

# Gibbard–Satterthwaite Games

**Edith Elkind**

University of Oxford  
United Kingdom  
eelkind@gmail.com

**Umberto Grandi**

University of Toulouse  
France  
umberto.grandi@irit.fr

**Francesca Rossi**

University of Padova  
and Harvard University  
frossi@math.unipd.it

**Arkadii Slinko**

University of Auckland  
New Zealand  
a.slinko@auckland.ac.nz

## Abstract

The Gibbard–Satterthwaite theorem implies the ubiquity of manipulators–voters who could change the election outcome in their favor by unilaterally modifying their vote. In this paper, we ask what happens if a given profile admits several such voters. We model strategic interactions among Gibbard–Satterthwaite manipulators as a normal-form game. We classify the 2-by-2 games that can arise in this setting for two simple voting rules, namely Plurality and Borda, and study the complexity of determining whether a given manipulative vote weakly dominates truth-telling, as well as existence of Nash equilibria.

## 1 Introduction

Voting is a common method of preference aggregation, which enables the participating agents to identify the best candidate given the individual agents’ rankings of the candidates. However, a common concern of electoral designers is that no “reasonable” voting rule is immune to manipulation: as shown by Gibbard [1973] and Satterthwaite [1975], for three or more candidates, every onto, non-dictatorial voting rule admits a preference profile (a collection of voters’ rankings) where some voter would be better off by submitting a ranking that differs from his truthful one. This cornerstone result is partially mitigated by the observation, developed in the last decade in the research area of computational social choice [Brandt *et al.*, 2015], that voters will not necessarily respond to an incentive of manipulation, in case such an action presents heavy computational costs.

An additional factor that may prevent strategic behavior in voters is that would-be manipulators must interact between themselves. The analysis of this interaction has been a very active area of research in political science, where several theoretical works have attempted to describe the micrologic behind strategic voting (e.g., [Cain, 1978; Cox, 1997; Myatt, 2007]). Most of these works concerns elections using the Plurality rule; for other voting rules the situation may be even more complex. On the one hand, some voters may

not be able to manipulate on their own, but have an opportunity to take a defensive action which will destroy somebody’s opportunity to manipulate. On the other hand, such counter-manipulators are uncertain whether the manipulation, which they are going to counter, will take place or not. The voters in most elections are typically acting independently and in the absence of communication, hence strategic voters must play a simultaneous one-shot game in a non-cooperative setting.<sup>1</sup>

Apart from the practical difficulties of manipulating, some voters may be *ideological* and interested in stating their preference no matter what. Practical observations suggest that the number of ideological voters is significant. 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 the overwhelming majority of Nader supporters preferred Gore to Bush. At the other end of the spectrum, truly *strategic* voters are interested in optimizing the outcome of the election and may or may not submit their insincere vote depending on strategic considerations. It is very hard to detect the number of strategic voters in a given election, but the percentage of those who actually manipulated may be easier to obtain. For instance, [Kawai and Watanabe, 2013] estimate the number of such voters in Japanese elections between 2,5% and 5,5%. Moreover, [Benjamin *et al.*, 2013] show that preference misrepresentation is related to cognitive skills, and [Choi *et al.*, 2014] demonstrate that decision-making ability in laboratory experiments correlates strongly with socio-economic status and wealth. Therefore, it is reasonable to assume that only a small fraction of voters in an election would act strategically when given an opportunity to do so.

The goal of our paper is to develop a basic understanding of how multiple manipulators can help or hurt each other by acting simultaneously, under some simple voting rules. In our model non-manipulators do not act strategically, since we assume that the strategic behavior of voters is prompted by the possibility of manipulating—see, however, our Exam-

<sup>1</sup>[Forsythe *et al.*, 1993] found in laboratory that strategic voting is, indeed, more likely to occur if pre-election coordination devices such as polls and shared voting histories are available; in the model of [Myerson and Weber, 1993], which assume such coordination devices in the form of a common prior, the set of voting equilibria is always nonempty.

ple 7, where we show that countermanipulations can make the game much more complex. We model the interaction among the manipulators as a normal-form game, where the players — called *GS-manipulators*— are a subset of the manipulators in a given profile, each player’s set of actions consists of her truthful vote and a subset of her manipulative votes, and players’ preferences over action profiles are determined by who wins in the resulting elections (see Section 3 for formal definitions). We call such games the *GS-games*.

The paper is organized as follows. Section 2 presents some basic concepts of voting theory. In Section 3 we formally define the Gibbard–Satterthwaite games. Section 4 focuses on 2-by-2 games. Section 5 deals with computing weakly dominant strategies. Section 6 presents our results on the existence of Nash equilibria. Section 7 concludes. Some proofs are omitted due to space constraints.

**Related Work** There is a substantial body of research dating back to [Farquharson, 1969] that analyzes non-truthful voting as a strategic game; see, e.g., [Moulin, 1979; Myerson and Weber, 1993; Dhillon and Lockwood, 2004; Feddersen *et al.*, 1990]. The algorithmic aspects of such models have recently received some attention as well [Desmedt and Elkind, 2010; Xia and Conitzer, 2010; Thompson *et al.*, 2013; Obraztsova *et al.*, 2013], with a strong focus on the problem of equilibrium selection. Indeed, a major difficulty in the study of general voting games is that they admit undesirable Nash equilibria such as, for instance, a situation in which all voters vote for the same candidate which none of them likes. Our analysis can also be seen from this perspective: by limiting the set of players, and restricting their action spaces, we eliminate counterintuitive behaviors.

