Cognitive Hierarchy and Voting Manipulation*

Edith Elkind[1], Umberto Grandi[2], Francesca Rossi[3], and Arkadii Slinko[4]

[1] *University of Oxford, elkind@cs.ox.ac.uk*
[2] *University of Toulouse, umberto.grandi@irit.fr*
[3] *University of Padova, frossi@math.unipd.it*
[4] *The University of Auckland, a.slinko@auckland.ac.nz*

August 27, 2018

Abstract

By the Gibbard–Satterthwaite theorem, every reasonable voting rule for three or more alternatives is susceptible to manipulation: there exist elections where one or more voters can change the election outcome in their favour by unilaterally modifying their vote. When a given election admits several such voters, strategic voting becomes a game among potential manipulators: a manipulative vote that leads to a better outcome when other voters are truthful may lead to disastrous results when other voters choose to manipulate as well. We consider this situation from the perspective of a boundedly rational voter, and use the cognitive hierarchy framework (Camerer et al., 2004) to identify good strategies. We then investigate the associated algorithmic questions under the k-approval voting rule, $k \geq 1$. We obtain positive algorithmic results for $k = 1, 2$ and NP- and coNP-hardness results for $k \geq 4$.

1 Introduction

Imagine that you and your friends are choosing a restaurant to go to for dinner. Everybody is asked to name their two most preferred cuisines, and the cuisine named most frequently will be selected (this voting rule is known as 2-approval). Your favourite cuisine is Japanese and your second most preferred cuisine is Indian. Indian is quite popular among your friends and you know that if you name it among your favourite two cuisines, it will be selected. On the other hand, you also know that only a few of your friends like Chinese food. Will you vote for Japanese and Chinese to give Japanese cuisine a chance?

This example illustrates that group decision-making is a complex process that represents an aggregation of individual preferences. Individual decision-makers would like to influence the final decision in a way that is beneficial to them, and hence they may be strategic in communicating their individual choices. Moreover, it is essentially impossible to eliminate strategic behavior by changing the voting rule: the groundbreaking result of Gibbard (1973) and Satterthwaite (1975)

*Some of the results in this paper were presented at the 24th International Joint Conference on Artificial Intelligence (IJCAI-15) (Elkind et al., 2015a). We are grateful to the IJCAI-15 reviewers, as well as to the audiences of the Workshop on Iterative Voting and Voting Games in Padova in 2014, the 5th International Workshop on Computational Social Choice (COMSOC-2014) in Pittsburgh, and to the participants of seminars in Toulouse in 2014 and in Dagstuhl in 2015 for their useful feedback.

states that, under any onto and non-dictatorial social choice rule, there exist situations where a voter can achieve a better outcome by casting a strategic vote rather than the sincere one, provided that everyone else votes sincerely; in what follows, we will call such voters *Gibbard–Satterthwaite (GS) manipulators.*

The Gibbard–Satterthwaite theorem alerts us that strategic behavior of voters cannot be ignored, but it does not tell us under which circumstances it actually happens. Of course, if there is just a single GS-manipulator at a given profile, and he[1] is fully aware of other voters' preferences, it is rational for him to manipulate. However, even in this case this voter may prefer to vote truthfully, simply because he assigns a high value to communicating his true preferences; such voters are called *ideological.* Moreover, if there are two or more GS-manipulators, it is no longer easy for them to make up their mind in favour of manipulation: while the Gibbard–Satterthwaite theorem tells us that each of these voters would benefit from voting strategically assuming that all other voters remain truthful, it does not offer any predictions if several voters may be able to manipulate simultaneously. The issues faced by GS-manipulators in this case are illustrated by the following example.

Example 1. Suppose four people are to choose among three alternatives by means of 2-approval, with ties broken according to the order $a > b > c$. Let the profile of sincere preferences be as in Table 1. There are two voters who prefer b to c to a, one voter who prefers a to c to b, and one voter who prefers c to b to a. If everyone votes sincerely, then c gets 4 points, b gets 3 points and a gets 1 point, so c is elected. Voters 1 and 2 are Gibbard–Satterthwaite manipulators. Each of them can make b the winner by voting bac, ceteris paribus. Let us consider this game from the first voter's perspective, assuming that he is strategic; let A_i denote the strategy set of voter i, $i = 1, 2, 3, 4$. The strategy set of voter 1 can then be assumed to be $A_1 = \{bca, bac\}$ (clearly, under 2-approval cba is indistinguishable from bca, abc is indistinguishable from bac, and the two votes that do not rank b first are less useful than either bca and bac). Voter 1 has a good reason to believe that voters 3 and 4 will vote sincerely, as voter 3 cannot achieve an outcome that he would prefer to the current outcome and voter 4 is fully satisfied.

voter 1	voter 2	voter 3	voter 4
b	b	a	c
c	c	c	b
a	a	b	a

Table 1: A preference profile. The most preferred candidates are on top, followed by the less preferred candidates in a complete ranking.

Case 1. If voter 1 believes that voter 2 is ideological, then he is analysing the game where $A_2 = \{bca\}$, $A_3 = \{acb\}$ and $A_4 = \{cba\}$. In this case he just votes bac and expects b to become the winner.

Case 2. Suppose now that voter 1 believes that voter 2 is also strategic. Now voter 1 has to analyse the game with $A_1 = A_2 = \{bca, bac\}$, $A_3 = \{acb\}$ and $A_4 = \{cba\}$. If either one of the strategic players—voter 1 or voter 2—manipulates and another stays sincere, b will be the winner. However, if they both manipulate, their worst alternative a will become the winner.

[1] We use 'he' to refer to voters and 'she' to refer to candidates.

Thus, in this case voter 1's manipulative strategy does not dominate his sincere vote, and if voter 1 is risk-averse, he should refrain from manipulating.

A popular approach (see Section 1.1) is to view voting as a strategic game among the voters, and use various game-theoretic solution concepts to predict the outcomes. The most common such concept is Nash equilibrium, which is defined as a combination of strategies, one for each player, such that each player's strategy is a best response to other players' strategies. In these terms, the Gibbard–Satterthwaite theorem says that under every reasonable voting rule there are situations where truthful voting is not a Nash equilibrium. As a further illustration, the game analysed in Example 1 (Case 2) has two Nash equilibria: in the first one, voter 1 manipulates and voter 2 remains truthful, and in the second one the roles are switched.

However, the principle that players can always be expected to choose equilibrium strategies is not universally applicable. Specifically, if players have enough experience with the game in question (or with similar games), both theory and experimental results suggest that players are often able to learn equilibrium strategies (Fudenberg and Levine, 1998). However, it is also well-known since the early work of Shapley (1964) that learning dynamics may fail to converge to an equilibrium. Moreover, in many applications—and voting is one of them—players' interactions have only imperfect precedents, or none at all so no learning is possible. If equilibrium is justified in such applications, it must be via strategic thinking of players rather than learning. However, in some games the required reasoning is too complex for such a justification of equilibrium to be behaviourally plausible (Harsanyi and Selten, 1988; Brandenburger, 1992). This is fully applicable to voting, where such reasoning, beyond very simple profiles, is impossible because of the number of voters involved.

In fact, a number of recent experimental and empirical studies suggest that players' responses in strategic situations often deviate systematically from equilibrium strategies, and are better explained by the structural nonequilibrium level-k (Nagel, 1995; Stahl and Wilson, 1994) or cognitive hierarchy (CH) models (Camerer et al., 2004); see also a survey by Crawford et al. (2013). In a level-k model players anchor their beliefs in a non-strategic initial assessment of others' likely responses to the game. Non-strategic players are said to be level-0 players. Level 1 players believe that all other players are at level 0, and they give their best response on the basis of this belief. Level 2 players assume that all other players belong to level 1, and, more generally, players at level k give their best response assuming that all other players are at level $k-1$. The cognitive hierarchy model is similar, but with an essential difference: in this model players of level k respond to a mixture of types from level 0 to level $k-1$. It is frequently assumed that other players' levels are drawn from a Poisson distribution. Some further approaches based on similar ideas are surveyed by Wright and Leyton-Brown (2010).

The aim of our work is to explore the applicability of these models to voting games. We believe that specifics of voting, and, in particular, the heterogeneity of types of voters in real electorates, make the cognitive hierarchy framework more appropriate for our purposes. In more detail, an important feature that distinguishes voting from many other applications of both level-k and CH models is the role of level-0 players. Specifically, level-0 (non-strategic) players are typically assumed to choose their strategy at random, and this type practically does not appear in real games at all. In contrast, in applications to voting it is natural to associate level-0 players with ideological voters, who have a significant presence in real elections. For instance, in the famous Florida vote (2000), where Bush won over Gore by just 537 votes, 97,488 Nader supporters voted for Nader—even though in such a close election every strategic voter should have voted either for Gore or for Bush (and an overwhelming majority of Nader supporters preferred Gore to Bush). However, in the level-k analysis voters of level 2 assume that all other voters have level 1, i.e., level-k models cannot be used to accommodate ideological voters. We therefore focus on the cognitive hierarchy approach. Moreover, we limit ourselves to considering the first three levels of the hierarchy (i.e., level-0, level-1, and level-2 players), as it seems plausible that very few voters are capable of higher-level reasoning (see the survey by Crawford et al. (2013) for some evidence in support of this assumption).

Adapting the cognitive hierarchy framework to voting games is not a trivial task. First, it does not make sense to assume that voters' levels follow a specific distribution. Second, in the standard model of social choice voters' preferences over alternatives are ordinal rather than cardinal. The combination of these two factors means that, in general, for voters at level 2 or higher their best response may not be well-defined. We therefore choose to focus on strategies that are not weakly dominated according to the voter's beliefs. We present our formal definitions and the reasoning that justifies them in Section 3.

To develop a better understanding of the resulting model, we instantiate it for a specific family of voting rules, namely, k-approval with $k \geq 1$. We develop a classification of level-1 strategies under k-approval and clarify the relationship between level-1 reasoning and the predictions of the Gibbard–Satterthwaite theorem (Section 4). We then switch our attention to level-2 strategies, and, in particular, to the complexity of computing such strategies. For k-approval with $k = 1$ (i.e., the classic Plurality rule) we describe an efficient algorithm that decides whether a given strategy weakly dominates another strategy; as a corollary of this result, we conclude that under the Plurality rule level-2 strategies can be efficiently computed and efficiently recognised (Section 5). We obtain a similar result for 2-approval under an additional assumption of *minimality* (Section 6). Briefly, this assumption means that the level-2 player expects all level-1 players to manipulate by making as few changes to their votes as possible. However, for larger values of k finding level-2 strategies becomes computationally challenging: we show that this problem is NP-hard for k-approval with $k \geq 4$ (Section 7). As the problem of finding a level-1 strategy under k-approval is computationally easy for any value of $k \geq 1$ (this follows immediately by combining our characterization of level-1 strategies with the classic results of Bartholdi et al. (1989)), this demonstrates that higher levels of voters' sophistication come with a price tag in terms of algorithmic complexity.

1.1 Related work

There is a substantial body of research in social choice theory and in political science that models non-truthful voting as a strategic interaction, with a strong focus on Plurality voting; this line of work dates back to Farquharson (1969) and includes important contributions by Cain (1978), Feddersen et al. (1990) and Cox (1997), to name a few.

