What You Know About People's Preferences Matters: *Investigating simpler notions of partial information in the context of strategic manipulation in voting*

MSc Thesis *(Afstudeerscriptie)*

written by

Lucy van Oostveen
(born October 21st, 1991 in Nijmegen, the Netherlands)

under the supervision of **Dr. Ronald de Haan** and **Dr. Jakub Szymanik**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
April 26th, 2018	Prof. Dr. Ronald de Wolf (chair)
	Dr. Ronald de Haan (supervisor)
	Dr. Jakub Szymanik (supervisor)
	Dr. Alexandru Baltag
	Dr. Fernando Velazquez Quesada
	Zoi Terzopoulou, MSc

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

The Gibbard-Satterthwaite Theorem tells us that most voting rules are susceptible to strategic manipulation, which means that a voter can benefit from voting something other than their true preference. In this theorem, however, it is assumed that voters have full information about the preferences of other voters, which is not always realistic in practice. While models of partial information of the preferences of other voters exist in the literature, the notion of partial information that is used is often very general and includes complicated instances that play an important role in the complexity analysis of voting. In this thesis, simpler and more plausible models of partial information in the context of voting are developed and investigated. We formalized four structures of partial information and looked at both the manipulability and the difficulty of manipulation of these structures in combination with the k-approval, Borda, and Copeland voting rules. We found that restricting partial information to simpler instances does not prevent manipulation and show there is a difference between knowing a little about the preferences of every voter and knowing everything about the preferences of some. Moreover, we show that for k-approval manipulation is computationally easy for these structures and argue why we believe this might also be the case for other voting rules. We believe that this thesis shows that the way in which partial information is modeled influences important properties of manipulability and we therefore believe that it is an important factor that should be taken into account.

Acknowledgements

First of all, I would like to thank my supervisors, Ronald and Jakub. I enjoyed our fruitful meetings and always left more confident, in a better mood, and with a clearer view of where I was going. You make a great team. Ronald, thanks for your contagious optimism and always making time for me when I was stuck with something but wasn't really sure whit what. Jakub, thanks for making sure I focused on my structure and story from the start and reminding me I don't have to write down everything I know.

I would like to thank Ulle for introducing me to computational social choice and Benno for being a great academic mentor the last two years. Moreover, thanks to the master of logic and the ILLC office, Maria, Tanja, Jenny, Karine, for caring this much about students and for your support whenever we knocked on your doors with crazy ideas.

I would also thank my committee members for taking the time to read and commenting on this thesis.

Thanks to the lovely logicians. First of all, the complexity people, I would definitely not have survived the last two and a half years without you. Jana, Esteban, Anthi, Jakob, and Melina, thanks for taking me dancing when I needed it most. Laura, I am happy I found someone who is as crazy about organizing as me. Natalia, Jelle, and all the other MoLs, thank you for the two years filled with spar-pizza and solving cake-cutting problems, it would not have been the same without you.

Thanks to Jenske, Milinda, Laura, and my dad for proofreading parts of this thesis.

Thanks to my family and non-logic friends, for supporting me when I decided to do another master, writing another thesis, and for putting up with me when I tried to explain 'easy things' for yet two more years. Nikki and Milinda, thanks for all your advice about life and taking trips to sunny places. Finally, thanks to Daan for making everything better.

Contents

Chapter 1

Introduction

Suppose you are attending a conference and you and a group of colleagues decide to go out for dinner. One of the local organizers suggests three places to eat: a sushi restaurant, a place where they serve burgers, and some local restaurant. To decide where to go, you and your colleagues all order the restaurants from most favorite to least favorite. Every time a restaurant is someone's favorite it receives two points, when a restaurant is at the second place it gets one point, and you will have dinner at whatever place receives the most points. Since you are really into sushi, you want to put this restaurant on the top of your list. However, you know that three out of four of your colleagues really dislike fish and will definitely put sushi on the lowest place of their list. After doing some quick calculations, you realize that even if the fourth colleague is a big sushi fan, either the burger or the local place will receive more points which means you will never actually go to eat sushi. Since you are not a fan of the local cuisine, you decide to report that you like burgers over sushi over the local establishment, to lower the possibility you end up in the local place.

This example introduces the three main topics of this thesis: *voting*, *uncertainty about the preferences of other voters*, and *strategic manipulation*: voting something different than your true preference in order to improve the outcome. In this thesis, we will study all three phenomena from the perspective of computational social choice, which studies the mechanisms of collective decision making. In this context, voting is defined as the aggregation of individual preferences to elect a (set of) winner(s) from a set of alternatives. While voting is often associated with political elections, it includes every setting in which a group of people, or artificial agents, with different personal preferences need to collectively decide upon an alternative, such as a place to eat, in the example above (see e.g. Brandt et al., 2016b).

In the early 70s, Gibbard (1973) and Satterthwaite (1975) independently showed that most reasonable voting rules are not *strategy-proof*, which means there always exists a situation in which some voter can benefit from misrepresenting their preference. As illustrated by the example above, there can be situations in which misrepresenting your preferences changes the outcome of the voting rule, and as such the mechanisms of a voting rule. Together with the fact that people manipulate in real-life voting situations, makes it a phenomenon that is of interest of many researchers in computational social choice (see e.g. Conitzer and Walsh, 2016).

In the Gibbard-Satterthwaite Theorem it is assumed, however, that all voters have *full information* about the preferences of the others. This is not always a realistic assumption in practice. For this reason, researchers have investigated what happens when voters are uncertain, or have *partial knowledge*, about the preferences of the other voters (see e.g. Conitzer et al., 2011; Reijngoud and Endriss, 2012; Meir et al., 2014).

In this thesis, we will investigate how partial knowledge about the preferences of other voters influences the manipulability of several voting rules. We focus on the voting rules k-approval, Borda, and Copeland, three rules that are widely studied in the literature (see e.g. Brandt et al., 2016b). We will approach manipulability from several perspectives, including a mathematical perspective in which we compare different kinds of uncertainty with respect to different voting rules and a computational perspective in which we investigate how difficult it is to manipulate.

The main contribution of this thesis is that we investigate simpler models of uncertainty. Existing frameworks of uncertainty are often very general, allowing for very specific and complicated instances of voting situations that play an important role in the analysis of partial information. In this thesis, we investigate what happens when we restrict the partial models to very simple instances instead, such as knowing only the least preferred alternatives of all voters similar to idea of the example above. This allows us to, on the one hand, compare several different types of partial information. In particular, we show that knowing something about all voters differs from the situation in which the manipulator knows everything about the preferences of some of the other voters. On the other hand, it allows us to investigate what effect restricting the models has on the computational complexity of manipulation.

1.1 Outline of this Thesis

The structure of this thesis will be as follows:

In **Chapter 2**, we introduce the theoretical landscape of voting theory and set the stage for the rest of this thesis by introducing the voting rules we will study, and show how manipulation is defined.

In **Chapter 3**, we investigate the notion of *incomplete information*, by comparing different models of partial information, introduce the model we will use in the rest of this thesis, and prove several mathematical properties of this model.

In **Chapter 4**, we show that with all three voting rules we will consider, k-approval, Borda, and Copeland, with most of the models defined in Chapter 3, voters still have an incentive to manipulate. Moreover, we investigate if there is a difference between what type of uncertainty a manipulator has when they are trying to manipulate.

In **Chapter 5**, we investigate the computational complexity of manipulation. In particular, we show that for k-approval manipulation for a simplified class of instances can be done in polynomial time. Moreover, we explain why we believe that showing similar results for other voting rules, including Borda and Copeland, is far from straightforward.

In **Chapter 6**, we investigate other ways the notion *disadvantage* could be formalized.

Finally, in **Chapter 7**, we consider alternative ways our research question could be investigated and suggest potential future research directions.

Chapter 2

Voting Theory

In this chapter, we will introduce the theoretical landscape of voting theory. For reasons of space, we will only focus on theory that is directly applicable to this thesis.For a complete theoretical overview of voting theory, see the work of Brandt et al. (2016b). We start with introducing the basic framework of voting theory, then introduce the concept of strategic manipulation and argue why it is an important concept to look at. The main aim of this chapter is to provide the reader with the technical machinery to be able to understand the rest of this thesis. A list of symbols and notion can be found before the Appendix. We will also provide arguments why we chose to focus on k-approval, Borda, and Copeland in the rest of this thesis and argue why we believe that manipulation is an important and relevant concept to investigate.

2.1 The Basic Framework of Voting

Let $N = \{1, ..., n\}$ be a set of n voters and $\mathcal{X} = \{X_1, X_2, ..., X_m\}$ be a set of m *alternatives* (or *candidates*). To vote, each voter submits a *ballot* R_i. In this thesis, a ballot is a linear order over the set of alternatives. That is, if $\mathcal{L}(\mathcal{X})$ is the set of all linear orders over the set of alternatives (all relations over $\mathcal{X}$ that are total, transitive, and anti-symmetric), we have that $R_i \in \mathcal{L}(\mathcal{X})$ for all voters i. If we have two alternatives $X, Y \in \mathcal{X}$ and voter i voted X above Y, so $(X, Y) \in R_i$, we will denote this by writing $X \succ_i Y$, $X \succ_{R_i} Y$, or leave out the i or R_i if this is clear from the context. Moreover, a *profile* $\boldsymbol{R} = (R_1, ..., R_n) \in \mathcal{L}(\mathcal{X})^n$ is a vector of ballots, containing one ballot for each voter.

Now, a *voting rule* (or a *social choice function*) F is a function from a profile to a non-empty subset of alternatives, the winners of the election:

$$F : \mathcal{L}(\mathcal{X})^n \to \wp(\mathcal{X}) \setminus \{\emptyset\}$$

Here $\wp(\mathcal{X})$ denotes the power set of $\mathcal{X}$. Several different voting rules exist. Below we introduce the three rules we will focus on in this thesis.

- *k-approval:* The top k alternatives of every ballot R_i receive one point and the alternative with most points wins. If k is one, only the top element of each voter receives a point, this is also known as the *plurality* rule. If k is $m-1$ (recall that m refers to the number of alternatives), every alternative besides the bottom element receives a point and this is also known as *anti-plurality* or *veto*.

- *Borda:* The alternative ranked on the jth position receives $m-j$ points. We do this for every ballot R_i and the alternative with the most points wins.

- *Copeland:* If we have two alternatives X and Y, then X wins a pairwise majority context if more than half of the voters prefer X over Y (more than half of the voters vote $X \succ Y$). Now, an alternative X receives one point for every pairwise majority contest won and loses one point for every contest lost and the alternative with the most points wins.

Example 2.1. *Recall the example from the introduction in which colleagues vote where to go for dinner. The alternatives where sushi (S), burgers (B), and local food (L). This means that $\mathcal{X} = \{S, B, L\}$ and consider the following profile $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^5$ consisting of the following 5 votes:*

$$\begin{aligned} \textit{3 voters:} \quad & L \succ B \succ S \\ \textit{2 voters:} \quad & B \succ S \succ L \end{aligned}$$

If we use plurality voting, the colleagues will go for local food, since L gets three points, B two, and S receives zero points. While if we use the anti-plurality rule, the colleagues will go for burgers, since B receives five points, S two, and L three.

If the winner is determined using Borda, alternative B will be elected. This is because S receives $3 \cdot 0 = 0$ points from the first 3 voters and $2 \cdot 1 = 2$ from the last 2 voters. Alternative B receives $3 \cdot 1 + 2 \cdot 2 = 7$ points and L receives $3 \cdot 2 + 2 \cdot 0 = 6$ and thus B has the most points and is elected.

Finally, if we consider the Copeland rule, we have to consider all pairwise majority contests to determine who is winning. Now, with S vs B, burgers wins, S vs L, sushi wins, and B vs L, the local cuisine wines. So, L has 2 points, B 0 points, and S has -2 points. So L is elected by the Copeland rule.

For most voting rules, including the ones described above, there are situations in which multiple winners are chosen. Voting rules that always select just one winner are called *resolute*:

Definition 2.1. *A voting rule F is* **resolute** *whenever for every profile $\boldsymbol{R}$, $|F(\boldsymbol{R})| = 1$. In other words, F always elects a unique winner.*

Since most voting rules are irresolute, in practice these rules are often combined with a *tie-breaking procedure* that picks a unique winner from the set of winners elected by the voting rule. Tie-breaking comes in different flavors, including random procedures and using any choice function $T : \wp(\mathcal{X}) \to \mathcal{X}$ (for an overview and discussion of different procedures, see Freeman et al. (2015)). In this thesis, we will consider *lexicographic tie-breaking functions*, also known as *fixed order tie-breaking.* These functions induce a linear order over the alternatives: $X_1 \triangleright X_2 \triangleright \ldots \triangleright X_m$ that determines which alternative is elected. This means, if $X \triangleright Y$ and $F(\boldsymbol{R}) = \{X, Y\}$, then X will win the election when we combine F with the lexicographic tie-breaking function.

2.1.1 Three Types of Rules: Condorcet Extensions, Positional Scoring Rules, and Tournaments

We have introduced k-approval, Borda, and Copeland here, but these are just three examples out of a whole range of different voting rules. In the literature, voting rules are often divided into three types: *Condorcet extensions*, *positional scoring rules*, and *tournaments* (Zwicker, 2016).

Condorcet extensions

Condorcet extensions are voting rules that satisfy the *Condorcet principle.* This principle says that if there is some alternative that is preferred to every other alternative by a majority of the voters, then it should win the election. Such an alternative is called the *Condorcet Winner* and does not always exist. Copeland is an example of a Condorcet extension. Another well-known Condorcet extension is the *maximin rule*, also known as the *Simpson rule.* Let $Net_p(X > Y)$ be the net difference between alternative X and Y, that is:

$$Net_P(X > Y) = |\{i \in \mathcal{N} \mid X \succ_i Y\}| - |\{i \in \mathcal{N} \mid Y \succ_i X\}|$$

Then, the alternative who has the maximum value for $min\{Net_P(X \succ Y) \mid Y \in \mathcal{X} \setminus \{X\}\}$, wins the election. If a Condorcet winner does exist, Copeland and maximin clearly elect the same alternative, if no Condorcet winner exists, maximin takes the margin of victory into account, while the Copeland rule chooses a winner based on the number of victories of an alternative.

Positional scoring rules

Positional scoring rules are voting rules that are not characterized by the type of winner they should elect, but by the structure of the rule. Formally, a positional scoring rule can be defined by a *scoring vector* $s = (s_1, ..., s_m)$ of real numbers, such that $s_1 \geq s_2 \geq \ldots \geq s_m$, and in particular that $s_1 > s_m$. Now, if a voter places an alternative on position i, this alternative receives s_i points, and the alternative with the most points is elected. Both k-approval (for $k > 0$ and $k < m$) and Borda are positional scoring rules. Another example of a positional scoring rule is the points system used in the Formula One Grand Prix. Here, every race can be seen as a voter and the point system as a positional scoring rule with the scoring vector $s = (25, 18, 15, 12, 10, 8, 6, 4, 2, 1, 0, ..., 0)$. This means that a racer receives 25 points when they win a race, 18 when they are second, and so forth, and the racer that has the most points at the end of the season wins (Zwicker, 2016, p. 37).

Tournaments

The third type of voting rules is based on the idea that the least favorite alternative should drop out until one alternative receives the support of the majority. These rules are called tournaments or *run-offs*. One of the most well-known voting rules of this type is *single transferable vote (STV)*. This rule proceeds in stages, at each stage the alternative that receives the least plurality points is removed from the ballots of all voters. This proceeds until there is one alternative that receives a majority of the votes and therefore wins the election. Another tournament rule is the *cup rule*, which can be defined by a binary tree where every leaf node is labeled by an alternative. The alternative that continues to the parent node is the alternative that wins in a majority contest. The alternative that makes it to the root, wins the election (Brandt et al., 2016a).

Structure versus type of winner

Positional scoring rules and tournaments are both characterized by their structure, while Condorcet extensions are defined by the type of winner that is elected. Nevertheless, the set of tournaments and the set of Condorcet extensions overlap, which means that some tournament rules also satisfy the Condorcet principle. An example of this it the *cup rule*, in which a Condorcet winner will always beat every other alternative in a majority contest and will thus by definition make it to the root of every cup rule. Positional scoring rules, on the other hand, never satisfy the Condorcet principle.[1] It can, therefore, be argued that positional scoring rules and Condorcet extensions represent a fundamentally different view on which alternative should be elected.

