial Intelligence (AAAI)*.

[34] Pini, M. S., Rossi, F., Venable, K. B., and Walsh, T. (2007). Incompleteness and incomparability in preference aggregation. In *Proceedings of the 20nd International Joint Conference on Artificial Intelligence (IJCAI)*, volume 7, pages 1464–1469.

[35] Walsh, T. (2007). Uncertainty in preference elicitation and aggregation. In *Proceedings of the International Conference on Artificial Intelligence (AAAI)*, volume 22, page 3.

[36] Weihe, K. (1998). Covering trains by stations or the power of data reduction. *Proceedings of Algorithms and Experiments, ALEX*, pages 1–8.

[37] Xia, L. and Conitzer, V. (2008). Determining possible and necessary winners under common voting rules given partial orders. In *Proceedings of the International Conference on Artificial Intelligence (AAAI)*, volume 8, pages 196–201.

[38] Xia, L., Zuckerman, M., Procaccia, A. D., Conitzer, V., and Rosenschein, J. S. (2009). Complexity of unweighted coalitional manipulation under some common voting rules. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, volume 9, pages 348–352.

[39] Yang, Y. (2014). Election attacks with few candidates. In *ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014)*, pages 1131–1132.

[40] Yang, Y. and Guo, J. (2013). Exact algorithms for weighted and unweighted borda manipulation problems. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, pages 1327–1328, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.