More recently, voting games and their equilibria have also received a considerable amount of attention from computer science researchers, with a variety of approaches used to eliminate counterintuitive Nash equilibria. For instance, some authors assume that voters have a slight preference for abstaining or for voting truthfully when they are not pivotal (Battaglini, 2005; Dutta and Sen, 2012; Desmedt and Elkind, 2010; Thompson et al., 2013; Obraztsova et al., 2013; Elkind et al., 2015b; Obraztsova et al., 2015a). Other works consider refinements of Nash equilibrium, such as subgame-perfect Nash equilibrium (Desmedt and Elkind, 2010; Xia and Conitzer, 2010), strong equilibrium (Messner and Polborn, 2007) or trembling-hand equilibrium (Obraztsova et al., 2016), or model the reasoning of voters who have incomplete or imperfect information about each others' preferences (Myerson and Weber, 1993; Myatt, 2007; Meir et al., 2014). Dominance-based solution concepts have been investigated as well (Moulin, 1979; Dhillon and Lockwood, 2004; Buenrostro et al., 2013; Dellis, 2010; Meir et al., 2014), albeit from a non-computational perspective. All the aforementioned papers do not impose any restrictions on the voters' reasoning ability, de facto assuming that they are fully rational. Boundedly rational voters are considered by Grandi et al. (2017); however, their work focuses on strategic interactions among Gibbard–Satterthwaite manipulators, and studies conditions that ensure existence of pure strategy Nash equilibria in the resulting games. In contrast, in this paper we go further and formally define the degree of voters' rationality by using the cognitive hierarchy approach.

Level-k models and the cognitive hierarchy framework have been long used to model a variety of strategic interactions; we refer the reader to the survey of Crawford et al. (2013). Nevertheless, to the best of our knowledge, ours is the first attempt to apply these ideas in the context of voting.

A topic closely related to voting games is voting dynamics, where players change their votes one by one in response to the current outcome (Meir et al., 2010; Reijngoud and Endriss, 2012; Reyhaneh and Wilson, 2012; Obraztsova et al., 2015b; Endriss et al., 2016; Lev and Rosenschein, 2016; Koolyk et al., 2017); see also a survey by Meir (2017). However, this line of work assumes the voters to be myopic.

Our work can also be seen as an extension of the model of safe strategic voting proposed by Slinko and White (2014). However, unlike us, Slinko and White focus on a subset of GS-manipulators who (a) all have identical preferences and (b) choose between truthtelling and using a specific manipulative vote, and on the existence of a weakly dominant strategic vote in this setting (such votes are called *safe strategic votes*). In contrast, our decision-maker takes into account that manipulators may have diverse preferences and have strategy sets that contain more than one strategic vote. It is therefore not surprising that computing safe strategic votes is easier than finding level-2 strategies: Hazon and Elkind (2010) show that safe strategic votes with respect to k-approval can be computed efficiently for every $k \geq 1$, whereas we obtain hardness results for $k \geq 4$.

One of our contributions is a classification of manipulative votes under k-approval with lexicographic tie-breaking. Peters et al. (2012) propose a similar classification for several approval-based voting rules. However, they view k-approval as a non-resolute voting rule, and therefore their results do not apply in our setting.

Paper outline. The paper is organised as follows. We introduce the basic terminology and definitions in Section 2. Section 3 presents the adaptation of the cognitive hierarchy framework to the setting of voting games. We then focus on the study of k-approval. Section 4 describes the structure of level-1 strategies under k-approval. In Section 5 we provide an efficient algorithm for identifying level-2 strategies with respect to the Plurality rule. Section 6 contains our results for 2-approval, and in Section 7 we present our hardness results for k-approval with $k \geq 4$. Section 8 summarises our results and suggests directions for future work.

2 Preliminaries

In this section we introduce the relevant notation and terminology concerning preference aggregation and normal-form games.

2.1 Preferences and Voting Rules

We consider n-voter elections over a candidate set $C = \{c_1, \dots, c_m\}$; in what follows we use the terms *candidates* and *alternatives* interchangeably. Let $\mathcal{L}(C)$ denote the set of all linear orders over C. An election is defined by a *preference profile* $V = (v_1, \dots, v_n)$, where each v_i, $i \in [n]$, is a linear order over C; we refer to v_i as the *sincere vote*, or *preferences*, of voter i. For two candidates $c_1, c_2 \in C$ we write $c_1 \succ_i c_2$, if voter i ranks c_1 above c_2, and say that voter i *prefers* c_1 to c_2. For brevity we will sometimes write $ab \dots z$ to represent a vote v_i such that $a \succ_i b \succ_i \dots \succ_i z$. We denote the top candidate in v_i by $\mathrm{top}(v_i)$. Also, we denote the set of top k candidates in v_i by $\mathrm{top}_k(v_i)$; note that $\mathrm{top}_1(v_i) = \{\mathrm{top}(v_i)\}$ and $a \succ_i b$ for all $a \in \mathrm{top}_k(v_i)$ and $b \in C \setminus \{\mathrm{top}_k(v_i)\}$.

A (resolute) *voting rule* $\mathcal{R}$ is a mapping that, given a profile V, outputs a candidate $\mathcal{R}(V) \in C$, which we call the *winner of the election defined by* V, or simply the *winner at* V. In this paper we focus on the family of voting rules known as *k-approval*. Under k-approval, $k \in [m-1]$, each candidate receives one point from each voter who ranks her in top k positions; the k-approval score of a candidate c, denoted by $\mathrm{sc}_k(c, V)$, is the total number of points that she receives. The winner is chosen among the candidates with the highest score according to a fixed tie-breaking linear order $>$ on the set of candidates C: specifically, the winner is the highest-ranked candidate with respect to this order among the candidates with the highest score. The 1-approval voting rule is widely used and known as Plurality. We will denote the k-approval rule (with tie-breaking based on a fixed linear order $>$) by $\mathcal{R}_k$. We say that a candidate x *beats* candidate y at V with respect to $\mathcal{R}_k$ and the tie-breaking order $>$ if $\mathrm{sc}_k(x, V) > \mathrm{sc}_k(y, V)$ or $\mathrm{sc}_k(x, V) = \mathrm{sc}_k(y, V)$ and $x > y$.

2.2 Strategic Voting

Given a preference profile $V = (v_1, \dots, v_n)$, and a linear order $v'_i \in \mathcal{L}(C)$, we denote by (V_{-i}, v'_i) the preference profile obtained from V by replacing v_i with v'_i; for readability, we will sometimes omit the parentheses around (V_{-i}, v'_i) and write V_{-i}, v'_i. We will often use this notation when voter i submits a strategic vote v'_i instead of his sincere vote v_i.

Definition 1. *Consider a profile $V = (v_1, \dots, v_n)$, a voter i, and a voting rule $\mathcal{R}$. We say that a linear order v'_i is a* manipulative vote *of voter i at V with respect to $\mathcal{R}$ if $\mathcal{R}(V_{-i}, v'_i) \succ_i \mathcal{R}(V)$. We say that i manipulates* in favour of candidate c *by submitting a vote v'_i if c is the winner at $\mathcal{R}(V_{-i}, v'_i)$. A voter i is a* Gibbard–Satterthwaite manipulator*, or a* GS-manipulator*, at V with respect to $\mathcal{R}$ if the set of his manipulative votes at V with respect to $\mathcal{R}$ is not empty. We denote the set of all GS-manipulators at V by $N(V, \mathcal{R})$.*

Note that a voter may be able to manipulate in favour of several different candidates. Let $F_i = \{c \in C \mid \mathcal{R}(V_{-i}, v'_i) = c \text{ for some } v'_i \in \mathcal{L}(C)\}$; we say that the candidates in F_i are *feasible for i at V with respect to $\mathcal{R}$*. Note that $F_i \neq \emptyset$ for all $i \in [n]$, as this set contains the $\mathcal{R}$-winner at V under truthful voting. We say that two votes v and v' over the same candidate set C are *equivalent* with respect to a voting rule $\mathcal{R}$ if $\mathcal{R}(V_{-i}, v) = \mathcal{R}(V_{-i}, v')$ for every voter $i \in [n]$ and every profile V_{-i} of other voters' preferences. It is easy to see that v and v' are equivalent with respect to k-approval if and only if $\mathrm{top}_k(v) = \mathrm{top}_k(v')$.

2.3 Normal-form Games

A *normal-form game* $(N, (A_i)_{i \in N}, (\succeq_i)_{i \in N})$ is defined by a set of *players* N, and, for each $i \in N$, a set of *strategies* A_i and a preference relation $\succeq_i$ defined on the space of *strategy profiles*, i.e., tuples of the form $\mathbf{s} = (s_1, \dots, s_n)$, where $s_i \in A_i$ for all $i \in N$[2]. For each pair of strategy profiles $\mathbf{s}, \mathbf{t}$ and a player $i \in N$, we write $\mathbf{s} \succ_i \mathbf{t}$ if $\mathbf{s} \succeq_i \mathbf{t}$ and $\mathbf{t} \not\succeq_i \mathbf{s}$. A normal-form game is viewed as a game of complete and perfect information, which means that all players are fully aware of the structure of the game they are playing.

Given a strategy profile $\mathbf{s} = (s_1, \dots, s_n)$ and a strategy $s'_i \in A_i$, we denote by $(\mathbf{s}_{-i}, s'_i)$ the strategy profile $(s_1, \dots, s'_i, \dots, s_n)$, which is obtained from $\mathbf{s}$ by replacing s_i with s'_i. We say that a strategy $s_i \in A_i$ *weakly dominates* another strategy $s'_i \in A_i$ if for every strategy profile $\mathbf{s}_{-i}$ of other players we have $(\mathbf{s}_{-i}, s_i) \succeq_i (\mathbf{s}_{-i}, s'_i)$ and there exists a profile $\mathbf{s}_{-i}$ of other players' strategies such that $(\mathbf{s}_{-i}, s_i) \succ_i (\mathbf{s}_{-i}, s'_i)$.

3 The Model

As suggested in Section 1, our goal is to analyse voting as a strategic game and consider it from the perspective of the cognitive hierarchy model. As we reason about voters' strategic behavior, we consider games where players are voters, their strategies are ballots they can submit, and their preferences over strategy profiles are determined by election outcomes under a given voting rule. We then use the cognitive hierarchy framework to narrow down the players' strategy sets.

3.1 Cognitive Hierarchy Framework for Voting Games

Recall that, in the general framework of cognitive hierarchy, players at level 0 are typically assumed to choose their action at random. This is because in general normal-form games a player who is unable to deliberate about other players' actions usually has no reason to prefer one strategy over another. In contrast, in the context of voting, there is an obvious focal strategy, namely, truthful voting. Thus, in our model we associate level-0 voters with ideological voters, i.e., voters who always vote according to their true preferences.

At the next level of hierarchy are level-1 voters. These voters assume that all other voters are ideological (i.e., are at level 0), and choose their vote so as to get the best outcome they consider possible under this assumption. That is, voter i votes so as to make his most preferred candidate in F_i the election winner (in particular, if F_i is a singleton, voter i votes truthfully). We say that a vote v'_i of a voter i is a *level-1 strategy at profile V with respect to $\mathcal{R}$* if $\mathcal{R}(V_{-i}, v'_i) \succ_i c$ for all $c \in F_i \setminus \{\mathcal{R}(V_{-i}, v'_i)\}$. Note that a level-1 voter that is not a Gibbard–Satterthwaite manipulator has no reason to vote non-truthfully, as he does not expect to be able to change the election outcome according to his tastes; hence we assume that such voters are truthful.