This also corresponds to the view of the inventor of the Condorcet principle, Marquis de Condorcet. In his essay *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Marquis de Condorcet (1785) argued that the only right way to choose the correct alternative is to assume that the majority of society is more often right than wrong. So, he believed that the only best alternative is the one beats all others; the Condorcet winner. Condorcet had quite a low esteem of another pioneer in election theory at the time, Jean Charles de Borda, who proposed the Borda method in 1770. Some researchers even argue that that Condorcet deliberately shifted his ideas to make sure they were fundamentally different from Borda because he saw him as his personal enemy (Young, 1988, p. 1238).

Since positional scoring rules and Condorcet extensions represent a different perspective on who should win an election, both theoretically and historically, we believe it is important to take both types of rules into consideration. This is one of the reasons we take both Borda and the Copeland rule into account in this thesis.

[1] This can be easily seen if we consider an example with three alternatives and seven voters in which three voters vote $A \succ B \succ C$, two voters vote $B \succ C \succ A$, one voter votes $B \succ A \succ C$, and the last voter votes $C \succ A \succ B$. Now, A is a Condorcet winner, but any positional scoring rule will make B (also) win, since B receives $3 \cdot s_1 + 3 \cdot s_2 + s_3$ points, A gets $3 \cdot s_1 + 2 \cdot s_2 + 2 \cdot s_3$ points, and C receives $1 \cdot s_1 + 2 \cdot s_2 + 4 \cdot s_3$ and it must hold that $s_1 \geq s_2 \geq s_3$ (Endriss, 2017).

2.1.2 Characterizing Voting Rules using Axioms

Most voting rules elect the alternative that is most 'fair'. However, as we have also seen in the debate between Condorcet and Borda, there are several ways 'being most fair' can be defined. A way to characterize different voting rules using this perspective is by applying the *axiomatic method*. With this method, you precisely formulate certain intuitive desirable or undesirable properties as axioms and characterize rules by which axioms they do or do not satisfy. The Condorcet principle, introduced above, is an example of such an axiom. The axiomatic method is often seen as one of the most important tools within social choice theory (Endriss, 2011, p. 336). Important to note here is that in the context of voting theory, axioms often have a normative character, stating which properties a good social choice function *should* posses. This is in contrast with mathematics and logic, in which axioms generally have a more descriptive character and talk about which properties mathematical objects *do* posses (Endriss, 2011).

The axiomatic method does not only entail formulating axioms but also proving theorems that tell us something about combinations of axioms. The famous *impossibility theorem* by Arrow (1963) is an example of this.[2] The axiomatic method also gave rise to the *Gibbard-Satterthwaite Theorem*, which tells us that strategic manipulation is possible with most common voting rules. This theorem shows that strategic manipulation is something that almost cannot be avoided and is, therefore, one of the theoretical starting points for this thesis. Before we can fully explain what Gibbard and Satterthwaite (independently) proved, some more formal concepts are needed. We will introduce basic notions below and go into strategic manipulation and the Gibbard-Satterthwaite Theorem in the next section.

Some fundamental properties of voting rules

Three fundamental properties that most voting rules satisfy are *anonymity*, which implies that the order of the voters does not matter, *neutrality*, which implies that the name or order of the alternatives does not matter, and *surjectivity*,[3] which means that every alternative has a chance to win. Borda, k-approval and the Copeland rule are all anonymous, neutral and surjective. Formally, we can define these properties as follows:

Definition 2.2. *A voting rule F is* **anonymous** *whenever for each permutation $P : \mathbb{N} \to \mathbb{N}$ of the natural numbers and profile $\boldsymbol{R} = (R_1, ..., R_n) \in \mathcal{L}(\mathcal{X})^n$ it holds that $F(R_1, ..., R_n) =$*

[2] The impossibility theorem of Arrow states that if there are three or more alternatives, there can never be a voting rules that can convert the ranked preferences of all voters into a community-wide ranking that also satisfies a set of criteria. Roughly, these can be summerized as follows: (1) if every voter prefers alternative X over Y, then the group prefers X over Y, (2) if the preference of each voter between X and Y remains the same, the group preference will also stay the same, and (3) there is no dictator that determines on their own.

[3] In the literature, being *non-imposed* is sometimes used as a synonym for surjectivity. In this thesis, we will say a voting rule is non-imposed when for every alternative X there exists a profile $\boldsymbol{R}$ such that $F(\boldsymbol{R}) = \{X\}$, while with surjectivity we mean that every alternative can be elected, but it does not need to be a unique winner. This is a substantial difference when F is irresolute, while if F is resolute these two concepts coincide.

$F(R_{P(1)}, ..., R_{P(n)})$.

Definition 2.3. *A voting rule F is* **neutral** *when for each permutation $P : \mathcal{X} \to \mathcal{X}$ of alternatives and every profile $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^n$ it holds that $F(P(\mathcal{R})) = P(F(\mathcal{R}))$.*

Definition 2.4. *A voting rule F is* **surjective** *if for every element $X \in \mathcal{X}$ there exists some $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^n$ such that $X \in F(\boldsymbol{R})$.*

Since voting rules are used to make a collective decision, a property that we often do not want voting rules to have is that it is a *dictatorship*, in which *one* voter determines who wins the election. Throughout this thesis, we will use the function $top : \mathcal{L}(\mathcal{X}) \to \mathcal{X}$ to refer to the alternative that is on the top position of some ballot R_i and $top_k : \mathcal{L}(\mathcal{X}) \to \wp(\mathcal{X})$ (where $|\, top_k(R_i) \,| = k$) to the top k alternatives of a ballot R_i.

Definition 2.5. *A resolute voting rule F is a* **dictatorship** *if there is some voter $i \in \mathcal{N}$ such that for every profile $\boldsymbol{R} = (R_1, ..., R_n)$ it holds that $F(\boldsymbol{R}) = top(R_i)$.*

Borda, k-approval and Copeland are all clearly no dictatorships.

2.2 Strategic Manipulation

So far we have only considered the ballot R_i that a voter submits. It is often also assumed that a voter has some *true preference* over the set of alternatives that is not necessarily the same as their ballot. This idea corresponds to the concept of *strategic manipulation*, in which voters misrepresent their preferences to improve the outcome of an election for themselves. Throughout this thesis, we will assume that a true preference of a voter is also a linear order over the set of alternatives. This means that if we only consider *honest* voters, every voter i has a true preference R_i which she then submits as a ballot R_i to participate in the voting procedure. Sometimes, however, a voter can benefit from voting something different than her true preference. Example 2.2 illustrates this.

Example 2.2. *Suppose we have a set of 4 alternatives $\mathcal{X} = \{A, B, C, D\}$ and a profile $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^5$ consisting of 5 voters:*

$$\begin{array}{rl} \textit{voter 1 and 2:} & C \succ D \succ A \succ B \\ \textit{voter 3:} & D \succ B \succ C \succ A \\ \textit{voter 4:} & B \succ D \succ A \succ C \\ \textit{voter 5:} & A \succ B \succ C \succ D \end{array}$$

If we use the Borda rule to determine the winner, we see that D wins with 9 points (A,B and C receive 6, 7 and 8 points respectively). Now, consider what happens if voter 5 submits an alternative ballot $R_5' : C \succ A \succ B \succ D$. Now we get the following profile $\boldsymbol{R}'$:

$$\begin{array}{ll} \textit{voter 1 and 2:} & C \succ D \succ A \succ B \\ \textit{voter 3:} & D \succ B \succ C \succ A \\ \textit{voter 4:} & B \succ D \succ A \succ C \\ \textit{voter 5:} & C \succ A \succ B \succ D \end{array}$$

Now, D still receives 9 points but since C now receives 10 points, alternative C gets elected. Since $C \succ_5 D$, it is beneficial for voter 5 to manipulate, to report to vote something that does not correspond to her truthful preference.

In the situation described in Example 2.2, we say that voter 5 has an *incentive to manipulate.* In the context of manipulation, we will often look at what happens if one voter changes her vote while the rest of the profile stays the same. To talk about this more easily in a formal setting, we will use the following notation: If we have some profile $\boldsymbol{R} = (R_1, ..., R_n)$ and some ballot $R'_i \neq R_i$, we will use $(\boldsymbol{R}_{-i}, R'_i)$ to denote the profile in which R_i is replaced by R'_i. Since almost all voting rules that we will consider are anonymous, we can also describe this situation by first fixing the votes of all other voters in a profile $\boldsymbol{R'} = (R_1, ..., R_{n-1}) \in \mathcal{L}(\mathcal{X})^{n-1}$ and separately add the ballot of the manipulator to the profile. We will write $(\boldsymbol{R'}, R)$ to refer to the situation where a ballot $R \in \mathcal{L}(\mathcal{X})$ is added to $\boldsymbol{R'}$.

Now, we can formally define manipulation or strategyproofness as follows:

Definition 2.6. *A resolute voting rule F is* **strategyproof** *(or immune to manipulation) if for no voter $i \in \mathcal{N}$ there exists a profile $\boldsymbol{R} = (R_1, ..., R_n) \in \mathcal{L}(\mathcal{X})^n$ and some untruthful ballot $R'_i \in \mathcal{L}(\mathcal{X})$ with $R_i \neq R'_i$ such that $F(\boldsymbol{R}_{-i}, R'_i) \succ_i F(\boldsymbol{R})$.*

Nevertheless, Gibbard (1973) and Satterthwaite (1975) independently showed that most voting rules are not strategyproof:

Theorem 2.1 (Gibbard-Satterthwaite)**.** *Any resolute voting rule for more than 3 alternatives that is surjective and strategyproof is a dictatorship.*

There are several proofs of this theorem, for a simple proof see (Benoît, 2000).

In this thesis, we only focus on resolute rules. One may wonder whether irresoluteness solves this problem. Duggan and Schwartz (2000) showed, however, that there are similar impossibility results for irresolute rules.

Some researchers in computational social choice see the Gibbard-Satterthwaite Theorem as a negative result and investigate ways to circumvent manipulability (see e.g. Endriss, 2011; Brandt et al., 2016b). On the other hand, there are also researchers that take manipulability for granted and see voting as a game in which voters strategize to get what they want. Meir et al. (2014), for example, take this approach. Independently of whether you believe manipulation is something that should be prevented or taken for granted, the Gibbard-Satterthwaite Theorem tell us manipulation is something that cannot be avoided if we want to study natural voting rules. Moreover, if a voters manipulate successfully, the outcome of a voting rule changes compared to the situation in which the voters vote truthfully. This means that manipulation influences the mechanisms of collective decision making. So, we believe that it is interesting and relevant to investigate under which circumstances voters can or cannot manipulate.

2.2.1 Barriers Against Strategic Manipulation

As a response to the Gibbard-Satterthwaite Theorem, researchers have tried to find barriers against manipulation, such as restricting the domain, considering the complexity of manipulation or restricting the amount of information the voters have. We will present these barriers in more depth below. These barriers can be seen as settings in which the Gibbard-Satterthwaite theorem does not apply and we might find some rules that are reasonable and strategy-proof if we accept these settings. Similar to the approach in this thesis, these barriers are also often studied in different combinations.

Domain restriction

In the Gibbard-Satterswaite theorem, voters are allowed to have any possible preference order, as long as it is a linear order of the preferences. If we *restrict the domain*, however, voters are only allowed to have preference orders of a certain type. One of the most well-known domain restrictions is the domain of *single-peaked* preferences (Black, 1948). If preferences are single-peaked, it is assumed that alternatives can be ordered by some linear order $<$. For example, when the voters have to decide on a certain number, such as the age when people are allowed to drive a car, a natural order world be the usual order of natural numbers $<$. Furthermore, in politics, it is not unreasonable to assume that political parties or candidates can be ordered on a spectrum from progressive to conservative.

Assuming this order $>$ over the alternative is fixed, the preference order of a voter with A as their favourite alternative is single-peaked whenever for any two alternatives B and C such that B is between A and C with respect to the order $>$ (i.e. either $A > B > C$ or $C > B > A$) it must hold that the voter thinks that $B \succ C$, see Figure 2.1. Intuitively, this means that voters always prefer alternatives that are 'closer' to their favorite alternative over alternatives that are 'further'. As also argued above, in some contexts it makes sense to assume voters preferences are aligned with some external linear order. For example, when a voter thinks 18 is the right age to drive a car, it is reasonable to believe that the voter would then not then prefer 16 over 17. The downside of this approach is that if preferences are not single-peaked, there is no way of making them single-peaked, so these voters cannot vote their true preference orders.

Dummett and Farquharson (1961) showed that the *median voter rule* is strategy-proof in the single-peaked domain. When applying this rule, voters are ordered by their most preferred alternative. Then, the candidate preferred by the voter in the middle, the median, is elected. This rule is also a Condorcet extension.[4] This shows that when we restrict the domain to instances that have certain properties, reasonable voting rules that are strategy-proof exist.

[4]Moulin (1980) generalized this result by showing that the class of anonymous, Pareto efficient and strategy-proof voting rules is much larger when only considering single-peaked preferences. A voting rule is Pareto efficient whenever it always selects an alternative such that for any other alternative at least one voter must be worse off. In other words, there is no alternative that is preferred by all voters to the one elected.

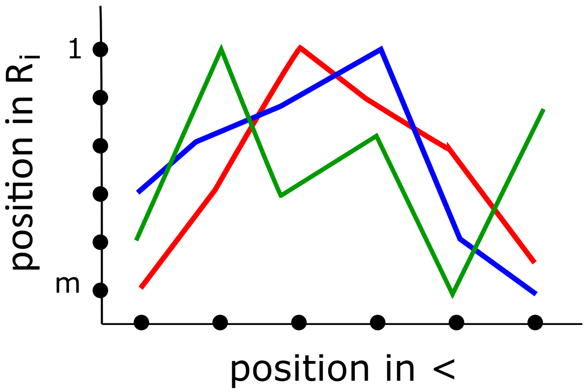

Figure 2.1: Visualization of single-peaked preference orders. The blue and red preference orders are single-peaked, the green is not. On the x-axis is the order $<$ of the alternatives, on the y-axis the position of an alternative in the preference order of a voter i, from m to 1.

Computational barriers

Even though it might be possible to manipulate, it can be computationally too expensive to find what to manipulate with giving their true preference. This concept is used for the idea of a computational barrier against manipulation. Often, this barrier is defined in terms of classical complexity theory. We assume that the reader is familiar with basic notions of this field, if this is not the case, see e.g. Arora and Barak (2007). In this context, the terms 'computationally too expensive', 'intractable', or 'resistant' usually mean that the problem is NP-hard, while 'easy' or 'tractable' means solvable in polynomial time. We will consider manipulation from this computational perspective in Chapter 5.

The idea that complexity can be a barrier to manipulation originates from an influential paper by Bartholdi et al. (1989). In this paper, Bartholdi et al. showed that with most voting rules it is easy to manipulate, including plurality, Borda, maximin, and Copeland. Moreover, they showed that a rule called *second order Copeland* is hard to manipulate. This is a rule that determines the winner in a similar way as the original Copeland rule, but if there is a tie the alternative whose defeated competitors have the largest sum of Copeland scores, wins. In a follow-up result, Bartholdi and Orlin (1991) showed that STV is also resistant to manipulate. This approach has led to a subfield of voting theory, including more complexity results. For an overview see e.g. Conitzer and Walsh (2016).

Informational barriers

In the Gibbard-Satterthwaite Theorem, it is assumed that all the voters have *full information* about what other voters are voting. This is not always a realistic assumption. The final barrier against strategic manipulation we will discuss here is to loosen this assumption. Several researchers have been looking at the setting in which voters have *incomplete information* about what others are voting. The first important question to answer is how to model incomplete information and what it means to be strategy-proof in this setting. This is one of the main topics of this thesis and we will discuss this situation in depth in Chapter 4. Before we do this, we will first formally define the model of partial information we will adopt in the rest of this research.

