Computational Complexity of Verifying the Group No-show Paradox

Farhad Mohsin[1] , **Qishen Han**[1] , **Sikai Ruan**[1] , **Pin-Yu Chen**[2] , **Francesca Rossi**[2] and **Lirong Xia**[1]
[1]Rensselaer Polytechnic Institute
[2]IBM Research

Abstract

The *(group) no-show paradox* [Fishburn and Brams, 1983] refers to the undesirable situation where a group of agents have incentive to abstain from voting to make the winner more favorable to them. To understand whether it is a critical concern in practice, in this paper, we take a computational approach by examining the computational complexity of verifying whether the group no-show paradox exists given agents' preferences and the voting rule. We prove that, unfortunately, the verification problem is NP-hard to compute for some commonly studied voting rules, i.e., Copeland, maximin, single transferable vote, and all Condorcetified positional scoring rules such as Black's rule. We propose integer linear programming-based algorithms and a depth first search algorithm for the verification problem for different voting rules. Experimental results on synthetic data illustrate that the former work better for a small number of alternatives, and the latter work better for a small number of agents. With the help of these algorithms, we observe that group no-show paradoxes rarely occur on real-world data.

1 Introduction

In social choice theory, the *no-show paradox*, first observed by Fishburn and Brams [1983], generally refers to the counter-intuitive event where a group of agents have an incentive to abstain from voting to make the winner more favorable to them. This is undesirable because when it occurs, agents can manipulate the result just by not showing up, which is much easier (thus much more threatening) than strategic manipulation [Gibbard, 1977; Satterthwaite, 1975]. The no-show paradox also discourages voters from participating in the election, reducing turnout and undermining democracy.

The significance of the no-show paradox urges researchers to study its existence under different voting rules. Unfortunately, even the *single-voter no-show paradox*, which means that there exists a group of one voter who has incentive to abstain from voting, always exists under a wide range of voting rules, including all Condorcet rules [Moulin, 1988]. Consequently, to understand its practical relevance, there is an extensive literature on verifying the frequency of various kinds of no-show paradox under specific distributions over voters for commonly-studied voting rules. For example, Ray [1986] studied the likelihood of a *group no-show paradox* for three alternatives under STV. Lepelley and Merlin [2001] generalized the concept from Ray [1986] and computed the frequency of various kinds of group no-show paradoxes under popular distributions for scoring run-off methods. Brandt *et al.* [2019] theoretically characterized the likelihood of the no-show paradox on Condorcet rules via Ehrhart theory. Xia [2021] characterized the likelihood that the single-voter no-show paradox does not exist under a semi-random model.

While verifying the existence of single-voter no-show paradox for many commonly-studied voting rules can be done in polynomial time, by simply enumerating the possible absentee, the following natural question remains open: **what is the computational complexity of verifying group no-show paradox under commonly studied voting rules?**

The question is not only interesting from a theoretical perspective, but also significant from a practical viewpoint. A low complexity is desirable, as it would allow us to verify whether the group no-show paradox is a significant concern in practice, and therefore, provides a computational basis for selecting a most robust voting rule against voter abstention.

Our contributions. We characterize the computational complexity of verifying group no-show paradox (GNSP) under several commonly studied voting rules: Copeland, maximin, STV, and all Condorcetified positional scoring rules, including Black's rule (Section 3). We prove that, unfortunately, the verification problem is NP-complete under all of them (Theorem 1–4). To computationally solve the problem, we propose integer linear program (ILP)-based algorithms and a depth first search (DFS) algorithm for verifying GNSP for these voting rules (Section 4). We perform experiments on both synthetic data and PrefLib election data [Mattei and Walsh, 2013] (Section 5). The results on synthetic data illustrate that the ILP algorithms work better for a small number of alternatives, and the DFS algorithm works better for a small number of agents. The results on PrefLib data suggest that no-show paradoxes rarely occur in real-world elections.

1.1 Related Works and Discussion

The no-show paradox. Although it is widely acknowledged that the no-show paradox refers to agent(s) having incentives to abstain from voting, its mathematical definition varies in the literature. Fishburn and Brams [1983], who introduced the paradox, described it as a group of agents having incentive to abstain from voting, but restricted the votes to be "identical" and the full-vote winner being "ranked last". On the other hand, Moulin [1988] restricted the no-show paradox to a single agent while relaxing the "ranked last" constraint, and call the non-existence of such singe-agent no-show paradox satisfaction of *partitipation*. Lepelley and Merlin [2001] investigated the group version as Fishburn and Brams [1983] did, while further redefining the paradox to four specific types. In this paper, we adopt a definition which inherits the spirits of all three papers: we consider the group version of no-show paradox as done by Fishburn and Brams [1983] and Lepelley and Merlin [2001], while not requiring the votes to be identical or the alternative to be ranked at a certain place, as done by Moulin [1988]. There is a large body of literature on the likelihood of no-show paradox. [Brandl *et al.*, 2015] shows that every Pareto Optimal majoritarian votes is prone to the no-show paradox. [Brandt *et al.*, 2016] follows [Moulin, 1988] and tighten the bound of the number of agents in a no-show paradox in Condorcet-consistent rules. [Pérez, 2001] and [Duddy, 2014] studied strong versions of no-show paradox's likelihood in Condorcet rules. Kamwa *et al.* [2021] revisited [Lepelley and Merlin, 2001]'s setting under a single-peaked preference, and found that no-show paradox occurs with a much lower probability under this restriction.

Manipulation and control. The no-show paradox is closely related to *manipulation* in voting, where some agents want to change the outcome by strategically misreporting their preferences. The no-show paradox can be viewed as performing "manipulation by abstention" [Brandt *et al.*, 2021], which a special case of "sincere truncation", i.e., agents partially reveal their truthful preference [Fishburn and Brams, 1984]. Technically, verifying GNSP is related to the complexity of the COALITIONAL MANIPULATION (CM) problem in voting [Bartholdi and Orlin, 1991; Conitzer and Walsh, 2016]. In CM, we are given the voting rule, the preferences of the sincere agents, a group of manipulators, and their favorable alternative c. We are asked whether the manipulators can cast votes to make c the winner. One might be tempted to think that verifying GNSP is easier than CM. However, we don't see a formal relationship between the two problems, because, first, when verifying GNSP, the group of "manipulators" are not fixed; and second, all absentees must prefer the new winner to the old winner. The no-show paradox is also related to *control* in voting, or more specifically, control by deleting agents [Bartholdi *et al.*, 1992; Faliszewski and Rothe, 2016], where an adversary aims at achieving a goal, e.g., make a favorable alternative win or an unfavorable alternative lose, by deleting agents, often under a budget constraint. The main differences between GNSP and control are similar to the differences between GNSP and CM: in GNSP, the size of the "deleted" agents is unbounded, and only agents who prefer the new winner to the old winner can be deleted.

2 Preliminaries

For any $m \in \mathbb{N}$, let $\mathcal{A}$ denote the set of $m \geq 3$ *alternatives*. Let $\mathcal{L}(\mathcal{A})$ denote the set of all linear orders or rankings over $\mathcal{A}$. Let $n \in \mathbb{N}$ denote the number of agents. Each agent uses a linear order $R \in \mathcal{L}(\mathcal{A})$ to represent his or her preferences, called a *vote*, where $a \succ_R b$ means that the agent prefers alternative a to alternative b. The vector of n agents' votes, denoted by P, is called a *(preference) profile*, sometimes called an n-profile. In this paper we focus on *resolute voting rules*, which always choose a unique winner. Formally, a *resolute voting rule* $r : \mathcal{L}(\mathcal{A})^* \to \mathcal{A}$ maps a profile to a single alternative in $\mathcal{A}$. A voting rules is *anonymous* if the winner is insensitive to identities of agents.

(Un)weighted majority graphs and Condorcet winners. For any profile P and any pair of alternatives a, b, let $P[a \succ b]$ denote the total number of votes in P where a is preferred to b. Let $\text{WMG}(P)$ denote the *weighted majority graph* of P, whose vertices are $\mathcal{A}$ and whose weight on edge $a \to b$ is $w_P(a,b) = P[a \succ b] - P[b \succ a]$. Let $\text{UMG}(P)$ denote the *unweighted majority graph*, which is the unweighted directed graph that is obtained from $\text{WMG}(P)$ by keeping the edges with strictly positive weights. The *Condorcet winner* of a profile P, denoted by $\text{CW}(P)$, is the alternative that only has outgoing edges in $\text{UMG}(P)$.

Integer positional scoring rules. An *(integer) positional scoring rule* $r_{\vec{s}}$ is characterized by an integer scoring vector $\vec{s} = (s_1, \ldots, s_m) \in \mathbb{Z}^m$ with $s_1 \geq s_2 \geq \cdots \geq s_m$ and $s_1 > s_m$. For any alternative a and any linear order $R \in \mathcal{L}(\mathcal{A})$, let $\vec{s}(R, a) = s_i$, where i is the rank of a in R. Given a profile P with weights $\vec{\omega}_P$, the positional scoring rule $r_{\vec{s}}$ chooses the alternative a with maximum $\sum_{R \in P} \omega_R \cdot \vec{s}(R, a)$. For example, the scoring vector is $(1, 0, \ldots, 0)$ for *plurality* and $(m-1, m-2, \ldots, 0)$ for *Borda*. When multiple alternatives achieve the maximum total score, a *tie-breaking mechanism* is applied to choose the winner.

STV. The single transferable vote (STV) rule chooses the winner in at most $m-1$ rounds. For each $1 \leq i \leq m-1$, in round i, a loser (an alternative with the lowest score) under plurality is eliminated. If multiple alternatives are tied as losers in any round, we will apply a tie-breaking mechanism to choose the (single) loser to be eliminated.

Copeland. The Copeland rule is parameterized by a number $0 \leq \alpha \leq 1$, and is therefore denoted by Cd_α. For any profile P, an alternative a gets 1 point for each other alternative it beats in head-to-head competitions, and gets α points for each tie. Cd_α chooses the alternative with the highest Copeland score as the winner, sometimes using a tie-breaking mechanism.

Maximin. For each alternative a, its *min-score* is defined to be $\text{MS}_P(a) = \min_{b \in \mathcal{A}} w_P(a, b)$. Maximin, denoted by MM, chooses the alternative with max min-score as the winner, sometimes using a tie-breaking mechanism.

Condorcetified (integer) positional scoring rules. The rule is defined by an integer scoring vector $\vec{s} \in \mathbb{Z}^m$ and is denoted by $\text{Cond}_{\vec{s}}$, which selects the Condorcet winner when it exits, and otherwise uses $r_{\vec{s}}$ to select the winner. For example, *Black's rule* [Black, 1958] is the Condorcetified Borda rule.

Group no-show Paradox. As discussed in Section 1.1, we adopt the following definition that inherits the spirits of [Fishburn and Brams, 1983] and [Moulin, 1988].