Voting dynamics is a closely related topic, where players change their votes one by one in response to the current outcome [Meir *et al.*, 2010; Lev and Rosenschein, 2012; Reijngoud and Endriss, 2012; Reyhaneh and Wilson, 2012; Obraztsova *et al.*, 2015; Rabinovich *et al.*, 2015]. With the exception of [Meir *et al.*, 2010], in all of these papers the set of players consists of all voters, i.e., a player is allowed to vote non-truthfully even if he would be unable to manipulate the election on his own. Note that a convergence result in voting dynamics implies the existence of a Nash equilibrium in a GS-game, and indeed our Theorem 13 can be obtained as a consequence of a result by [Meir *et al.*, 2010].

Our work can also be viewed as an extension of the model of safe strategic voting proposed by Slinko and White [2008; 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 weakly dominant non-truthful votes in this setting (such votes are called *safe strategic votes*). In contrast, we allow manipulators to have diverse preferences and to use strategy sets that contain more than one non-truthful vote.

## 2 Preliminaries

We consider  $n$ -voter elections over a candidate set  $C = \{c_1, \dots, c_m\}$ . An election is defined by a *preference profile*  $V = (v_1, \dots, v_n)$ , where each  $v_i$ ,  $i = 1, \dots, n$ , is a total

order over  $C$ ; we refer to  $v_i$  as the *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  $c_2$ ; if this is the case, we say that voter  $i$  *prefers*  $c_1$  to  $c_2$ . For brevity we will sometimes write  $ab \dots z$  to represent a full ordering of the candidates  $v_i$  with  $a \succ_i b \succ_i \dots \succ_i z$ . We denote by  $\text{top}_k(v_i)$  the set of top  $k$  candidates in  $v_i$ , and abbreviate  $\text{top}_1(v_i)$  to  $\text{top}(v_i)$ .

Given a preference profile  $V = (v_1, \dots, v_n)$ , we denote by  $(V_{-i}, v'_i)$  the preference profile obtained from  $V$  by replacing  $v_i$  with  $v'_i$ . Given two lists of candidates  $X = (x_1, \dots, x_\ell)$ ,  $Y = (y_1, \dots, y_\ell)$  we denote by  $v[X; Y]$  the vote obtained by swapping  $x_j$  with  $y_j$  for  $j = 1, \dots, \ell$  in vote  $v$ . If  $X$  and  $Y$  are singletons, i.e.,  $X = (x)$ ,  $Y = (y)$ , we omit the parentheses, and simply write  $v[x; y]$ .

A *voting rule* is a mapping  $\mathcal{R}$  that, given a profile  $V$ , outputs a candidate  $\mathcal{R}(V) \in C$ . We consider the following voting rules in this paper.

- $k$ -approval,  $1 \leq k \leq m - 1$ : under this rule, each candidate receives one point from each voter that ranks her in top  $k$  positions; 1-approval is also known as Plurality.
- Borda: under this rule, each candidate gets  $m - j$  points from each voter that ranks her in position  $j$ .

Under both rules, the *score* of a candidate is the total number of points she receives, and the winner is the candidate with the highest score. Ties are broken according to a fixed order  $>$  over  $C$ . We denote the  $k$ -approval score of  $c$  in a profile  $V$  by  $sc_k(c, V)$ , and  $c$ ’s Borda score in  $V$  by  $sc_B(c, V)$ .

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 profile  $V$ . It is easy to see that  $v$  and  $v'$  are equivalent with respect to  $k$ -approval if and only if  $\text{top}_k(v) = \text{top}_k(v')$ , and  $v$  and  $v'$  are equivalent with respect to Borda if and only if  $v = v'$ .

## 3 The Model

Our goal is to investigate strategic situations that arise when one or more voters could improve the outcome of the election from their perspective by unilaterally changing their votes. We will now define such situations formally.

**Definition 1.** We say that a voter  $i$  is a Gibbard–Satterthwaite manipulator, or a GS-manipulator, in a profile  $V = (v_1, \dots, v_n)$  with respect to a voting rule  $\mathcal{R}$  if there exists a vote  $v'_i \neq v_i$  such that  $i$  prefers  $\mathcal{R}(V_{-i}, v'_i)$  to  $\mathcal{R}(V)$ . We denote the set of all GS-manipulators in a profile  $V$  with respect to a voting rule  $\mathcal{R}$  by  $N(V, \mathcal{R})$ . A vote  $v'_i$  is called a GS-manipulation of voter  $i$  if  $i$  prefers  $\mathcal{R}(V_{-i}, v'_i)$  to  $\mathcal{R}(V)$ , and, moreover, for every  $v''_i$  it holds that either  $\mathcal{R}(V_{-i}, v'_i) = \mathcal{R}(V_{-i}, v''_i)$  or  $i$  prefers  $\mathcal{R}(V_{-i}, v'_i)$  to  $\mathcal{R}(V_{-i}, v''_i)$ .

We say that  $i$  manipulates in favor of  $p$  by submitting a vote  $v'_i \neq v_i$  if  $p$  is the winner in  $\mathcal{R}(V_{-i}, v'_i)$ . Note that a voter can manipulate in favor of several different candidates; however, in the definition of a GS-manipulation we require the voter to focus on her most preferred candidate among the ones she can make the election winner.

Recall that a *normal-form game* is defined by a set of players  $N$ , and, for each  $i \in N$ , a set of actions  $A_i$  and a preference relation  $\succeq_i$  defined on the space of *action profiles*,

i.e., tuples of the form  $(a_1, \dots, a_n)$ , where  $a_i \in A_i$  for all  $i \in N$  (while one could define normal-form games in terms of utility functions or in terms of preference relations, the latter approach is more suitable for our setting, as we only have ordinal information about the voters' preferences).

For each preference profile  $V$  and voting rule  $\mathcal{R}$ , we consider a family of normal-form games defined as follows. Recall that  $N(V, \mathcal{R})$  is the set of GS-manipulators for  $V$  and  $\mathcal{R}$ .