2.3 Conclusion

In this chapter, we introduced the theoretical basis of voting theory. We started with explaining what a social choice function is, pointed out the difference between Condorcet extensions, positional scoring rules, and tournaments, and introduced the axiomatic method. Second, we introduced strategic manipulation, which is one of the main topics of this thesis, and presented several barriers against manipulation suggested in the literature.

Besides providing the reader with the needed theoretical background, we also set the stage for the rest of this thesis by introducing which phenomenon and voting rules we will study: manipulation under k-approval, Borda, and Copeland. We have argued that since manipulation influences collective decision making and the Gibbard-Satterthwaite Theorem shows it cannot be avoided, we believe it is an interesting and relevant phenomenon to study. In the next chapter, we will explain and formalize the context in which we will study manipulation: partial information about the preferences of other voters.

Chapter 3

Incomplete Information

In this thesis, we investigate how partial knowledge about the preferences of other voters influences the manipulability of different voting rules. To do so, a formal framework needs to be defined in which partial knowledge in combination with manipulation can be expressed. Several frameworks for this concept already exist. In this chapter, we will introduce and compare three frameworks of voting with incomplete information and then argue why we think the model based on partial orders fits our situation best. Nevertheless, we believe that some alterations should be made to this framework. In the final section of this chapter, we will introduce the formal framework we will work with and analyze several properties of it.

3.1 A Set of Possible Profiles

A popular approach to modeling incomplete information in a non-probabilistic manner is to replace a single profile with a set of possible profiles. This idea is illustrated by the following example:

Example 3.1. *Suppose we have some manipulator with a true preference $A \succ B \succ C \succ D$ that does not exactly know what the preferences of the other voters are. However, one of the following two situations must be the case:*

Situation 1		*Situation 2*
2 voters:	$C \succ D \succ B \succ A$	
1 voter:	$D \succ A \succ C \succ B$	*4 voters:* $D \succ A \succ C \succ B$
1 voter:	$A \succ D \succ B \succ C$	

When the Borda rule is used to select the winner, alternative D will win in both situations if the manipulator votes her true preference. However, if she decides to manipulate and vote preference $C \succ B \succ A \succ D$ instead, alternative C will win in situation 1 and the outcome does not change in situation 2. This seems like a good condition for the manipulator, since she might be better off and is at least never worse off.

Now, suppose the manipulator considers a third possible situation:

Situation 3	
2 voters:	$A \succ D \succ C \succ B$
1 voter:	$D \succ B \succ C \succ A$
1 voter:	$B \succ D \succ A \succ C$

In this situation, it is no longer safe for the manipulator to vote $C \succ B \succ A \succ D$ instead of her true preference, since this will change the outcome of the third situation from A to D. This means manipulating possibly leads to a worse result.

In this thesis, we will adopt the terminology of Conitzer et al. (2011) and refer to the set of possible profiles as the *information set* $E \subseteq \mathcal{L}(\mathcal{X})^{n-1}$. Note that only the ballots of the other voters are included in the information set, since we assume manipulators always know their own preference orders. Moreover, a manipulator considers every profile in the information set to be equally likely candidates for the actual situation. Now, a manipulator has an *incentive to manipulate* if there exists some non-truthful ballot with which she is never worse off, and least one time strictly better off. We can formally define this as follows.

Definition 3.1. *Given some information set E, a manipulator with a true preference R has an* **incentive to manipulate** *with respect to E if there exists some $R' \in \mathcal{L}(\mathcal{X})$ with $R \neq R'$ such that:*

(1) there exists a profile $\boldsymbol{R}^* \in E$ such that voting R' leads to a strictly better result. In other words, $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$, and

(2) there is no profile $\boldsymbol{R} \in E$ in which voting R' leads to a worse result. In other words, there is no $\boldsymbol{R} \in E$ such that $F(\boldsymbol{R}, R) \succ_R F(\boldsymbol{R}, R')$.

3.2 Existing Frameworks of Partial Information

There exist multiple frameworks that formalize incomplete information using a set of possible profiles. Below, three frameworks will be described. The main difference between these models is the structure of the information set E. As we will also argue below, we believe that these models are not necessarily competing for one single truth, but rather modeling strategic manipulation with incomplete information from a different perspective. In this thesis, we want to learn how knowledge about the preferences of other voters influences the manipulability of different voting rules. In the next section, we will argue why we chose to expand on one of these models, the one based on partial information by Conitzer et al. (2011).

Partial orders

Conitzer et al. (2011) start by defining information similar to the one we defined above: a set that consists of different profiles, without any additional assumptions about the structure of E. In the rest of their paper, however, they assume that E is induced by a partial order. That is, if $\mathcal{P}(\mathcal{X})$ are all partial orders over $\mathcal{X}$ (every relation that is reflexive, anti-symmetric and transitive, but not necessarily total), then they say that every E is induced by some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ and contains every $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ that extends $\boldsymbol{R}$. Formally, if $\boldsymbol{R} = (R_1, ..., R_{n-1}) \in \mathcal{P}(\mathcal{X})^{n-1}$, we define the information set $E_{\boldsymbol{R}}$ as follows:

$$E_{\boldsymbol{R}} = \{ (R'_1, ..., R'_{n-1}) \mid R_i \subseteq R'_i \text{ and } R'_i \in \mathcal{L}(\mathcal{X}) \text{ for all } 0 < i < n \}$$

In the rest of this thesis we will say that $\boldsymbol{R}' = (R'_1, ..., R'_{n-1})$ *extends* $\boldsymbol{R}$, write $\boldsymbol{R}' \supseteq \boldsymbol{R}$ when $R'_i \supseteq R_i$ for all $0 < i < n$.

Conitzer et al. (2011) mainly focus on the computational aspect of manipulation. For example, they show that manipulation with partial orders is NP-hard for Borda and Copeland.

Information polls

Reijngoud and Endriss (2012) take a different approach; they assume that voters have no direct information about the preferences of others but instead have access to the outcome of an *opinion poll*, which they define relatively unconstrained as 'a piece of information'. Formally, they let $\mathcal{I}$ be the set of all possible pieces of poll information that could be communicated to voters. Now, a *poll information function (PIF)* is a function $\pi : \mathcal{L}(\mathcal{X})^n \to \mathcal{I}$ that given a specific profile, returns an element of $\mathcal{I}$ that contains this information. Examples of PIF's they consider are the winner PIF, that given a profile returns which alternative is winning, the majority graph PIF (MG-PIF) that given a profile returns the corresponding majority graph, and the profile PIF, which simply returns the initial profile.

A PIF also induces an information set that contains all profiles that are consistent with the piece of information a voter received from the poll. Reijngoud and Endriss (2012) define this formally as follows: given a voter i, a PIF π and a profile $\boldsymbol{R}$, the information set containing all possible profiles can be defined as:[1]

$$\mathcal{W}_i^{\pi(\boldsymbol{R})} := \{\boldsymbol{R}'_{-i} \in \mathcal{L}(\mathcal{X})^{n-1} \mid \pi(\boldsymbol{R}'_{-i}, R_i) = \pi(\boldsymbol{R})\}$$

First, Reijngoud and Endriss focus on the *single-poll setting*, in which they show which PIFs in combination with which voting rules are susceptible or immune to manipulation. For instance, they show that when we consider the winner PIF, $m \geq 3$, and $n \geq 4$, any positional scoring rule, paired with lexicographic tie-breaking is susceptible to manipulation. Endriss et al. (2016) extend this work by showing that k-approval (for $k \leq m - 2$), Copeland, and Maximin voting rules are also susceptible to manipulation when considering the winner PIF. In the second part of their work, Reijngoud and Endriss focus on the *repeated-poll setting*. They show, using both analytical and experimental methods, how different groups of voting rules respond to different assumptions on how voters respond to poll information.

Distance based uncertainty

Meir et al. (2014) use the concept of distance to define uncertainty. In their model, the information set consists out of all profiles that are within a certain distance of one specific

[1] We have adapted their notation to match the one used in this thesis.

profile. That is, if $d : \mathcal{L}(\mathcal{X})^{n-1} \times \mathcal{L}(\mathcal{X})^{n-1} \to \mathbb{N}$ is a distance measure on profiles (for example, how many adjacent swaps are needed to obtain one profile from the other), then the information set includes all profiles that have at most a certain distance to one specific profile. Formally:[2]

$$E_{\boldsymbol{R}}(x) = \{\boldsymbol{R}' \mid d(\boldsymbol{R}, \boldsymbol{R}') \leq x\}$$

As said before, Meir et al. mainly focus on the voting dynamics if we assume that all voters manipulate. They treat voting as an iterative game in which voters can strategize and change their ballot, and define a *voting equilibrium* as a state in which no voter wants to change their vote. They show, for example, that if voters start with their truthful ballot and plurality voting is used that such an equilibrium is reached relatively quickly.

Describing different type of situations

While in all three models the notion of an information set plays an important role, the differences in the structure of this set lead to three conceptually different models of incomplete information. Both the model of Conitzer et al. (2011) and the model of Meir et al. (2014) focus on the situation where you have direct knowledge about the preferences of other voters,[3] while in the case of Reijngoud and Endriss (2012), this knowledge is obtained indirectly from information obtained by polls. Moreover, the model based on partial orders and the model based on information polls both show that there are more things that the manipulator knows for certain besides the information set. For example, who wins or that voter j has alternative X on position l. However, this is not always the case when distance-based uncertainty is used. In particular, it could be that the manipulator does not know any piece of information, or has no certainty about which voter has which alternative on which position.

We believe that these three models are describing a different situation, rather than competing to be a 'true' model of incomplete information. The model of partial orders, for example, can be used to describe the case where you know certain facts about the preferences of others. At the same time, the distance-based uncertainty model can be used to describe the situation in which you have certain intuitions about the preferences of others, but are not completely certain about it. The model by Reijngoud and Endriss can, obviously, be used to model the situation in which you have information from polls. In other words, which model is the best, depends on the situation you want to capture.

3.3 Our Formal model

Since we focus on partial knowledge of the preferences orders of the voters, the model of Conitzer et al. (2011) will be used as a theoretical basis of this thesis information sets will be defined using partial orders. We believe, however, that some alterations or adaptations

[2] Here, we also adapted the notion to match the one used in this thesis.

[3] In the model of Meir et al. this obviously depends on which distance measure used, but this holds for almost every distance measure they define in their work.

need to be made. This will be discussed in this section. Furthermore, we will formally define the framework that is used in the rest of this thesis.

A more fine-grained notion of partial information

In contrast to Conitzer et al., we will not use arbitrary partial orders but focus partial orders with a certain 'shape' only. For example, the case where you know the top and bottom alternative of every voter, or the top t alternatives of every voter. The reason for this is twofold. First of all, we believe that allowing for arbitrary partial orders is not very cognitively plausible. Although arguments about the cognitive plausibility of incomplete information are difficult to support scientifically, we believe that in this case, it is reasonable to assume that there are some boundaries to what people can remember. In particular, we believe that the instances that Conitzer et al. use to show NP-hardness are so complicated, that we are slightly hesitant to believe that these cases actually occur in practice. Since these instances led to the NP-hardness results, we believe it is interesting to investigate what kind of result can be obtained when some restrictions are put on the partial orders. Chapter 5 will present that limiting the shape of partial orders leads to polynomial time complexity results in the case of k-approval. Moreover, we will show that the NP-hardness proofs cannot be translated to these structures or any structures with a relatively general but restricted structure.

Second, since we would like to know how partial knowledge about preferences influences manipulability, it could be interesting to compare different structures of partial orders. As we will show in Chapter 4, for example, knowing a little about the preference of every voter differs from knowing everything about the preferences of some of the voters.

An approach worth mentioning here is the work by Briskorn et al. (2016) (see also Erdélyi and Reger, 2016). Despite their focus on bribery and control instead of manipulation, they also look at partial information in a more fine-grained way.[4] In their work, they perform an exhaustive study of how different ways partial knowledge of the voters can be modeled influences the complexity of bribery and control. They argue that different models of partial ballots yield different complexity results for the same problems. In particular, they state that using all partial orders often result in hardness, while other very general models give polynomial results. This corresponds to the findings of this thesis, that hardness results seem to rely on very specific partial orders that we believe might not be very plausible in practice.

3.3.1 Different Structures of Partial Orders

In this thesis, we want to consider partial orders that have a certain 'shape'. Ideally, we would like to have convincing cognitive arguments supporting which partial orders are likely to actually occur. Unfortunately, as argued before, these type of claims are difficult to back

[4]Control models the situation in which the chair, or some other organizer, has the ability to change the structure of the election, while in the case of bribery the structure of the election remains the same. In this case there is, however, some outside agent that tries to bribe voters to change their vote.

up scientifically. For this reason, we will start with very simple scenarios that we believe are plausible and investigate what kind of results can be obtained. We want to stress that we do not believe that these are the only structures that are cognitively plausible.

Recall the example of the introduction about the colleagues who vote to decide where to go for dinner. Here, the manipulator only had information about what the other voters really disliked. This can also be turned around by letting the manipulator have information about what the other voters really like. In other words, the manipulator knows some of the top or bottom (or both) alternatives of the voters. These structures can be used to model situations in which the manipulator has conversations with the other voters in which they discuss what their favorite or least favorite alternatives are, but do not communicate their full profile. Second, we will consider the situation where the manipulator knows the full preferences of some voters and nothing about the preferences of others. In the dinner example, this could be the case when you know some of the colleagues really well and you know exactly what they like, while you have no idea about the preferences the others.

We have formalized these two intuitions into four different structures. Note that $A \sim_R B$ means that not $A \succ_R B$ and not $B \succ_R A$. Moreover, if $\mathcal{X}' \subseteq \mathcal{X}$ and $A \in \mathcal{X}$, we will write $A \succ \mathcal{X}'$ when we mean that $A \succ B$ for every $B \in \mathcal{X}'$ and similarly $\mathcal{X}' \succ A$ when it holds that $B \succ A$ for every $B \in \mathcal{X}'$.

Formally, we can define these notions as follows and let $\boldsymbol{R} = (R_1, ..., R_{n-1}) \in \mathcal{P}(\mathcal{X})^{n-1}$ be a partial order of $n-1$ voters. In this thesis, we will consider the following four structures that $\boldsymbol{R}$ can have (see Figure 3.3.1 for a visual representation):

- TOPBOT: The manipulator only knows the most and least preferred alternative of all other voters. Formally, we say that the information set induced by $\boldsymbol{R}$ is an instance of TOPBOT (e.g. $E_{\boldsymbol{R}} \in$ TOPBOT) if for every voter $i \in \{1, ..., n-1\}$ there exist $A, B \in \mathcal{X}$ (A being the top and B the bottom alternative) such that for every $C \in \mathcal{X} \setminus \{A, B\}$ it holds that $A \succ_i C$ and $C \succ_i B$ and for every $D \in \mathcal{X} \setminus \{A, B\}$ it holds that $C \sim_i D$.

- TOP(t): This is the situation in which the manipulator only knows the top t alternatives of every voter i. Formally, we say $E_{\boldsymbol{R}} \in$ TOP(t) if for each voter $i \in \{1, ..., n-1\}$ there exists a subset $\mathcal{X}' \subseteq \mathcal{X}$ of alternatives with $|\mathcal{X}'| = t$ such that for every $A, B \in \mathcal{X}'$ with $A \neq B$ that either $A \succ_i B$ or $B \succ_i A$, and for every $C, D \in \mathcal{X} \setminus \mathcal{X}'$ we have $A \succ_i C$ and $C \sim_i D$.

- BOT(t): In this situation, the manipulator only knows the bottom t alternatives of every voter i. Formally, $\boldsymbol{R} \in$ BOT(t) when for each voter i there exists a subset $\mathcal{X}' \subseteq \mathcal{X}$ of alternatives with $|\mathcal{X}'| = t$ such that for every $A, B \in \mathcal{X}'$ with $A \neq B$ either $A \succ_i B$ or $B \succ_i A$, and for every $C, D \in \mathcal{X} \setminus \mathcal{X}'$ we have $C \succ_i A$ and $C \sim_i D$.

- COMPL(t): Here, the manipulator knows from t the complete preference order and nothing from the other voters. Formally, we say that $E_{\boldsymbol{R}} \in$ COMPL(t) whenever there exists a subset of voters $\mathcal{N}' \subseteq \{1, ..., n-1\}$, such that $|\mathcal{N}'| = t$ and for every $i \in \mathcal{N}'$ we have $R_i \in \mathcal{L}(\mathcal{X})$ and for every voter $j \in \{1, ..., n-1\} \setminus \mathcal{N}'$ it holds that $R_j = \emptyset$.