Definition 1 (**Group no-show Paradox).** *A* group no-show paradox, *denoted by* GNSP, *occurs in a profile P under a resolute voting rule r, if there exists a subset of agents $P' \subseteq P$, each of which prefers $r(P - P')$ to $r(P)$, thus giving them an incentive to abstain from voting.*

Example 1. Let P denote the profile of 14 votes with 6 votes of $[2 \succ 1 \succ 3]$, 4 votes of $[1 \succ 3 \succ 2]$, and 4 votes of $[3 \succ 2 \succ 1]$. As illustrated in Figure 1, $\text{Copeland}_{0.5}(P) = 1$ when the *lexicographic tie-breaking* $1 \rhd 2 \rhd 3$ is used. If group P' consisting of 2 votes of $[3 \succ 2 \succ 1]$ abstain from voting, then $\text{Copeland}_{0.5}(P - P') = 2$. Notice that $2 \succ 1$ for both agents in P'. This means that no-show paradox occurs in $\text{Copeland}_{0.5}$ at P. □

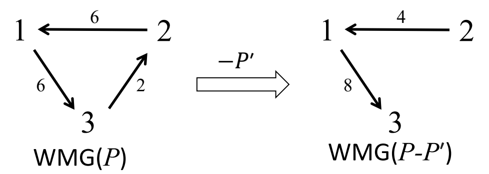

Figure 1: GNSP under $\text{Copeland}_{0.5}$.

3 Complexity of Verifying GNSP

In this section, we investigate the computational complexity of computing the existence of group no-show paradox for Copeland, Maximin, Condocetified positional scoring rules and STV. No-show paradoxes trivially do not occur for positional scoring rules, so we do not discuss them here. The problem is formally defined as GNSP-r as follows:

Definition 2 (GNSP-$\boldsymbol{r}$**).** *Given a voting rule r,* GNSP-r *is the computational problem taking a profile P as an input and outputs whether there exists a subset of agents $P' \subseteq P$, each of which prefers $r(P - P')$ to $r(P)$.*

The definition of GNSP-r immediately implies the following easiness result for fixed m.

Proposition 1. *For any fixed m and anonymous voting rule r,* GNSP-r *can be solved in polynomial time.*

The proposition holds because all possible anonymous profiles (counting how many votes each linear order receives) can be enumerated in polynomial time.

In the remainder of this section, we assume a variable m, and investigate the complexity of GNSP-r for voting rules using the following tie-breaking mechanisms. The *lexicographic tie-breaking* (LEX) breaks ties alphabetically. Fixed-agent (FA) tie-breaking chooses a fixed agent's preference (w.l.o.g., agent 1) to break ties. Most popular singleton ranking tie-breaking (MPSR) [Xia, 2020] first tries to use the linear order that uniquely occurs most often in the profile, and if such linear order does not exist, a backup tie-breaking mechanism is used. For example, MPSR+LEX uses LEX as the backup. See Appendix A for the formal definitions.

We are now ready to present the main results of this paper. Due to the space constraint, we only present the proof or a sketch under LEX. The full proofs can be found in Appendix B.

Theorem 1 (**Copeland).** *For any $0 \leq \alpha \leq 1$,* GNSP-Cd_α *is NP-complete to compute, where the tie-breaking mechanism is* LEX, FA, MPSR+LEX *or* MPSR+FA.

Proof sketch. It is easy to check that the problem is in NP because it is easy to check whether a given set of agents have an incentive to abstain from voting. The NP-hardness is proved by a reduction from RXC3, which is a restriction of EXACT 3 COVER that requires every element to be in exactly three sets and is proved to be NP-complete [Gonzalez, 1985].

In a RXC3 instance, we are given (1) a set of q-elements, denoted by $X = \{x_1, \ldots, x_q\}$, where q is divisible by 3; (2) q sets $\mathcal{S} = \{S_1, \ldots, S_q\}$ such that for every $j \leq q$, $S_j \subseteq X$ and $|S_j| = 3$. Moreover, for every $i \leq q$, x_i is in exactly three sets in $\mathcal{S}$. W.l.o.g. we assume that q is an even number. If q is odd, then we use an instance with duplicate X and $\mathcal{S}$.

We first prove the NP-hardness for $\alpha < 1$. For any RXC3 instance $(X, \mathcal{S})$ with $\alpha < \frac{q-4}{q-3}$, we construct a GNSP-Cd_α instance and $q+2$ alternatives as follows.

Alternatives: there are $q+2$ alternatives $\{1, 2, 3, \ldots, q+2\}$, where for every $3 \leq i \leq q+2$, alternative i corresponds to x_{i-2}. For convenience, we will use i and x_{i-2} interchangeably.

Profile: Let profile $P = P_1 \cup P_2 \cup P_3 \cup P_4$ consist of the following four parts, whose WMGs are illustrated in Figure 2.

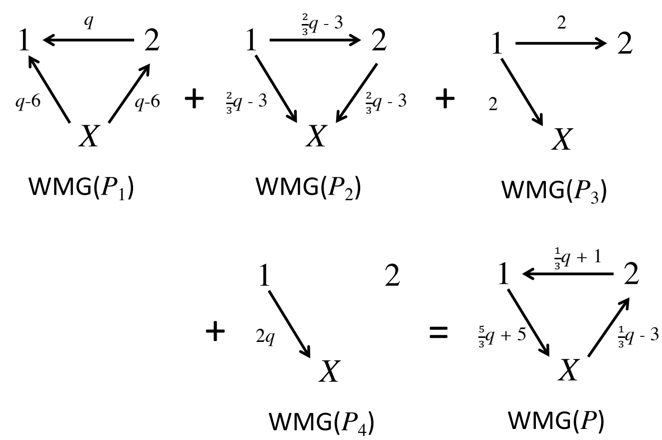

Figure 2: WMG of P for for Cd_α with $\alpha < 1$.

• **Part 1:** P_1 consists of q votes that correspond to the sets in $\mathcal{S}$: for every $j \leq q$, there is a vote R_S defined as follows

$$R_S = (X \setminus S_j) \succ 2 \succ 1 \succ S_j,$$

where alternatives in $(X \setminus S_j)$ and in S_j are ranked alphabetically. More precisely, $P_1 = \{R_S : S \in \mathcal{S}\}$.

• **Part 2:** P_2 consists of the $\frac{2}{3}q - 3$ copies of $[1 \succ 2 \succ X]$.

• **Part 3:** P_3 consists of the the following pair of votes

$$\{[1 \succ X \succ 2], [1 \succ 2 \succ X]\}$$

• **Part 4:** P_4 consists of q copies of the following pair of votes

$$\{[1 \succ X \succ 2], [2 \succ 1 \succ X]\}$$

It is not hard to verify that $\text{Cd}_\alpha(P) = 1$ due to tie-breaking.

Suppose the RXC3 instance is a yes instance, i.e., it has a solution $\mathcal{S}^*$. Then, GNSP-Cd_α is a yes instance, because after agents in P_1 that correspond to the 3-sets in $\mathcal{S}^*$ abstain from voting, alternative 2 becomes the Condorcet winner. Notice that all agents in P_1 prefer 2 to 1. This constitutes a group no-show paradox.

Suppose GNSP-Cd_α is a yes instance and a group of agents, denoted by P^*, have an incentive to abstain from voting. We will show that the RXC3 instance is a yes instance in the following steps.

First, $\text{Cd}_\alpha(P - P^*) = 2$. When the alternative winner is $a \neq 2$, only agents in P_1 have an incentives to abstain. Notice that no matter how many votes in P_1 are removed, 1 beats all alternatives in X in their head-to head competition and gets Copeland score of at least q, while a beaten by 1 gets at most q. Therefore, a cannot be the winner because the tie-breaking mechanism favors 1.

Second, $|P^*| \leq \frac{q}{3}$. If this is not true, $\text{WMG}(P - P^*)$ is non-negative, because the removal of any vote in P^* increases the weight on $1 \to 2$ by 1. Notice that $P^* \subseteq P_1 \cup P_4$. Therefore, 1 beats all alternatives in X in $\text{WMG}(P-P^*)$ and gets Copeland score of at least $q + \alpha$, while 2 gets Copeland score of at most $q + \alpha$. Therefore, 2 cannot beat 1 due to the tie breaking mechanism, which is a contradiction.

Third, $P^* \subseteq P_1$ and corresponds to an exact cover of X. Notice if there exists $a \in X$ such that $2 \succ X$ in more than one vote in P^*, then a beats 2 in $P - P^*$, and 2's Copeland score is not strictly larger than 1's score. Therefore, all votes in P^* come from P_1, and for any two votes in P^*, whose corresponding sets denoted as S_i and S_j, $S_i \cap S_j = \emptyset$. Now, if $|P^*| \leq \frac{q}{3} - 1$, then the number of alternatives 2 is defeated by or tied with in $\text{WMG}(P - P^*)$ is at least $|X| - 3 = q - 3$. This means that the Copeland score of 2 in $P - P^*$ is at most $4 + \alpha(q - 3)$, while the score of 1 is at least q. Recall that we assumed $\alpha < \frac{q-4}{q-3}$. This means that $\text{Cd}_\alpha(P - P^*) \neq 2$, which is a contradiction. Therefore, we have $|P^*| = \frac{q}{3}$, and for every $a \in X$, $2 \succ a$ in at most one vote in P^*. This means that $P^* \subseteq P_1$ and $\mathcal{S}^* = \{S \in \mathcal{S} : R_S \in P^*\}$ constitutes a 3-cover of X, which means that the RXC3 instance is a yes instance.

$\boldsymbol{\alpha = 1.}$ The profile P is similar to that in the proof for the $\alpha < 1$ except (1) 1 and 2 are switched in all votes, and (2) P_2 consist of $\frac{2}{3}q - 2$ copies of the vote, and (3) P_3 consists of $\{[2 \succ X \succ 1], [X \succ 2 \succ 1]\}$. We have $\text{Cd}_\alpha(P) = 2$, and in order for 1 to win, exactly $\frac{q}{3}$ agents in P_1 that correspond to an exact cover of X need to abstain from voting, so that in the new WMG, 1 is tied with every alternative, and therefore beats 2 according to the tie-breaking mechanism.

The full proof (for other tie-breaking mechanisms) can be found in Appendix B.1. □

The next theorem (about Condorcetified integer positional scoring rules) works for any tie-breaking mechanism.

Theorem 2 (**Condorcetified rules).** *For any Condorcetified integer positional scoring rule $\text{Cond}_{\vec{s}}$ and any tie-breaking rule,* GNSP-$\text{Cond}_{\vec{s}}$ *is NP-complete to compute.*

Proof. The proof is similar to the proof of the $\alpha < 1$ case of Theorem 1. Specifically, the NP hardness is proved by constructing a profile $P = P_1 \cup P_2 \cup P_3 \cup P_4$, where the main difference is the definition of P_4. Let σ^i denote the permutation $i \to (i+1) \to \cdots \to (q+2) \to 3 \to 4 \to \cdots \to (i-1)$, i.e., 1 and 2 are not changed. Let P_4 denote the $2q^2$-profile that consists of q copies of the following $2q$ votes:

$$\{\sigma^i(1 \succ X \succ 2), \sigma^i(2 \succ 1 \succ X) : 1 \leq i \leq q\}.$$

It is not hard to check that there is no Condorcet winner, and alternative 1 is the winner under the positional scoring rule with the unique highest score (see Appendix C for detailed proof). For any alternative $a \neq 1$, removing a vote where $a \succ 1$ does not reduce the score difference between 1 and a. Therefore, if group no-show paradox happens by removing $P^* \subseteq P$, then the winner in $P - P^*$ must be a Condorcet winner. Following a similar reasoning as in the proof of Theorem 1, this happens if and only if 2 is the Condorcet winner in $P - P^*$, and P^* corresponds to a 3-cover of the RXC3 instance. □

Theorem 3 (**Maximin).** GNSP-*MM is NP-complete to compute, where the tie-breaking mechanism is* LEX, FA, MPSR+LEX *or* MPSR+FA.

Proof sketch. The proof is similar to the proof of the $\alpha < 1$ case of Theorem 1. Specifically, the NP hardness is proved by the following reduction to RXC3. The full proof can be found in Appendix B.2.