We are now ready to discuss level-2 voters. These voters believe that all other voters are at levels 0 or 1 of the cognitive hierarchy. We will further assume that level-2 voters are agnostic about other voters' levels; thus, from their perspective every other voter may turn out to be a level-0 voter (which in our setting is equivalent to being sincere) or a level-1 voter. Thus, from the point of view of a level-2 voter, a voter who is not a GS-manipulator will stick to his truthful vote, whereas a GS-manipulator will either choose his action among level-1 strategies or (in case

[2]While one usually defines normal-form games in terms of utility functions, defining them in terms of preference relations is more appropriate for our setting, as preference profiles only provide ordinal information about the voters' preferences.

he is actually a level-0 voter) vote truthfully. Thus, when selecting his vote strategically, a level-2 voter takes into account the possibility that other voters—namely, the GS-manipulators—may be strategic as well.

We further enrich the model by assuming that a level-2 voter may be able to identify, for each other voter i, a subset of level-1 strategies such that i always chooses his vote from that subset, i.e., a level-2 voter may be able to rule out some of the level-1 strategies of other voters. There are several reasons to allow for this possibility. First, the set of all level-1 strategies for a given voter can be very large, and a voter may be unable or unwilling to identify all such votes. For example, our level-2 voter may know or believe that other voters use a specific algorithm (e.g., that of Bartholdi et al. (1989)) to find their level-1 strategies; in this case, his set of strategies for each voter i would consist of the truthful vote v_i and the output of the respective algorithm. Also, voters may be known not to choose manipulations that are (weakly) dominated by other manipulations. Finally, voters may prefer not to change their vote beyond what is necessary to make their target candidate the election winner, either because they want their vote to be as close to the true preference order as possible (see the work of Obraztsova and Elkind (2012)), or for fear of unintended consequences of such changes in the complex environment of the game. Thus, a preference profile together with a voting rule define not just a single game, but a family of games, which differ in sets of actions available to GS-manipulators.

3.2 Gibbard–Satterthwaite Games

We will now describe a formal model that will enable us to reason about the decisions faced by a level-2 voter. For convenience, we assume that voter 1 is a level-2 voter and describe a normal-form game that captures his perspective of strategic interaction, i.e., his beliefs about the game he is playing.

Fix a voting rule $\mathcal{R}$, let V be a profile over a set of candidates C, let $N = N(V, \mathcal{R})$ be the set of GS-manipulators at V with respect to $\mathcal{R}$, and set $N_1 = N \cup \{1\}$. We consider a family of normal-form games defined as follows. In each game the set of players is N_1, i.e., voter 1 is a player irrespective of whether he is actually a GS-manipulator. For each player $i \in N_1 \setminus \{1\}$, i's strategy set A_i consists of his truthful vote and a (possibly empty) subset of his level-1 strategies; for voter 1 we have $A_1 = \mathcal{L}(C)$, i.e., 1 can submit an arbitrary ballot. It remains to describe the voters' preferences over strategy profiles. For a strategy profile $V^* = (v_i^*)_{i \in N_1}$, where $v_i^* \in A_i$ for $i \in N_1$, let $V[V^*] = (v'_1, \dots, v'_n)$ be the preference profile such that $v'_i = v_i$ for $i \notin N_1$ and $v'_i = v_i^*$ for $i \in N_1$. Then, given two strategy profiles V^* and V^{**} and a voter $i \in N_1$, we write $V^* \succeq_i V^{**}$ if and only if i prefers $\mathcal{R}(V[V^*])$ to $\mathcal{R}(V[V^{**}])$ or $\mathcal{R}(V[V^*]) = \mathcal{R}(V[V^{**}])$. We refer to any such game as a *GS-game*.

We denote the set of all GS-games for V and $\mathcal{R}$ by $\mathcal{GS}(V, \mathcal{R})$. Note that an individual game in $\mathcal{GS}(V, \mathcal{R})$ is fully determined by the GS-manipulators' sets of strategies, i.e., $(A_i)_{i \in N(V, \mathcal{R})}$ (player 1's set of strategies is always the same, namely, $\mathcal{L}(C)$). Thus, in what follows, we write $G = (V, \mathcal{R}, (A_i)_{i \in N(V, \mathcal{R})})$; when V and $\mathcal{R}$ are clear from the context, we simply write $G = (A_i)_{i \in N}$. We refer to a strategy profile in a GS-game as a *GS-profile*, and we will sometimes identify the GS-profile $V^* = (v_i^*)_{i \in N_1}$ with the preference profile $V[V^*]$. We denote the set of all GS-profiles in a game G by $\mathcal{GSP}(G)$.

We will now argue that games in $\mathcal{GS}(V, \mathcal{R})$ reflect the perspective of voter 1 when he is at the second level of the cognitive hierarchy. Fix a game $G \in \mathcal{GS}(V, \mathcal{R})$. Note first that, since voter 1 believes that all other voters belong to levels 0 and 1 of the cognitive hierarchy, he expects all voters who are not GS-manipulators to vote truthfully, i.e., he does not need to reason about their strategies at all. This justifies having $N_1 = N \cup \{1\}$ as our set of players. On the other hand, consider a voter $i \in N \setminus \{1\}$. Voter 1 considers it possible that i is a level-0 voter, who votes truthfully. Voter 1 also entertains the possibility that i is a level-1 voter, in which case i's vote has to be a level-1 strategy; as argued above, voter 1 may also be able to rule out some of i's level-1 strategies. Consequently, the set A_i, which, by definition, contains v_i, consists of all strategies that voter 1 considers possible for i. Thus, voter 1's view of other voters' actions is captured by G.

We are now ready to discuss level-2 strategies. In game-theoretic literature, it is typical to assume that a level-2 player is endowed with probabilistic beliefs about other players' types as well as a utility function describing his payoffs under all possible strategy profiles. Under these conditions, it makes sense to define level-2 strategies as those that maximise player 2's expected payoff. However, in the absence of numerical information, as in the case of voting games, we cannot reason about expected payoffs. We can, however, compare different strategies pointwise, and remove strategies that are weakly dominated by other strategies. On the other hand, if a strategy v is not weakly dominated, a level-2 player may hold beliefs that make him favour v, so no such strategy can be removed from consideration without making additional assumptions about the behavior of players in $N(V, \mathcal{R})$. This reasoning motivates the following definition of a level-2 strategy.

Definition 2. *Given a GS-game $G = (V, \mathcal{R}, (A_i)_{i \in N(V,\mathcal{R})})$, we say that a strategy $v \in A_1$ of player 1 is a* level-2 strategy *if no other strategy of player 1 weakly dominates v.*

We note that being weakly undominated is not a very demanding property: a strategy can be weakly undominated even if it fares badly in many scenarios. This is illustrated by the following example.

Example 2. Consider the 4-voter profile over $\{a, b, c, d\}$ given in Table 2. Suppose that the voting rule is Plurality and the tie-breaking rule is $a > b > c > d$. As always, we assume that voter 1 is the level-2 voter. Voters 2, 3, and 4 are GS-manipulators; their most preferred feasible candidates are, respectively, d, b, and c. Consider the GS-game where $A_2 = \{bdac, dbac\}$, $A_3 = \{cbad, bcad\}$, $A_4 = \{dcab, cdab\}$. In this game every vote that does not rank d first is a level-2 strategy for the first player. Indeed, a vote that ranks a first is optimal when all other players submit their sincere votes; a vote that ranks b first is optimal when players 2 and 3 stay sincere, but player 4 votes for c; and a vote that ranks c first is optimal when player 2 votes for d, but players 3 and 4 stay sincere. Note, in particular, that, by changing his vote from $abcd$ (his sincere vote) to $cabd$, player 1 changes the outcome from a (his top choice) to c (his third choice) when other players vote truthfully; however, this behavior is rational if player 1 expects players 3 and 4 (but not player 2) to vote sincerely.

voter 1	voter 2	voter 3	voter 4
a	b	c	d
b	d	b	c
c	a	a	a
d	c	d	b

Table 2: A profile where voter 1 has three distinct level-2 strategies under Plurality voting.

Example 2 illustrates that level-2 strategies are not 'safe': there can be circumstances where a level-2 strategy results in a worse outcome than sincere voting. Now, a cautious level-2 player may prefer to stick to his sincere vote unless he can find a manipulative vote which leads to an outcome that is at least as desirable as the outcome under truthful voting, for any combination of actions of other players that our level-2 player considers possible. The following definition, which is motivated by the concept of safe strategic voting (Slinko and White, 2014), describes the set of strategies that even a very cautious level-2 player would prefer to sincere voting.

Definition 3. *Given a GS-game* $G = (V, \mathcal{R}, (A_i)_{i \in N(V,\mathcal{R})})$*, we say that a strategy* $v \in A_1$ *of player 1 is an* improving strategy *if* v *weakly dominates player 1's sincere strategy* v_1*.*

We note that a level-2 strategy may fail to be an improving strategy, and conversely, an improving strategy is not necessarily a level-2 strategy. For instance, in Example 1 the strategy *bac* is a level-2 strategy, but not an improving strategy, and none of the level-2 strategies in Example 2 is improving. However, it is easy to see that if a player has an improving strategy, he also has an improving strategy that is a level-2 strategy. Moreover, an improving strategy exists if and only if sincere voting is not a level-2 strategy.

One can also ask if a given strategy weakly dominates all other (non-equivalent) strategies. However, while strategies with this property are highly desirable, from the perspective of a strategic voter it is more important to find out whether his truthful strategy is weakly dominated. Indeed, the main issue faced by a strategic voter is whether to manipulate at all, and if a certain vote can always ensure an outcome that is at least as good, and sometimes better, as that guaranteed by his truthful vote, this is a very strong incentive to use it, even if another non-truthful vote may be better in some situations. This issue is illustrated by Example 3 below, which describes a profile where a player has two incomparable improving strategies.

Example 3. Let the profile of sincere preferences be as in Table 3, and assume that the voting rule is Plurality and the tie-breaking order is given by $w > d > c > b > a$. The winner at the sincere profile is w. All level-1 strategies of voter 2 are equivalent to *cbdwa*, whereas all level-1 strategies of voter 3 are equivalent to *dcbwa*; voters 4 and 5 are not GS-manipulators. Consider the GS-game where for $i \in \{2, 3\}$ the set of strategies of player i consists of his truthful vote and all of his level-1 strategies. Voter 1, who is our level-2 player, can manipulate either in favour of b or in favour of d, by ranking the respective candidate first. Indeed, for player 1 both *badwc* and *dabwc* weakly dominate truthtelling. However, neither of these strategies weakly dominates the other: *badwc* is preferable if no other player uses a level-1 strategy, whereas *dabcw* is preferable if player 2 uses his level-1 strategy, but player 3 votes sincerely.

voter 1	voter 2	voter 3	voter 4	voter 5
a	b	c	d	w
b	c	d	w	a
d	d	b	c	b
w	w	w	a	c
c	a	a	b	d

Table 3: Player 1 has two incomparable improving strategies.

We note that a level-2 voter may find it useful to act as a counter-manipulator (Pattanaik, 1976; Grandi et al., 2017), i.e., to submit a vote that is not manipulative with respect to the truthful profile, but minimises the damage from someone else's manipulation.

Example 4. Let the profile of sincere preferences be as in Table 4, and assume that the voting rule is Plurality and the tie-breaking order is given by $a > b > c$. Under truthful voting a wins, so voter 6 is the only GS-manipulator: if he changes his vote to bca then b wins, and $b \succ_6 a$. Therefore, for voter 1 voting acb is preferable to voting truthfully: this insincere vote has no impact if voter 6 votes truthfully, but prevents b from becoming a winner when voter 6 submits a manipulative vote.

Thus, in this example acb is an improving strategy, and truthful voting is not a level-2 strategy as it is weakly dominated by voting acb. In contrast, acb is a level-2 strategy, as no other strategy weakly dominates it.

voter 1	voter 2	voter 3	voter 4	voter 5	voter 6
c	a	a	b	b	c
a	b	b	a	a	b
b	c	c	c	c	a

Table 4: Countermanipulation under Plurality.