In the formalization, TOPBOT, TOP(t), and BOT(t), correspond to the scenario in which the manipulator only knows which alternatives the other voters like or dislike and COMPL(t) corresponds to the idea that the manipulator knows the complete profile from some voters and nothing about the preferences of the others.

Finally, we let STRUCT be the set that contains all structures:

Definition 3.2. *The set of all structures: STRUCT* $\subseteq \wp(\mathcal{L}(\mathcal{X})^{n-1})$ *is defined as follows:*

$$STRUCT = \{TOPBOT\} \cup \{BOT(t), TOP(t), COMPL(t) \mid t \in \mathbb{N}\}$$

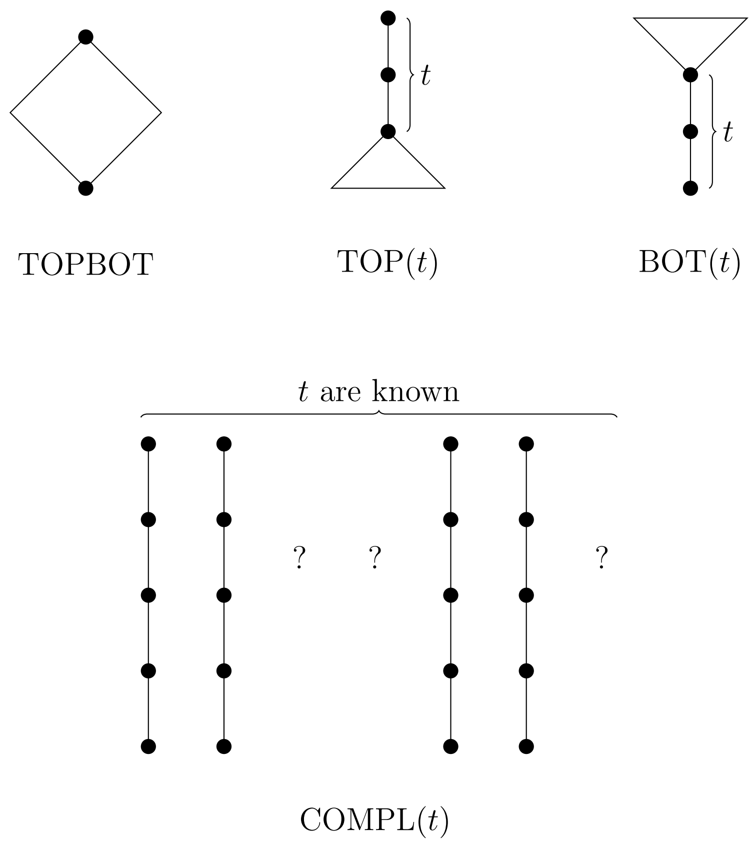

Figure 3.1: A visual representations of the five structures of partial orders. With TOPBOT, TOP(t) and BOT(t), we have drawn one preference order, the profile consists of n preferences with the same structure. In the case of COMPL(t) the picture represents the whole profile.

3.3.2 Properties of STRUCT

The structures introduced above have several properties. These properties are sued throughout this thesis. Since these properties also provide some more insight into the differences between the structures, they are already introduced in this section.

Every profile belongs to an information set

The first observation is that every full profile $\boldsymbol{R}$ belongs to some information set. This follows directly from the definitions of the structures. Since this observation is used multiple times throughout this thesis, we make this explicit with the following Lemma:

Lemma 3.1. *For every $\Sigma \in$ STRUCT, every $\mathcal{X}$ and every n, it holds that for every $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ there is some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}} \in \Sigma$ and $\boldsymbol{R}' \in E_{\boldsymbol{R}}$.*

Proof. Follows directly from the definition of every $\Sigma \in$ STRUCT. □

Partitions

The structures TOPBOT, TOP(t), and BOT(t) are partitions of $\mathcal{L}(\mathcal{X})^n$. This is because if there are two partial profiles $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{P}(\mathcal{X})^n$ with the same structure, that is $E_{\boldsymbol{R}}, E_{\boldsymbol{R}'} \in \Sigma$ for some $\Sigma \in$ {TOPBOT, BOT(t), TOP(t)}, then it holds that if some full profile $\boldsymbol{R}''$ is a member of both information sets (e.g $\boldsymbol{R}'' \in E_{\boldsymbol{R}}$ and $\boldsymbol{R}'' \in E_{\boldsymbol{R}'}$) it must hold that $\boldsymbol{R} = \boldsymbol{R}''$. In other words, some profile can not be part of two different information sets, but every profile does belong to some information set. This is exactly the definition of a partition. This does not hold for COMPL(t), however, since in this case one profile could easily be part of different information sets.

IP- and PI-transitivity

Intuitively, a partial order is both *IP-transitive* and *PI-transitive* whenever for any two alternatives with an undefined order, it holds that if there is a third alternative that is either above or below one of them, this third alternative must also be above or below the other one. All structures $\Sigma \in$ STRUCT have this property. The notions IP- and PI-transitivity are used by researchers that study formal structures of preferences (see e.g. Hansson, 2001). Since our partial knowledge is mathematically similar to these preferences, we can apply these notions to our structures. Chapter 5, we demonstrate this property might play an important role in the complexity of manipulation.

Definition 3.3. *A partial preference $R \in \mathcal{P}(\mathcal{X})$ is* **IP-transitive** *whenever $A \sim_R B$ and $B \succ_R C$ implies that $A \succ_R C$.*

Definition 3.4. *A partial preference $R \in \mathcal{P}(\mathcal{X})$ is* **PI-transitive** *whenever $A \succ_R B$ and $B \sim_R C$ implies that $A \succ_R C$.*

Lemma 3.2. *For any $\boldsymbol{R} = (R_1, ..., R_{n-1})$ such that $E_{\boldsymbol{R}} \in \Sigma$ for any $\Sigma \in$ STRUCT, every $R_i \in \boldsymbol{R}$ is both PI- and IP-transitive.*

Proof. Follows directly from the definition of Σ. □

Amount of knowledge is explicit

Most structures have a parameter t that explicitly indicates how much information the manipulator knows for certain. For example, we define COMPL(t) in which the manipulator knows the ballots of exactly t voters. The reason we make this explicit is that besides looking at the complexity of our problems, in which this might not play a very major role, we also investigate when you can manipulate and the amount of information that a certain structure gives the manipulator. To answer these questions it is useful to talk about these notions in a more quantitative setting. As we will show below, for some values of t it is possible to manipulate with COMPL(t), while for other values of t this is not possible.

3.3.3 Relation to Other Models

In the model by Meir et al. (2014), a distance measure can defined that results into the same partition as TOP(t), BOT(t) or TOPBOT, this measure returns zero when two profiles are in the same information set and one if they are not. However, we do not believe this is how Meir et al. intended their model. A more straightforward comparison is to see how COMPL(t) relates to information sets based on discrete distance. The discrete difference (d_d) between two profiles is the number of voters that have a different preference. This seems very similar to COMPL(t) but there is one major difference. If we have some $E_{\boldsymbol{R}} \in$ COMPL(t), it does hold that for every two $\boldsymbol{R}', \boldsymbol{R}'' \in E_{\boldsymbol{R}}$ we have that $d_d(\boldsymbol{R}', \boldsymbol{R}'') \leq t$. However, if there is another $\boldsymbol{R}^*$ such that $d_d(\boldsymbol{R}^*, \boldsymbol{R}')$ or $d_d(\boldsymbol{R}^*, \boldsymbol{R}'')$ it does not need to be the case that $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$. This is because besides stating how many preference orders can be different, as determined by the discrete distance, a partial order $\boldsymbol{R}$ also determines *which* preference orders are fixed and which are not. Nevertheless, Meir et al. do not go into the question for which distance measure and which distance you can manipulate and for which you cannot. Reijngoud and Endriss, on the other hand, do ask these types of questions. In their work, they look at with which PIF you can manipulate with which voting rule.

3.4 Discussion

There are several other ways we could have formalized the intuitive ideas about partial information suggested in this chapter. For example, Briskorn et al. (2016), who similarly defined a more fine-grained notion of partial information but in the context of bribery and control, use more general structures. In particular, all of our five structures are special cases of the nine models that they consider. We believe that the main reason for this is that we started with some simple intuitions of situations and tried to translate them into formal definitions, while Briskorn et al. started with some more general models and tried to simplify them. For instance, Briskorn et al. use and simplify the notion of double truncated ballots as defined by Baumeister et al. (2012), which is intended to model the situation in which voters can submit ballots where they are indifferent about some alternatives. We believe that for future work, it would be very interesting to consider other, in particular, slightly more complex structures of partial information.

Moreover, we explicitly model how much the manipulator knows for certain by using the variable t. As argued above, we believe this has the advantage that we have a way of quantifying the amount of knowledge the manipulator has and make precise statements about *how much* knowledge a manipulator needs to be able to manipulate. One drawback of explicitly defining t is that this implies that we assume that the manipulator has the same amount of information about all other voters. It can for example not be the case she knows the top three of one voter and the top five of another, while this situation seems to be realistic in practice. As we will see later, our complexity results do, however, take these type of structures into account.

3.5 Conclusion

In this chapter, the notion of partial information is introduced. After presenting the initial notion of an information set, we compared three different models of partial information and concluded that the model of Conitzer et al. (2011) suits our purposes best. The main reason for this is that this model allows us to model fixed partial knowledge about the preferences of other voters directly using partial orders. At the same time, the other models either model knowledge about the outcome of the election (e.g. who is winning) which means that knowledge about the preferences is modeled indirectly. Since the goal is to investigate how partial knowledge about the preferences of others influences manipulability of voting rules, we believe this model fits our purposes best. However, it is discussed that we think the model should be restricted since we believe not every arbitrary partial order is cognitively plausible. In the final section of this chapter, we introduced the restricted model which forms the theoretical basis of the rest this thesis. In the next chapter, the possibility of manipulation within this framework is investigated. In Chapter 5 we will investigate how difficult it is to manipulate in this context.

Chapter 4

When Can Voters Manipulate?

In the previous chapter, we argued that we believe the structure of partial orders matters. In this chapter, we will show that this is indeed is the case when we consider whether voters can manipulate or not. In particular, we will show that considering k-approval, Borda and Copeland, knowing a little about the preferences of all voters leads to different results than knowing everything about the preferences of some voters. In the case of k-approval, for example, we show that a manipulator can have an incentive to manipulate when only knowing the top alternative of every voter (also when $k > 1$), while if the voter knows the complete preference of t voters, they cannot manipulate unless t is at least half of the other voters.

The aim of this chapter is twofold. First of all, these results show that manipulators can have an incentive to manipulate when the partial information has the form of our structures. If this was not the case, Chapter 5 that concerns the complexity of manipulation would not have been relevant. Secondly, we show that there are differences between manipulability given the structures for all voting rules we consider. We believe this shows that what manipulators know about other voters matters, at least for k-approval, Borda, and Copeland. We will first formally define manipulation considering different structures and then show results for k-approval, Borda, and Copeland.

4.1 Σ-manipulation

In Chapter 3, we defined STRUCT to contain all four structures we will consider:

$$\text{STRUCT} = \{\text{TOPBOT}\} \cup \{\text{BOT}(t), \text{TOP}(t), \text{COMPL}(t) \mid t \in \mathbb{N}\}$$

Given these four structures, we need to formally define what manipulation means in this context. Recall from Definition 3.1 that given some partial profile $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ and the corresponding information set $E_{\boldsymbol{R}}$, we say that a manipulator with a true preference R has *an incentive to manipulate* with respect to $E_{\boldsymbol{R}}$ if there exists some $R' \in \mathcal{L}(\mathcal{X})$ with $R \neq R'$ such that:

(1) there exists a profile $\boldsymbol{R}^* \in E$ such that voting R' leads to a strictly better result: In other words, $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$, and

(2) there is no profile $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ in which voting R' leads to a worse result. Formally, there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$

In this thesis, we want to distinguish between being manipulable given different structures. Therefore, we introduce the notion of being susceptible when the manipulator has partial information in the form of a certain structure Σ:

Definition 4.1. *If $\Sigma \in$ STRUCT, then we say that a voting rule F is* **susceptible to Σ-manipulation** *whenever there is some $E_{\boldsymbol{R}} \in \Sigma$ and a true preference $R \in \mathcal{L}(\mathcal{X})$ such that the manipulator has an incentive to manipulate with respect to $E_{\boldsymbol{R}}$.*

If a voting rule is *not* susceptible to Σ-manipulation, we say that the voting rule is *immune* to Σ-manipulation.

There can be situations in which the manipulator can easily infer what the actual situation is. For example, when $t = m - 1$ when the manipulator knows TOP(t). In this case, there is *one* $\boldsymbol{R}'$ that extends every partial profile $\boldsymbol{R}$. This means that there is no longer any difference between the full and partial information. As a result, we can apply the Gibbard-Satterthwaite Theorem to obtain the following lemma:

Lemma 4.1. *If $m \geq 3$ and it holds that $|E_{\boldsymbol{R}}| = 1$ for all $E_{\boldsymbol{R}} \in \Sigma$, then any resolute voting rule F that is surjective and non-dictatorial, is susceptible to Σ-manipulation*

Proof. Let F be a resolute voting rule that is surjective, nondictorial, take any $m \geq 3$ and such an n, t and $\Sigma \in$ STRUCT such that $|E_{\boldsymbol{R}}| = 1$ for every $E_{\boldsymbol{R}} \in \Sigma$. Since $m \geq 3$ and F is both surjective and non-dictorial, it follows from the Gibbard-Satterthwaite Theorem that there exists some full profile $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ and two preferences R and R' such that $F(\boldsymbol{R}', R') \succ_R F(\boldsymbol{R}', R)$. Now, by Lemma 3.1 there must be exist some $E_{\boldsymbol{R}} \in \Sigma$ such that $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. By our assumption this means that $E_{\boldsymbol{R}} = \{\boldsymbol{R}'\}$. Hence, there is some $E_{\boldsymbol{R}} \in \Sigma$ and R with which the manipulator has an incentive to manipulate. Thus, F is susceptible to Σ-manipulation. $\square$

In the following sections, we will prove both susceptibility and immunity results for structures $\Sigma \in$ STRUCT in combination with k-approval, Borda, and Copeland. Proofs in which we show susceptibility results will often contain an example of an information set and a truthful profile with which the manipulator has an incentive to manipulate. Immunity results, on the other hand, will show why such an information set can never exist. In particular, we will regularly show that if there is a profile in which you can improve the outcome by an untruthful vote, there is always also a profile in the information set with which you will be worse off with this untruthful vote. Moreover, we will focus on the case where $m, n \geq 3$ only, since we believe these cases are the most interesting. For readability, some (parts) of the proofs of this chapter can be found in the Appendix.

4.2 k-approval

The first voting rule we consider is k-approval. A summary of the results can be found in Table 4.1. It can be seen that the case of TOPBOT and TOP(t) are comparable, which shows that knowing only the top element only of every voter already allows the manipulator to manipulate in some case. Knowing only the BOT(t) element is different, however. In this case, you need to know either as many alternatives as k or more than half of the $m - k$ alternatives ($t > m - k - t$) to be able to manipulate. The immunity bounds are slightly complicated but they are true whenever $t \leq m - k - t$, $t < k$, and n is 'large enough'. This shows that in the case of k-approval, there is small difference between knowing only the top and knowing only the bottom elements of all voters.

In the case of COMPL(t), the manipulator needs to know at least half of the preferences of the other voters to manipulate, compared to only the top alternative of all the voters in the case of TOP(t). This indicates that there is a difference between something everything about the preferences of all voters, and knowing everything about the preferences of some of the voters. We will show later that this is also the case when we consider Borda and Copeland.