Given a RXC3 instance $X, \mathcal{S}$, we construct the following GNSP-MM with $q + 4$ alternatives $\{1, 2, 3, 4, 5, \ldots, q + 4\}$, where for every $4 \leq i \leq q + 4$, alternative i corresponds to x_{i-4}. The alternatives consists of four parts, whose WMG is illustrated in Figure 3.

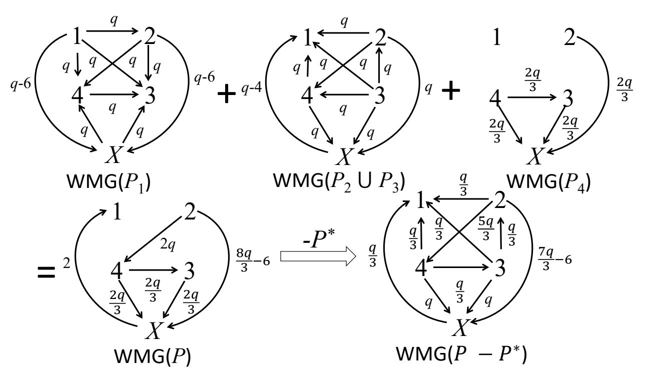

Figure 3: WMG of P for maximin.

Specifically, P_1 consists of q votes that correspond to the sets in $\mathcal{S}$: for every $j \leq q$, there is a vote R'_S defined as

$$R'_S = S_j \succ 1 \succ 2 \succ (X \setminus S_j) \succ 4 \succ 3.$$

It is not hard to verify that $\text{MM}(P) = 2$, whose min-score is 0 (via alternatives 1 and 3).

When the RXC3 instance is a yes instance with solution $\mathcal{S}^*$, let $P^* \subset P_1$ denote the votes that correspond to $\mathcal{S}^*$. Then the WMG for $(P - P')$ is illustrated in Figure 3. It's no hard to find that 1,2, and 3 share the max min-score of $-\frac{q}{3}$. Due to the lexicographic tie-breaking, $\text{MM}(P - P^*) = 1$, which is a GNSP yes instance. When GNSP-MM is a yes instance with abstention group P^*, following a similar reasoning as in the proof of Theorem 1, we can find that this happens only when $\text{MM}(P - P^*) = 1$ and P^* corresponds to a 3-cover of the RXC3 instance. □

Theorem 4 (STV). *For any $0 \leq \alpha \leq 1$,* GNSP-*STV is NP-complete to compute, where the tie-breaking mechanism is* LEX, FA, MPSR+LEX *or* MPSR+FA.

Proof sketch. The hardness is proved by a reduction from RXC3 that is similar to the reduction in the hardness proof for the manipulation problem under STV [Bartholdi and Orlin, 1991]. For any RXC3 instance $(X, \mathcal{S})$, where $X = \{x_1, \ldots, x_q\}$ and $\mathcal{S} = \{S_1, \ldots, S_q\}$, we construct the following GNSP-STV instance. The full proof can be found in Appendix B.3.

Alternatives: there are $3q + 3$ alternatives $\{w, c\} \cup \{d_0, d_1, \ldots, d_q\} \cup \{b_1, \bar{b}_1, \ldots, b_q, \bar{b}_q\}$. We assume that b_i has higher priority than $\bar{b}_i$, and $d_1 \succ d_2 \succ \cdots \succ d_q$ in tie-breaking.

Profile: The profile P consists of the following votes, where the top preferences are specified and the remaining alternatives (denoted by X) are ranked arbitrarily. Both i and j in the table are in $\{1, 2, \ldots, q\}$.

	Number of agents	Votes
P_1	$12q$	$c \succ w \succ X$
P_2	$12q - 1$	$w \succ c \succ X$
P_3	$10q + \frac{2q}{3}$	$d_0 \succ w \succ c \succ X$
P_4	$12q - 2$ for each j	$d_j \succ w \succ c \succ X$
P_5	$6q + 4i - 2$ for each i	$b_i \succ \bar{b}_i \succ w \succ c \succ X$
	2 for each i	$b_i \succ d_0 \succ w \succ c \succ X$
P_6	$6q + 4i - 6$ for each i	$\bar{b}_i \succ b_i \succ w \succ c \succ X$
	2 for each $j \in S_i$	$\bar{b}_i \succ d_j \succ w \succ c \succ X$
P_7	1 for each i	$\bar{b}_i \succ b_i \succ c \succ w \succ X$

Note that $\text{STV}(P) = w$. In the first q rounds, the order of elimination is $b_1, b_2, \ldots, b_q$ (whose votes are transferred to $\bar{b}_1, \ldots, \bar{b}_q$ and d_0). At the beginning of round $q + 1$, d_q is eliminated, whose votes transfer to w. In the remaining rounds w is never eliminated and will become the winner.

Suppose the RXC3 is a yes instance (with solution $\mathcal{S}^*$). Let $I = \{i \leq q : S_i \in \mathcal{S}^*\}$, then agents in P_7 whose top choices are b_i such that $i \in I$ have an incentive to (jointly) abstain from voting. After they abstain from the voting, in the first q rounds, for each $i \leq q$, $\bar{b}_i$ is eliminated if and only if $i \in I$, otherwise b_i is eliminated. In the beginning of round $q + 1$, the plurality scores of the remaining alternatives are as in the following table. Therefore, w is eliminated in round $q + 1$, whose votes transfer to c. Finally c will be the winner.

Rd.	w	c	b_i or $\bar{b}_i$	d_0	d_j
$q + 1$	$12q - 1$	$12q$	$12q + 8i - 1$ or $12q + 8i - 5$	$12q$	$12q$

Suppose the GNSP-STV instance is a yes instance. We prove that the RXC3 instance is a yes instance in the following two steps.

First, notice that the new winner must be c. Notice that c and w are adjacent in all votes. Therefore, q is eliminated, all of its votes are transferred to c;, and c is ranked at the top in at least $24q - 1$ votes (in $P_1 \cup P_2$). This make c the new winner.

Second, the absent votes in P_7 constitutes a solution to the RXC3 instance. First note that only agents in P_1 and P_7 have an incentives to abstain. Let I denote the indices i's of agents abstain from voting in P_7 whose top-ranked preferences are $\bar{b}_i$. In round $q + 1$, if $\mathcal{S}^* = \{S_i : i \in I\}$ is not a RXC3 solution, d_j will be eliminated for some j not in any S_i. Therefore $\mathcal{S}^*$ must be a solution. □

4 Algorithms for Verifying GNSP

We first present a general depth first search (DFS) algorithm for any rule (with any tie-breaking mechanism), giving us a baseline. Then, we present integer linear programming (ILP)-based algorithms for different anonymous rules.

DFS algorithm. The DFS algorithm (described in detail in Algorithm 1 in Appendix D) enumerates combinations of group abstentions in a depth-first manner. In the worst-case, this algorithm explores 2^n states for n agents, enumerating all possible combinations of groups, making it inefficient to use for a large n.

ILP-based algorithms. To circumvent the computational inefficiency faced in the DFS approach for large n, we consider ILP-based algorithms for determining the GNSP-r problem. We present ILP formulations of the group participation problem for four different voting rules – Copeland, maximin, Black's rule, and STV. The ILP formulations also output the smallest group size that can abstain from voting to change the outcome for any P.

First, we define the variables for the ILP formulations. Assume, that for preference profile P and voting rule r, $r(P) = a$. To verify if GNSP for any alternative $b \neq a$, we need to check all agents j with preference $b \succ_{V_j} a$. For any linear order $R_i \in \mathcal{L}(\mathcal{A})$, assume that $|\{j \mid V_j = R_i\}| = n_i$. That is, n_i is the number of agents with preference R_i. Let $V_{b \succ a} = \{R \mid b \succ_R a\}$. Now, for any linear order $R_i \in V_{b \succ a}$ agents may strategically abstain from voting. We denote x_i as the actual number of agents who vote R_i, with $x_i \leq n_i$. Thus, $\mathbf{x} = \{x_i\}_{R_i \in V_{b \succ a}}$ are the variables for the ILPs. We call P_x the alternate preference profile that is created by x_i agents voting for each linear order, $R_i \in V_{b \succ a}$. For a preference profile P, the group no-show paradox occurs for voting rule r, when there is a solution, $\mathbf{x}$ such that $r(P_x) = b$ for some $b \neq a$. Minimizing the objective function $n - \sum_{R_i \in V_{b \succ a}} x_i$ gives the smallest group size with incentive to abstain.

Based on these variables, we can define specific ILPs for each voting rule that we mentioned. We describe the ILPs for Copeland here and only discuss the rest briefly.

Copeland. For any alternative $b \neq a$, if there exists a feasible solution $\mathbf{x}$ for any UMG such that $b = \text{Cd}_\alpha(P_x)$, that means GNSP has occurred. Now, In any UMG, for any alternatives c, d, there can be three scenarios: (1) $P_x[c \succ d] >$

$P_x[d \succ c]$, (2) $P_x[c \succ d] < P_x[d \succ c]$, (3) $P_x[c \succ d] = P_x[d \succ c]$. $P_x[c \succ d]$ for any c, d can be defined as follows-

$$P_x[c \succ d] = \sum_{R_i \in V_{b \succ a}} x_i \mathbb{1}_{c \succ d}(R_i) + \sum_{R_i \notin V_{b \succ a}} n_i \mathbb{1}_{c \succ d}(R_i) \tag{1}$$

Here $\mathbb{1}_{c \succ d}(R_i)$ is an indicator random variable for whether $c \succ d$ in linear order R_i.

Thus, any UMG gives a set of linear constraints. We enumerate all UMGs where a is not the Copeland winner, construct the corresponding ILP, and check for a solution. The process is summarized in Algorithm 2 in Appendix D.

With this formulation, for each P, we need to consider $O(3^{\frac{m(m-1)}{2}})$ ILPs, each with $O(m^2)$ constraints.

Other voting rules. For maximin voting rule, we again get constraints based on the pairwise comparisons between alternatives,. For Black's rule, for UMGs with Condorcet winners, the ILP is the same as for Copeland, and in other cases there are additional constraints for the Borda winner. ILP formulations for no-show paradoxes was developed by Lepelley *et al.* [1996] for 3 alternatives. We generalize the idea of having a set of constraints for each round of voting in a multi-round voting rule for any number of alternatives. That gives us the ILP formulations for STV voting rule. The ILP formulations and corresponding algorithms for these voting rules can be found in Appendix D.

5 Experiments

In this section, we present results on the performance of the two algorithms and the likelihood of GNSP in both synthetic and real world election data. All experiments were implemented in Python 3 and were run on a Windows laptop with 3.2 GHz AMD Ryzen 7 5800 CPU and 16 GB memory.

Synthetic datasets. We sample 1,000 preference profiles under the i.i.d. uniform distribution (known as *Impartial Culture*, IC for short) and Mallows distributions with different dispersion values ϕ, for $n \in [10, 100]$ and $m \in [3, 15]$. Each instance has a 20 minutes timeout limit. Due to the space limitation, we only present results for Copeland$_{0.5}$, while the conclusions for other voting rules are similar and can be found in Appendix E.

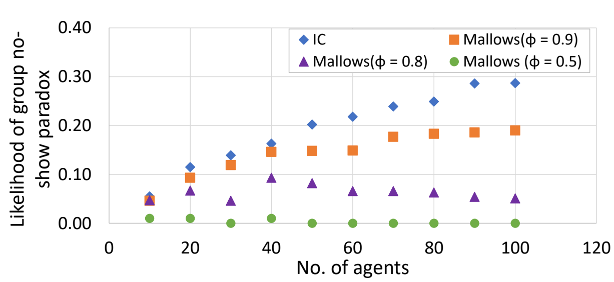

Figure 4: likelihood of GNSP under Copeland$_{0.5}$.