**Definition 2.** Given a preference profile  $V$  and a voting rule  $\mathcal{R}$ , a GS-game for  $V$  and  $\mathcal{R}$  is a normal-form game  $\langle N, \{A_i\}_{i \in N}, \{\succeq_i\}_{i \in N} \rangle$  where:

- $N = N(V, \mathcal{R})$ ;
- for each player  $i$ , his set of actions  $A_i$  consists of his truthful vote and a subset of his GS-manipulations;
- the preference relation of player  $i$  is determined by the outcome of  $\mathcal{R}$  on the preference profile that corresponds to a given action profile. Given an action profile  $V^* = (v_i^*)_{i \in N}$ , let  $(V_{-N}, V^*) = (v_1', \dots, v_n')$  be the preference profile such that  $v_i' = v_i$  for  $i \notin N$  and  $v_i' = v_i^*$  for  $i \in N$ . Given two action profiles  $V^*$  and  $V^{**}$ , we write  $V^* \succeq_i V^{**}$  if and only if player  $i$  prefers  $\mathcal{R}(V_{-N}, V^*)$  to  $\mathcal{R}(V_{-N}, V^{**})$  or  $\mathcal{R}(V_{-N}, V^*) = \mathcal{R}(V_{-N}, V^{**})$ .

Observe that different choices of the actions sets  $A_i$  correspond to different games in the family. Denote the set of all GS-games for  $V$  and  $\mathcal{R}$  by  $\mathcal{GS}(V, \mathcal{R})$ . All games in  $\mathcal{GS}(V, \mathcal{R})$  have the same set of players, namely,  $N(V, \mathcal{R})$ , so an individual game in  $\mathcal{GS}(V, \mathcal{R})$  is fully determined by the players' sets of actions, i.e.,  $(A_i)_{i \in N(V, \mathcal{R})}$ . 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 an action profile in a GS-game as a *GS-profile*; we will sometimes identify the GS-profile  $V^* = (v_i^*)_{i \in N}$  with the preference profile  $(V_{-N}, V^*)$ . We denote the set of all GS-profiles in a game  $G$  by  $\mathcal{GSP}(G)$ .

**Discussion.** We emphasize a number of important features of this definition. First, we restrict the set of players to Gibbard–Satterthwaite manipulators. That is, we assume that, if a player cannot benefit from voting non-truthfully in the original profile, he will not be interested in exploring the possible benefits of strategic behavior when other voters manipulate as well. This attitude appears to be quite plausible: a voter may only be attracted to investigating the strategic aspects of the situation if manipulative behavior is profitable.

Second, we allow the players to limit themselves to subsets of their GS-manipulations, rather than consider a single game where each player's set of actions consists of his truthful vote and *all* of his GS-manipulations. There are several reasons for that. First, the space of all GS-manipulations for a given voter can be very large, and a player may be unable to identify all such votes. Second, the player may use a specific algorithm, such as that of Bartholdi *et al.* [1989], to find a GS-manipulation; in this case, her set of actions would consist of her truthful vote and the output of this algorithm. Third, the player may choose to ignore manipulations that are (weakly) dominated by others. Finally, a player may prefer not to change his vote beyond what is necessary to make his

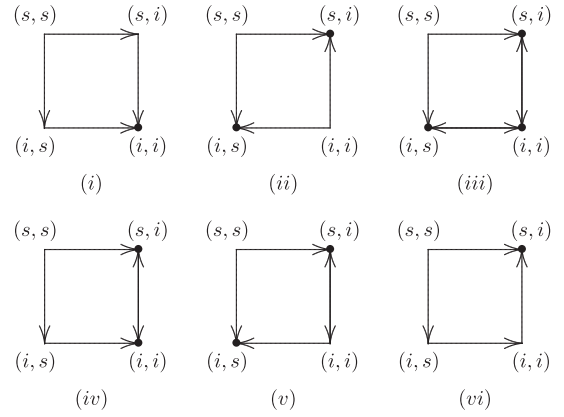


Figure 1: Diagrams for 2-by-2 GS-games

target candidate the election winner, because he does not enjoy lying [Obraztsova and Elkind, 2012], or fears unintended consequences of his actions in the complex environment of the game.

## 4 2-by-2 GS Games

In this section, we investigate which 2-by-2 games (i.e., games with two players and two actions per player) can be represented as GS-games. To address this question, we need a suitable classification of 2-by-2 games. Note first that every such game corresponds to 4 action profiles, and is fully described by giving both players' preferences over these profiles. By considering all possible pairs of preference relations over domains of size 4, Fraser and Kilgour [1986] show that there are 724 distinct 2-by-2 games.

However, this classification is too fine-grained for our purposes. Instead, we represent every 2-by-2 game by a diagram with 4 vertices and 4 directed edges, where an edge is directed from a less preferred profile to a more preferred profile (a bidirectional edge indicates indifference). For each player, let  $s$  denote his truthful vote and let  $i$  denote his manipulative vote; thus, the vertices of our diagram are  $(s, s)$ ,  $(i, s)$ ,  $(s, i)$ , and  $(i, i)$ . For two edges of this diagram their direction is determined by the fact that  $i$  is a GS-manipulation: namely, both of the edges adjacent to  $(s, s)$  are directed away from  $(s, s)$ . Thus, by renaming the players if necessary, we can represent any 2-by-2 GS-game by one of the 6 diagrams in Figure 1. Observe that an action profile in a 2-by-2 game is a Nash equilibrium if and only if the corresponding vertex in the diagram of the game has two incoming edges. The following proposition is immediate from Figure 1.