3.3 Algorithmic Questions

From an algorithmic perspective, perhaps the most natural questions suggested by our framework are how to decide whether a given strategy is a level-2 strategy, or how to compute a level-2 strategy. A related question is whether a given strategy is an improving strategy and whether an improving strategy can be efficiently computed. These questions offer an interesting challenge from an algorithmic perspective: the straightforward algorithm for deciding whether a given strategy is a level-2 strategy or an improving strategy relies on considering all combinations of other players' strategies, and hence has exponential running time. It is therefore natural to ask whether for some voting rules exhaustive choice can be avoided. We explore this question in Sections 5–7; for concreteness, we focus on k-approval, for various values of k.

4 Level-1 Strategies Under k-Approval

The goal of this section is to understand and classify level-1 strategies under the k-approval voting rule; this will help us reason about level-2 strategies in subsequent sections. In what follows, we fix a linear order $>$ used for tie-breaking. We start with a simple, but useful lemma.

Lemma 1. *Fix $k \geq 1$. Consider a profile V over C, let w be the k-approval winner at V, and let x be an alternative in $C \setminus \{w\}$. Then any manipulative vote by a voter i in favour of x at V falls under one of the following two categories:*

Type 1 *Voter i increases the score of x by 1 without decreasing the score of w. In this case $w, x \notin \text{top}_k(v_i)$, $x \succ_i w$, and the manipulative vote v'_i satisfies $x \in \text{top}_k(v'_i)$, $w \notin \text{top}_k(v'_i)$. In such cases voter i will be referred to as a* promoter *of x.*

Type 2 *Voter i decreases the score of w (and possibly that of some other alternatives) by 1 without increasing the score of x. In this case $w, x \in \text{top}_k(v_i)$, $x \succ_i w$, and the manipulative vote v_i' satisfies $x \in \text{top}_k(v_i')$, $w \notin \text{top}_k(v_i')$. In such cases voter i will be referred to as a* demoter *of w. Manipulations of type 2 only exist for $k \geq 2$.*

Proof. Suppose that voter i manipulates in favour of x. If i can increase the score of x, then $x \notin \text{top}_k(v_i)$. However, i must rank x higher than w (otherwise, this would not be a manipulation). Thus, $w \notin \text{top}_k(v_i)$ and therefore voter i cannot decrease w's score. Moreover, if $w \in \text{top}_k(v_i')$, then w would beat x under k-approval in (V_{-i}, v_i'); thus, $w \notin \text{top}_k(v_i')$.

On the other hand, suppose that i cannot increase the score of x. This means that $x \in \text{top}_k(v_i)$ and hence i is left with reducing the scores of some of x's competitors including the current winner w. For this to be possible, it has to be the case that $w \in \text{top}_k(v_i)$ and $x \succ_i w$. Also, we have $w \notin \text{top}_k(v_i')$, as otherwise w would beat x under k-approval in (V_{-i}, v_i'). Finally, as $w \neq x$, we can only have $x, w \in \text{top}_k(v_i)$ if $k \geq 2$. □

The classification in Lemma 1 justifies our terminology: a promoter promotes a new winner and a demoter demotes the old one. Under Plurality, i.e., when $k = 1$, we only have promoters.

Let $X = \{x_1, \ldots, x_\ell\}$ and $Y = \{y_1, \ldots, y_\ell\}$ be two disjoint sets of candidates. Given a linear order v over C, we denote by $v[X; Y]$ the vote obtained by swapping x_j with y_j for $j \in [\ell]$. If the sets X and Y are singletons, i.e., $X = \{x\}$, $Y = \{y\}$, we omit the curly braces, and simply write $v[x; y]$. Clearly, under k-approval any manipulative vote of voter i is equivalent to a vote of the form $v_i[X; Y]$, where $X \subseteq \text{top}_k(v_i)$, $Y \subseteq C \setminus \text{top}_k(v_i)$. We can now state a corollary of Lemma 1, which characterises the possible effects of a manipulative vote under k-approval.

Corollary 2. *Let w be the k-approval winner at a profile V, let $v_i^* = v_i[X; Y]$, where $X \subseteq \text{top}_k(v_i)$ and $Y \subseteq C \setminus \text{top}_k(v_i)$. Let $V' = (V_{-i}, v_i^*)$, and let $w' \neq w$ be the k-approval winner at V'. Then either $w \in X$ or $w' \in Y$ but not both.*

Proof. Lemma 1 implies that either the new winner is promoted or the old winner is demoted, but not both. □

Consider a manipulative vote $v_i[X; Y]$ of voter i at V under k-approval; we say that $v_i[X; Y]$ is *minimal* if for every other manipulative vote v_i' of voter i there is a vote $v_i[X'; Y']$ that is equivalent to v_i' and satisfies $|X'| \geq |X|$. That is, a manipulative vote is minimal if it performs as few swaps as possible. Arguably, minimal manipulative votes are the main tool that a rational voter would use, as they achieve the desired result in the most straightforward way possible.

We now introduce some useful notation. Fix a profile V. Let w be the k-approval winner at V, and let $t = \text{sc}_k(w, V)$. Set

$$
\begin{aligned}
S_1(V, k) &= \{c \in C \mid \text{sc}_k(c) = t, w > c\}, \\
S_2(V, k) &= \{c \in C \mid \text{sc}_k(c) = t - 1, c > w\},
\end{aligned}
$$

and set $S(V, k) = S_1(V, k) \cup S_2(V, k)$.

We say that a candidate c is *k-competitive* at V if $c \in S(V, k)$. The following proposition explains our choice of the term: only k-competitive candidates can become k-approval winners as a result of a manipulation.

Proposition 1. *Suppose that some voter can manipulate in favour of a candidate $p \in C$ at a profile V with respect to k-approval. Then $p \in S(V, k)$.*

Proof. Let w be the k-approval winner at V; clearly, $w \neq p$. Suppose that voter i can manipulate in favour of p at V by submitting a vote v'_i; let $V' = (V_{-i}, v'_i)$. Set $t = \mathrm{sc}_k(w, V)$; then $\mathrm{sc}_k(p, V) \leq t$. Note that if $\mathrm{sc}_k(p, V) = t$, it has to be the case that $w > p$, since otherwise p would beat w at V. Thus, in this case $p \in S_1(V, k)$. Now, suppose that $\mathrm{sc}_k(p, V) = t-1$. By Corollary 2 we have either $\mathrm{sc}_k(w, V') = \mathrm{sc}_k(p, V') = t$ (if p was promoted) or $\mathrm{sc}_k(w, V') = \mathrm{sc}_k(p, V') = t-1$ (if w was demoted). In both cases we have to have $p > w$, as otherwise w would beat p at V'. Therefore, in this case $p \in S_2(V, k)$. Finally, note that it cannot be the case that $\mathrm{sc}_k(p, V) \leq t-2$, since in this case by Corollary 2 we would have either $\mathrm{sc}_k(w, V') \geq t-1$, $\mathrm{sc}_k(p, V') \leq t-2$ or $\mathrm{sc}_k(w, V') = t$, $\mathrm{sc}_k(p, V') \leq t-1$, i.e., w would beat p at V'. □

Suppose that $S(V, k) \neq \emptyset$. If $S_1(V, k) \neq \emptyset$, then by $p^*(V, k)$ we denote the top-ranked candidate in $S_1(V, k)$ with respect to $>$; otherwise, we denote by $p^*(V, k)$ the top-ranked candidate in $S_2(V, k)$ with respect to $>$. Thus, $p^*(V, k)$ beats all candidates other than w at V, and would become a winner if it were to gain one point or if w were to lose one point. We omit V and k from the notation when they are clear from the context.

We are now ready to embark on the computational complexity analysis of level-2 strategies under k-approval, for various values of k.

5 Plurality

Plurality voting rule is k-approval with $k = 1$. For this rule we only have manipulators of type 1, and all manipulative votes of voter i in favour of candidate c are equivalent: in all such votes c is placed in the top position.

The main result of this section is that the problem of deciding whether a given strategy of voter 1 weakly dominates another strategy of that voter admits a polynomial-time algorithm. Note that, since under Plurality there are only m votes that are pairwise non-equivalent, this means that we can check if a given strategy is a level-2 strategy or an improving strategy, or find a level-2 strategy or an improving strategy (if it exists) in polynomial time; we formalise this intuition in Corollary 5 at the end of this section.

Fix a preference profile V over a candidate set C and consider a GS-game $(V, \mathcal{R}_1, (A_i)_{i \in N})$, where $N = N(V, \mathcal{R}_1)$. Let w be the Plurality winner at V. As argued above, for each $i \in N \setminus \{1\}$ the set A_i consists of v_i and possibly a number of pairwise equivalent manipulative votes; without loss of generality, we can remove all but one manipulative vote, so that $|A_i| \leq 2$ for all $i \in N \setminus \{1\}$. We will now explain how, given two votes v'_1 and v''_1, voter 1 can efficiently decide if one of these votes weakly dominates the other.

We will first describe a subroutine that will be used by our algorithm.

Lemma 3. *There is a polynomial-time procedure*

$$Alg = Alg(G, r, r', x, y, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]})$$

that, given a GS-game $G = (V, \mathcal{R}_1, (A_i)_{i \in N(V, \mathcal{R}_1)})$ with $|V| = n$, two integers $r, r' \in \{0, \ldots, n\}$, two distinct candidates $x, y \in C$, and a partition of candidates in $C \setminus \{x, y\}$ into $C^{[1]}$, $C^{[0]}$, $C^{[-1]}$ and $C^{[-2]}$, decides whether there is a strategy profile V^* in G such that

- $\mathrm{sc}_1(x, V[V^*]_{-1}) = r$,
- $\mathrm{sc}_1(y, V[V^*]_{-1}) = r'$, *and*

- *for each $c \in C \setminus \{x, y\}$ and each $\ell \in \{1, 0, -1, -2\}$ if $c \in C^{[\ell]}$ then* $\mathrm{sc}_1(c, V[V^*]_{-1}) \le r + \ell$.

Proof. We proceed by reducing our problem to an instance of network flow with capacities and lower bounds, as follows. We construct a source, a sink, a node for each voter $i \in [n] \setminus \{1\}$ and a node for each candidate in C. There is an arc from the source to each voter node; the capacity and the lower bound of this arc are set to 1, i.e., it is required to carry one unit of flow. Also, there is an arc with capacity 1 and lower bound 0 from voter i to candidate c if $i \in N(V, \mathcal{R}_1) \setminus \{1\}$ and $c = \mathrm{top}(v)$ for some $v \in A_i$ or if $i \in [n] \setminus (N(V, \mathcal{R}_1) \cup \{1\})$ and $c = \mathrm{top}(v_i)$. Finally, there is an arc from each candidate c to the sink. The capacity of this arc is set to $r + \ell$ if $c \in C^{[\ell]}$ for some $\ell \in \{1, 0, -1, -2\}$; the lower bounds for these arcs are 0. For x, both the capacity and the lower bound of the arc to the sink are set to r, and for y they are both set to r'. We note that some of the capacities may be negative, in which case there is no valid flow. It is immediate that an integer flow that satisfies all constraints corresponds to a strategy profile in G where all candidates have the required scores; it remains to observe that the existence of a valid integer flow can be decided in polynomial time. □

We are now ready to describe our algorithm.