Likelihood of GNSP. Figure 4 shows the likelihood of GNSP under Copeland$_{0.5}$. It can be seen that more consensus among the agents (ϕ decreases—notice that IC is equivalent to Mallows with $\phi = 1$) leads to lower likelihood of GNSP. We also observe in our experiments that Copeland has the highest likelihood of GNSP, followed by STV, and then Black's rule, and maximin has the lowest likelihood, see Figure 8 in Appendix E for the comparison under IC.

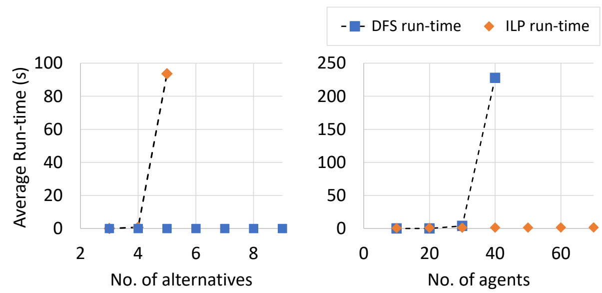

Figure 5: Run-time for the DFS and ILP algorithm. Left: Run-time vs No. of alternatives for $n = 10$. Right: Run-time vs No. of agents for $m = 4$.

Run-time of algorithms. Figure 5 (left) shows that the run-time of the DFS algorithm does not increase significantly with the number of alternatives, m, when n is small. On the other hand, the ILP-based algorithm becomes inefficient even for $m = 5$. Figure 5 (right) shows that the run-time for the ILP-based algorithm does not increase significantly with n for small m, while the DFS algorithm's run-time increases exponentially in n, and becomes prohibitively high even for $n = 50$. Additional experimental results (for other rules) can be found in Appendix E. These results illustrate that the ILP-based algorithms work better for a small number of alternatives, and the DFS algorithm works better for a small number of agents.

Real world data from PrefLib [Mattei and Walsh, 2013]. We use all available election data which have strict order-complete lists (SOC) for all agents. In total, there are 315 such profiles on PrefLib. We note that 208 of them have lesser than or equal to 40 agents, on which we could use the DFS algorithm, despite the often large alternative sets (up to 263 alternatives in a profile). On a different set of 98 profiles, we have lesser than or equal to 6 alternatives. but high number of agents (up to 14,000 agents in a profile). Out of the 306 observed profiles, only one profile each violates group participation for Copeland, Black's rule and STV, and we found no violations for maximin. This indicates that GNSP occurrences are rare in real election data.

6 Discussion and Future Work

We prove that group no-show paradox is computationally hard to verify for Copeland, maximin, STV, and Condorcetified integer positional scoring rules, and provided a DFS algorithm and ILP-based algorithms for computing them. Natural directions for future work include designing more efficient algorithms, e.g., based on AI/ML techniques, for verifying GNSP, and studying the complexity and algorithms for other undesirable events, such as (group) manipulation. In general, research in this direction helps establish a data-centric foundation for evaluating and designing voting rule.

References

John Bartholdi, III and James Orlin. Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8(4):341–354, 1991.

John J. Bartholdi, Craig A. Tovey, and Michael A. Trick. How hard is it to control an election? *Mathematical and Computer Modelling*, 16(8):27–40, 1992.

Duncan Black. *The Theory of Committees and Elections*. Cambridge University Press, 1958.

Florian Brandl, Felix Brandt, and Johannes Hofbauer. Incentives for Participation and Abstention in Probabilistic Social Choice. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, pages 1411–1419, Istanbul, Turkey, 2015.

F. Brandt, Christian Geist, and Dominik Peters. Optimal bounds for the no-show paradox via sat solving. *Math. Soc. Sci.*, 90:18–27, 2016.

Felix Brandt, Johannes Hofbauer, and Martin Strobel. Exploring the No-Show Paradox for Condorcet Extensions Using Ehrhart Theory and Computer Simulations. In *Proceedings of AAMAS*, pages 520–528, 2019.

Felix Brandt, Johannes Hofbauer, and Martin Strobel. Exploring the No-Show Paradox for Condorcet Extensions. In Mostapha Diss and Vincent Merlin, editors, *Evaluating Voting Systems with Probability Models*. Springer, 2021.

Vincent Conitzer and Toby Walsh. Barriers to Manipulation in Voting. In Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel Procaccia, editors, *Handbook of Computational Social Choice*, chapter 6. Cambridge University Press, 2016.

Conal Duddy. Condorcet's principle and the strong no-show paradoxes. *Theory and decision*, 77(2):275–285, 2014.

Piotr Faliszewski and Jörg Rothe. Control and bribery in voting. In *Handbook of Computational Social Choice*, chapter 7. Cambridge University Press, 2016.

Peter C. Fishburn and Steven J. Brams. Paradoxes of Preferential Voting. *Mathematics Magazine*, 56(4):207–214, 1983.

Peter C Fishburn and Steven J Brams. Manipulability of voting by sincere truncation of preferences. *Public Choice*, 44(3):397–410, 1984.

Allan Gibbard. Manipulation of schemes that mix voting with chance. *Econometrica*, 45:665–681, 1977.

Teofilo F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985.

Eric Kamwa, Vincent Merlin, and Faty Mbaye Top. Scoring Run-off Rules, Single-peaked Preferences and Paradoxes of Variable Electorate. working paper or preprint, February 2021.

Dominique Lepelley and Vincent Merlin. Scoring run-off paradoxes for variable electorates. *Economic Theory*, 17:53–80, 2001.

Dominique Lepelley, Frederic Chantreuil, and Sven Berg. The likelihood of monotonicity paradoxes in run-off elections. *Mathematica Slocial Science*, 31:133–146, 1996.

Nicholas Mattei and Toby Walsh. PrefLib: A Library of Preference Data. In *Proceedings of Third International Conference on Algorithmic Decision Theory*, Lecture Notes in Artificial Intelligence, 2013.

Hervé Moulin. Condorcet's principle implies the no show paradox. *Journal of Economic Theory*, 45(1):53–64, 1988.

Joaquín Pérez. The Strong No Show Paradoxes are a common flaw in Condorcet voting correspondences. *Social Choice and Welfare*, 18(3):601–616, 2001.

Depankar Ray. On the practical possibility of a 'no show paradox'under the single transferable vote. *Mathematical Social Sciences*, 11(2):183–189, 1986.

Mark Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.

Lirong Xia. The Smoothed Possibility of Social Choice. In *Proceedings of NeurIPS*, 2020.

Lirong Xia. The Semi-Random Satisfaction of Voting Axioms. In *Proceedings of NeurIPS*, 2021.

Supplementary material for Computational Complexity of Verifying the Group No-show Paradox

A Tie-breaking mechanisms

In this paper, we investigate three tie-breaking mechanisms. Each tie-breaking mechanism is defined by a pre-determined linear order, which indicates the order for ties to be broken.

First, the lexicographic tie-breaking method, denoted by LEX, breaks ties between alternatives alphabetically. So for a tie between 1 and 2, LEX will break the tie in favor of 1.

The *fixed-agent tie-breaking*, denoted by FA_j, is parameterized by $j \leq n$, which uses the j-th agent's vote to break ties. In this paper, we'll use always use agent 1's vote for fixed-agent tie-breaking (denoted as FA) w.l.o.g.

The *most popular singleton ranking tie-breaking* [Xia, 2020], denoted by MPSR, is parameterized by another "backup" tie-breaking mechanism. We define the most popular singleton ranking as follows.

Definition 3 (Most Popular Singleton Ranking)**.** *Given a profile P, we define its most popular singleton ranking as $MPSR(P) = \arg\max_R(P[R] : \nexists W \neq R \text{ s.t } P[R] = P[W])$, where $P[R]$ is the number of votes that linear order R receives in P.*

The MPSR tie-breaking method first tries calculates the MPSR. If such a ranking (called the most popular singleton ranking) exists, then it is used to break ties. If a unique MPSR does not exist for a profile P, then a backup tie-breaking mechanism $\mathcal{T}$ is used. In this paper we mainly consider two MPSR tie-breaking mechanisms: MPSR+LEX and MPSR+FA, where lexicographic and fixed-agent tie-breaking mechanisms are used as the backup, respectively.

Here we give an example to illustrate how these tie-breaking mechanisms work.

Example 2 (Tie breaking mechanisms)**.** Suppose there are 3 alternatives in a vote with Plurality rule. The profile P consists of 6 votes: $P = \{V_1, V_2, V_3, V_3, V_4, V_5\}$, where $V_1 = [2 \succ 1 \succ 3]$, $V_2 = [2 \succ 3 \succ 1]$, $V_3 = [3 \succ 2 \succ 1]$, $V_4 = [1 \succ 2 \succ 3]$, and $V_5 = [1 \succ 3 \succ 2]$. We notice that there is a three-way tie in which alternative 1, 2, and 3 are all voted top twice.

- When the tie-breaking mechanism is LEX, the tie is broken in order $1 \rhd 2 \rhd 3$, and alternative 1 is the winner.
- When the tie-breaking mechanism is FA_1, the tie is broken in order of V_1, i.e., $2 \rhd 1 \rhd 3$. Then alternative 2 is the winner.
- When the tie-breaking mechanism of is MPSR+LEX or MPSR+FA, we first find the most popular singleton ranking in P. Note that V_3 occurs twice while other votes occurs only once. Therefore, the tie is broken in order $3 \rhd 2 \rhd 1$, and alternative 3 will be the winner.

Then we consider another case where the profile $P' = \{V_1, V_1, V_3, V_3, V_4, V_4\}$. While the tie-breaking results of LEX and FA_1 are the same with those in P, P' does not have a most popular singleton ranking since every votes appear twice. Therefore, when the tie breaking mechanism is MPSR, a backup mechanism is applied. For example, MPSR+LEX will choose 1 to be the winner, and MPSR+FA will choose 2.

B Full Proof of the Theorems

B.1 Full Proof of Theorem 1(Copeland)

Theorem 1 (**Copeland**)**.** *For any $0 \leq \alpha \leq 1$,* GNSP-Cd_α *is NP-complete to compute, where the tie-breaking mechanism is* LEX, FA, MPSR+LEX *or* MPSR+FA.

Proof. It is easy to check that the problem is in NP because it is easy to check wither a given set of agents have incentive to abstain from voting. The NP-hardness is proved by a reduction from RXC3, which is a restriction of EXACT 3 COVER that requires every element to be in exactly three sets and is proved to be NP-complete [Gonzalez, 1985].

We first give the proof for LEX tie-breaking mechanism, and will comment on how to modify the proof for other tie-breaking mechanisms afterwards. In a RXC3 instance, we are given (1) a set of q-elements, denoted by $X = \{x_1, \dots, x_q\}$, where q is divisible by 3; (2) q sets $\mathcal{S} = \{S_1, \dots, S_q\}$ such that for every $j \leq q$, $S_j \subseteq X$ and $|S_j| = 3$. Moreover, for every $i \leq q$, x_i is in exactly three sets in $\mathcal{S}$. W.l.o.g. we assume that q is an even number. If q is odd, then we use an instance with duplicate X and $\mathcal{S}$.