	Susceptible	Immune
TOPBOT	$m, n \geq 3$	-
TOP(t)	$m, n \geq 3$ and $t > 0$	$t = 0$ and $n \geq \left\lceil \frac{m-2}{m-k} \right\rceil$
BOT(t)	$m, n \geq 3$, $t > 0$ and	$m, n \geq 3$ and both
	$t \geq k$ or $t > m - k - t$	$\left\lfloor \frac{(n-1) \cdot t}{k} \right\rfloor < \left\lfloor \frac{(n-1) \cdot (m-k-t)}{m-k} \right\rfloor$
		$\left\lfloor \frac{(n-1) \cdot t}{m-k} \right\rfloor < \left\lfloor \frac{(n-1) \cdot (m-k-t)}{k} \right\rfloor$
COMPL(t)	$m, n \geq 3$ and $t \geq \left\lceil \frac{n-1}{2} \right\rceil$	$t < \left\lceil \frac{n-1}{2} \right\rceil$ and $k \leq \frac{m}{2}$

Table 4.1: Results of k-approval.

Below, we will first introduce the notion of a *k-equivalence class* and then prove theoretical bounds for both susceptibility and immunity.

The k-equivalence class

In Lemma 4.1, we have shown that when $E_{\boldsymbol{R}}$ always contains a single profile, there always exists a information set in which the manipulator has an incentive to manipulate. With k-approval, we can extend this idea, since the order of the top k alternatives of some voter does not influence who wins the election. The same holds for the order of bottom $m - k$ alternatives. For example, when k is 3 it does not matter whether a voter votes $A \succ B \succ C \succ D \succ E$ or $C \succ A \succ B \succ E \succ D$. In both cases, A, B, and C receive one point, while

D and E get zero points. To make this intuition formal, we will introduce the notion of a k-*equivalence class* and generalize this idea to profiles.

Definition 4.2. *Given some preference order $R \in \mathcal{L}(\mathcal{X})$ and some integer k, we define the k-equivalence class $[R]_k$ of R as:*

$$[R]_k = \{R' \in \mathcal{L}(X) \,|\, top_k(R') = top_k(R)\}$$

Definition 4.3. *Given some profile $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^{n-1}$ and some integer k, we define the k-equivalence class of a profile $[\boldsymbol{R}]_k$ of $\boldsymbol{R}$ as:*

$$[\boldsymbol{R}]_k = \{\boldsymbol{R'} \in \mathcal{L}(\mathcal{X})^{n-1} \,|\, \textit{for all } i \in \{1, ..., n-1\} \textit{ it holds that } [R'_i]_k = [R''_i]_k\}$$

Now, we can use these definitions to show that knowing the equivalence class of all voters leads to the fact that there must exist some information set with which a manipulator has the incentive to manipulate. Similar to Lemma 4.1, we could have shown this using the Gibbard-Satterthwaite Theorem, but in this specific case it is more straightforward to give a direct proof.

Lemma 4.2. *If $n, m \geq 3$ and for every $E_{\boldsymbol{R}} \in \Sigma$ there is some $\boldsymbol{R'} \in E_{\boldsymbol{R}}$ with $E_{\boldsymbol{R}} \subseteq [\boldsymbol{R'}]_k$, then k-approval in combination with lexicographic tie-breaking is susceptible to Σ-manipulation.*

Proof. Take some arbitrary $n, m \geq 3$, let $\mathcal{X} = \{A, B, ..., Z\}$ with $|\mathcal{X}| = m$ and the ties be broken by $A \rhd B \rhd \cdots \rhd Z$. Now, suppose that for any $E_{\boldsymbol{R}} \in \Sigma$ we know that there is some $\boldsymbol{R'} \in E_{\boldsymbol{R}}$ it holds that $E_{\boldsymbol{R}} \subseteq [\boldsymbol{R'}]_k$. This implies that for any $\boldsymbol{R}$ with $E_{\boldsymbol{R}} \in \Sigma$, we know that all $\boldsymbol{R'} \in E_{\boldsymbol{R}}$ are in the same profile equivalence class.

First, suppose $k > 1$ and let F denote the k-approval rule. Now, we consider the $\boldsymbol{R}$ with $E_{\boldsymbol{R}} \in \Sigma$, such that for any $\boldsymbol{R'} \in E_{\boldsymbol{R}}$ we have that $A, B \in top_k(R'_i)$ and $Z \notin top_k(R_i)$ of any $R_i \in \boldsymbol{R'}$. Suppose the manipulator has a truthful profile $R : B \succ A \succ \cdots \succ Z$. Note that if the manipulator votes R, we know that in any $\boldsymbol{R'} \in E_{\boldsymbol{R}}$ both A and B will receive n points, and since $A \rhd B$, $F(\boldsymbol{R'}, R) = A$. However, if the manipulator votes R' which is equal to R but in which A and Z are swapped, in every $\boldsymbol{R'} \in E_{\boldsymbol{R}}$, we know that A receives $n-1$ points, while B still receives n points. Since B is higher than all other alternatives in the tie-breaking order, it holds that $F(\boldsymbol{R'}, R') = B$. Since $B \succ_R A$, the manipulator is better of voting R' than R and never worse off. Hence, the manipulator has an incentive to manipulate.

The case that $k = 1$ must be checked separately, this part of the proof can be found in the Appendix. $\square$

4.2.1 Susceptibility Results

In this section, we will show susceptibility results of k-approval. Some of the proofs can be found in the Appendix.

TOPBOT and TOP(t)

The idea behind the proofs of both TOPBOT and TOP(t) is that if the manipulator know some alternatives in the bottom k cannot win, it can be profitable to swap your lowest alternative in the top k with this alternative, to improve the outcome in some cases. Note that in the case of TOP(t), the extra assumption that $t > 0$ is needed to make sure the manipulator knows at least the top alternative. The proof of Theorem 4.2 can be found in the Appendix.

Theorem 4.1. *For all $m \geq 3$ and $n \geq 3$, k-approval in combination with lexicographic tie-breaking is susceptible to* TOPBOT-*manipulation.*

Proof. When $k = 1$, we have that for every $E_{\boldsymbol{R}} \in$ TOPBOT, there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $E_{\boldsymbol{R}} \subseteq [\boldsymbol{R}']_k$. Thus, from Lemma 4.2 we know that in this case k-approval in combination with lexicographic tie-breaking is susceptible to TOPBOT-manipulation.

Now, suppose $k > 1$. We now need to show that for every $m \geq 3$ and $n \geq 3$, we can find some true preference R for the manipulator and a partial profile $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}} \in$ TOPBOT and the manipulator has an incentive to manipulate. Take some arbitrary $m \geq 3$ and $n \geq 3$ and let $\mathcal{X} = \{A, B, ..., Z\}$ with $|\mathcal{X}| = m$, where $A \rhd B \rhd \cdots \rhd Z$. Let F denote the k-approval rule with this tie-breaking scheme and suppose the manipulator has the following true preference: $R : B \succ \cdots \succ A \succ \cdots \succ Z$ such that for all $C \succeq_R A$ it holds that $C \in top_k(R)$ and for all Y such that $A \succ_R Y$ we have $Y \notin top_k(R)$. Since we assumed $k > 1$, such a preference always exists.

Now, let the partial profile $\boldsymbol{R}$ consists of only voters with the partial preference $R_1 : A \succ \mathcal{X} \setminus \{A, Z\} \succ Z$. It clearly holds that $E_{\boldsymbol{R}} \in$ TOPBOT. Consider the situation in which the manipulator uses the untruthful ballot R' which is equal to R except that A and Z are switched from position, so the manipulator gives Z one point instead of A. Now, we will show that (1) there exists a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$ and (2) that there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

(1) We can extend $\boldsymbol{R}$ to $\boldsymbol{R}^*$ as follows: we extend R_1 to R_1' such that $B \in top_k(R_1')$. Since we assumed $k > 1$, this is always possible. The rest of the profile we can fix randomly as it does not matter for the proof. Now, if the manipulator votes R both A and B have n points, which is the maximal amount of points an alternative could get. Since $A \rhd X$ for all $X \in \mathcal{X}$, $F(\boldsymbol{R}^*, R) = A$. If the manipulator votes R', however, alternative A will have $(n-1)$ points while B still has n points. Since B now has the maximal amount of points and $B \rhd X$ for all $X \in \mathcal{X} \setminus \{A\}$ we must have that $F(\boldsymbol{R}^*, R') = B$. Since $B \succ_R A$, we have found an $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$.

(2) First note that when the manipulator votes R', we know that for all $X \in \mathcal{X}$ with $A \succ_R X$ that for every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ the amount of points X gets can never be more than $(n-1)$. For all $X \neq Z$ this is trivial, and Z has 1 point which is less than $n-1$ since we assumed $n \geq 3$. Now, since alternative A always $(n-1)$ points in every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ and

we have that $A \rhd X$, we can never have that $F(\boldsymbol{R}', R') = X$ for any X with $A \succ_R X$ and any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ ($\triangle$).

Now, note that if the manipulator votes R, A receives the maximal amount of points (n) in every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ and since $A \rhd Y$ for all $Y \in \mathcal{X} \setminus \{A\}$, we have that $F(\boldsymbol{R}', R) = A$, for all $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. Now, from ($\triangle$) we know that it can never be the case that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$ for any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. $\square$

Theorem 4.2. *For all $t > 0$, $m \geq 3$ and $n \geq 3$, k-approval in combination with lexicographic tie-breaking is susceptible to TOP(t)-manipulation.*

BOT(t)

As said above, when the manipulator only knows the BOT(t), the situation is slightly different. The idea behind the proofs in this subsection are that a manipulator has enough information to know that either all their top k elements or all their bottom $m-k$ elements cannot win. We first show that the Theorem holds when $t \geq k$ in the Lemma below, and use this Lemma in Theorem 4.3 to show the susceptibility results for k-approval. The proof of Lemma 4.3 can be found in the Appendix

Lemma 4.3. *When $m \geq 3$, $n \geq 3$ then for any $t \geq k$, k-approval in combination with lexicographic tie-breaking is susceptible to BOT(t)-manipulation.*

Theorem 4.3. *When $m \geq 3$, $n \geq 3$ and $t > 0$ then if either $t \geq k$ or $t > m-k-t$, k-approval in combination with lexicographic tie-breaking is susceptible to BOT(t)-manipulation.*

Proof. We take some arbitrary $m, n \geq 3$ and $t > 0$. If $t \geq k$, we know from Lemma 4.3 that the Theorem holds. Moreover, if $t \geq m-k$, we know that $E_{\boldsymbol{R}} \subseteq [\boldsymbol{R}']_k$ for any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, so from Lemma 4.2, we again know the Theorem holds.

So, we suppose that $t < k$, $t < m-k$, and $t > m-k-t$ and let $\mathcal{X} = \{A, ..., Z\}$ be the set of m alternatives with the usual tie-breaking $A \rhd \cdots \rhd Z$. Now, we can distinguish two cases, either $k < m-k$ or $k \geq m-k$:

Case 1: Suppose $k < m-k$. Now, similar to Lemma 4.3, we let the manipulator have a truthful preference $R : Z \succ \cdots \succ A$ and let X be the highest ranked alternative that is not in the top k of R. Let $\boldsymbol{R}$ be defined such that the top k is evenly distributed over the bot t of every R_i. Since $t < k$ this is always possible.

Note that if we take an arbitrary $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, it can never be the case that $F(\boldsymbol{R}', R) = D$ for any $D \in top_k(R)$. In the worst case, all $m-k$ bottom alternatives are divided equally over the $m-k-t$ positions of the voters. However, since $k < m-k$ and $t > m-k-t$ there must always be some $C \notin top_k(R)$ that is in the top k of $\boldsymbol{R}'$ more often than any $D \in top_k(R)$. And since $C \rhd D$ by construction, this C will always beat D in $(\boldsymbol{R}', R)$.

Now, let R' be equal to R but in which X receives a point instead of Z. With a similar argument as in Lemma 4.3 the observation that no argument in the top k of R can win, implies there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$. Moreover, we can similarly define a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') = X$ and $F(\boldsymbol{R}^*, R) = A$. So, the manipulator has an incentive to manipulate with R'.

Case 2: Suppose $k \geq m - k$. Now, we let $R = B \succ \cdots A \succ \cdots \succ Z$ which is almost similar to $\rhd$, except that A is the lowest ranked alternative in the top k instead of the highest. Moreover, we let $\boldsymbol{R}$ be the preference order such that all bottom alternatives of R are divided equally among the bot t alternatives of $\boldsymbol{R}$.

With a similar argument as in Case 1, we know that there cannot be any $D \notin top_k(R)$ such that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}, R') = D$. This because we know that $t > m - k - t$ and $m - k < k$, so even if we equally divide all top k elements among the $m - k - t$ positions, there must be some $C \in top_k(R)$ that is more in the top k than any $D \notin top_k(R)$ and since $C \rhd D$ by construction, $F(\boldsymbol{R}, R') = D$ can never happen.

Towards a contradiction assume that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = E$ and $F(\boldsymbol{R}', R') = E'$ and $E \succ_R E'$. From the structure of k-approval this means that either $E \in top_k(R)$ and $E \notin top_k(R')$ or $E' \in top_k(R')$ and $E' \notin top_k(R)$. We just showed that the second case can never happen. From the first, we have that $E = X$ but this again implies that $E' \notin top_k(R)$, but than E' cannot win when the manipulator votes R'. Contradiction. So, such an $\boldsymbol{R}'$ can never exist.

What is left to show is that there exist a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. From $k \geq m - k$ and $m \geq 3$ it follows that $k \geq 2$. So, we let $\boldsymbol{R}^*$ be the profile in which both B and A are in the top k of every voter. Clearly, we have that $F(\boldsymbol{R}^*, R) = A$ and $F(\boldsymbol{R}^*, R') = B$. □

COMPL(t)

What is left, is to show that with COMPL(t) a manipulator also has an incentive to manipulate. Note that the proof of this theorem is slightly strange, since it shows that even if all voters completely agree with the manipulator, they still have an incentive to manipulate. This is due to the fact that a voter cannot distinguish in their ballot between elements in their top k, since they all receive the same amount of points.

Theorem 4.4. *If $t \geq \lceil \frac{n-1}{2} \rceil$, $m \geq 3$ and $n \geq 3$, k-approval in combination with lexicographic tie-breaking is susceptible to COMPL(t)-manipulation.*

Proof. Take some arbitrary $m \geq 3$ and $n \geq 3$ and let $\mathcal{X} = \{A, B, ..., Z\}$ with $|\mathcal{X}| = m$, where $A \rhd B \rhd \cdots \rhd Z$ and let F denote k-approval with this tie-breaking scheme. Moreover, let R and R' be similar as in Lemma 4.1. Now, we let $\boldsymbol{R}$ be such that t voters vote R, the true preference of the manipulator and show that (1) there exists a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$ and (2) that there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

(1) We can extend $\boldsymbol{R}$ to $\boldsymbol{R}^*$ by extending the other $n - 1 - t$ preferences also to R. If the manipulator also votes R, all alternatives that are in the top k of R receive n points, and since $A \rhd X$ for all other X, we have that $F(\boldsymbol{R}, R) = A$. If the manipulator votes

R', however, A receives $n-1$ points. Since $B \rhd X$ for all $X \in \mathcal{X} \setminus \{A\}$, it now holds that $F(\boldsymbol{R'}, R') = B$.

(2) Take an arbitrary $\boldsymbol{R'} \in E_{\boldsymbol{R}}$. If B is in the top k of the manipulator, we know that B has at least $t+1$ points. Moreover, every alternative X with $A \succ_R X$ has at most $n-t-1$ points in $\boldsymbol{R'}$. Clearly we cannot have that $F(\boldsymbol{R'}, R) = X$, since we assumed $t \geq \lceil \frac{n-1}{2} \rceil$. If the manipulator votes R', the same holds for all $X \neq Z$ with $A \succ_R X$. But, Z receives at most $n-t$ points in $(\boldsymbol{R'}, R')$ and since we assumed that $t \geq \lceil \frac{n-1}{2} \rceil$ and $B \rhd Z$ we can also not have that $F(\boldsymbol{R'}, R') = Z$. So, $F(\boldsymbol{R'}, R) \neq X$ and $F(\boldsymbol{R'}, R') \neq X$ for all X with $A \succ_R X$ and $\boldsymbol{R'} \in E_{\boldsymbol{R}}$ $(\triangle)$.