Theorem 4. *Given a GS-game $G = (V, \mathcal{R}_1, (A_i)_{i \in N(V, \mathcal{R}_1)})$ and two strategies $v_1', v_1'' \in \mathcal{L}(C)$ of player 1 we can decide in polynomial time whether v_1' weakly dominates v_1''.*

Proof. We will design a polynomial-time procedure that, given two strategies u, v of player 1, decides if there exists a profile V_{-1}^* of other players' strategies such that $\mathcal{R}_1(V[V_{-1}^*, u]) \succ_1 \mathcal{R}_1(V[V_{-1}^*, v])$; by definition, v_1' weakly dominates v_1'' if this procedure returns 'yes' for $u = v_1'$, $v = v_1''$ and 'no' for $u = v_1'', v = v_1'$.

Let $a = \mathrm{top}(u)$, $b = \mathrm{top}(v)$. We can assume without loss of generality that $a \neq b$, since otherwise u and v are equivalent with respect to Plurality. Consider an arbitrary profile V_{-1}^* of other players' strategies, and let $V^u = V[V_{-1}^*, u]$, $V^v = V[V_{-1}^*, v]$, $w^u = \mathcal{R}_1(V^u)$, $w^v = \mathcal{R}_1(V^v)$. We note that $w^u \neq a$ implies $w^v \neq a$: if w^u beats a at V^u, this is also the case at V^v. Similarly, if $w^v \neq b$ then also $w^u \neq b$. Now, suppose that $w^u \neq a$ and $w^v \neq b$. We claim that in this case $w^u = w^v$. Indeed, suppose for the sake of contradiction that $w^u \neq w^v$. As $w^u \neq a$, $w^v \neq b$, the argument above shows that $\{w^u, w^v\} \cap \{a, b\} = \emptyset$. Thus, both w^u and w^v have the same Plurality score at V^u and V^v; as w^u beats w^v at V^u, this must also be the case at V^v, a contradiction.

Note that $\mathcal{R}_1(V[V_{-1}^*, u]) \succ_1 \mathcal{R}_1(V[V_{-1}^*, v])$ if and only if $w^u \succ_1 w^v$. By the argument in the previous paragraph, this can happen in one of the following three cases: (i) $w^u = a$, $w^v = b$ and $a \succ_1 b$; (ii) $w^u = a$, $w^v = w$ for some $w \neq b$, $a \succ_1 w$; (iii) $w^u = w$, $w^v = b$ for some $w \neq a$, $w \succ_1 b$. (We note that we can merge case (i) into case (ii) or case (iii); we choose not to do so for the sake of clarity of presentation.) We will now explain how to check if there exists a profile V_{-1}^* that corresponds to any of these three situations.

Case (i): $w^u = a$, $w^v = b$.

Suppose first that $a > b$. Then a desired profile V_{-1}^* exists if and only if there is some value $t \in [n]$ such that $\mathrm{sc}_1(a, V^u) = t$ and

(a) $\mathrm{sc}_1(b, V^u) = t$, $\mathrm{sc}_1(c, V^u) \le t$ for all $c \in C \setminus \{a, b\}$ with $a > c$, $\mathrm{sc}(c, V^u) \le t - 1$ for all $c \in C \setminus \{a, b\}$ with $c > a$, or

(b) $\mathrm{sc}_1(b, V^u) = t - 1$, $\mathrm{sc}_1(c, V^u) \le t$ for all $c \in C \setminus \{a, b\}$ with $b > c$, and $\mathrm{sc}(c, V^u) \le t - 1$ for all $c \in C \setminus \{a, b\}$ with $c > b$.

Note that $\mathrm{sc}_1(a, V^u_{-1}) = \mathrm{sc}_1(a, V^u) - 1$ and $\mathrm{sc}_1(c, V^u_{-1}) = \mathrm{sc}_1(c, V^u)$ for $c \in C \setminus \{a\}$. Thus, to check if condition (a) is satisfied for some $t \in [n]$, we set $C^{[1]} = \{c \in C \setminus \{a, b\} \mid a > c\}$, $C^{[0]} = \{c \in C \setminus \{a, b\} \mid c > a\}$, $C^{[-1]} = C^{[-2]} = \emptyset$ and call

$$Alg(G, t-1, t, a, b, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]}).$$

Similarly, to determine whether condition (b) is satisfied for some $t \in [n]$, we set $C^{[1]} = \{c \in C \setminus \{a, b\} \mid b > c\}$, $C^{[0]} = \{c \in C \setminus \{a, b\} \mid c > b\}$, $C^{[-1]} = C^{[-2]} = \emptyset$ and call

$$Alg(G, t-1, t-1, a, b, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]}).$$

The answer is 'yes' if one of these calls returns 'yes' for some $t \in [n]$.

For the case $b > a$ the analysis is similar. In this case, we need to decide whether there exists a value of $t \in [n]$ such that $\mathrm{sc}_1(a, V^u) = t$ and

(a) $\mathrm{sc}_1(b, V^u) = t-1$, $\mathrm{sc}_1(c, V^u) \le t$ for all $c \in C \setminus \{a, b\}$ with $a > c$, and $\mathrm{sc}(c, V^u) \le t-1$ for all $c \in C \setminus \{a, b\}$ with $c > a$, or

(b) $\mathrm{sc}_1(b, V^u) = t-2$, $\mathrm{sc}_1(c, V^u) \le t-1$ for all $c \in C \setminus \{a, b\}$ with $b > c$, and $\mathrm{sc}(c, V^u) \le t-2$ for all $c \in C \setminus \{a, b\}$ with $c > b$.

Again, this can be decided by calling the procedure *Alg* with appropriate parameters; we omit the details.

Case (ii): $w^u = a$, $w^v = w$ for some w with $a \succ_1 w$. In this case, we go over all candidates $w \in C \setminus \{a, b\}$ with $a \succ_1 w$ and all values of $t \in [n]$ and call *Alg* with appropriate parameters.

Specifically, if $a > w$, we start by setting $r = t-1, r' = t$, and

$$C^{[1]} = \{c \in C \setminus \{a, w, b\} \mid w > c\}, C^{[0]} = \{c \in C \setminus \{a, w, b\} \mid c > w\}, C^{[-1]} = C^{[-2]} = \emptyset.$$

We then place b in $C^{[0]}$ if $w > b$ and in $C^{[-1]}$ otherwise; our treatment of b reflects the fact that she gets an extra point at V^v.

If $w > a$ we start by setting $r = t-1, r' = t-1$, and

$$C^{[1]} = \emptyset, C^{[0]} = \{c \in C \setminus \{a, w, b\} \mid w > c\}, C^{[-1]} = \{c \in C \setminus \{a, w, b\} \mid c > w\}, C^{[-2]} = \emptyset.$$

We then place b in $C^{[-1]}$ if $w > b$ and in $C^{[-2]}$ otherwise.

Finally, we call

$$Alg(G, r, r', a, w, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]}).$$

The answer is 'yes' if one of these calls returns 'yes' for some $t \in [n]$ and some w with $a \succ_1 w$.

Case (iii): $w^u = w$, $w^v = b$ for some w with $w \succ_1 b$. The analysis is similar to the previous case; we omit the details. □

Theorem 4 immediately implies that natural questions concerning level-2 strategies and improving strategies are computationally easy.

Corollary 5. *Given a GS-game $G = (V, \mathcal{R}_1, (A_i)_{i \in N(V,\mathcal{R}_1)})$ and a strategy $v_1' \in \mathcal{L}(C)$ of player 1 we can decide in polynomial time whether v_1' is a level-2 strategy or an improving strategy. Moreover, we can decide in polynomial time whether player 1 has a level-2 strategy or an improving strategy in G.*

Proof. Let $a = \mathrm{top}(v_1')$. To decide whether v_1' is an improving strategy, we use the algorithm described in the proof of Theorem 4 to check whether v_1' weakly dominates v_1. Similarly, to decide whether v_1' is a level-2 strategy, for each $c \in C \setminus \{a\}$ we construct a vote v^c with $\mathrm{top}(v^c) = c$ and check whether v^c weakly dominates v_1' using the algorithm from the proof of Theorem 4. As every strategy of player 1 is equivalent either to v_1' or to one of the votes we constructed, v_1' is a level-2 strategy if and only if it is not weakly dominated by any of the votes v^c, $c \in C \setminus \{a\}$.

Similarly, to decide whether v_1 has a level-2 strategy (respectively, an improving strategy), we consider all of his m pairwise non-equivalent strategies, and check if any of them is a level-2 strategy (respectively, an improving strategy), as described in the previous paragraph. □

6 2-Approval

In this section, we study the computational complexity of identifying level-2 strategies and improving strategies in GS-games under 2-approval. We show that if the level-2 player believes that level-1 players can only contemplate minimal manipulations, he can efficiently compute his level-2 strategies as well as his improving strategies. As argued in Section 4, minimality is a reasonable assumption, as level-1 players have no reason to use complex strategies when simple strategies can do the job.

Specifically, we prove that, under the minimality assumption, given two strategies v' and v'', the level-2 player can decide in polynomial time whether one of these strategies weakly dominates the other; just as in the case of Plurality, this implies that he can check in polynomial time whether a given strategy is a level-2 (respectively, improving) strategy or identify all of his level-2 (respectively, improving) strategies.

The following observations play a crucial role in our analysis.

Proposition 2. *Consider a GS-game $G = (V, \mathcal{R}_2, (A_i)_{i \in N(V,\mathcal{R}_2)})$. Let w be the 2-approval winner at V. Then for each player $i \in N(V, \mathcal{R}_2) \setminus \{1\}$ such that $w \in \mathrm{top}_2(v_i)$ it holds that $\mathrm{top}(v_i) \neq w$ and the candidate $\mathrm{top}(v_i)$ is ranked in top two positions in every vote $v \in A_i$.*

Proposition 2 concerns voters who are demoters, and follows immediately from Lemma 1; note also that it does not depend on the minimality assumption.

Proposition 3. *Consider a GS-game $G = (V, \mathcal{R}_2, (A_i)_{i \in N(V,\mathcal{R}_2)})$. Let w be the 2-approval winner at V. Consider a player $i \in N(V, \mathcal{R}_2) \setminus \{1\}$ such that $w \notin \mathrm{top}_2(v_i)$ and the set A_i consists of i's truthful vote and a subset of i's minimal manipulations. Let $\mathrm{top}_2(v) = \{a, a'\}$. Then there is a candidate $c \in C \setminus \{a, a'\}$ such that for each $v \in A_i$ we have $\mathrm{top}_2(v) \in \{\{a, a'\}, \{a, c\}, \{a', c\}\}$.*

Proof. Player i cannot lower the score of w by changing his vote. However, he can raise the scores of some candidates in $C \setminus \mathrm{top}_2(v_i)$ by moving these candidates into top two positions. In general, i can do that for two candidates simultaneously; however, the minimality assumption implies that i only moves one candidate into the top two positions. Thus, i is a promoter (see Section 2). For a vote v_1' to be a level-1 strategy the promoted candidate has to be i's most preferred candidate in $S(V, 2) \setminus \mathrm{top}_2(v_i)$ (let us denote this candidate by p). Thus, in this case voter i has 3 options: (1) to vote truthfully, (2) to swap p with the candidate that he ranks first or (3) to swap p with the the candidate he ranks second. This completes the proof □

Propositions 2 and 3 enable us to establish an analogue of Lemma 3 for 2-approval under the minimality assumption.