We first prove the NP-hardness for $\alpha < 1$. For any RXC3 instance $(X, \mathcal{S})$ with $\alpha < \frac{q-4}{q-3}$, we construct a GNSP-Cd_α instance with $q + 2$ alternatives as follows.

Alternatives: there are $q + 2$ alternatives $\{1, 2, 3, \dots, q + 2\}$, where for every $3 \leq i \leq q + 2$, alternative i corresponds to x_{i-2}. For convenience, we will use i and x_{i-2} interchangeably.

Profile: Let profile $P = P_1 \cup P_2 \cup P_3 \cup P_4$ consist of the following four parts, whose WMGs are illustrated in Figure 2.

- **Part 1:** P_1 consists of q votes that correspond to the sets in $\mathcal{S}$: for every $j \leq q$, there is a vote R_S defined as follows

$$R_S = (X \setminus S_j) \succ 2 \succ 1 \succ S_j,$$

where alternatives in $(X \setminus S_j)$ and in S_j are ranked alphabetically. More precisely, $P_1 = \{R_S : S \in \mathcal{S}\}$.

- **Part 2:** P_2 consists of the $\frac{2}{3}q - 3$ copies of $[1 \succ 2 \succ X]$.
- **Part 3:** P_3 consists of the the following pair of votes

$$\{[1 \succ X \succ 2], [1 \succ 2 \succ X]\}$$

- **Part 4:** P_4 consists of q copies of the following pair of votes

$$\{[1 \succ X \succ 2], [2 \succ 1 \succ X]\}$$

It is not hard to verify that the WMG(P) contains the following edges:

- An edge $2 \to 1$ with weight $\frac{q}{3} + 1$.
- For every $x \in X$, there is an edge $1 \to x$ with weight $\frac{5}{3}q + 5$ and an edge $x \to 2$ with weight $\frac{q}{3} - 3$.

The weights on edges among alternatives in X are not relevant to the proof. It follows that $\text{Cd}_\alpha(P) = 1$ due to tie-breaking.

Suppose the RXC3 instance is a yes instance, i.e., it has a solution $\mathcal{S}^*$. Then, GNSP-Cd$_\alpha$ is a yes instance, because after voters in P_1 that correspond to the 3-sets in $\mathcal{S}^*$ abstain from voting, alternative 2 becomes the Condorcet winner. Notice that all voters in P_1 prefer 2 to 1. This constitutes a group no-show paradox.

Suppose GNSP-Cd$_\alpha$ is a yes instance and a group of voters, denoted by P^*, have incentive to abstain from voting. We will show that the RXC3 instance is a yes instance in the following steps.

First, $\text{Cd}_\alpha(P - P^*) = 2$. Suppose for the sake of contradiction this is not true. Let $a = \text{Cd}_\alpha(P - P^*)$ and $a \neq 2$. Because all voters in P^* prefer a to 1, we have $P^* \subseteq P_1$. Notice that no matter how many votes in P_1 are removed, 1 beats all alternatives in X in their head-to-head competitions. Therefore, the Copeland score of 1 is at least q. On the other hand, the Copeland score of a is at most q (because 1 beats a). Therefore, a cannot be the winner in $P - P^*$ because the tie-breaking mechanism favors alternative 1.

Second, $|P^*| \leq \frac{q}{3}$. Suppose for the sake of contradiction this is not true. Then, the weight on the edge $1 \to 2$ in WMG$(P - P^*)$ is non-negative, because the removal of any vote in P^* increases the weight on $1 \to 2$ by 1. Notice that $P^* \subseteq P_1 \cup P_4$, where 2 is ranked right above 1. Therefore, 1 beats all alternatives in X in WMG$(P - P^*)$. This means that the Copeland score of 1 is at least $q + \alpha$, which is not strictly smaller than the Copeland score of 2. This contradicts $\text{Cd}_\alpha(P - P^*) = 2$.

Third, $P^* \subseteq P_1$ and corresponds to an exact cover of X. We note that if 2 loses to any alternative in their head-to-head competition in $P - P^*$, then 2's Copeland score is not strictly larger than 1's Copeland score in $P - P^*$, which means that $\text{Cd}_\alpha(P - P^*) \neq 2$. Therefore, if there exists $a \in X$ such that $2 \succ X$ in more than one vote in P^*, then we have

$$\text{WMG}_{P-P^*}(2 \to a) \leq |P^*| - 4 - (\frac{q}{3} - 3) \leq -1,$$

which means that $\text{Cd}_\alpha(P - P^*) \neq 2$, which is a contradiction. This means all votes in P^* come from P_1.(Otherwise if P^* contains any vote in P_4, it cannot contain any other vote, which is impossible.) Moreover, for any two votes in P^*, whose corresponding sets denoted as S_i and S_j, we have $S_i \cap S_j = \emptyset$. Now, if $|P^*| \leq \frac{q}{3} - 1$, then the number of alternatives 2 is defeated by or tied with in WMG$(P - P^*)$ is at least $|X| - 3 = q - 3$. This means that the Copeland score of 2 in $P - P^*$ is at most $4 + \alpha(q - 3)$, and the Copeland score of 1 is at least q. Recall that we assumed $\alpha < \frac{q-4}{q-3}$. This means that $\text{Cd}_\alpha(P - P^*) \neq 2$, which is a contradiction.

Therefore, we have $|P^*| = \frac{q}{3}$, and for every $a \in X$, $2 \succ a$ in at most one vote in P^*. This means that $P^* \subseteq P_1$ and $\mathcal{S}^* = \{S \in \mathcal{S} : R_S \in P^*\}$ constitutes a 3-cover of X, which means that the RXC3 instance is a yes instance. This finished the NP-hardness proof for $\alpha < 1$.

$\boldsymbol{\alpha = 1.}$ The profile P is similar to that in the proof for the $\alpha < 1$ except (1) 1 and 2 are switched in all votes, and (2) WMG(P) is as indicated in Figure 6, which is achieved by letting P_2 consist of $\frac{2}{3}q - 2$ copies of the vote, and letting P_3 consists of $\{[2 \succ X \succ 1], [X \succ 2 \succ 1]\}$. We have $\text{Cd}_\alpha(P) = 2$, and in order for 1 to win, exactly $\frac{q}{3}$ voters in P_1 that correspond to an exact cover of X need to abstain from voting, so that in the new WMG, 1 is tied with every alternative, and therefore beats 2 according to the tie-breaking mechanism.

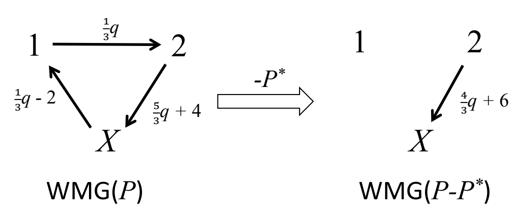

Figure 6: WMG of P for Cd_α with $\alpha = 1$.

Other tie-breaking mechanisms. For other tie-breaking mechanisms, we modify the construction to get a same tie-breaking order as LEX. Our modification for MPSR guarantees to find a most-popular preference, so we write MPSR-LEX and MPSR-FA together as MPSR.

FA. Let every voter of P_2 have the preference of $1 \succ 2 \succ 3 \succ \cdots \succ q+2$. Then switch P_1 and P_2. Then voter 1 becomes the first voter in P_2, with a order same as LEX.

MPSR Let every voter in P_2 has the preference of $1 \succ 2 \succ 3 \succ \cdots \succ q+2$. For other voters, notice that the preference of X has $q!$ different permutations. Therefore, we can assign every other voter a different permutation. (First, we assign each voter in P_1 a permutation of X that in accord with their current preference. Then for P_3 and P_4, we assign each of them a different permutation from the rest.) Therefore $1 \succ 2 \succ 3 \succ \cdots \succ q+2$ is the most popular voting and becomes the tie-breaking order.

Notice that P_2 voters have no incentives to abstain from the voting. Therefore, neither the first voter or the most popular voting won't change, so the tie-breaking order remain still before and after abstaining. □

B.2 Full Proof of Theorem 3 (Maximin)

Theorem 3 (Maximin). GNSP-*MM is NP-complete to compute, where the tie-breaking mechanism is* LEX, FA, MPSR+LEX *or* MPSR+FA.

Proof. It is easy to verify that the problem is in NP. The NP-hardness is proved by a reduction from RXC3. Given a RXC3 instance $X, \mathcal{S}$, we construct the following GNSP-MM with $q+4$ alternatives as follows.

Alternatives: there are $q+4$ alternatives $\{1, 2, 3, 4, 5, \ldots, q+4\}$, where for every $4 \leq i \leq q+4$, alternative i corresponds to x_{i-4}. For convenience, we will use i and x_{i-4} interchangeably.

Profile: Let profile $P = P_1 \cup P_2 \cup P_3 \cup P_4$ consist of the following four parts, whose WMGs are illustrated in Figure 3.

- **Part 1:** P_1 consists of q votes that correspond to the sets in $\mathcal{S}$: for every $j \leq q$, there is a vote R'_S defined as follows
$$R'_S = S_j \succ 1 \succ 2 \succ (X \setminus S_j) \succ 4 \succ 3,$$
where alternatives in $(X \setminus S_j)$ and in S_j are ranked alphabetically. More precisely, $P_1 = \{R'_S : S \in \mathcal{S}\}$.
- **Part 2:** P_2 consists of the $q-2$ copies of $[3 \succ 2 \succ 4 \succ X \succ 1]$.
- **Part 3:** P_3 consists of two copies of $[3 \succ 2 \succ 4 \succ 1 \succ X]$.
- **Part 4:** P_4 consists of $\frac{q}{3}$ copies of the following pair of votes
$$\{[1 \succ 4 \succ 3 \succ 2 \succ X], [2 \succ 4 \succ 3 \succ X \succ 1]\}$$

The weights on edges among alternatives in X are not relevant to the proof. It is not hard to verify that MM$(P) = 2$, whose min-score is 0 (via alternatives 1 and 3).

Suppose the RXC3 instance is a yes instance, i.e., it has a solution $\mathcal{S}^*$. Then, we show that GNSP-MM is a yes instance by letting $P' \subset P_1$ denote the votes that correspond to $\mathcal{S}^*$. That is,
$$P' = \{R'_S : S \in \mathcal{S}^*\}$$
Then, we make the following observations about the min-scores of alternatives in $P - P'$, whose WMG is also illustrated in Figure 3.

- The min-scores of $1, 2, 3$ are $-\frac{q}{3}$.
- The min-scores of 4 is $-\frac{5q}{3}$ (via 2).
- For any $x \in X$, the min-scores of x is at most $-\frac{7q}{3} + 6$ (via 2 or alternatives in X).

Therefore, due to the lexicographic tie-breaking, MM$(P - P') = 1$. Notice that all voters in P' prefer 1 to 2. This constitutes a group no-show paradox.

Suppose GNSP-MM is a yes instance and a group of voters, denoted by P^*, have incentive to abstain from voting. We will show that the RXC3 instance is a yes instance in the following steps.

First, MM$(P - P^*) = 1$. Equivalently, we prove that MM$(P - P^*) \notin (\{3, 4\} \cup X)$.