**Proposition 3.** Every 2-by-2 GS-game has at least one Nash equilibrium in pure strategies.

**Example 4.** Consider the GS-game for the preference profile  $(abc, bac, cab, cba)$  under the Plurality voting rule, with ties broken according to  $a > b > c$ . In this game players 1 and 2 are the GS-manipulators; we can assume that their GS-manipulations are  $v_1^* = v_1[a; b]$  and  $v_2^* = v_2[b; a]$ , respectively. Note that if both GS-manipulators vote insincerely,  $c$

remains the election winner. Thus, this game corresponds to diagram (ii) in Figure 1.

We say that diagram  $D$  is *realizable* by voting rule  $\mathcal{R}$  if there exist a profile  $V$  and a 2-by-2 game  $G \in \mathcal{GS}(V, \mathcal{R})$  such that  $D$  is the diagram for  $G$ . Our next goal is to study which diagrams are realizable by common voting rules.

**Theorem 5.** *The only diagrams realizable by Plurality are (ii), (iii), (iv) and (v).*

*Proof.* Consider a profile  $V$ , and assume that voters 1 and 2 are the GS-manipulators in  $V$ . Suppose that the Plurality winner in  $V$  is  $w$ , voter 1 manipulates in favor of  $a$ , and voter 2 manipulates in favor of  $b$ . Since voters 1 and 2 are GS-manipulators, we have  $w \neq \text{top}(v_1)$  and  $w \neq \text{top}(v_2)$ . Thus, we can assume that their GS-manipulations are given by  $v_1^* = v_1[\text{top}(v_1); a]$ ,  $v_2^* = v_2[\text{top}(v_2); b]$ . Let  $V^1 = (V_{-1}, v_1^*)$ ,  $V^2 = (V_{-2}, v_2^*)$ ,  $V^{12} = (v_1^*, v_2^*)$ . The winner in  $V^{12}$  can be  $w$ ,  $a$ , or  $b$ , and the case where it is  $w$  corresponds to diagram (ii). Further, if  $a = b$ , then  $a$  is the winner at  $V^1$ ,  $V^2$ , and  $V^{12}$ , corresponding to diagram (iii). Now, suppose that  $a \neq b$ . As the winner at  $V^{12}$  is either  $a$  or  $b$ , one of the arrows adjacent to  $(i, i)$  must be bidirectional, ruling out diagrams (i), (ii) and (vi) for this case. We have argued that diagrams (i) and (vi) are not realizable by Plurality. Example 4 realizes diagram (ii). We omit the examples for the remaining two cases due to lack of space.  $\square$

Note that Theorem 5 implies that the Prisoner's dilemma is not realizable by Plurality. In contrast, for Borda all six diagrams are realizable.

**Theorem 6.** *Diagrams (i)–(vi) are all realizable by Borda.*

*Proof.* We provide an example for diagram (i), the remaining cases are omitted. Assume that ties are broken alphabetically. Let  $V = (dbcae, dbcae, acebd, acebd)$ . The Borda winner in  $V$  is  $a$ . The first two voters are the GS-manipulators, and for each of them the vote  $v^* = bdcea$  is a GS-manipulation in favor of  $c$ . Further, in  $(v^*, v^*, v_3, v_4)$  the Borda winner is  $b$ .  $\square$

To conclude the section, let us introduce the possibility of counter-manipulation, i.e., let us expand the notion of GS-games by letting sincere players respond to hypothetic manipulations by strategic voters.

**Example 7.** We show that a 2-by-2 game with one manipulator and one counter-manipulator may not have a Nash equilibrium. Suppose that the voting rule is 2-approval and the profile is  $V = (adcb, bdca)$  with ties broken according to  $a > b > c > d$ . For voter 1 switching  $c$  and  $d$  is a manipulation in favor of  $a$ . To counter-manipulate, voter 2 can also switch  $c$  and  $d$ , making  $c$  win. But then voter 1 would be better off switching  $c$  and  $d$  back, after which the same move will be beneficial for voter 2.

## 5 Weak Dominance

In this section, we consider the complexity of checking whether a given GS-manipulation is always at least as good as the truthful vote. This problem is captured by the game-theoretic notion of *weak dominance*.

**Definition 8.** *Given a GS-game  $G = (V, \mathcal{R}, (A_i)_{i \in N(V, \mathcal{R})})$  and two strategies  $v_i^*, v_i^{**} \in A_i$ , we say that  $v_i^*$  weakly dominates  $v_i^{**}$  if for every profile  $V' \in \mathcal{GSP}(G)$  either  $\mathcal{R}(V'_{-i}, v_i^*) = \mathcal{R}(V'_{-i}, v_i^{**})$  or voter  $i$  prefers  $\mathcal{R}(V'_{-i}, v_i^*)$  to  $\mathcal{R}(V'_{-i}, v_i^{**})$ . We say that  $v_i^*$  is weakly dominant if  $v_i^*$  weakly dominates  $v_i^{**}$  for every  $v_i^{**} \in A_i$ .*

In what follows, we will mostly be interested in determining whether a given GS-manipulation  $v_i^*$  weakly dominates the truthful vote  $v_i$ . A related problem is whether  $v_i^*$  is weakly dominant. However, in the context of GS-games the former problem appears to be more relevant than the latter. Indeed, the main issue faced by a GS-manipulator is whether to manipulate or not, and if a certain vote can always ensure an outcome that is no worse than that of truthful voting, then this is a strong incentive to use it, even if another non-truthful vote may sometimes be better. Our approach is consistent with the one taken in the study of safe manipulation [Slinko and White, 2008; 2014], where different manipulations are not compared to each other either.

**Preliminary Observations** In the rest of the paper, we limit our attention to  $k$ -approval rules. We will now state some definitions and observations that apply to this class.