Now, suppose $F(\boldsymbol{R'}, R) = A$, from $(\triangle)$ it follows that $F(\boldsymbol{R'}, R') \succeq_R A$. Otherwise from $(\triangle)$ it must hold that $F(\boldsymbol{R'}, R) = C$ for some $C \succ_R A$. Then, since C receives the same amount of points from the manipulator when voting R' or R, and by $(\triangle)$ Z cannot win, it must hold that $F(\boldsymbol{R'}, R') = C$. So, there is no $\boldsymbol{R'} \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R'}, R) \succ_R F(\boldsymbol{R'}, R')$. □

4.2.2 Immunity Results

In this subsection, we show when a manipulator never has an incentive to manipulate. Reijngoud and Endriss (2012) showed that when $n \geq 2m-2$ any positional scoring rule is susceptible to manipulation when the poll gives the manipulator gives zero information. We slightly improve their result here for TOP(t). The proof of this theorem can be found in the Appendix.

Theorem 4.5. *When $t = 0$ and both $n \geq 3$ and $n \geq \lceil \frac{m-2}{m-k} \rceil$, k-approval in combination with lexicographic tie-breaking is immune to TOP(t)-manipulation.*

These results also hold for COMPL(t) and BOT(t), but we will show below that we can prove stricter results for these two structures. In the case of BOT(t), we know from the two bounds that at least one alternative in the top k and one in the bottom $m-k$ of the manipulator is a potential winner. This helps to show that a manipulator can never improve their vote in one situation without making the outcome worse in another. The proof of this Theorem is similar to the susceptibility proof of BOT(t) above and can be found in the Appendix.

Theorem 4.6. *For all $m \geq 3$ and $n \geq 3$, if both*

$$\left\lfloor \frac{(n-1) \cdot t}{k} \right\rfloor < \left\lfloor \frac{(n-1) \cdot (m-k-t)}{m-k} \right\rfloor$$
$$\left\lfloor \frac{(n-1) \cdot t}{m-k} \right\rfloor < \left\lfloor \frac{(n-1) \cdot (m-k-t)}{k} \right\rfloor$$

then k-approval in combination with lexicographic tie-breaking is immune to BOT(t)-manipulation.

Finally, we show that with COMPL(t) that if you know the votes of less than half of the other voters and k is small enough, the manipulator cannot have an incentive to manipulate. This proof is based on a proof in Damanik et al. (2017).

Theorem 4.7. *For all $m \geq 3$ and $n \geq 3$, if $t < \lceil \frac{n-1}{2} \rceil$ and $k \leq \frac{m}{2}$ then k-approval in combination with lexicographic tie-breaking is immune to COMPL(t)-manipulation.*

Proof. Take some arbitrary $m, n \geq 3$ and suppose $t < \lceil \frac{n-1}{2} \rceil$ and $k \leq \frac{m}{2}$. Let F denote k-approval in combination with lexicographic tie-breaking. Now, towards a contradiction, suppose that F is susceptible to COMPL(t)-manipulation. This means that there is some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}} \in$ COMPL(t) and two $R, R' \in \mathcal{L}(\mathcal{X})$ such that a manipulator with a true preference R has an incentive to manipulate with R'. First, note that since there exists some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that voting R or R' gives a different result, it must hold that there exists a pair $A, B \in \mathcal{X}$ such that $A \in top_k(R)$ and $A \notin top_k(R')$, and $B \notin top_k(R)$ and $B \in top_k(R')$. So, it holds that $A \succ_R B$. Since k-approval is anonymous, w.l.o.g. we may assume that for all R_i with $i \leq t$ it holds $R_i \in \mathcal{L}(\mathcal{X})$ and for all R_i with $i > t$ it holds that $R_i = \emptyset$. Now, we can distinguish between two cases:

Case 1: Suppose $k > 1$. We will show that this implies that there must also be some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = A$ and $F(\boldsymbol{R}', R') = B$. Since $A \succ_R B$, this is a contradiction with the assumption that k-approval is susceptible to manipulation, and thus proves this theorem.

Now, we can extend $\boldsymbol{R}$ to $\boldsymbol{R}'$ as follows: For every R_i with $i \leq t$, we extend R_{t+i} to R'_{t+i} such that A and B are both in $top_k(R'_{t+i})$ if either $A, B \in top_k(R_i)$ or $A, B \notin top_k(R_i)$. If either A or B is in the top k of R_i, only the other alternative is in the top k of R'_{t+i}. Moreover, all other alternatives that are in the top k of R'_{t+i} are not in the top k of R_i. This is always possible since we assumed $k \leq \frac{m}{2}$.

Now, since $t < \lceil \frac{n-1}{2} \rceil$, there is at least one R_j left that is not yet extended. Note that until now A and B have an equal amount of points, have both at least t points, and there can be at most k alternatives that also have t points (since to get t points these alternatives should have been in $top_k(R_i)$ for every $i \leq t$ or in the $top_k(R_{i'})$ for every $t < i' < 2t$). Hence, we extend every R_j with $j > 2t$ such that $A, B \in top_k(R_j)$ and every alternative that also has t points not in the top k. Again, since $k \leq \frac{m}{2}$ this is always possible.

By construction, we now know that A and B have an equal amount of points in $\boldsymbol{R}'$ and there is no other alternative that has an equal or more points than both A and B. Thus, is must hold that $F(\boldsymbol{R}', R) = A$ and $F(\boldsymbol{R}', R') = B$. Contradiction!

Case 2: Suppose $k = 1$. Since we assumed that k-approval is susceptible to COMPL(t)-manipulation, we know that there must be some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') = B$ and $F(\boldsymbol{R}^*, R) = C$ with $B \succ_R C$. This implies that either $top(R') = B$ or $top(R) = C$, but since $B \succ_R C$, $top(R) = C$ cannot be the case, so $top(R') = B$ must hold. We also know that $top(R) \neq top(R')$ and we will call $top(R) = A$.

Now, let X be the alternative that is on top most often in R_i for $i \leq t$ (except if that is A, than the second most often). Now, we show that there is some $\boldsymbol{R}'$ such that $F(\boldsymbol{R}', R') = X$

while $F(\boldsymbol{R}', R) = A$. We extend $\boldsymbol{R}$ to $\boldsymbol{R}'$ in a similar way as when $k > 1$. For every R_i with $i \leq t$ such that X on top, we extend R_{t+i} such that A is on top, and for every R_i such that X is on top we extend R_{t+i} to R'_{t+i} such that A is on top. For every R_j with $j > t$ that are not extended yet we alternate between X and A such that If $X \rhd A$, A and X tie in $\boldsymbol{R}'$ and if $A \rhd X$, we make sure that X has 1 more point than A in $\boldsymbol{R}'$. This is always possible, since after extending R_{t+t}, it holds that A and X tie, and now there is at least one R_j left that is not extended yet. If we need to extend one R_j with $j > 2t$ such that X and A should not be on top we can take an alternative that is least on top in R_i for $i \leq t$. Now, by construction we created $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = A$ and $F(\boldsymbol{R}', R') = X$. Since $A \succ_R X$, we again get a contradiction.

Hence, for all $m \geq 3$ and $n \geq 3$, if $t < \lceil \frac{n-1}{2} \rceil$ and $k \leq \frac{m}{2}$ then k-approval in combination with lexicographic tie-breaking is immune to COMPL(t)-manipulation. $\square$

4.3 Borda

In this section, we show susceptibility and immunity results regarding Borda. A summary of these results can be found in Table 4.2. Similarly to k-approval, we can see that the results of COMPL(t) differ from the results of TOP(t), BOT(t), and TOPBOT. Which shows that everything about some of the voters leads to different results than knowing something about all of the voters. Contrary to the case of k-approval, there seems to be no big difference between knowing the top or the bottom alternatives of each voter.

	Susceptible	Immune
TOPBOT	$m, n \geq 3$	-
TOP(t)	$m \geq 4, n \geq 3$, and $t > 1$	$t = 0$ and $m, n \geq 3$
BOT(t)	$m \geq 4, n \geq 3$, and $t > 0$	$t = 0$ and $m, n \geq 3$
COMPL(t)	$m \geq 4$, $n \geq 3$, and $t > \lceil \frac{n-1}{2} \rceil$	$t \leq \lceil \frac{n-3}{2} \rceil$

Table 4.2: Results of Borda.

4.3.1 Susceptibility Results

We will first show some susceptibility results regarding Borda. Again, the proofs of several theorems can be found in the Appendix.

TOPBOT, BOT(t) and COMPL(t)

The intuition behind the TOPBOT, BOT(t) and COMPL(t) proofs is that the manipulator knows their top alternative cannot win. This makes it safe to swap the first and second alternative. The proof of TOPBOT is given below, the ones of BOT(t) and COMPL(t) can be found in the Appendix.

Theorem 4.8. *For all $m \geq 3$ and $n \geq 3$, Borda in combination with lexicographic tie-breaking is susceptible to* TOPBOT-*manipulation.*

Proof. We need to show that for every $m \geq 3$ and $n \geq 3$, there exists some partial profile $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}} \in$ TOPBOT, and a manipulator with a true preference R that has an incentive to manipulate. First, suppose that $m = 3$, this implies that for every $E_{\boldsymbol{R}} \in$ TOPBOT, we have $|E_{\boldsymbol{R}}| = 1$. From Lemma 4.1, we know the theorem holds.

So, we take some arbitrary $m \geq 4$ and $n \geq 3$ and let $\mathcal{X} = \{A, B, C, ..., Z\}$ with $|\mathcal{X}| = m$, where $A \triangleright B \triangleright C \triangleright \cdots \triangleright Z$. Now, let F denote the Borda voting rule combined with $\triangleright$ and let $R = \triangleright$. The partial profile $\boldsymbol{R}$ be structured as follows:

- $\left\lceil \frac{1}{2}(n-2) \right\rceil$ voters with the partial preference order $R_1 : C \succ \mathcal{X} \setminus \{A, C\} \succ A$.
- One voter with partial preference order $R_2 : Z \succ \mathcal{X} \setminus \{A, Z\} \succ A$
- $\left\lfloor \frac{1}{2}(n-2) \right\rfloor$ voters with partial preference order $R_3 : B \succ \mathcal{X} \setminus \{A, B\} \succ A$. Note that if $n = 3$ there are no such voters, since $\left\lfloor \frac{1}{2} \right\rfloor = 0$.

Consider the following non-truthful preference of the manipulator $R' : B \succ A \succ C \succ \cdots \succ Z$. We will first show that there exists a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. We extend $\boldsymbol{R}$ to the full profile $\boldsymbol{R}^*$ as follows:

- If n is odd, extend every partial preference R_1 to the full preference $R_1' : C \succ B \succ ... \succ Z \succ A$. If n is even, extend one voter with the partial preference R_1 to the full preference $R_1'' : C \succ Z \succ B \succ ... \succ A$ and the rest of the partial preferences R_1 to R_1'.
- extend the partial preference R_2 to the full preference $R_2' : Z \succ C \succ B \succ \cdots \succ A$
- extend every partial preference R_3 to the full preference $R_3' : B \succ C \succ \cdots \succ Z \succ A$

It is easy to check that given either R or R', alternatives B and C have more points than all other alternatives, so either B or C wins. Now if n is odd and we do not consider the manipulator, alternative B has $\lceil \frac{1}{2}(n-2) \rceil \cdot (m-2) + (m-3) + \lfloor \frac{1}{2} \cdot (n-2) \rfloor \cdot (m-1)$ points and C has $\lceil \frac{1}{2}(n-2) \rceil \cdot (m-1) + (m-2) + \lfloor \frac{1}{2} \cdot (n-2) \rfloor \cdot (m-2)$ points. Because n is odd, it holds that $\lceil \frac{1}{2}(n-2) \rceil - \lfloor \frac{1}{2}(n-2) \rfloor = 1$ and thus without the vote of the manipulator, C has 2 points more than B. So, we get that $F(\boldsymbol{R}^*, R) = C$. But, since $B \triangleright C$ we get that $F(\boldsymbol{R}^*, R') = B$. In the case that n is even, we also know that without the ballot of the manipulator, C has two points more than B. So similarly to when n is odd, we get that $F(\boldsymbol{R}^*, R) = C$ and $F(\boldsymbol{R}^*, R') = B$. Since $B \succ_1 C$, we have shown that there is a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$.

What is left to show is that no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$. First note, that for any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, we cannot have that $F(\boldsymbol{R}', R) = A$, since C always has more Borda points than A. If there is some $\boldsymbol{R}'$ such that $F(\boldsymbol{R}', R) = B$, we must have that $F(\boldsymbol{R}', R') = B$, since B has one point more in this profile, so if it had the most Borda points, it will still have the most Borda points. Finally, if $F(\boldsymbol{R}, R) = Y$ for any $Y \in \mathcal{X} \setminus \{A, B\}$, we must have that either $F(\boldsymbol{R}', R) = Y$ or $F(\boldsymbol{R}', R') = B$, since the amount of Borda points for Y is the same and B is the only alternative that got more points. Since $B \succ_R Y$ for all $Y \in \mathcal{X} \setminus \{A, B\}$, there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$. □

Theorem 4.9. *If $t > 0$, $m \geq 4$ and $n \geq 3$, Borda in combination with lexicographic tie-breaking is susceptible to BOT(t)-manipulation.*

Theorem 4.10. *If $t \geq \lceil \frac{n-1}{2} \rceil$, $m \geq 4$ and $n \geq 3$, Borda in combination with lexicographic tie-breaking is susceptible to COMPL(t)-manipulation.*

TOP(t)

The situation of TOP(t) is slighly more complicated. If t is one, it this depends on the difference between m and n and the tie-breaking order. We left this situation out of consideration. When $t > 1$ and m and n are large enough, you can always manipulate. This is because there are situations in which you know that two alternatives are very close and one of the two must win.

Theorem 4.11. *If $t > 1$, $m \geq 4$ and $n \geq 3$, Borda in combination with lexicographic tie-breaking is susceptible to TOP(t)-manipulation*

Proof. Take any $t > 1$, $m \geq 4$ and $n \geq 3$, and let $\mathcal{X} = \{A, B, C, D...\}$ with $|\mathcal{X}| = m$ and $A \rhd B \rhd C \rhd D \rhd \cdots$. Furthermore let F denote k-approval with this tie-breaking scheme.

First, suppose $t = 1$ and consider the case in which the manipulator has the true preference $R : A \succ B \succ C \succ \cdots$ and wants to manipulate with preference $R' : A \succ C \succ B \cdots$, that is swapping B and C. First, we can define a partial order $\boldsymbol{R}$ as follows:

- 1 voter votes $R_1 : B \succ C \succ \mathcal{X} \setminus \{B, C\}$
- $\lfloor \frac{n-2}{2} \rfloor$ vote $R_2 : B \succ A \succ \mathcal{X} \setminus \{A, B\}$
- $\lceil \frac{n-2}{2} \rceil$ vote $R_3 : A \succ B \succ \mathcal{X} \setminus \{A, B\}$

Clearly, $E_{\boldsymbol{R}} \in$ TOP(t). Now, we will show that (1) there exists some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$ and (2) there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}^*, R')$.

(1) We extend $\boldsymbol{R}$ to $\boldsymbol{R}^*$ by extending R_1 to $R_1' : B \succ C \succ D \succ A \succ \cdots$, when n is odd and to $R_1'' : B \succ C \succ A \succ D \succ \cdots$ when n is even. The ballots of the other voters can be extend arbitrarily. Now, it is easy to check that when the manipulator votes R', A and B exactly tie and all other alternatives have less points. So since $A \rhd B$, we get that $F(\boldsymbol{R}^*, R') = A$, while $F(\boldsymbol{R}^*, R) = B$. So (1) holds.