- Suppose for the sake of contradiction that MM$(P - P^*) = 3$. Then, because everyone in P^* prefers 3 to 2, P^* must be contained in $P_2 \cup P_3$ and the $\frac{q}{3}$ copies of $[1 \succ 4 \succ 3 \succ 2 \succ X]$ in P_4. Let $n_1 = |P^* \cap (P_2 \cup P_3)|$ and $n_2 = |P^* \cap P_4|$. It follows that the min-score of 2 in $P - P^*$ is $n_2 - n_1$ (via alternative 1), and the min-score of 3 in $P - P^*$ is at most $n_2 - n_1 - \frac{2q}{3}$ (via alternative 4). This means that 3 cannot be the maximin winner, which is a contradiction.
- Suppose for the sake of contradiction that MM$(P - P^*) = 4$. Then, because everyone in P^* prefers 4 to 2, P^* must be contained in the $\frac{q}{3}$ copies of $[1 \succ 4 \succ 3 \succ 2 \succ X]$ in P_4. Let $n^* = P^* \cap P_4$. It follows that the min-score of 2 in $P - P^*$ is n^* (via 1 and 3), and the min-score of 4 in $P - P^*$ is at most $-n^* - 2q$ (via alternative 2). This means that 4 cannot be the maximin winner, which is a contradiction.

- Suppose for the sake of contradiction that $\mathrm{MM}(P - P^*) = x \in X$. Then, because everyone in P^* prefers x to 2, $P^* \subseteq P_1$ and $|P^*| \leq 3$ because x appears in 3 sets in $\mathcal{S}$. Clearly the min-score of 2 is strictly larger than the min-score of x in $P - P^*$, which is a contradiction.

Second, w.l.o.g., $P^* \subseteq P_1$. Because everyone in P^* prefers 1 to 2, P^* is a subset of P_1 and the $\frac{q}{3}$ copies of $[1 \succ 4 \succ 3 \succ 2 \succ X]$ in P_4. Notice that removing a vote of $[1 \succ 4 \succ 3 \succ 2 \succ X]$ always reduces the min-score of 1 by one. Therefore, the min-score of 1 is not lower than the min-score of 2 in $P - (P^* \cap P_4)$, which mean that we can assume that $P^* \subset P_1$.

Third, P^* corresponds to a solution to the RXC3 instance. Let $\mathcal{S}^* = \{S : R'_S \in P^*\}$. We first note that $|P^*| \leq \frac{q}{3}$. Otherwise, the min score of 3 is $|P^*| - \frac{2q}{3}$ (via 4), which is strictly larger than $-|P^*|$, which is not smaller than the min-score of 1 (via 2 or alternatives in X). This contradicts the assumption that $\mathrm{MM}(P - P^*) = 1$. Now, suppose for the sake of contradiction that $\mathcal{S}^*$ is not a solution to the RXC3 instance, which means that an alternative x is not contained in any set in $\mathcal{S}^*$. Then, the min-score of 1 is $-|P^*| - 2$ (via x), which is strictly smaller than the min-score of 2 (via 3). This contradicts the assumption that $\mathrm{MM}(P - P^*) = 1$ and concludes the proof.

Other tie-breaking mechanisms. For Maximin, we cannot find a same order as LEX in the profile. Therefore, we find another order in the profile that satisfies all tie-breaking. Note that the only tie-breaking in the proof is $r(P - P^*)$, where 1 needs to beat 2 and 3 in the tie-breaking. Therefore, we choose $1 \succ 4 \succ 3 \succ 2 \succ X$ in P_4.

FA. Switch P_1 and P_4, let the agent with preference $1 \succ 4 \succ 3 \succ 2 \succ X$ become agent 1.

MPSR For convenience let us denote those in P_4 with preference $1 \succ 4 \succ 3 \succ 2 \succ X$ as P_4^1, others as P_4^2. Let every voter in P_4^1 have preference $1 \succ 4 \succ 3 \succ 2 \succ 5 \succ 6 \succ 7 \succ \cdots \succ q + 4$. For every other voter, $q!$ permutations in X can guarantee every voter have a different permutation, thus different preference. Therefore, $1 \succ 4 \succ 3 \succ 2 \succ 5 \succ 6 \succ 7 \succ \cdots \succ q+4$ with $\frac{q}{3}$ votes is the most popular.

In *Second* part we prove that w.l.o.g $P^* \subseteq P_1$. Note that here our goal is to show a yes instance for RXC3. Since we can find one with $P^* \subseteq P_1$, we can recognize that P_4 voters will not abstain from voting. Therefore, the tie-breaking order does not change. ☐

B.3 Full Proof of Theorem 4(STV)

Theorem 4 **(STV).** *For any* $0 \leq \alpha \leq 1$, GNSP-*STV is NP-complete to compute, where the tie-breaking mechanism is* LEX, FA_1, MPSR+LEX *or* MPSR+FA_1.

Proof. Membership in NP is straightforward. The hardness is proved by a reduction from RXC3 that is similar to the reduction in the hardness proof for the manipulation problem under STV [Bartholdi and Orlin, 1991]. For any RXC3 instance $(X, \mathcal{S})$, where $X = \{x_1, \ldots, x_q\}$ and $\mathcal{S} = \{S_1, \ldots, S_q\}$, we construct the following GNSP-STV instance.

Alternatives: there are $3q + 3$ alternatives $\{w, c\} \cup \{d_0, d_1, \ldots, d_q\} \cup \{b_1, \bar{b}_1, \ldots, b_q, \bar{b}_q\}$. We assume that b_i has higher priority than $\bar{b}_i$, and $d_1 \succ d_2 \succ \cdots \succ d_q$ in tie-breaking.

Profile: The profile P consists of the following votes, where the top preferences are specified and the remaining alternatives ("others") are ranked arbitrarily.

- $\boldsymbol{P_1}$: There are $12q$ votes of $[c \succ w \succ \text{others}]$.
- $\boldsymbol{P_2}$: There are $12q - 1$ votes of $[w \succ c \succ \text{others}]$.
- $\boldsymbol{P_3}$: There are $10q + 2q/3$ votes of $[d_0 \succ w \succ c \succ \text{others}]$.
- $\boldsymbol{P_4}$: For every $j \in \{1, \ldots, q\}$, there are $12q - 2$ votes of $[d_j \succ w \succ c \succ \text{others}]$.
- $\boldsymbol{P_5^1}$: For every $i \in \{1, \ldots, q\}$, there are $6q+4i-2$ votes of $[b_i \succ \bar{b}_i \succ w \succ c \succ \text{others}]$; and $\boldsymbol{P_5^2}$: for every $i \in \{1, \ldots, q\}$, there are two votes of $[b_i \succ d_0 \succ w \succ c \succ \text{others}]$.
- $\boldsymbol{P_6^1}$: For every $i \in \{1, \ldots, q\}$, there are $6q+4i-6$ votes of $[\bar{b}_i \succ b_i \succ w \succ c \succ \text{others}]$; and $\boldsymbol{P_6^2}$: for every $i \in \{1, \ldots, q\}$ and every $j \in S_i$, there are two votes of $[\bar{b}_i \succ d_j \succ w \succ c \succ \text{others}]$.
- $\boldsymbol{P_7}$: For every $i \in \{1, \ldots, q\}$, there is a vote $[\bar{b}_i \succ b_i \succ c \succ w \succ \text{others}]$.

We first show STV$(\boldsymbol{P}) = \boldsymbol{w}$. In the first round, the plurality scores of the alternatives are as in the following table.

Rd.	w	c	b_i	$\bar{b}_i$	d_0	d_j
1	$12q - 1$	$12q$	$6q + 4i$	$6q + 4i + 1$	$10q + \frac{2q}{3}$	$12q - 2$
$q + 1$	$12q - 1$	$12q$	Removed	$12q + 8i - 1$	$12q + \frac{2q}{3}$	$12q - 2$

In the first q rounds, the order of elimination is $b_1, b_2, \ldots, b_q$ (whose votes are transferred to $\bar{b}_1, \ldots, \bar{b}_q$ and d_0). At the beginning of round $q + 1$, the plurality scores of the remaining alternatives are also shown in previous table. Then, d_q is eliminated, whose votes transfer to w. In the remaining rounds w is never eliminated and will become the winner.

Suppose the RXC3 **is a yes instance,** and let $\mathcal{S}^* \subseteq \mathcal{S}$ denote the solution. Let $I = \{i \le q : S_i \in \mathcal{S}^*\}$. We prove that the GNSP-STV instance is a yes instance by showing that agents in P_7 whose top choices are b_i such that $i \in I$ have incentive to (jointly) abstain from voting. It is not hard to verify that after they abstain from voting, in the first q rounds, for each $i \le q$, $\bar{b}_i$ is eliminated if and only if $i \in I$, otherwise b_i is eliminated. Notice that if b_i is eliminated, then $6q + 4i - 2$ of its votes (in P_5^1) transfer to $\bar{b}_i$, and two of its votes (in P_5^2) transfer to d_0. If $\bar{b}_i$ is eliminated, then $6q + 4i - 5$ of its votes (in P_6^1) transfer to b_i, and six of its votes (in P_6^2) are distributed evenly among d_j whose indices are the three alternatives in S_i. Therefore, in the beginning of round $q + 1$, the plurality scores of the remaining alternatives are as in the following table.

Rd.	w	c	b_i or $\bar{b}_i$	d_0	d_j
$q+1$	$12q-1$	$12q$	$12q+8i-1$ or $12q+8i-5$	$12q$	$12q$

Therefore, w is eliminated in round $q + 1$, whose votes transfer to c. It is not hard to verify that c will be the winner, and everyone in P_7 (including the agents who abstain from voting) prefers c to w.

Suppose the GNSP-STV **instance is a yes instance.** We prove that the RXC3 instance is a yes instance by proving that the new winner must be c and the absent votes in P_7 constitutes a solution to the RXC3 instance.

First, we prove that the new winner must be c. Suppose for the sake of contradiction this is not true. Notice that c and w are adjacent in all votes. Therefore, if in any round c is eliminated, then all of its votes are transferred to w; and vice versa. Consider the round right after c or w is eliminated. Then, the remaining alternative in $\{c, w\}$ is ranked at the top in at least $24q - 1$ votes (in $P_1 \cup P_2$). Moreover, given that one of $\{c, w\}$ is not eliminated, in any round we have the following upper bounds on the plurality scores of other alternatives (which are no more than the number of votes they ranked higher than c and w).

- For every $i \in \{1, \dots, q\}$, the plurality score of b_i or $\bar{b}_i$ is at most $21q$ (all votes in $P_5^1 \cup P_5^2 \cup P_6^1 \cup P_6^2 \cup P_7$ where b_i or $\bar{b}_i$ is ranked at the top).
- The plurality score of d_0 is at most $(12 + \frac{2}{3})q$ (all votes in $P_3 \cup P_6^2$).
- For every $j \in \{1, \dots, q\}$, the plurality score of d_j is at most $12q + 4$ (all votes in $P_4 \cup P_5^2$ where d_j is ranked higher than c).

Notice that all upper bounds are lower than $24q - 1$. Therefore, the winner is c or w, which contradicts the assumption.

Given that the new winner is c, because only agents in P_1 and P_7 rank c above w, only they have incentive to abstain from voting. Let I denote the indices i's of agents abstain from voting in P_7 whose top-ranked preferences are $\bar{b}_i$. It is not hard to verify that at the beginning of round $q + 1$, the plurality score of d_0 is $10q + \frac{2q}{3} + 2(q - |I|)$, and for every $1 \le j \le q$, the plurality score of d_j is $12q - 2$ if and only if d_j is not in any S_i with $i \in I$. Suppose for the sake of contradiction that $\mathcal{S}^* = \{S_i : i \in I\}$ is not a solution to the RXC3 instance. Then, in round $q + 1$, d_j for some $j \in \{0, 1, \dots, q\}$ is removed. Once d_j is removed, all of its votes will transfer to w, and subsequently, w will not lose in any rounds, which contradicts the assumption that the winner is c. Therefore, the RXC3 instance is a yes instance. This completes the proof.