Note first that under  $k$ -approval any GS-manipulation of voter  $i$  is equivalent to a vote of the form  $v_i[X; Y]$ , where  $X \subseteq \text{top}_k(v_i)$ ,  $Y \subseteq C \setminus \text{top}_k(v_i)$ . Consider a GS-manipulation  $v_i[X; Y]$  of voter  $i$  in  $V$  under  $k$ -approval; we say that  $v_i[X; Y]$  is *minimal* if for every other GS-manipulation  $v_i'$  of voter  $i$  it holds that  $v_i' = v_i[X'; Y']$ , where  $|X'| \geq |X|$ . That is, a GS-manipulation is minimal if it performs as few swaps as possible.

Consider a profile  $V$  such that the  $k$ -approval winner in  $V$  is  $w$ , and  $\text{sc}_k(w, V) = t$ . Recall that  $>$  is the tie-breaking order over candidates. Let

$$S^1(V, k) = \{c \in C \mid \text{sc}_k(c, V) = t, w > c\},$$

$$S^2(V, k) = \{c \in C \mid \text{sc}_k(c, V) = t - 1, w < c\},$$

and set  $S(V, k) = S^1(V, k) \cup S^2(V, k)$ . We say that a candidate  $c$  is  *$k$ -competitive* in  $V$  if  $c \in S(V, k)$ . Consider a GS-manipulator  $i$  who has a GS-manipulation in favor of some candidate  $p$ . As  $i$  prefers  $p$  to  $w$ , it cannot be the case that, by submitting a non-truthful vote, he both decreases the score of  $w$  and increases the score of  $p$ . Thus, his manipulation increases the score of  $p$  or decreases the score of  $w$ , but not both. It follows that under  $k$ -approval every GS-manipulation in  $V$  is in favor of some candidate  $p \in S(V, k)$ ; in particular, if  $V$  admits a GS-manipulator under  $k$ -approval, then  $S(V, k)$  is not empty.

There are two types of GS-manipulators: those who rank  $w$  in top  $k$  positions, and those who do not. The GS-manipulators of the first type can only manipulate in favor of candidates they already approve of, so in their manipulation they rank  $w$  in position  $k + 1$  or lower; we will refer to such voters as *demoters*. The GS-manipulators of the second type manipulate by promoting some candidate they rank above  $w$  into top  $k$  positions; we refer to such voters as *promoters*. We say that a candidate  $c$  is  *$k$ -plausible* for  $i$  in  $V$  if there exists a promoter  $j \neq i$  and a GS-manipulation  $v_j^*$  such that  $c$  is the  $k$ -approval winner in  $(V_{-j}, v_j^*)$ ; we denote the set of all

candidates that are  $k$ -plausible for  $i$  in  $V$  by  $S_i^+(V, k)$ , and set  $S^+(V, k) = \cup_{i \in N} S_i^+(V, k)$ .

**Plurality** Consider a GS-manipulator  $i$  in a profile  $V$  under Plurality, and let  $v_i^*$  be his GS-manipulation. Note that  $i$  is necessarily a promoter. Let  $p_i$  be  $i$ 's favorite candidate in  $S(V, k) \setminus \{\text{top}(v_i)\}$ . It is clear that  $v_i^*$  is equivalent to  $v_i[\text{top}(v_i); p_i]$  and  $p_i$  is the winner in  $(V_{-i}, v_i^*)$ , i.e., all GS-manipulations of voter  $i$  in  $V$  are equivalent to each other. Hence, there is essentially a unique GS-game that corresponds to  $V$ , namely, the one where  $A_i = \{v_i, v_i[\text{top}(v_i); p_i]\}$  for each player  $i$ ; we will denote this game by  $G_1^*(V)$ . We will now characterize GS-manipulations that are weakly dominant in this game (as each player  $i$  has two strategies, her GS-manipulation  $v_i^*$  is weakly dominant if and only if it weakly dominates  $v_i$ ).

**Theorem 9.** *Let  $v_i^*$  be the GS-manipulation of voter  $i$  in  $G_1^*(V)$ , and let  $c = \text{top}(v_i)$ . Then  $v_i^*$  is weakly dominant in  $G_1^*(V)$  if and only if  $c \notin S_i^+(V, 1)$ .*

*Proof.* Suppose that  $c \in S_i^+$ , i.e., there exists a promoter  $j \neq i$  with GS-manipulation  $v_j^*$  such that  $c$  wins in  $V^1 = (V_{-j}, v_j^*)$ . Note that  $c = \text{top}_1(v_j^*)$ . Consider the profile  $V^2 = (V_{-i}^1, v_i^*)$ . In this profile both  $c$  and the initial winner  $w$  have the same score as in  $V$ , so  $c$  loses to  $w$ . As  $c = \text{top}(v_i)$ , this shows that  $v_i^*$  is not weakly dominant.

Conversely, suppose that  $c \notin S_i^+$ . Then for any profile  $V' = (v'_1, \dots, v'_n) \in \mathcal{GSP}(G_1^*(V))$  such that  $v'_i = v_i$  we have  $\text{sc}_1(c, V') \leq \text{sc}_1(c, V)$ . On the other hand, since no GS-manipulator ranks  $w$  first, we have  $\text{sc}_1(w, V') = \text{sc}_1(w, V)$ . Thus,  $c$  is not the Plurality winner in  $V'$ . Let  $w'$  be the Plurality winner in  $V'$ ; note that  $w' \in \{w\} \cup S$ . Let  $V'' = (V'_{-i}, v_i^*)$ . Suppose that  $v_i^* = v[c; p]$ . It is easy to see that this implies that the Plurality winner in  $V''$  is either  $p$  or  $w'$ . Since  $i$  weakly prefers  $p$  to every other candidate in  $\{w\} \cup S$ , it follows that  $v_i^*$  is weakly dominant.  $\square$

Theorem 9 illustrates that every GS-game for Plurality in the absence of dominant strategies is essentially a coordination game: the GS-manipulators have to coordinate to ensure that their efforts do not cancel out. We obtain the following:

**Corollary 10.** *There is a polynomial-time algorithm for checking whether a given GS-manipulation  $v_i^*$  of voter  $i$  weakly dominates his truthful vote  $v_i$  in  $G_1^*(V)$ .*

**2-Approval: An Algorithm** We will now focus on the 2-approval rule. Let  $w$  be the 2-approval winner in  $V$ , and recall that  $p^*$  is the candidate in  $S$  that beats every other candidate in  $S$  in profile  $V$  under 2-approval. It can be shown that all demoters rank the same candidate  $p^*$  first, and have the same set of GS-manipulations. Now, suppose that  $w \notin \text{top}_2(v_i)$ , i.e.,  $i$  is a promoter. Then  $i$  can raise the scores of some candidates in  $C \setminus \text{top}_2(v_i)$  by moving these candidates into top two positions; note that  $i$  can do that for two candidates simultaneously. Clearly, a promoter's vote is a GS-manipulation if and only if one of the promoted candidates is  $i$ 's most preferred candidate in  $S \setminus \text{top}_2(v_i)$  (let us denote this candidate by  $p$ ) and the other promoted candidate  $p'$  is such that promoting  $p'$  does not prevent  $p$  from becoming a winner. We emphasize

that  $p$  is not necessarily  $i$ 's most preferred 2-competitive candidate: it may be the case that  $i$  ranks another 2-competitive candidate in top two positions.

Given a profile  $V$ , let  $G_2^*(V)$  be the GS-game where each player's set of actions consists of his truthful vote and all his minimal manipulations. The argument above shows that for demoters the minimality assumption imposes no additional constraints, whereas for promoters it excludes manipulations where two candidates are swapped into top two positions simultaneously.

We can now state the main result of this section, whose proof is omitted in the interest of space.

**Theorem 11.** *There is a polynomial-time algorithm for checking whether a given GS-manipulation  $v_i^*$  of voter  $i$  weakly dominates his truthful vote  $v_i$  in  $G_2^*(V)$ .*

**4-Approval: A Hardness Proof** We were unable to determine the complexity of checking whether a given GS-manipulation weakly dominates truth-telling under 3-approval. We conjecture that this problem is computationally hard. Indeed, this is the case for 4-approval: there is a way to select one manipulative action per player so that for the resulting game this problem is coNP-complete.

We reduce from the classic NP-complete problem EXACT COVER BY 3-SETS (X3C). An instance of this problem is given by a ground set  $\Gamma = \{g_1, \dots, g_{3\nu}\}$  and a collection  $\Sigma = \{\sigma_1, \dots, \sigma_\mu\}$  of 3-element subsets of  $\Gamma$ . 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.

**Theorem 12.** *The problem of deciding whether a given strategy  $v_i^*$  weakly dominates truth-telling in a GS-game  $(V, 4\text{-App}, (A_i)_{i \in N(V, 4\text{-App})})$  is coNP-complete.*

*Proof sketch.* It is immediate that this problem is in coNP. To prove coNP-hardness, consider an instance  $I = (\Gamma, \Sigma)$  of X3C with  $|\Gamma| = 3\nu$ ,  $|\Sigma| = \mu$ . We can assume without loss of generality that  $\sigma_1 = \{g_1, g_2, g_3\}$  and no other set in  $\Sigma$  contains  $g_1$ ; if this is not the case, we can modify  $I$  by adding three new elements and a single set containing them. We will now construct an instance of our problem. In what follows, when writing  $X \succ Y$  in describing 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  $Y$  can be arbitrary.

There is a set of candidates  $C' = \{c_1, \dots, c_{3\nu}\}$  that correspond to elements of  $\Gamma$ , three special candidates  $w, p, c$ , and a set of dummy candidates  $D = \bigcup_{i=0}^{\mu} D_i$ , where  $|D_i| = 4$  for  $i = 0, \dots, \mu$ . We define the order  $\succ$  by setting  $w \succ c \succ p \succ c_1 \succ \dots \succ c_{3\nu} \succ D$ .

For each set  $\sigma_i \in \Sigma$  we construct a vote  $v_i$  of the form

$$D_i \succ \sigma_i \succ c \succ \dots,$$

where candidates in  $\sigma_i$  are ordered according to  $\succ$ , and set

$$v_0 = D_0 \succ p \succ c_1 \succ c \succ C' \setminus \{c_1\} \succ \dots$$

We also construct additional voters so that in the resulting profile  $V$  we have  $\text{sc}_4(w, V) = \text{sc}_4(p, V) = \nu + 1$ ,  $\text{sc}_4(c_i, V) = \nu + 1$  for all  $i = 1, \dots, 3\nu$ ,  $\text{sc}_4(c, V) = 1$ , the 4-approval score of each candidate in  $D$  is at most 1, and the only GS-manipulators in  $V$  are voters  $0, 1, \dots, \mu$ .

Note that  $w$  is the 4-approval winner in  $V$ , and  $S = C' \cup \{p\}$ . We now define a GS-game for this profile by constructing the players' sets of actions as follows. Let  $D_0 = \{d_1, d_2, d_3, d_4\}$ . Then  $v_0^* = v_0[\{d_1, d_2\}; \{p, c\}]$  is a GS-manipulation for voter 0, which makes  $p$  the winner with  $\nu + 2$  points. Similarly, for each  $i = 1, \dots, \mu$  the vote  $v_i^* = v_i[D_i; \sigma_i \cup \{c\}]$  is a GS-manipulation which makes  $i$ 's top candidate in  $\sigma_i$  the winner with  $\nu + 2$  points (note that  $i$  orders  $\sigma_i$  in the same way  $>$  does, so tie-breaking favors  $i$ 's most preferred candidate in  $\sigma_i$ ). We set  $A_i = \{v_i, v_i^*\}$  for  $i = 0, \dots, \mu$ . This completes the description of our game. Clearly, we can construct the profile  $V$  and the players' sets of strategies in polynomial time given  $I$ . It can be shown that  $v_0^*$  weakly dominates  $v_0$  if and only if we started with a "no"-instance of X3C; we omit the proof.  $\square$

## 6 Nash Equilibrium

In this section, we study existence and uniqueness of Nash equilibria in GS-games for  $k$ -approval with  $k = 1, 2, 3, 4$ . In what follows, when considering a GS-game, we assume that its set of GS-manipulators is not empty, since otherwise a Nash equilibrium exists trivially.