(2) If we take some arbitrary $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, we know that when the manipulator votes either R or R', B always has as least as many points as any $X \in \mathcal{X} \setminus \{A\}$. This means that if $F(\boldsymbol{R}', R) \neq F(\boldsymbol{R}', R')$, this can only happen when in the one case A is elected and the other case B. If $F(\boldsymbol{R}', R) = A$ however, then it must also hold that $F(\boldsymbol{R}', R') = A$ since A has an equal amount of points and C can never beat A. Hence, if $F(\boldsymbol{R}', R) \neq F(\boldsymbol{R}', R')$ then it must hold that $F(\boldsymbol{R}', R) = B$ and $F(\boldsymbol{R}', R') = A$. Since $A \succ_R B$, it cannot be the case that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

Now, for any $t > 2$ there must exist some $\boldsymbol{R}'' \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}''} \in \text{TOP}(t)$ and $\boldsymbol{R}^* \in E_{\boldsymbol{R}''}$. Since it also holds that $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ and $t > 2$, we know that $E_{\boldsymbol{R}''} \subseteq E_{\boldsymbol{R}}$. So, in this case we also know that the manipulator has an incentive to manipulate. $\square$

4.3.2 Immunity Results

Below we show the immunity results for Borda. In the case of COMPL(t), if t is small enough, every alternative can still win. The same holds when $t = 0$ in the case of BOT(t) and TOP(t). The proof of COMPL(t) can be found below, the other in the Appendix. Morever, the proof of COMPL(t) is based on a proof in Damanik et al. (2017).

Theorem 4.12. *For all $m \geq 3$ and $n \geq 3$, if $t \leq \lceil \frac{n-3}{2} \rceil$, then Borda in combination with lexicographic tie-breaking is immune to COMPL(t)-manipulation.*

Proof. Take some arbitrary $m, n \geq 3$ and suppose $t \leq \lceil \frac{n-3}{2} \rceil$. Now, towards a contradiction, suppose that Borda in combination with lexicographic tie-breaking is susceptible to COMPL(t)-manipulation. This means that there is some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}} \in COMPL(t)$ and two $R, R' \in \mathcal{L}(\mathcal{X})$ such that a manipulator with a true preference R has an incentive to manipulate with R'. First, note that this implies there exists some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that voting R or R' gives a different result. This means that there exists a pair $A, B \in \mathcal{X}$ such that $A \succ_R B$ and $B \succ_{R'} A$. Since Borda is anonymous, we may assume without loss of generalization that for all R_i with $i \leq t$ it holds $R_i \in \mathcal{L}(\mathcal{X})$ and for all R_i with $i > t$ it holds that $R_i = \emptyset$. Now, we will show that there must exists some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = A$ and $F(\boldsymbol{R}', R') = B$, which is a contradiction with the assumption that Borda is susceptible to manipulation and proves the theorem.

We will extend $\boldsymbol{R}$ to such $\boldsymbol{R}'$ as follows. If n is odd, for every $i \leq t$, we extend R_{t+i} to R'_{t+i} which is defined as the opposite of R_i. Now we have an even number (and at least 2) R_j left with $j > 2t$. Here we alternative between extending it to R'_j with $A \succ B \succ \cdots$ and $R''_j : B \succ A \succ \cdots$ such that the alternatives below B and A are in reversed order in R''_j compared to R'_j. Note that in R'_1 till R'_{2t}, all alternatives have exactly the same amount of points and in all R'_j with $j > 2t$, A and B receive $2m - 1$ points per two R'_j, while all other alternative only receive $m - 3$ points per 2 R'_js.

This means, that no matter how many points an alternative receives in R or R', the only two alternatives that can win are A or B, because, for any $l > 0$:

$$l(2m - 1) > l(m - 3) + (m - 1)$$

Now, since by construction A and B receive exactly the same amount of points in $\boldsymbol{R'}$ we know since $A \succ_R B$ and $B \succ_{R'} A$ that $F(\boldsymbol{R'}, R) = A$ and $F(\boldsymbol{R'}, R') = B$.

If n is even, we do the exact same thing as when n is odd, but then we have R_{n-1} left. Now, if $A \triangleright B$ we extend this profile to $B \succ A \succ \cdots$ and otherwise to $A \succ B \succ \cdots$. Now, with a similar argument as when n is odd, we know that $F(\boldsymbol{R'}, R) = A$ and $F(\boldsymbol{R'}, R') = B$. □

Theorem 4.13. *When $t = 0$ and both $n, m \geq 3$ then Borda in combination with lexicographic tie-breaking is immune to both TOP(t)-manipulation and BOT(t)-manipulation.*

4.4 Copeland

The final voting rule we consider is Copeland. The results can be found in Table 4.3. Here, we again see that the difference between knowing everything about some of the voters leads to different results than knowing something about all of the voters.

	Susceptible	Immune
TOPBOT	$m, n \geq 3$	-
TOP(t)	$m, n \geq 3$ and $t > 1$	$t = 0$ and $m, n \geq 3$
BOT(t)	$m, n \geq 3$ and $t > 1$	$t = 0$ and $m, n \geq 3$
COMPL(t)	$m, n \geq 3$ and $t > \frac{n-1}{2}$	$t < \lceil \frac{n-3}{2} \rceil$

Table 4.3: Results of Copeland

4.4.1 Susceptibility Results

We will show that with TOPBOT and TOP(t) and BOT(t) with $t > 1$ the manipulator has an incentive to manipulate. The structure of the three proofs is similar, in each case the manipulator swaps their two top alternatives which makes their second alternative beat the third in some cases. The proofs of TOP(t) and BOT(t) can be found in the Appendix.

Theorem 4.14. *When $m, n \geq 3$, Copeland in combination with lexicographic tie-breaking is susceptible to TOPBOT-manipulation.*

Proof. First note that if $m = 3$, we have that $|E_{\boldsymbol{R}}| = 1$ for any $E_{\boldsymbol{R}} \in$ TOPBOT. Since Copeland with any lexicographic tie-breaking rule is surjective, we know from Lemma 4.1, that the theorem holds. So, we take an arbitrary $m \geq 4$, $n \geq 3$ where n is odd and let $\mathcal{X} = \{A, B, ..., Y, Z\}$ with $|\mathcal{X}| = m$, where $A \triangleright B \triangleright \cdots \triangleright Y \triangleright Z$. We let the true preference of the manipulator be the opposite of $\triangleright$, that is $R = Z \succ Y \succ \cdots \succ B \succ A$. Now, we will assume that n is odd, the case that n is even can be found in the Appendix.

We let the partial profile $\boldsymbol{R}$ be structured as follows:

- $\frac{1}{2}(n-1)$ voters with the partial preference order $R_1 : Y \succ \mathcal{X} \setminus \{Y, Z\} \succ Z$, and
- $\frac{1}{2}(n-1)$ voters with the partial preference order $R_2 : A \succ \mathcal{X} \setminus \{A, B\} \succ B$.

Consider the non-truthful preference order $R' : Y \succ Z \succ \cdots \succ B \succ A$, this is the preference order where the first two alternatives are swapped and the rest of the profile remains the same. We will show that (1) there exists a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$ and (2) that there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

(1) We extend the partial profile $\boldsymbol{R}$ to the full profile $\boldsymbol{R}^*$ by extending

- R_1 to $R_1' : Y \succ A \succ B \succ \cdots \succ Z$
- R_2 to $R_2' : A \succ Z \succ Y \succ \cdots \succ B$

Now suppose the manipulator votes R. Then if we consider all pairwise majority contests, Z beats everything except A, Y beats everything except Z, and A beats everything except Y. So A, Y and Z have $m-3$ points (they beat all $m-2$ other alternatives and lose one pairwise majority contest), while other alternatives have at most $m-7$ points (they beat at most $m-4$ alternatives and always lose from 3 alternatives). Since $A \rhd Y$ and $A \rhd Z$ we have that $F(\boldsymbol{R^*}, R) = A$. Whenever the manipulator votes R' instead, Y is no longer beaten by Z in a pairwise majority contest and Y and becomes the Condorcet winner. Since Copeland is a Condorcet extension, we now haven that $F(\boldsymbol{R}^*, R') = Y$ and since $Y \succ_R A$, we know (1) holds.

(2) Let's take an arbitrary $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. If the manipulator votes R' we know Y is Condorcet extension, so we must have that $F(\boldsymbol{R}', R') = Y$. This means that the only way we could have that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$ is when $F(\boldsymbol{R}', R) = Z$. Now consider the case where the manipulator votes R. Now, Y is above all alternatives except Z in more than half of the votes. This means that Y must beat every alternative besides Z in a pairwise majority contest. Z, on the other hand, must be beaten by A in a pairwise majority contest since A is above Z in all votes but one. This means that Z has at most the same amount of points as Y. Since $Y \rhd Z$, we can never have that $F(\boldsymbol{R}', R) = Z$. $\square$

Theorem 4.15. *When $m, n \geq 3$ and $t > 1$, then Copeland in combination with lexicographic tie-breaking is susceptible to TOP(t)-manipulation.*

Theorem 4.16. *When $m, n \geq 3$ and $t > 1$, then Copeland in combination with lexicographic tie-breaking is susceptible to BOT(t)-manipulation.*

COMPL(t)

Theorem 4.17. *When $m, n \geq 3$ and $t > \frac{n-1}{2}$, then Copeland in combination with lexicographic tie-breaking is susceptible to COMPL(t)-manipulation.*

Proof. We take any $m, n \geq 3$ and $t > \frac{n-1}{2}$ and let $\mathcal{X}$ be the set of m alternatives with the usual tie-breaking order $\rhd$. Now, let R, the true preference of the manipulator be the opposite of $\rhd$, that is $R = Z \succ Y \succ X \succ \cdots$. We suppose n is odd, the case for n is even can be found in the Appendix. Now, we let the partial profile $\boldsymbol{R}$ be defined as follows:

- $\frac{1}{2}(n-1)$ voters vote $X \succ Z \succ Y \succ \cdots$
- The rest of voters vote $Y \succ X \succ Z \succ \cdots$

Now, we first show that the manipulator has an incentive to manipulate with the untruthful preference $R' : Y \succ Z \succ X \succ \cdots$. Let $\mathbf{R}^*$ be such that the other voters also vote $Y \succ X \succ Z \succ \cdots$. Now, if the manipulator votes R, X beats everything besides Y, Y everything besides Z, and Z everything besides X. So since $X \rhd Y \rhd Z$, we have $F(\mathbf{R}^*, R) = X$. If the manipulator votes R' instead, Y is the Condorcet winner so $F(\mathbf{R}^*, R') = Y$.

What is left to show is that for any $\boldsymbol{R}'$, the manipulator is not worse off by voting R'. Now, first note that if there is any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$ it must hold that $F(\boldsymbol{R}', R) = Z$. However, since $t > \frac{n-1}{2}$, we know that X must beat everything besides Y in $\boldsymbol{R}'$, while Z at best beats everything besides X. Since $X \rhd Y$, this means we can never have $F(\boldsymbol{R}', R) = Z$, so such $\boldsymbol{R}'$ does not exist. $\square$

4.4.2 Immunity Results

Reijngoud and Endriss (2012) show that when n is odd and $n \geq 3$, every Condorcet extention is immune to manipulation when the manipulator receives no information from the polls. From this we can conclude that when $t = 0$ and $n \geq 3$, Copeland is immune to both BOT(t)-manipulation and TOP(t)-manipulation.

What is left to show is the case of COMPL(t). Similar to Borda and k-approval, this proof is based on a proof in (Damanik et al., 2017). The idea behind this proof is that any alternative can still be the Condorcet winner.

Theorem 4.18. *For all $m \geq 3$ and $n \geq 3$, if $t < \left\lceil \frac{n-3}{2} \right\rceil$ then Copeland in combination with lexicographic tie-breaking is immune to COMPL(t)-manipulation.*

Proof. Towards a contradiction suppose that we have an $n, m \geq 3$ and $t < \left\lceil \frac{n-3}{2} \right\rceil$ and there is some $E_{\boldsymbol{R}} \in$ COMPL(t) such that the manipulator has some true preference R and an incentive to manipulate with a non-truthful preference R'. This implies there must be some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\mathbf{R}^*, R)$. So, there must be some $A, B \in \mathcal{X}$ such that $A \succ_R B$ and $B \succ_{R'} A$. Now, we will show that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = A$ and $F(\boldsymbol{R}', R') = B$. This gives us a contradiction and proves the theorem.

Since Copeland is anonymous, w.l.o.g. we may assume that $\boldsymbol{R} = (R_1, ..., R_{n-1})$ such that R_i for every $i \leq t$ we have $R_i \in \mathcal{L}(\mathcal{X})$ and R_i for $i > t$ that $R_i = \emptyset$. Now, we can construct $\boldsymbol{R}'$ as follows: For every R_i for $i \leq t$, we extend R_{t+i} to $R_a : A \succ B \succ \cdots$ when $B \succ_{R_i} A$ and $R_b : B \succ A \succ \cdots$ otherwise. Now, we have at least two empty votes left, so we alternate between R_a and R_b. If n is even, we have one voter left. We extend this vote to R_b when $A \rhd B$, and R_a otherwise.

Now, A and B clearly beat all other alternatives in pairwise majority contest, no matter if the manipulator votes R or R'. Now, if n is odd, A is the Condorcet winner when the manipulator votes R, while B is the Condorcet winner when the manipulator votes R'. So, since $A \succ_R B$ we have that $F(\boldsymbol{R}', R) \succ F(\boldsymbol{R}', R') = B$. If n is even, we have by construction that $F(\boldsymbol{R}', R) \succ F(\boldsymbol{R}', R') = B$ also holds. □

Since the proof of the case that n is odd is based on Condorcet winners only, we can generalize it to the following result:

Corollary 4.1. *If $m \geq 3$ and $n \geq 3$, if $t < \lceil \frac{n-3}{2} \rceil$ and n is odd then then every Condorcet extension immune to COMPL(t)-manipulation.*

4.5 Discussion

The results of all three voting rules have some space between the lower bound and upper bounds for susceptibility and immunity. In other words, for some combinations of values of parameters, the question whether you can manipulate is still an open problem. Often, these situations depend heavily on the combination of the parameters and it will probably not be possible to show general results. Since in this chapter we want to show that you can still m@anipulate with these structures and there are differences between different structures, we believe that for our purposes some space between the bounds is not a major problem.

Moreover, in most proofs above, tie-breaking plays an important role. To a certain extent, we believe this is unavoidable since the profile in which one voter can change the outcome is per definition a situation in which at least two alternatives are both very close to each other. These are exactly the situations in which tie-breaking also plays an important role. However, all the proofs in this chapter are very general since they hold for every $n \geq q$ and $m \geq q'$ for some q and q'. For some specific m and n it is in most cases possible to give an information set for which the manipulator can manipulate without using the tie-breaking rule. However, since the information sets are slightly different for different values of m and n with the same structure and voting rules, the proofs that show susceptibility in general without tie-breaking will be certainly very long and complicated proofs that contain many case distinctions. It would be interesting for future work, however, if we could find similar results that depend less heavily on tie-breaking. We will go into this in some more depth in Chapter 7.

4.6 Conclusion

In this chapter, we have shown both susceptibility and immunity results for all structures in combination with three voting rules. On the one hand, we can see that partial information in the form of the structures we considered is no barrier against manipulation, since it is still possible to manipulate for some m, n, and t. In some cases, however, some knowledge about the preferences of others does not give the manipulator enough information to manipulate. We have seen that this is the case with COMPL(t), where the manipulator needs the information about roughly half of the voters before manipulation is possible. On the other hand, we found that that for k-approval TOP(t)-manipulation is possible when $t > 0$ and for Borda and Copeland, TOP(t)-manipulation is possible when $t > 1$. For BOT(t) we found in almost all cases similar constant values for t and TOPBOT-manipulation was also always possible. This shows that, at least in the case of the three voting rules we considered, knowing a little about the preferences of all voters differs from knowing everything about the preferences of some.

Chapter 5

The Computational Complexity of Manipulation

As said in the introduction, researchers looked at the complexity of several social choice problems, including manipulation. In this chapter, we will show that for k-approval, every structure $\Sigma \in$ STRUCT, the manipulator can decide whether to manipulate in polynomial time. We have generalized this proof to hold for all structures that are IP- and PI-transitive. Moreover, we will show that for other natural voting rules, such as Copeland and Borda, there is some evidence to believe that only with partial orders that are not IP- and PI-transitive it is difficult to manipulate. In other words, we believe it could be the case that for all structures that are both IP- and PI-transitive, manipulation with other natural voting rules besides k-approval can also be done in polynomial time.