Other tie-breaking mechanisms . For STV, we still use the strategy that constructing agents that have the same tie-breaking order as LEX. We use voters in P_2 to achieve this.

FA$_1$. Let every voter of P_2 have the preference $w \succ c \succ b_1 \succ \bar{b}_1 \succ b_2 \succ \bar{b}_2 \succ \cdots \succ b_q \succ \bar{b}_q \succ d_0 \succ d_1 \succ d_2 \succ \cdots \succ d_q$. Then switch P_1 and P_2 and let agent 1 to be the first agent of P_2.

MPSR Let every voter of P_2 have the preference $w \succ c \succ b_1 \succ \bar{b}_1 \succ b_2 \succ \bar{b}_2 \succ \cdots \succ b_q \succ \bar{b}_q \succ d_0 \succ d_1 \succ d_2 \succ \cdots \succ d_q$. For other groups P_i, let every two voter in the same group have different permutation of their "other" part. This can be easily achieved since no group has more than $12q$ voters, but the 'other' part in a group has at least $q!$ permutations. In this way, the tie-breaking order we want has a population of $12q$ while other orders has at most 7 voters.

For STV, agents in P_2 will not abstain from voting as well. Therefore the tie-breaking order won't change.

□

C Supplemental Proof for Theorem 2 (Condorcetified rule): Why 1 is the winner in P.

Here we give an illustration in Condorcetified rules. Specifically, we explain why 1 is the winner under the positional scroing rule. Recall the profile P in the proof:

- **Part 1:** P_1 consists of q votes that correspond to the sets in $\mathcal{S}$: for every $j \le q$, there is a vote R_S defined as follows

$$R_S = (X \setminus S_j) \succ 2 \succ 1 \succ S_j,$$

where alternatives in $(X \setminus S_j)$ and in S_j are ranked alphabetically. More precisely, $P_1 = \{R_S : S \in \mathcal{S}\}$.

- **Part 2:** P_2 consists of the $\frac{2}{3}q - 3$ copies of $[1 \succ 2 \succ X]$.

- **Part 3:** P_3 consists of the the following pair of votes

$$\{[1 \succ X \succ 2], [1 \succ 2 \succ X]\}$$

- **Part 4:** P_4 consists of q copies of the following $2q$ votes:

$$\{\sigma^i(1 \succ X \succ 2), \sigma^i(2 \succ 1 \succ X) : 1 \leq i \leq q\},$$

where σ^i denote a cyclic permutation among X where i is ranked at the first, i.e., $\sigma^i : i \to (i+1) \to \cdots \to (q+2) \to 3 \to 4 \to \cdots \to (i-1)$.

We count the number of each rank for each alternatives in each part of the profile in Table 1.

Alternatives	P_1(rank: times)	P_2	P_3	P_4
1	$q-1:q$	$1:\frac{2}{3}q-3$	$1:2$	$1:q^2$ $2:q^2$
2	$q-2:q$	$2:\frac{2}{3}q-3$	$q+2:1$ $2:1$	$q+2:q^2$ $1:q^2$
Upper bound of $a \in X$	$1:q$	$3:\frac{2}{3}q-3$	$2:1$ $3:1$	$i:q$ for $i=2,3,\cdots,q+1$. $j:q$ for $j=3,4,\cdots,q+2$.

Table 1: Number of ranks for alternatives in each part of the profile.

Note that the third line is an upper bound of ranks which an $a \in X$ can achieve. For example, when X are ranked alphabetically, and alternative 3 is not contained in any S_j, such upper bound will be reached. Otherwise an $a \in X$ will get only strictly lower ranks.

Now we compare the score of each alternative. Denote the scoring vector as $\vec{s} = (s_1, \ldots, s_m) \in \mathbb{Z}^m$. Recall that we have $s_1 \geq s_2 \geq \cdots \geq s_m$ and $s_1 > s_m$.

- **1's score is higher than 2's.** It's not hard to verify that alternative 1 dominates alternative 2 in every parts, which means 1's score is not lower than 2's. Moreover, in P_3 1 is ranked 1st in both two votes, while 2 is ranked the last $(q+2)$ in one vote. Therefore, 1's score is strictly higher than 2's.
- **1's score is higher than any $a \in X$'s.** We'll show that 1's score is higher than the upper bound. For simplicity, suppose $a \in X$ reach the bound, and we prove that 1's score is strictly higher than a's score. Firstly, 1's rank dominates as-rank in P_2 and P_3. Then we look P_1 and P_4 together. In P_1, a's score exceeds 1's score by $q(s_1 - s_{q-1})$. However, in P_4, 1's score exceeds a's score by

$$\sum_{i=2}^{q+1} q(s_1 - s_i) + \sum_{j=3}^{q+2} q(s_1 - s_j) > q(s_2 - s_{q-1}).$$

The strictness of the inequality is because $(s_1 - s_{q+1}) + (s_2 - s_{q+2}) = (s_1 - s_{q+2}) + (s_2 - s_{q+1}) > 0$.

Therefore, we prove that 1's integer positional score is strictly higher than any other alternative. Therefore, 1 is the winner in P.

D Algorithms for Verification of GNSP

D.1 DFS Algorithm

The DFS Algorithm (Algorithm 1) solves the GNSP-r problem by doing a depth first search. Depth first search obviously becomes too expensive where there might be no solution (i.e., no GNSP), thus unrealistic for large n.

D.2 Integer Linear Programming (ILP)-based Algorithms

In this section, we present the ILP formulations and ILP-related algorithms for four voting rules. Copeland, maximin, Black's rule and STV. The variable formulation is same for all voting rules, as explained in Section 4. The chosen objective function actually returns the smallest group size, k, for which there exists a group of agents who has incentive to abstain from voting. If there is no feasible solution for any of the formulated ILPs, obviously GNSP does not occur.

Next, we discuss how to construct specific ILPs for each of the voting rules mentioned. For all voting rules, we assume a is the original winner for preference profile P and lexicographic tie-breaking is used. The presented algorithms can very easily adapted to other tie-breaking methods as well.

Copeland.

Algorithm 2 summarizes the algorithm for verifying GNSP-Cd$_\alpha$. As mentioned in Section 4, any UMG gives a set of linear constraints for the variables x which determines the constraints for Algorithm 2 Additionally, we always have the constraints $x_i \leq n_i$ for integer x_i. So, we can choose to minimize the objective function $\sum_{R_i \in V_{b \succ a}} x_i$ to find the smallest group.

Algorithm 1: DFS algorithm for GNSP-r

```
Input: A profile P
Output: Whether there is a GNSP occurrence for P
Set a ← r(P)
for every alternative b ∈ A − {a} do
    Initialize a stack F with initial state P
    while F is not empty do
        Pop state Q from F
        if r(Q) = b then
            return false
        end
        for all voters j ∈ Q with b ≻_j a do
            Add Q − {j} to stack F
        end
    end
end
return true
```

Algorithm 2: ILP-GNSP-Cd$_\alpha$

```
Input: A profile P
Output: Minimum k for which GNSP-Cd_α if such k exists.
Set a ← Cd_α(P)
Set k ← ∞
for every alternative b ∈ A − {a} do
    for all possible UMG with b as winner do
        Create system of linear constraints using the pairwise relations in the UMG
        if optimal solution x found then
            if ∑_{R_i ∈ V_{b≻a}} < k then
                k = ∑_{R_i ∈ V_{b≻a}}
            end
        end
    end
end
if solution found then
    return k
else
    return No solution found
end
```

Thus, we enumerate all UMGs such that a is not the Copeland winner, construct the ILP, and check if there is a solution. We now give an example for a specific UMG to illustrate this. For $m = 3$ alternatives, a possible UMG is $1 \to 2, 1 \to 3, 2 \to 3$, with Copeland winner 1. Thus the constraints would be-

$$P_x[1 \succ 2] \geq P_x[2 \succ 1]$$
$$P_x[1 \succ 3] \geq P_x[3 \succ 1]$$
$$P_x[2 \succ 3] \geq P_x[3 \succ 2]$$
$$x_i \leq n_i \forall R_i \in V_{b \succ a}$$

The objective is to minimize $n - \sum_{R_i \in V_{b\succ a}} x_i$. If the original winner was not alternative 1, and we find a feasible solution for this ILP, it means that the group no-show paradox has occurred.

Maximin.

For maximin, again we start by constructing the ILP. Alternative b is the maximin winner when $\mathrm{MS}_{P_x}(b) \geq \mathrm{MS}_{P_x}(c)$ for all $c \neq b$ i.e. $\min_{d\in\mathcal{A}} w_{P_x}(b,d) \geq \min_{d\in\mathcal{A}} w_{P_x}(c,d)$. For b, if $b' \in \arg\min_{d\in\mathcal{A}} w_{P_x}(b,d)$, then we have $P[b \succ b'] \leq P[b,d], \forall d \neq b$. Similarly, assume that alternatives $c', d', \ldots$ lead to the maximin score for alternatives $c, d, \ldots$. That is, $P[c \succ c'] \leq P[c,d], \forall d \neq c$. Additionally, $P[b \succ b'] \geq P[c \succ c']$ where $c' \in \arg\min_{d\in\mathcal{A}} w_{P_x}(c,d)$ because of the maximin property. So, to summarize, we have the following system of constraints for any combination of $(b', c', \ldots) \in (\mathcal{A}-\{b\}) \times (\mathcal{A}-\{c\}), \ldots$.

$$\begin{aligned} P_x[b \succ b'] &\leq P_x[b \succ d] \ \ \forall d \neq b \\ P_x[c \succ c'] &\leq P_x[c \succ d] \ \ \forall d \neq c, \forall c \in \mathcal{A} \end{aligned} \tag{2}$$

Algorithm 3: ILP-GNSP-MM

```
Input: A profile P
Output: Minimum k for which GNSP-MM if such k exists.
Set a ← MM(P)
Set k ← ∞
for every alternative b ∈ A − {a} do
    for every possible combination of (b′, c′, . . .) ∈ (𝒜 − {b}) × (𝒜 − {c}), . . .. do
        Check for feasible solution for system of constraints in 2 and3
        if optimal solution x found then
            if ∑_{R_i ∈ V_{b≻a}} < k then
                k = ∑_{R_i ∈ V_{b≻a}}
            end
        end
    end
end
if solution found then
    return k
else
    return No solution found
end
```

Finally, in order for a to be the winner given lexicographic tie-breaking-

$$\begin{aligned} P_x[b \succ b'] &\geq P_x[c \succ c'] \;\; \forall c \in \mathcal{A}, b \rhd c \\ P_x[b \succ b'] &> P_x[c \succ c'] \;\; \forall c \in \mathcal{A}, c \rhd b \end{aligned} \tag{3}$$

where $b \rhd c$ indicates dominance in the lexicographic order.

This is summarized in Algorithm 3. With this formulation, for each P, we need to consider $O(m^m)$ different ILPs. And for each ILP, we have $O(m^2)$ different constraints.

STV.

Being a multi-round voting rule, the winner of STV is uniquely determined by order of elimination. Thus for each of the $m!$ possible order of eliminations, we can get constraints for each elimination. For STV, in each round the alternative with least plurality score is eliminated (with proper tie-breaking in case of tied least scores). In STV, when an alternative is eliminated, they are eliminated from each ranking and thus plurality score needs to be updated for all non-eliminated alternatives. Denote $s_P^k(a)$ as a's plurality score in round k for preference profile P.