We will first show that for Plurality, a Nash equilibrium always exists.

**Theorem 13.** *For any profile  $V$  the game  $G_1^*(V)$  has a Nash equilibrium in pure strategies.*

*Proof.* Consider the set  $S^+(V, 1)$ , and let  $p$  be the candidate with the highest score in this set (using the tie-breaking order eventually). Let  $i$  be some voter whose GS-manipulation is in favor of  $p$ , and let  $v_i^* = v_i[\text{top}(v_i); p]$ . Then  $(V_{-i}, v_i^*)$  is a Nash equilibrium with winner  $p$ .  $\square$

However, many preference profiles admit multiple Nash equilibria under Plurality. For instance, if there are three voters  $i, j, \ell$  such that  $v_r[\text{top}(v_r); c] \in A_r$  for each  $r \in \{i, j, \ell\}$  and some candidate  $c$ , then there is a NE where these three voters manipulate and everyone else votes truthfully; if there are several candidates for whom such a triple of voters exists, any of them could win in a Nash equilibrium.

For 2-approval, we can prove existence of Nash equilibria as long as every manipulator has at least one minimal GS-manipulation in her action set.

**Theorem 14.** *Any game  $G = (V, 2\text{-App}, (A_i)_{i \in N(V, 2\text{-App})})$ , where for each  $i \in N(V, 2\text{-App})$  the set  $A_i$  contains some minimal manipulation of voter  $i$ , has a Nash equilibrium.*

*Proof.* We can assume that  $N(V, 2\text{-App}) \neq \emptyset$ . Let  $w$  be the current winner of the election. Assume first that the set of promoters in  $G$  is not empty. Let  $p^+$  be the 2-plausible candidate that beats all other candidates in  $S^+(V, 2)$ . Let  $i$  be a promoter who manipulates in favor of  $p^+$ , and let  $v_i^*$  be some minimal GS-manipulation of  $i$ . Then  $V^* = (V_{-i}, v_i^*)$  is a Nash equilibrium with winner  $p^+$ . Indeed, by minimality of  $v_i^*$  we know that only the score of  $p^+$  has increased in  $V^*$ . Hence, to prevent  $p^+$  from winning in  $V^*$ , a manipulator  $j$  would have to lower  $p^+$ 's score, and he could only do that if  $p^+ \in \text{top}_2(v_j)$ . If  $j$  is a promoter, then he prefers  $p^+$  to any candidate that he could make a winner in this way. On

the other hand, if  $j$  is a demoter then  $p^+$  is his top candidate, hence he does not have incentives to change the outcome.

If all GS-manipulators are demoters, then a profile where any demoter, say  $j$ , submits a GS-manipulation and everyone else votes truthfully is a Nash equilibrium with winner  $\text{top}(v_j)$ . Indeed, a demoter can only manipulate in favor of her top candidate by decreasing the score of  $w$ . Hence all demoters manipulate in favor of the same candidate.  $\square$

A similar result holds for 3-approval, though under stronger assumptions. Given a voter  $i$  and a set of candidates  $Z$ , we denote by  $\text{bot}_i(Z)$  the least preferred candidate of  $i$  in  $Z$ . A manipulative vote  $v_i[X; Y]$  is called *greedy* if it is minimal and for every other minimal strategy  $v_i[X; Y']$  it holds that  $i$  prefers  $\text{bot}_i(Y)$  to  $\text{bot}_i(Y')$ . Clearly, there is a unique greedy manipulation for each manipulator  $i$ .

**Theorem 15.** *The game  $G = (V, 3\text{-App}, (A_i)_{i \in N(V, 3\text{-App})})$  where for each  $i \in N(V, 3\text{-App})$  the set  $A_i$  consists of  $i$ 's truthful vote and her greedy GS-manipulation has a Nash equilibrium.*

The restriction to greedy strategies in Theorem 15 is essential. Moreover, for 4-approval the existence of Nash equilibria is not guaranteed; we conjecture that this result extends to  $k$ -approval with  $k > 4$ .

**Theorem 16.** *Under 3-approval, there exists a game  $G = (V, 3\text{-App}, (A_i)_{i \in N(V, 3\text{-App})})$ , where for each player  $i \in N(V, 3\text{-App})$  the set  $A_i$  consists of  $i$ 's truthful vote and some minimal manipulations, s.t.  $G$  has no Nash equilibrium.*

Theorem 16 holds even if we restrict the set of strategies of each player to her greedy manipulations:

**Theorem 17.** *Under 4-approval, there exists a game  $G = (V, 4\text{-App}, (A_i)_{i \in N(V, 4\text{-App})})$ , where for each player  $i \in N(V, 4\text{-App})$  the set  $A_i$  consists of  $i$ 's truthful vote and her greedy GS-manipulation, s.t.  $G$  has no Nash equilibrium.*

## 7 Conclusions

In this paper we have studied the strategic interaction of manipulators in an electoral situation, providing a formal model in the form of a class of strategic-form games, which we called GS-games, and showing that these games may require extensive communication and a non-trivial computational effort to be played, which may serve as a barrier against manipulation. In particular, we have shown that for Plurality these games exhibit a fairly simple structure, which still allows for situations as complex as a coordination game. For Borda and  $k$ -approval, with  $k > 1$ , GS-games are quite complicated, and it may therefore be difficult for the players to coordinate their actions: indeed, for large enough  $k$  a Nash equilibrium may fail to exist, and even if it exists, it is not necessarily unique.