Lemma 6. *There is a polynomial-time procedure*

$$Alg' = Alg'(G, r, r', x, y, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]})$$

that, given a GS-game $G = (V, \mathcal{R}_2, (A_i)_{i \in N(V,\mathcal{R}_2)})$ *with* $|V| = n$*, where for each* $i \in N \setminus \{1\}$ *the set* A_i *consists of* i*'s truthful vote and a subset of* i*'s minimal manipulations, two integers* $r, r' \in \{0, \dots, n\}$*, two distinct candidates* $x, y \in C$*, and a partition of candidates in* $C \setminus \{x, y\}$ *into* $C^{[1]}$*,* $C^{[0]}$*,* $C^{[-1]}$ *and* $C^{[-2]}$*, decides whether there is a strategy profile* V^* *in* G *such that*

- $\mathrm{sc}_2(x, V[V^*]_{-1}) = r$*,*
- $\mathrm{sc}_2(y, V[V^*]_{-1}) = r'$*, and*
- *for each* $\ell \in \{1, 0, -1, -2\}$ *and each* $c \in C^{[\ell]}$ *it holds that* $\mathrm{sc}_2(c, V[V^*]_{-1}) \leq r + \ell$*.*

Proof. Let w be the 2-approval winner at V. If $S(V, 2) \neq \emptyset$, set $p^* = p^*(V, 2)$. We use essentially the same construction as in the proof of Lemma 3. Specifically, the set of nodes consists of a sink, a source, one node for each voter in $[n] \setminus \{1\}$, and one node for each candidate $c \in C$. For each $i \in N \setminus \{1\}$, the capacity and the lower bound of the arc from the source to node i are equal to 2, and the capacities and lower bounds of the arcs from candidates to the source are defined as in the proof of Lemma 3. It remains to describe the arcs connecting voters and candidates.

If $i \notin N$, we add an arc from i to c for each $c \in \mathrm{top}_2(v_i)$; the capacity and the lower bound of these arcs are 1, encoding the fact that i has to vote for his top 2 candidates.

Now, consider a voter $i \in N \setminus \{1\}$ who is a demoter; if such a voter exists, we have $S(V, 2) \neq \emptyset$ and hence p^* is well-defined. By Proposition 2 we have $\mathrm{top}_2(v_i) = \{p^*, w\}$ and $A_i = \{v_i[w; c] \mid c \in C_i\}$ for some $C_i \subset C \setminus \{p^*, w\}$. Then we introduce an arc from i to p^* whose capacity and lower bound are both set to 1, and arcs with capacity 1 and lower bound 0 from i to each $c \in C_i \cup \{w\}$.

Finally, consider a voter $i \in N \setminus \{1\}$ who is a promoter; let $\mathrm{top}_2(v_i) = \{a, a'\}$ and let p be i's most preferred candidate in $S(2, V) \setminus \{a, a'\}$. If A_i contains both a vote v' with $\mathrm{top}_2(v') = \{a, p\}$ and a vote v'' with $\mathrm{top}_2(v'') = \{a', p\}$, then by Proposition 3 it suffices to add arcs with capacity 1 and lower bound 0 that go from i to a, a', and p. If we have $\mathrm{top}_2(v) \in \{\{a, a'\}, \{a, p\}\}$ for each $v \in A_i$, we add an arc with capacity 1 and lower bound 1 from i to a and arcs with capacity 1 and lower bound 0 from i to a' and p. Similarly, if we have $\mathrm{top}_2(v) \in \{\{a, a'\}, \{a', p\}\}$ for each $v \in A_i$, we add an arc with capacity 1 and lower bound 1 from i to a' and arcs with capacity 1 and lower bound 0 from i to a and p.

It is clear from the construction that a valid integer flow in this network corresponds to a strategy profile V^* with the desired properties. □

We are now ready to prove the main result of this section.

Theorem 7. *Given a GS-game* $G = (V, \mathcal{R}_2, (A_i)_{i \in N(V,\mathcal{R}_2)})$*, where for each* $i \in N \setminus \{1\}$ *the set* A_i *consists of* i*'s truthful vote and a subset of* i*'s minimal manipulations, and two strategies* $v_1', v_1'' \in \mathcal{L}(C)$ *of player 1, we can decide in polynomial time whether* v_1' *weakly dominates* v_1''*.*

Proof. Just as in the proof of Theorem 4, it suffices to design a polynomial-time procedure that, given two strategies u, v of player 1, decides if there exists a profile V^*_{-1} of other players' strategies such that $\mathcal{R}_2(V[V^*_{-1}, u]) \succ_1 \mathcal{R}_2(V[V^*_{-1}, v])$. Let $\text{top}_2(u) = \{a, a'\}$, $\text{top}_2(v) = \{b, b'\}$. We can assume that $\{a, a'\} \neq \{b, b'\}$, and we will focus on the case where $\{a, a'\} \cap \{b, b'\} = \emptyset$; the case where $\{a, a'\} \cap \{b, b'\}$ is a singleton is similar (and simpler).

We use the same notation as in the proof of Theorem 4: given a profile V^*_{-1} of other players' strategies, we let $V^u = V[V^*_{-1}, u]$, $V^v = V[V^*_{-1}, v]$, $w^u = \mathcal{R}_2(V^u)$, $w^v = \mathcal{R}_2(V^v)$. Our goal then is to decide if there exists a profile V^*_{-1} such that $w^u \succ_1 w^v$. To this end, we go over all values of $t \in [n-1]$ and all candidates $w, w' \in C$ with $w \succ_1 w'$, and ask if there is a profile V^*_{-1} such that w wins at $V[V^*_{-1}, u]$ with t points, whereas w' wins at $V[V^*_{-1}, v]$. As in the proof of Theorem 4, for each triple (t, w, w') we have to consider a number of possibilities, depending on whether $w \in \{a, a'\}$, $w' \in \{b, b'\}$ as well as on the relative positions of w, w', a, a', b, and b' with respect to the tie-breaking order. The analysis is as tedious as it is straightforward; to illustrate the main points, we consider two representative cases.

$\boldsymbol{w = a, w' = b, a > a' > b > b'}$ In this case, a wins with t points at V^u if and only if $\text{sc}_2(a', V^u_{-1}) \leq t-1$ and for each $c \in C \setminus \{a, a'\}$ we have $\text{sc}_2(c, V^u_{-1}) \leq t$ if $a > c$ and $\text{sc}_2(c, V^u_{-1}) \leq t-1$ if $c > a$. Suppose that these conditions are satisfied. Then b can win at V^v with $t+1$ or t points. The former case is possible if and only if $\text{sc}_2(b, V^u_{-1}) = t$. The latter case is possible if and only if $\text{sc}_2(b, V^u_{-1}) = t-1$, $\text{sc}_2(b', V^u_{-1}) \leq t-1$, and for each $c \in C \setminus \{a, b, b'\}$ we have $\text{sc}_2(c, V^u_{-1}) \leq t$ if $b > c$ and $\text{sc}_2(c, V^u_{-1}) \leq t-1$ if $c > b$.

Thus, to decide whether this situation is possible, we have to call Alg' twice. For our first call, we set $C^{[1]} = \{c \in C \setminus \{a', b\} \mid a > c\}$, $C^{[0]} = \{c \in C \setminus \{a'\} \mid c > a\} \cup \{a'\}$, $C^{[-1]} = C^{[-2]} = \emptyset$ and call

$$Alg'(G, a, b, t-1, t, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]}).$$

For our second call, we set $C^{[1]} = \{c \in C \setminus \{b'\} \mid b > c\}$, $C^{[0]} = \{c \in C \setminus \{a\} \mid c > b\} \cup \{b'\}$, $C^{[-1]} = C^{[-2]} = \emptyset$ and call

$$Alg'(G, a, b, t-1, t-1, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]}).$$

$\boldsymbol{w \notin \{a, a', b, b'\}, w' = b, b' > b > a' > w > a}$ If w wins at V^u with t points, this means that $\text{sc}_2(w, V^u_{-1}) = t$, $\text{sc}_2(a', V^u_{-1}) \leq t-2$, $\text{sc}_2(a, V^u_{-1}) \leq t-1$, $\text{sc}_2(c, V^u_{-1}) \leq t$ for all $c \in C \setminus \{w, a, a'\}$ with $w > c$, and $\text{sc}_2(c, V^u_{-1}) \leq t-1$ for all $c \in C \setminus \{w, a, a'\}$ with $c > w$. Suppose that these conditions are satisfied. As w still receives t points at V^v, this means that b wins at V^v if and only if $\text{sc}_2(b, V^u_{-1}) = t-1$, $\text{sc}_2(b', V^u_{-1}) \leq t-2$. Thus, we set $C^{[1]} = \emptyset$, $C^{[0]} = \{c \in C \setminus \{a\} \mid w > c\}$, $C^{[-1]} = \{c \in C \setminus \{a', b, b'\} \mid c > w\}$, $C^{[-2]} = \{a', b'\}$ and call

$$Alg'(G, w, b, t, t-1, C^{[1]}, C^{[0]}, C^{[-1]}, C^{[-2]}).$$

□

Just as for Plurality, we obtain the following corollary that describes the complexity of finding and testing level-2 strategies and improving strategies under 2-approval.

Corollary 8. *Given a GS-game $G = (V, \mathcal{R}_2, (A_i)_{i \in N(V, \mathcal{R}_2)})$, where for each $i \in N$ the set A_i consists of i's truthful vote and a subset of his minimal manipulations, and a strategy $v'_1 \in \mathcal{L}(C)$ of player 1 we can decide in polynomial time whether v'_1 is a level-2 strategy and whether v'_1 is an improving strategy. Moreover, we can decide in polynomial time whether player 1 has a level-2 strategy or an improving strategy in G.*

We remark that the minimality assumption plays an important role in our analysis. Indeed, in the absence of this assumption a promoter i may manipulate by swapping two different candidates (one of which is his most preferred 2-competitive candidate p) into the top two positions. Let $v_i[\mathrm{top}_2(v_i); \{p, c\}]$ be some such manipulation. If we try to model this possibility via a network flow construction, we would have to add edges from i to both p and c; the lower bounds on these edges would have to be set to 0, to allow i to vote truthfully. However, there may then be a flow that uses the edge (i, c), but not (i, p), which corresponds to a vote that promotes c, but not p; such a vote is not a level-1 strategy.

Interestingly, a level-2 player may want to swap two candidates into the top two positions, even if he assumes that all level-1 players use minimal strategies. In fact, the following example shows that a strategy of this form may weakly dominate all other non-equivalent strategies.

Example 5. Let the profile of sincere preferences be as in Table 5, and assume that the voting rule is 2-approval and the tie-breaking order is given by $a > b > c > d > \dots$. Assume that player 1 is the level-2 player. The winner under 2-approval is a with two points; candidates b, c, and d also have two points each. Voters 1, 4 and 5 are GS-manipulators; voter 4 may manipulate by swapping c into top two positions, and voter 5 may manipulate by swapping d into top two positions.

Consider the GS-game where $N = \{1, 4, 5\}$, $A_4 = \{v_4, v_4[b; c]\}$, and $A_5 = \{v_5, v_5[b; d]\}$ (note that both A_4 and A_5 only contain a proper subset of the respective player's minimal manipulations; for instance, $v_4[u_1; c] \notin A_4$). We claim that $v_1[\{u_5, u_6\}; \{b, c\}]$ is a weakly dominant strategy for player 1. Indeed, consider the four possible scenarios:

- Players 4 and 5 are truthful. Then the best outcome that voter 1 can ensure is that b wins.
- Player 4 is truthful, but player 5 manipulates. Then the best outcome that voter 1 can ensure is that c wins.
- Player 4 manipulates, but player 5 is truthful. Then the best outcome that voter 1 can ensure is that c wins.
- Players 4 and 5 both manipulate. Then the best outcome that voter 1 can ensure is that c wins.