W.l.o.g. assume order of elimination $(a_{r_1}, \ldots, a_{r_{m-1}})$, where a_{r_k} is eliminated in round k. This gives us the following set of constraints.

$$\begin{aligned} s_{P_x}^k(a_{r_1}) &\leq s_{P_x}^k(c) \forall c \in \mathcal{A}, c \rhd a_{r_1} \forall k = 1, \ldots, m-1 \\ s_{P_x}^k(a_{r_1}) &< s_{P_x}^k(c) \forall c \in \mathcal{A}, a_{r_1} \rhd c \forall k = 1, \ldots, m-1 \end{aligned} \tag{4}$$

However, we can note that while the full system of equation considers all $m-1$ rounds, it can be somewhat beneficial to not construct the full ILP at the beginning. For example, for $m = 4$, consider two orders of eliminations $a \to b \to c, a \to c \to b$. In both cases, a is eliminated first. If we find that the constraints for a being eliminated first does not have a feasible solution, we do not need to check any of the $(m-1)!$ orders of elimination that has a eliminated first. Thus, we add the constraints, one round at a time. This leads to a larger number of ILP optimization problems, but allows to prune a number of larger systems of constraints early.

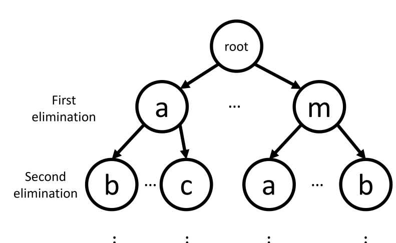

Figure 7: Permutation tree for traversal of order of elimination for Algorithm 4

Algorithm 4: ILP-GNSP-STV

```
Output: Minimum k for which GNSP-STV if such k exists.
Set a ← MM(P)
Set k ← ∞
Create permutation tree with all possible orders of elimination (as in Figure 7)
Do a breadth first traversal over the graph
for nodes in depth k = 1 to m − 1 do
    if node depth ≤ m − 1 then
        Add constraints for round k to the constraints from predecessor nodes according to Eq. 4
        if feasible solution does not exist then
            Prune branch from tree
        end
    end
    if node depth m − 1 then
        if winner of elimination order ≠ a then
            Add constraints for round m − 1
            if optimal solution x found then
                if ∑_{R_i ∈ V_{b≻a}} x_i < k then
                    k = ∑_{R_i ∈ V_{b≻a}} x_i
                end
            end
        end
    end
end
if solution found then
    return k
else
    return No solution found
end
```

We define the full process in Algorithm 4. With this formulation, if all permutations have to be checked with no pruning possible, then we need to consider $O(m!)$ different ILPs. And for each ILP, we have $O(m)$ different constraints in each round, so $O(m^2)$ constraints in total.

In the ILP formulation for all three voting rules, we can get a linear program(LP) relaxation by relaxing the integer requirement for the solution x. That allows us to discard some constraints when no feasible LP solution exists. We run experiments to check the performance of the algorithms, along with the pruning schemes (the round-based pruning in STV and LP-relaxation pruning) and present the results in Section 5.

Black's rule.

Black's rule, just like Copeland depends on the UMG. Whenever the UMG indicates that there is a Condorcet winner, the Copeland and Black's winner will be the same. The ILPs in these cases will also be the same for the two rules. However, when there is no Condorcet winner, Black's winner is determined by the Borda rule. Thus, for each such UMG, we have m new scenarios for each m alternative becoming the Borda winner. Assume the Borda score for alternative b is $s_B(b)$. Thus for v to be Borda winner, the additional constraints are

$$\begin{aligned} s_B(b) &\geq s_B(c) \quad \forall c \in \mathcal{A}, b \rhd c \\ s_B(b) &> s_B(c) \quad \forall c \in \mathcal{A}, c \rhd b \end{aligned} \tag{5}$$

These additional constraints give the full ILP to be solved for the UMGs where there is no Condorcet winner. This formulation gives $O(m \times 3^{\frac{m(m-1)}{2}})$ total scenarios and $O(m^2)$ constraints for the scenario of each ILP. The process is summarized in Algorithm 5.

Algorithm 5: ILP-GNSP-Black's rule

```
Input: A profile P
Output: Minimum k for which GNSP-Black if such k exists.
Set a ← Blacks(P)
Set k ← ∞
for all possible UMG do
    b ← Copeland winner if b = a then
        Continue
    end
    if Condorcet winner exists then
        Create system of linear constraints using the pairwise relations in the UMG
    else
        for each alternative b ≠ a do
            Create additional constraints for b ← Borda winner
            if optimal solution x found then
                if ∑_{R_i ∈ V_{b≻a}} < k then
                    k = ∑_{R_i ∈ V_{b≻a}}
                end
            end
        end
    end
    if optimal solution x found then
        if ∑_{R_i ∈ V_{b≻a}} < k then
            k = ∑_{R_i ∈ V_{b≻a}}
        end
    end
end
if solution found then
    return k
else
    return No solution found
end
```

E Additional Experiment Details

E.1 Experiment Details

Definition of Mallow's Model.

Definition 4 (Mallows Model with Fixed Dispersion)**.** *Given* $0 < \phi < 1$*, a linear order* $W \in \mathcal{L}(\mathcal{A})$ *and* $m \in \mathbb{N}$ *alternatives, Mallows model with fixed dispersion has the following probability for any full linear order* V*:* $\Pr_W(V) = \frac{1}{Z_{m,\phi}}\phi^{KT(V,W)}$*, where* $Z_{m,\phi}$ *is a normalizing factor and* $KT(V,W)$ *is the Kendall's Tau distance between* V, W*.*

Computational details. All experiments were run on the CPU without use of GPU. The CPU configuration is given below.

- Architecture: x86_64
- CPU(s): 16
- RAM: 16GB
- Processor: 3.2 GHz
- Thread(s) per core: 2
- Core(s) per socket: 8
- Socket(s): 1
- Model name: AMD Ryzen 7 5800H with Radeon Graphics

ILP Solver

For all but the maximin voting rule, we used the GLPK ILP solver. For maximin voting rule, we faced into some issues for some preference profiles, so we instead used the Gurobi ILP solver.

E.2 Experiment Results

Likelihood of group no-show paradox. Figure 8 gives the likelihood of group no-show paradox for different voting rules under IC distribution. We see that in general, Copeland rule has the highest likelihood, except for $n = 10$. STV has the second most likelihood, then Black's rule and finally maximin has the least likelihood.

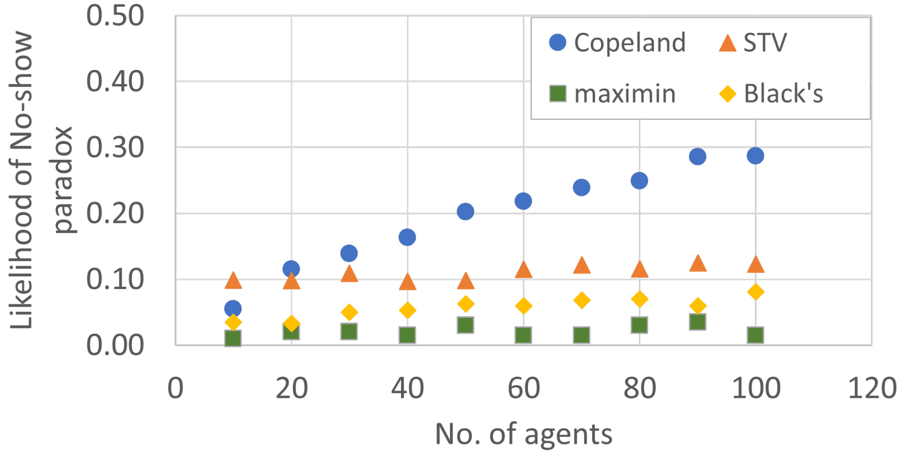

Figure 8: Likelihood of GNSP for Different Voting Rules under IC distribution for $m = 4$ alternatives

Then in Figure 9 and Figure 10, we see the likelihood of group no-show paradox for STV and maximin voting rules under different distributions. Similar to the results for Copeland, we notice that the likelihood seems highest for impartial culture (IC) and nearly zero for Mallows with $\phi = 0.5$ where there is high consensus among the agents.

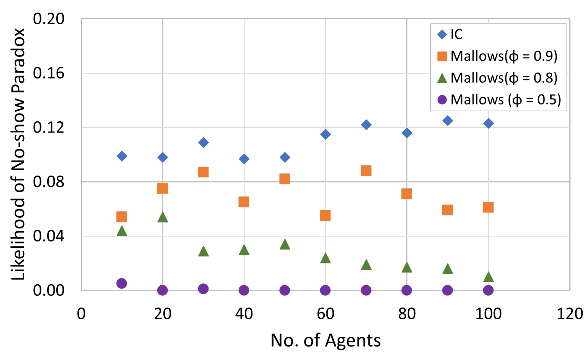

Figure 9: Empirical likelihood of GNSP for different distributions and different No. of agents for $m = 4$ for STV voting rule

Run-time for the DFS and ILP algorithms.

Tables 2 and 3 show the run-time of the two types of algorithms for different scenarios. As expected, for small n, DFS works well for any number of alternatives in the experiment, whereas ILP does not work for high m. On the other hand, for small m, the ILP algorithm works well for all the voting rules.

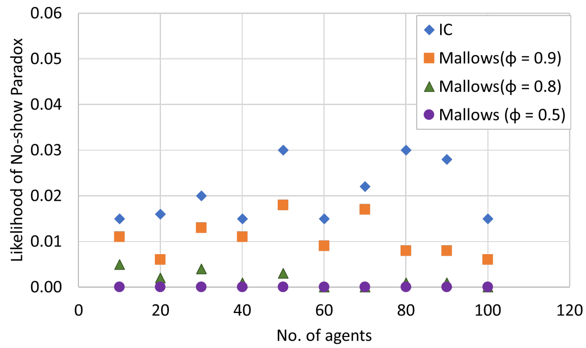

Figure 10: Empirical likelihood of GNSP for different distributions and different No. of agents for $m = 4$ for maximin voting rule

	Copeland		maximin		STV		Black's	
n	DFS	ILP	DFS	ILP	DFS	ILP	DFS	ILP
10	0.00	0.73	0.003751	0.89	0.01	0.02	0.01	3.24
20	0.09	1.05	0.154844	1.21	0.51	0.02	0.16	5.27
30	3.98	1.25	4.73052	1.25	16.94	0.02	7.37	7.48
40	227.71	1.39	187.262	1.31	302.88	0.02	315.89	6.22
50		1.48		1.72		0.02		6.62
60		1.53		1.89		0.03		6.89
70		1.56		1.91		0.03		7.10
80		1.59		2.02		0.03		7.21
90		1.61		2.09		0.03		7.32
100		1.61		2.10		0.03		7.41

Table 2: Run-time of DFS and ILP algorithm for fixed No. of alternatives ($m = 4$) and different number of agents. The missing entries indicate where the DFS algorithm did not finish running even a single instance within the designated time-frame.

	STV		Maximin		Black's	
m	DFS	ILP	DFS	ILP	DFS	ILP
4	0.02	0.02	0	0.89	0	3.255
6	0.05	0.57	0.01	3.255	0.01	
8	0.09	56.02	0.02		0.02	
10	0.13		0.03		0.03	
12	0.19		0.05		0.06	
14	0.26		0.09		0.09	
16	0.32		0.12		0.14	
18	0.4		0.18		0.18	
20	0.5		0.24		0.28	

Table 3: Run-time of DFS and ILP algorithm for fixed No. of agents ($n = 10$) and different number of alternatives