Many questions concerning GS-games remain open. The most immediate of them is to fully understand the role of minimality assumptions in our proofs. Further afield, it would be interesting to extend our study to other voting rules, and to identify reasonable restrictions on the manipulators' strategy spaces that lead to existence and uniqueness of Nash equilibria, and make it easy to compute manipulations that weakly dominate truth-telling.

## Acknowledgements

Part of this work was done when Arkadii Slinko was visiting the Institute for Mathematical Sciences (Singapore) and the University of Padova (Italy). Francesca Rossi and Umberto Grandi were partially supported by the strategic project “KIDNEY - Incorporating patients’ preferences in kidney transplant decision protocols” of the University of Padova.

## References

- [Bartholdi *et al.*, 1989] J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241, 1989.
- [Benjamin *et al.*, 2013] D. J. Benjamin, S. A. Brown, and J. M. Shapiro. Who is behavioral? cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11:1231–1255, 2013.
- [Brandt *et al.*, 2015] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2015.
- [Cain, 1978] B.E. Cain. Strategic voting in Britain. *American Journal of Political Science*, 22:639–655, 1978.
- [Choi *et al.*, 2014] S. Choi, S. Kariv, W. Muller, and D. Silverman. Who is (more) rational? *American Economic Review*, 104:1518–1550, 2014.
- [Cox, 1997] G.W. Cox. *Making Votes Count*. Cambridge University Press, 1997.
- [Desmedt and Elkind, 2010] Y. Desmedt and E. Elkind. Equilibria of plurality voting with abstentions. In *ACM EC’10*, pages 347–356, 2010.
- [Dhillon and Lockwood, 2004] A. Dhillon and B. Lockwood. When are plurality rule voting games dominance-solvable? *Games and Economic Behavior*, 46:55–75, 2004.
- [Farquharson, 1969] R. Farquharson. *Theory of Voting*. Yale University Press, 1969.
- [Feddersen *et al.*, 1990] 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.
- [Forsythe *et al.*, 1993] R. Forsythe, R. B. Myerson, T. A. Rietz, and R. J. Weber. An experiment on coordination in multi-candidate elections: The importance of polls and election histories. *Social Choice and Welfare*, 10:223–247, 1993.
- [Fraser and Kilgour, 1986] N.M. Fraser and D.M. Kilgour. Non-strict ordinal 2x2 games: A comprehensive computer-assisted analysis of the 726 possibilities. *Theory and Decision*, 20(2):99–121, 1986.
- [Gibbard, 1973] A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.
- [Kawai and Watanabe, 2013] K. Kawai and Y. Watanabe. Inferring strategic voting. *American Economic Review*, 103:624–662, 2013.
- [Lev and Rosenschein, 2012] O. Lev and J. S. Rosenschein. Convergence of iterative voting. In *AAMAS’12*, pages 611–618, 2012.
- [Meir *et al.*, 2010] R. Meir, M. Polukarov, J. S. Rosenschein, and N. R. Jennings. Convergence to equilibria of plurality voting. In *AAAI’10*, pages 823–828, 2010.
- [Moulin, 1979] H. Moulin. Dominance solvable voting schemes. *Econometrica*, 47:1337–1351, 1979.
- [Myatt, 2007] David P. Myatt. On the theory of strategic voting. *The Review of Economic Studies*, 74:255–281, 2007.
- [Myerson and Weber, 1993] R. Myerson and R. Weber. A theory of voting equilibria. *American Political Science Review*, 87(1):102–114, 1993.
- [Obraztsova and Elkind, 2012] S. Obraztsova and E. Elkind. Optimal manipulation of voting rules. In *AAMAS’12*, pages 619–626, 2012.
- [Obraztsova *et al.*, 2013] S. Obraztsova, E. Markakis, and D. R. M. Thompson. Plurality voting with truth-biased agents. In *SAGT’13*, 2013.
- [Obraztsova *et al.*, 2015] S. Obraztsova, E. Markakis, M. Polukarov, Z. Rabinovich, and N. Jennings. On the convergence of iterative voting: How restrictive should restricted dynamics be? In *AAAI’15*, 2015.
- [Rabinovich *et al.*, 2015] Z. Rabinovich, S. Obraztsova, O. Lev, E. Markakis, and J. S. Rosenschein. Analysis of equilibria in iterative voting schemes. In *AAAI’15*, 2015.
- [Reijngoud and Endriss, 2012] A. Reijngoud and U. Endriss. Voter response to iterated poll information. In *AAMAS’12*, pages 635–644, 2012.
- [Reyhaneh and Wilson, 2012] R. Reyhaneh and M. Wilson. Best reply dynamics for scoring rules. In *ECAI’12*, pages 672–677, 2012.
- [Satterthwaite, 1975] 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.
- [Slinko and White, 2008] A. Slinko and S. White. Non-dictatorial social choice rules are safely manipulable? In *COMSOC’08*, pages 403–414, 2008.
- [Slinko and White, 2014] A. Slinko and S. White. Is it ever safe to vote strategically? *Social Choice and Welfare*, 43(2):403–427, 2014.
- [Thompson *et al.*, 2013] D. R. M. Thompson, O. Lev, K. Leyton-Brown, and J. S. Rosenschein. Empirical analysis of plurality election equilibria. In *AAMAS’13*, pages 391–398, 2013.
- [Xia and Conitzer, 2010] L. Xia and V. Conitzer. Stackelberg voting games: Computational aspects and paradoxes. In *AAAI’10*, pages 805–810, 2010.