Now, it is clear that only votes that rank b and c in top two positions achieve all of these objectives simultaneously.

voter 1	voter 2	voter 3	voter 4	voter 5	voter 6	voter 7
u_5	a	a	b	b	c	c
u_6	d	d	u_1	u_2	u_3	u_4
b	$\dots$	$\dots$	c	d	a	a
c	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
d	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$

Table 5: The strategy $bc\dots$ of player 1 weakly dominates all non-equivalent strategies.

7 k-Approval for $k \geq 3$

Regrettably, our analysis of k-approval under the minimality assumption does not extend from $k = 2$ to $k = 3$. Specifically, the argument breaks down when we consider a potential demoter under 3-approval who can only help his top candidate by swapping his second and third candidate out of the top three positions. If he chooses to manipulate, he has to perform both of these swaps at once; he can also remain truthful and not perform any swaps. It is not clear how to capture this all-or-nothing behavior via network flows. We conjecture that finding a level-2 strategy under 3-approval is computationally hard, even under minimality assumption. We will now prove a weaker result, showing that this problem is NP-hard for k-approval with $k \geq 4$ (and without the minimality assumption). Moreover, we will also show that it is coNP-hard to decide whether a given strategy is improving.

Theorem 9. *For every fixed $k \geq 4$, given a GS-game $G = (V, \mathcal{R}_k, (A_i)_{i \in N})$ and a strategy v of voter 1, it is* NP*-hard to decide whether v is a level-2 strategy, and it is* coNP*-hard to decide whether v is an improving strategy.*

Proof. Our hardness proof proceeds by a reduction from the classic NP-complete problem Exact Cover by 3-Sets (X3C). An instance of this problem is given by a ground set $\Gamma = \{g_1, \ldots, g_{3\nu}\}$ and a collection $\Sigma = \{\sigma_1, \ldots, \sigma_\mu\}$ of 3-element subsets of Γ. It is a 'yes'-instance if there is a subcollection $\Sigma' \subseteq \Sigma$ with $|\Sigma'| = \nu$ such that $\cup_{\sigma \in \Sigma'} \sigma = \Gamma$, and a 'no'-instance otherwise.

We will first establish that our problems are hard for $k = 4$; towards the end of the proof, we will show how to extend our argument to $k > 4$.

Consider an instance $I^0 = (\Gamma^0, \Sigma^0)$ of X3C with $|\Gamma| = 3\nu'$. We will first modify this instance as follows. We add three new elements to Γ^0 and a set containing them to Σ^0. We then add $\nu' + 2$ triples x_i, y_i, z_i, $i \in [\nu' + 2]$, of new elements to Γ^0 and for each such triple we add the set $S_i = \{x_i, y_i, z_i\}$ to Σ^0. Finally, we add sets $S'_i = \{y_i, z_i, x_{i+1}\}$, $i \in [\nu' + 1]$, and $S'_{\nu'+2} = \{y_{\nu'+2}, z_{\nu'+2}, x_1\}$ to Σ^0. We then renumber the elements of the ground set so that the elements added at the first step are numbered g_1, g_2, g_3. We denote the resulting instance by (Γ, Σ), and let $\nu = |\Gamma|/3$, $\mu = |\Sigma|$. Clearly, $I = (\Gamma, \Sigma)$ is a 'yes'-instance of X3C if and only if $I = (\Gamma^0, \Sigma^0)$ is. We let $\widehat{\Sigma} = \{S_i, S'_i \mid i \in [\nu' + 2]\}$; we have $\nu = 2\nu' + 3$, $|\widehat{\Sigma}| = 2(\nu' + 2)$.

In what follows, when writing $X \succ Y$ in the description of an order $\succ$, we mean that all elements of X are ranked above all elements of Y, but the order of elements within X and within Y is not specified and can be arbitrary. We construct a GS-game as follows. We introduce a set of candidates $C' = \{c_1, \ldots, c_{3\nu}\}$ that correspond to elements of Γ, three special candidates w, p, c, and, finally, a set of dummy candidates

$$D = \bigcup_{i=0}^{\mu} D_i \cup D_c \cup \bigcup_{j=1}^{\nu+1} E_j \cup \bigcup_{i=1}^{3\nu} \bigcup_{j=1}^{\nu+1} F_{i,j},$$

where $|D_i| = 4$ for $i = 0, \ldots, \mu$, $|D_c| = 3$, and $|E_j| = 2$, $|F_{i,j}| = 3$ for $i \in [3\nu]$, $j \in [\nu + 1]$. Thus, the set of candidates is $C = \{w, p, c\} \cup C' \cup D$. We define the tie-breaking order $>$ on C by setting

$$w > c > p > c_1 > \cdots > c_{3\nu} > D.$$

For each $j \in [\mu]$, we let $C_j = \{c_i \mid g_i \in \sigma_j\}$. The profile V consists of $2 + \mu + (3\nu + 1)(\nu + 1)$

votes defined as follows:

$$\begin{aligned}
z_0 &= D_0 \succ p \succ c_1 \succ c \succ C' \setminus \{c_1\} \succ \dots, \\
z_i &= D_i \succ C_i \succ c \succ \dots, && i \in [\mu], \\
u &= c \succ D_c \succ w \succ \dots, \\
u_j &= w \succ p \succ E_j \succ \dots, && j \in [\nu+1], \\
u_{i,j} &= c_i \succ F_{i,j} \succ w \succ \dots && i \in [3\nu],\ j \in [\nu+1].
\end{aligned}$$

We have

$$\mathrm{sc}_4(w, V) = \mathrm{sc}_4(p, V) = \mathrm{sc}_4(c_i, V) = \nu + 1 \quad \text{for all } i \in [3\nu],$$

$\mathrm{sc}_4(c, V) = 1$, and $\mathrm{sc}_4(d, V) \le 1$ for each $d \in D$. Thus, w wins under 4-approval because of the tie-breaking rule.

We have $S(V, 4) = C' \cup \{p\}$. The set of GS-manipulators in this profile consists of the first $\mu + 1$ voters; we assume that the first voter (i.e., voter 0) is the level-2 voter. We now define a GS-game for this profile by constructing the players' sets of strategies as follows:

$$z'_i = z_i[D_i; C_i \cup \{c\}], \quad A_i = \{z_i, z'_i\} \text{ for all } i \in [\mu].$$

Observe that for each $i \in [\mu]$ the vote z'_i is a level-1 strategy for voter i, which makes i's top candidate in C_i the winner with $\nu + 2$ points (note that voter i orders C_i in the same way as $>$ does, so tie-breaking favours i's most preferred candidate in C_i). This completes the description of our game G.

Fix some $d, d' \in D_0$ and let

$$z'_0 = z_0[\{d, d'\}; \{p, c\}], \qquad z''_0 = z_0[d; p].$$

Note that both z'_0 and z''_0 are level-1 strategies for voter 0, which make p the winner with $\nu + 2$ points. Clearly, we can construct the profile V and the players' sets of strategies in polynomial time given I.

We will now argue that z'_0 is an improving strategy if and only if $I = (\Gamma, \Sigma)$ is a 'no'-instance of X3C, and that z''_0 is a level-2 strategy if and only if $I = (\Gamma, \Sigma)$ is a 'yes'-instance of X3C.

As a preliminary observation, consider some strategy z of voter 0 such that $\mathrm{top}_4(z)$ consists of p and three dummy candidates. By construction, for every profile of other players' strategies, z and z''_0 result in the same outcome. Moreover, if everyone except voter 0 votes truthfully, voter 0 strictly prefers z''_0 to every strategy $\hat{z}$ with $\mathrm{top}_4(\hat{z}) \subseteq D$. Thus, z''_0 can only be weakly dominated by a strategy that places at least one candidate in $C' \cup \{c, w\}$ in top 4 positions.

Suppose first that I is a 'yes'-instance of X3C. Fix a subcollection Σ' witnessing this, and consider a profile V' where the GS-manipulators that correspond to sets in Σ' vote strategically, whereas everyone else votes truthfully. We have $\mathrm{sc}_4(p, V') = \mathrm{sc}_4(c, V') = \mathrm{sc}_4(w, V') = \nu + 1$, $\mathrm{sc}_4(c_i, V') = \nu + 2$ for all $c_i \in C'$, so c_1 wins. However, if voter 0 changes his vote to z'_0, the winner would be c, and voter 0 prefers c_1 to c, so voter 0 strictly prefers voting z_0 over voting z'_0 in this case, i.e., z'_0 is not an improving strategy.

Now, if voter 0 changes her vote to z''_0 instead, p becomes the election winner, which is the best feasible outcome from voter 0's perspective. The only way for voter 0 to achieve this outcome is to rank p and some dummy candidates in the top 4 positions; any vote $\widehat{z}$ with $\mathrm{top}_4(\widehat{z}) \cap (C' \cup \{c, w\}) \neq \emptyset$ is strictly worse for voter 0, and hence cannot weakly dominate z''_0.

As we have already observed that no strategy $\widehat{z}$ with $\mathrm{top}_4(\widehat{z}) \cap (C' \cup \{c, w\}) = \emptyset$ can weakly dominate z''_0, it follows that if I is a 'yes'-instance of X3C then z''_0 is a level-2 strategy.

On the other hand, suppose that I is a 'no'-instance of X3C. Consider a strategy profile V^* in G, and let Σ'' be a subcollection of Σ that corresponds to players in $[\mu]$ who vote non-truthfully in V^*; we know that Σ'' is not an exact cover of Γ. We will argue that voter 0 weakly prefers z'_0 to both z_0 and z''_0 for every choice of Σ'', and there are choices of Σ'' for which this preference is strict.

If $\Sigma'' = \emptyset$, i.e., all voters in $[\mu]$ are truthful, then voter 0 benefits from changing his vote from z_0 to z'_0, as this vote makes p the winner. Similarly, suppose that all sets in Σ'' are pairwise disjoint (and hence $|\Sigma''| \leq \nu - 1$). Then candidate c gets at most ν points and the winner in $V[V^*_{-0}, z_0]$ is one of the candidates from C' (with $\nu + 2$ points). On the other hand, the winner in $V[V^*_{-0}, z'_0]$ is p (with $\nu + 2$ points), so voter 0 benefits from changing his vote to z'_0. In both of these cases, z''_0 has the same effect as z'_0.

Now, suppose that the sets in Σ'' are not pairwise disjoint. Let X be the set of elements that appear in the largest number of sets in Σ'', and let g_ℓ be the element of X with the smallest index. Note that $g_\ell \neq g_1$, since we modified our instance of X3C so that g_1 only occurs in one set. The winner in $V[V^*_{-0}, z_0]$ is either c_ℓ or c, and the winner's score is at least $\nu + 3$. Suppose that voter 0 changes his vote from z_0 to z'_0. If the winner in $V[V^*_{-0}, z_0]$ was c, this remains to be the case, and if the winner was c_ℓ then either c_ℓ remains the winner or c becomes the winner, and voter 0 prefers c to c_ℓ. Thus, in this case voting z'_0 is at least as good as voting z_0, and voting z''_0 has the same effect as voting z_0.

We conclude that whenever Σ'' is not an exact cover of Γ, voting z'_0 is at least as good as voting z_0 or z''_0. It remains to establish that z'_0 is sometimes strictly better than either of these strategies. To this end, suppose that $\Sigma'' = \widehat{\Sigma}$. If voter 0 votes z'_0, then the scores of the candidates covered by sets in $\widehat{\Sigma}$ are $\nu + 3$, the score of c is $1 + 2(\nu' + 2) + 1 = 2\nu' + 6 = \nu + 3$, and all other candidates have lower scores, so c wins. However, if voter 0 votes z_0 or z''_0, c's score is $\nu + 2$, and therefore the winner is a candidate in C'. Thus, in those circumstances, voter 0 strictly prefers z'_0 to both z_0 and z''_0. Hence, if I is a 'no'-instance of X3C, z'_0 weakly dominates z_0 and z''_0, and hence z''_0 is not a level-2 strategy. This completes the proof for $k = 4$.

For $k > 4$, we modify the construction by introducing $|V|$ additional groups of dummy candidates $H_1, \ldots, H_{|V|}$ of size $k - 4$ each. We renumber the voters from 1 to $|V|$ and modify the preferences of the i-th voter, $i \in [\mu]$, by inserting the group H_i in positions $5, \ldots, k$, and adding all other new dummy candidates at the bottom of his ranking. Then the k-approval scores of all candidates in C remain the same as in the original construction, and the k-approval score of each new dummy candidate is 1. The rest of the proof then goes through without change. □

We note that the strategies of level-1 players in our hardness proof are not minimal; determining whether our hardness result remains true under the minimality assumption is an interesting research challenge.

Remark 1. *Our complexity lower bounds are not tight: we do not know whether the computational problems we consider are in, respectively,* NP *and* coNP*. The following argument provides upper bounds on their complexity.*

For every n-player game $G = (N, (A_i)_{i \in N}, (\succeq_i)_{i \in N})$, where each relation $\succeq_i$ is represented by a polynomial-time computable function of its arguments, and for every pair of strategies u, v of player 1, the problem of deciding whether u weakly dominates v belongs to the complexity *class* DP *(difference polynomial-time) (Papadimitriou and Yannakakis, 1984; Wechsung, 1985). Indeed, u weakly dominates v if and only if*

(a) for every profile P_{-1} of other players' strategies we have $(P_{-1}, u) \succeq_1 (P_{-1}, v)$ (which can be checked in coNP*), and*

(b) for some profile P_{-1} of other players' strategies we have $(P_{-1}, u) \succ_1 (P_{-1}, v)$ (which can be checked in NP*),*

i.e., the language associated with our problem is an intersection of an NP*-language and a* coNP*-language.*

Thus, for every GS-game based on a polynomial-time voting rule (including k-approval) the problem of checking whether a given strategy is improving is in DP*. This also means that for k-approval with a fixed value of k the problem of checking whether a given strategy is a level-2 strategy belongs to the boolean hierarchy (Wechsung, 1985), as there are only $\binom{m}{k} \le m^k$ pairwise non-equivalent votes, and it suffices to check that none of these votes weakly dominates the given strategy.*

8 Conclusions and Further Research

We have initiated the analysis of voting games from the perspective of the cognitive hierarchy. We have adopted a distribution-free approach that uses the concept of weak dominance in order to reason about players' actions. The resulting framework is mathematically rich, captures some interesting behaviours, and presents a number of algorithmic challenges, even for simple voting rules. To illustrate this, we focused on a well-known family of voting rules, namely, k-approval with $k \ge 1$, and investigated the complexity of finding level-2 strategies and improving strategies under various rules in this family. For Plurality, i.e., for $k = 1$, level-2 strategies and improving strategies are easy to find, and for $k \ge 4$ these problems are computationally hard, but for $k = 2, 3$ we do not have a full understanding of the complexity of these problems. We identify a natural assumption (namely, the minimality assumption), which is sufficient to obtain an efficient algorithm for $k = 2$; however, it is not clear if it remains useful for larger values of k.

We list a few specific algorithmic questions that remain open:

- Is there a polynomial-time algorithm for computing level-2 strategies and improving strategies under 2-approval without the minimality assumption?
- Does Theorem 9 remain true under the minimality assumption?
- What can be said about 3-approval, with or without the minimality assumption?
- What can be said about other prominent voting rules, most importantly the Borda rule?

In our analysis, we have focused on level-1 and level-2 voters. It would also be interesting to extend our formal definitions to level-ℓ players for $\ell \ge 3$ and to investigate the associated algorithmic issues. While it is intuitively clear that the view of the game for these players will be more complex, it appears that for Plurality our algorithm can be extended in a straightforward manner; however, it is not clear if this is also the case for 2-approval. Another interesting question, which can be analysed empirically, is whether truthful voting is likely to be a level-2 strategy, or, more broadly, how many votes in $\mathcal{L}(C)$ are level-2 strategies; again, this question can also be asked for level-ℓ strategies with $\ell \geq 2$. Note that the analysis in our paper contributes algorithmic tools tools to tackle this issue.

A yet broader question, which can only be answered by combining empirical data and theoretical analysis, is whether the cognitive hierarchy approach provides a plausible description of strategic behavior in voting. While our paper makes the first steps towards answering it, there is more to be done to obtain a full picture.

References

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

M. Battaglini. Sequential voting with abstention. *Games and Economic Behavior*, 51:445–463, 2005.

A. Brandenburger. Knowledge and equilibrium in games. *The Journal of Economic Perspectives*, 6(4):83–101, 1992.

L. Buenrostro, A. Dhillon, and P. Vida. Scoring rule voting games and dominance solvability. *Social Choice and Welfare*, 40(2):329–352, 2013.

B. Cain. Strategic voting in Britain. *American Journal of Political Science*, 22:639–655, 1978.

C. F. Camerer, T.-H. Ho, and J.-K. Chong. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898, 2004.

G. Cox. *Making Votes Count*. Cambridge University Press, 1997.

V. P. Crawford, M. A. Costa-Gomes, and N. Iriberri. Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51 (1):5–62, 2013.

A. Dellis. Weak undominance in scoring rule elections. *Mathematical Social Sciences*, 59(1): 110–119, 2010.

Y. Desmedt and E. Elkind. Equilibria of plurality voting with abstentions. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC-2010)*, pages 347–356, 2010.

A. Dhillon and B. Lockwood. When are plurality rule voting games dominance-solvable? *Games and Economic Behavior*, 46:55–75, 2004.

B. Dutta and A. Sen. Nash implementation with partially honest individuals. *Games and Economic Behavior*, 74(1):154–169, 2012.

E. Elkind, U. Grandi, F. Rossi, and A. Slinko. Gibbard–Satterthwaite games. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI-2015)*, pages 533–539, 2015a.

E. Elkind, S. Obraztsova, E. Markakis, and P. Skowron. Equilibria of plurality voting: Lazy and truth-biased voters. In *Proceedings of the 8th International Symposium on Algorithmic Game Theory (SAGT-2015)*, pages 110–122, 2015b.

U. Endriss, S. Obraztsova, M. Polukarov, and J. S. Rosenschein. Strategic voting with incomplete information. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-2016)*, pages 236–242, 2016.

R. Farquharson. *Theory of Voting*. Yale University Press, 1969.

T. J. Feddersen, I. Sened, and S. G. Wright. Rational voting and candidate entry under plurality rule. *American Journal of Political Science*, 34(4):1005–1016, 1990.

D. Fudenberg and D. K. Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

U. Grandi, D. Hughes, F. Rossi, and A. Slinko. Gibbard-Satterthwaite games for k-approval voting rules. Technical report, arXiv:1707.05619, 2017.

J. C. Harsanyi and R. Selten. *A general theory of equilibrium selection in games*. MIT Press, 1988.

N. Hazon and E. Elkind. Complexity of safe strategic voting. In *Proceedings of the 3rd International Symposium on Algorithmic Game Theory (SAGT-2010)*, pages 210–221, 2010.

A. Koolyk, T. Strangeway, O. Lev, and J. S. Rosenschein. Convergence and quality of iterative voting under non-scoring rules. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-2017)*, 2017.

O. Lev and J. S. Rosenschein. Convergence of iterative scoring rules. *Jornal of Artificial Intelligence Research*, 57:573–591, 2016.

R. Meir. Iterative voting. In U. Endriss, editor, *Trends in Computational Social Choice*. AI Access, 2017.

R. Meir, M. Polukarov, J. S. Rosenschein, and N. R. Jennings. Convergence to equilibria of plurality voting. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI-2010)*, pages 823–828, 2010.

R. Meir, O. Lev, and J. S. Rosenschein. A local-dominance theory of voting equilibria. In *Proceedings of the 15th ACM Conference on Electronic Commerce (EC-2014)*, pages 313–330, 2014.

M. Messner and M. K. Polborn. Strong and coalition-proof political equilibria under plurality and runoff rule. *International Journal of Game Theory*, 35(2):287–314, 2007.

H. Moulin. Dominance solvable voting schemes. *Econometrica*, 47:1337–1351, 1979.

D. P. Myatt. On the theory of strategic voting. *The Review of Economic Studies*, 74:255–281, 2007.

R. Myerson and R. Weber. A theory of voting equilibria. *American Political Science Review*, 87(1):102–114, 1993.

R. Nagel. Unraveling in guessing games: An experimental study. *The American Economic Review*, 85(5):1313–1326, 1995.

S. Obraztsova and E. Elkind. Optimal manipulation of voting rules. In *Proceedings of 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2012)*, pages 619–626, 2012.

S. Obraztsova, E. Markakis, and D. R. M. Thompson. Plurality voting with truth-biased agents. In *Proceedings of the 6th International Symposium on Algorithmic Game Theory (SAGT-2013)*, pages 26–37, 2013.

S. Obraztsova, O. Lev, E. Markakis, Z. Rabinovich, and J. S. Rosenschein. Beyond plurality: Truth-bias in binary scoring rules. In *Proceedings of the 4th International Conference on Algorithmic Decision Theory (ADT-2015)*, pages 451–468, 2015a.

S. Obraztsova, E. Markakis, M. Polukarov, Z. Rabinovich, and N. R. Jennings. On the convergence of iterative voting: How restrictive should restricted dynamics be? In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI-2015)*, pages 993–999, 2015b.

S. Obraztsova, Z. Rabinovich, E. Elkind, M. Polukarov, and N. R. Jennings. Trembling hand equilibria of plurality voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-2016)*, pages 440–446, 2016.

C. H. Papadimitriou and M. Yannakakis. The complexity of facets (and some facets of complexity). *Journal of Computer and System Sciences*, 28(2):244–259, 1984.

P. K. Pattanaik. Threats, counter-threats, and strategic voting. *Econometrica: Journal of the Econometric Society*, pages 91–103, 1976.

H. Peters, S. Roy, and T. Storcken. On the manipulability of approval voting and related scoring rules. *Social Choice and Welfare*, 39(2-3):399–429, 2012.

A. Reijngoud and U. Endriss. Voter response to iterated poll information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2012)*, pages 635–644, 2012.

R. Reyhaneh and M. Wilson. Best reply dynamics for scoring rules. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI-2012)*, pages 672–677, 2012.

M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10 (2):187–217, 1975.

L. S. Shapley. Some topics in two-person games. *Advances in game theory, Annals of Mathematical Studies*, 52:1–29, 1964.

A. Slinko and S. White. Is it ever safe to vote strategically? *Social Choice and Welfare*, 43: 403–427, 2014.

D. O. Stahl and P. W. Wilson. Experimental evidence on players' models of other players. *Journal of economic behavior & organization*, 25(3):309–327, 1994.

D. R. M. Thompson, O. Lev, K. Leyton-Brown, and J. S. Rosenschein. Empirical analysis of plurality election equilibria. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-13)*, pages 391–398, 2013.

G. Wechsung. On the boolean closure of NP. In *Proceedings of the 4th International Conference on Fundamentals of Computation Theory (FCT-1985)*, pages 485–493, 1985.

J. R. Wright and K. Leyton-Brown. Beyond equilibrium: Predicting human behavior in normal-form games. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI-2010)*, pages 901–907, 2010.

L. Xia and V. Conitzer. Stackelberg voting games: Computational aspects and paradoxes. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI-2010)*, pages 805–810, 2010.