5.1 Complexity Theory and Voting

As also said in Chapter 2, researchers have been looking at manipulation in terms of its computational complexity. This idea is based on the ideas posed by Bartholdi et al. (1989), who showed that with full information manipulation can be solved in polynomial time for most voting rules. In particular, research focuses on distinguishing problems that are easy to solve, that is solvable in polynomial time, from problems that are not. Most notably, problems that are NP-hard are not polynomial time solvable (Assuming P $\neq$ NP) [1]. For a problem to be in P or NP-hard, we need to define the manipulation problem in terms of a decision problem. This means the problem is formulated as a question with a yes or no answer. We will do this for manipulation below.

[1] The assumption that NP $\neq$ P is common in theoretical complexity theory. We assume the reader is familiar with the basic conjectures and notions in this field, if not see (Arora and Barak, 2007)

5.1.1 The Manipulation Problem

In this thesis, we will define the complexity problem of manipulation similar to the way we use manipulation in the rest of this thesis. This is also related to what manipulation refers to in the Gibbard-Satterthwaite Theorem. We say that an instance, consisting of a partial profile $\boldsymbol{R}$ and a true preference R, is a yes-instance of manipulation if and only if the manipulator has an incentive to manipulate with respect to $E_{\boldsymbol{R}}$:

Manipulation(F, Σ)	
Input:	A partial profile $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ (e.g. a partial order over $\mathcal{X}$ for each voter) such that $E_{\boldsymbol{R}} \in \Sigma$ and a true preference $R \in \mathcal{L}(\mathcal{X})$ of the manipulator.
Question:	Does there exist a non-truthful preference $R' \neq R \in \mathcal{L}(\mathcal{X})$ such that there is some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$ and there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$

Related problems

In the literature, there are several problems that are related to Manipulation with partial information. Some problems have a different name but are essentially the same, some problems have the same name but refer to another question. This makes it sometimes difficult to get an overview of what is solved and what is not. Conitzer et al. (2011) define the Dominating Manipulation problem, which is identical to our Manipulation problem when Σ contains the information sets of every partial order, i.e. $\Sigma = \{E_{\boldsymbol{R}} \,|\, \boldsymbol{R} \in \mathcal{P}(\mathcal{R})\}$. They also define the Domination problem, which is given some non-truthful preference R' the question whether the manipulator has an incentive to manipulate with this specific R'. Interesting to note is that both problems have the same complexity results for the voting rules they consider.

Dey et al. (2016), on the other hand, define Weak Manipulation as the question whether the manipulator(s) can make some given alternative C win by manipulating their vote. The reason this is such a different question than ours, is because they also consider manipulating by coalitions, by Conitzer et al. (2007) also referred to as the Constructive Coalition Manipulation problem. Constructive manipulation refers to the situation in which a coalition of manipulators tries to get some alternative X elected. Besides the number of manipulators, the conceptual difference between Coalitional Manipulation and Manipulation is that in the first case deciding which alternative is made to win or lose is not part of the problem but determined beforehand, while in the second case deciding which alternative will benefit from the manipulation is part of the problem the manipulator needs to solve. In their work, Dey et al. also prove results of Weak Manipulation for a coalition of size one, which boils down to the question whether one manipulator can make X win.

The Constructive Coalition Manipulation or Weak Manipulation problem is similar to the Possible Winner problem, which given some set of partial of votes and alternative X asks the questions whether there is a full extension of the partial vote in which X wins. The main conceptual difference between the Possible Winner problem and all manipulation problems described above is that in the case of manipulation the partial orders are seen as partial information the manipulator has about the preferences of the other voters, while in the other case, the partial order represents the chairs information about the votes (Conitzer et al., 2011).

Intuitively, we believe it seems that our Manipulation problem is at least as difficult as the Possible Winner problem. Conitzer et al. (2011) show that this is at least partly true. They prove that a NP-hard fragment of all instances of the Possible Winner problem can be reduced in polynomial time to the Dominating Manipulation problem for all voting rules who's winner can be calculated from the weighted majority graph[2] (that is, for any two profiles $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^n$ that have the same weighted majority graph, it holds that $F(\boldsymbol{R}) = F(\boldsymbol{R'})$).[3] Since Dominating Manipulation and Manipulation are the same, this means that Manipulation is at least as difficult as (a difficult) part of the Possible Winner problem.

5.2 Related Complexity Results

Researchers have shown that both the Possible Winner and the Manipulation problem are NP-hard for Borda if we consider all partial profiles $\boldsymbol{R}$ (Betzler and Dorn, 2010; Conitzer et al., 2011; Xia and Conitzer, 2011) For k-approval it is shown, again for all $\boldsymbol{R}$ that Possible Winner is NP-hard when $k \neq 1$ and $k \neq (m-1)$, for plurality and veto it is solvable in polynomial time (Xia and Conitzer, 2011; Betzler and Dorn, 2010). For Manipulation and k-approval, it is only shown that plurality and veto can be solved in polynomial time (Conitzer et al., 2011).

First, note that in all proofs above it is assumed that both m and n are unbounded. In contrast, when we assume the number of alternatives is fixed, every voting rule for which the winner can be calculated in polynomial time, both the Possible Winner and Manipulation problem can be calculated in polynomial time. This is because for every partial order $\boldsymbol{R}$ there are at most n^{m-1} full profiles that extend $\boldsymbol{R}$ and only $m!$ possible profiles the manipulator can manipulate with. This means that it is possible to brute-force check all situations for all possible preferences to manipulate in polynomial time (Conitzer et al., 2007).

Secondly, after inspecting the difference between our proof and the proofs given in the

[2] A weighted majority graph is a graph that contains a vertex for every alternative and there is an edge from alternative A to an alternative B if the majority of the voters prefers A over B. Every edge has a weight with the net difference between X and Y.

[3] Note that since Borda rule can be calculated from the weighted majority graph (See e.g. Fischer et al., 2016, p.97), this construction also holds for Borda, even though Conitzer et al. do not explicitly state this and provide a separate NP-hardness proof for Manipulation of Borda

literature we made an interesting observation. While the proof of MANIPULATION for k-approval cannot be generalized to all partial profiles, it does hold for all structures that are both PI- and IP-transitive (recall from Chapter 3 that this means that if $A \sim B$ and $B \succ C$ then $A \succ C$ (IP-transitive) and if $A \succ B$ and $B \sim C$ then $A \succ C$ (PI-transitive)). However, all the proofs of the results of problems that hold for all partial profiles make a polynomial reduction to a partial profile that is *not* IP- and PI-transitive. We will show this for the three voting rules below:

- For k-approval we considered the proofs of Conitzer et al. (2011) and Betzler and Dorn (2010) that both showed that POSSIBLE WINNER is NP-hard. Betzler and Dorn did a reduction from multicolored clique[4] to k-approval, Conitzer et al. from 3-SAT. In the first case, they created partial preferences with a structures in which there are multiple pairs such that the partial order contained $A \succ A'$ and $B \succ B'$ but no order between elements of different pairs. But this means we have that $A \succ A'$ and $A' \sim B$, but not $A \succ B$, so this structure is not PI-transitive, and since $B \succ B'$ but not $A' \succ B'$ also not PI-transitive. The same holds for the reduction from Conitzer et al., in which they have partial preferences with the exact same structure.

- For both the Manipulation proof of Borda, Conitzer et al. argue that the proof is omitted due to space constraint but is similar to the proof of the NP-hardness of the Possible Winner Problem (Conitzer et al., 2011, p. 641). In the Possible Winner Problem, Xia and Conitzer (2011) create a partial order $P' : \mathcal{X}_1 \succ A \succ \mathcal{X}_2 \succ B \succ \mathcal{X}_3$ with $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3 \subseteq \mathcal{X}$ and $\mathcal{X}_2 \neq \emptyset$. Now they define a partial preference $R \in \mathbf{R}$ as $P' \setminus (\{A\} \times (\mathcal{X}_2 \cup \{B\}))$. Now, we again have that for some $Y \in \mathcal{X}_2$, $A \sim Y$ and $Y \succ B$ but not $A \succ B$ and $Y \succ B$ and $A \sim B$ but not $A \succ B$. So this instance is also not PI- and IP-transitive.

- In the case of Copeland, Conitzer et al. (2011) do a polynomial reduction from possible winner to manipulation, but they take input instance of the possible winner problem as part of the input of manipulation. In the possible winner proof, Xia and Conitzer (2011) use a similar construction as in the case of Borda. In this case, the instances are also not IP- and PI-transitive.

This observation shows why complexity results of Borda and Copeland limited to the structures we have defined are far from straightforward, since there is no way to translate the ideas in the proofs in the literature to this case. Moreover, it poses the following conjecture:

> *Does the NP-hardness of manipulation and the possible winner problem for all natural voting rules only hold for classes of instances that are not PI- and IP-transitive?*

We believe it is interesting to investigate this in future work. We will go into this more in the discussion at the end of this chapter. We will first show the complexity results that we found: a polynomial-time algorithm for manipulation with k-approval.

[4] The vertices of a graph G are colored with k colored. This instance is a yes-instance if and only if there exists a clique that contains a vertex from each color.

5.3 The Complexity of Manipulating with k-approval

In this section we will show that manipulation is polynomial-time solvable for k-approval, given that the information set is a member of some $\Sigma \in$ STRUCT, as introduced in Chapter 3. As said above, the proof holds for a more general class of structures, which we will also show. We will let IPPI be the set of all structures that are PI- and IP-transitive. That is:

$$\text{IPPI} = \{\, E_{\boldsymbol{R}} \,|\, \text{all } R_i \in \boldsymbol{R} \text{ are both PI- and IP-transitive} \,\}$$

In Chapter 3, we have shown that every $\Sigma \in$ STRUCT satisfies both the PI- and IP-transitivity, so $\Sigma \subseteq$ IPPI holds. Therefore, the proof of Manipulation(k-approval, IPPI) being polynomial, implies that Manipulation(k-approval, Σ) is also polynomial for every $\Sigma \in$ STRUCT.

We will show Manipulation(k-approval, IPPI) in several steps. We will first introduce the Flow Problem which we will use in our reduction. Then, we will discuss several properties that we will use to show the beats Lemma, from which the final proof follows in a relatively straightforward manner. Recall that we will consider k to be fixed, and thus not part of the input.

5.3.1 The Flow Problem

The flow problem was intended as a simplified model of the Soviet railway system and was first formulated by Harris and Ross (1955). Over the years, many variants of this problem were discovered and analyzed. The flow problem has a whole range of applications, ranging from the more technical problems in graph theory to practical applications such as airline scheduling. Below we will introduce the variants of the problem that are used in this thesis.

The original flow problem can be formalized as follows: given directed some network $N = (V, E)$ with a source $s \in V$ and a sink $t \in V$, we can define the *capacity* of edges $(u, v) \in E$ as a function $c : E \to \mathbb{R}^+$. The capacity represents the maximum amount of flow that can pass through an edge. Now, a *flow* is a mapping $f : E \to \mathbb{R}^+$ such that (1) it does not exceed the capacity of an edge: $f(u, v) \leq c(u, v)$ for all $(u, v) \in E$, and (2) the amount of flow that goes into a vertex is equal to the amount of flow that goes out of a vertex (except for the source s and sink t):

$$\sum_{u:(u,v)\in E} f(u, v) = \sum_{w:(v,w)\in E} f(v, w) \quad \text{for each } v \in V \setminus \{s, t\}$$

The *value* $|f|$ of a flow represents the amount of flow passing from the source to the sink, which is formally defined as follows:

$$|f| = \sum_{v:(s,v)\in E} f(s, v)$$

There are several properties of flow that are interesting to investigate, but here we will focus on problems related to the maximum flow, the Flow Problem and the Flow with

LOWER BOUNDS problem in particular. In these settings, we want to move as much flow as possible from the source to the sink, which boils down to maximizing $|f|$. So far we have only talked about optimization problems, but since we want to prove membership of different complexity classes we need to transform this problem into a decision problem. This can be done as follows:

THE FLOW PROBLEM	
Input:	A network $N = (V, E)$ with a source $s \in V$ and a sink $t \in V$, a capacity $c : E \to \mathbb{R}^+$, and a integer q
Question:	Does there exist a flow $f : E \to \mathbb{R}^+$ where $|f| \geq q$?

There is a whole range of algorithms that that solve the FLOW PROBLEM in polynomial time (See e.g. Ahuja et al. (1993) for an overview). For example, the Edmonds-Karp algorithm solves the flow problem in $O(|V| \cdot |E|^2)$ time (Edmonds and Karp, 1972). This means that if we can reduce a problem Q to the FLOW PROBLEM in polynomial time, the problem Q is also a member of the complexity class P. We will use this construction in the proof below.

An interesting version of the FLOW PROBLEM is when we introduce a lower bound $d : E \to \mathbb{R}^+$ on every edge. Besides not exceeding the capacity of an edge, a flow f should be as least as big as the lower bound: $f(u, v) \geq d(u, v)$ for all $(u, v) \in E$. In fact, the FLOW PROBLEM is a special case of the FLOW WITH LOWER BOUNDS, where $d(u, v) = 0$ for all $(u, v) \in E$. In this thesis we will refer to this problem as the FLOW WITH LOWER BOUNDS problem.

It is known that we can reduce the FLOW WITH LOWER BOUNDS problem to the FLOW PROBLEM in polynomial time (see e.g. Ahuja et al. (1993)). Since we also know that the FLOW PROBLEM can be solved in polynomial time, this means that the FLOW WITH LOWER BOUNDS problem is also solvable in polynomial time. As a result, we can use polynomial reductions to FLOW WITH LOWER BOUNDS to show that problems are solvable in polynomial time. We will use this construction in Lemma 5.4.

5.3.2 Properties of Manipulating with k-approval

To show that MANIPULATION(k-approval, IPPI) can be solved in polynomial time, we will exploit certain properties of manipulating with k-approval and partial profiles. Therefore, we will first give some intuitions about this setting, define these formally, and prove several properties. We will then use these properties in the final part of this section when we give the proof that k-approval is polynomial. We start with proving a useful lemma about the notion k-equivalence, introduced in Chapter 4, then we discuss solutions we think are maximal and finally introduce a function called BEATS.

The k-equivalence class

Recall from Definition 4.2 from Chapter 4 that given some preference order $R \in \mathcal{L}(\mathcal{X})$ and some integer k, we define the **k-equivalence class** $[R]_k$ of R as:

$$[R]_k = \{R' \in \mathcal{L}(X) \,|\, top_k(R') = top_k(R)\}$$

We will use this notion to prove that the order of the top k and bottom $k - m$ indeed does not influence the result. This intuition will help later in this thesis to show that we do not need to consider all preference orders R' when looking for a preference order to manipulate with.

Lemma 5.1. *Let F denote the k-approval rule for any $0 < k < m$. Then, for any two $R, R' \in \mathcal{L}(\mathcal{X})$ such that $[R]_k = [R']_k$ it must hold that $F(\boldsymbol{R}, R) = F(\boldsymbol{R}, R')$ for all $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^{n-1}$.*

Proof. Let $0 < k < m$ and take two arbitrary $R, R' \in \mathcal{L}(\mathcal{X})$ with $[R]_k = [R']_k$. Now, we know that $top_k(R) = top_k(R')$. By the definition of k-approval, we know that exactly the same alternatives receive one point when a voter submits either R or R'. Since with k-approval a winner is completely determined by the amount of points, we have that $F(\boldsymbol{R}, R) = F(\boldsymbol{R}, R')$ for any $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^{n-1}$. □

Fusing two profiles

The second property we will use is the idea that we can always fuse two linear extensions of the same partial profile in a clever way. This is the first moment where IP and PI-transitivity come into play, because we need this property to prove the following theorem:

Lemma 5.2. *For any $\boldsymbol{R} = (R_1, ..., R_{n-1})$ with $E_{\boldsymbol{R}} \in IPPI$ it holds that for any $R \in \boldsymbol{R}$ if we have some $X \in top_k(R')$ for $R' \supseteq R$ and $R' \in \mathcal{L}(\mathcal{X})$, then for any $R'' \supseteq R$ with $R'' \in \mathcal{L}(\mathcal{X})$ there is some $R^* \supseteq R$ such that $R^* \in \mathcal{L}(\mathcal{X})$ and for any $Z \in top_k(R^*)$ either $Z = X$ or $Z \in top_k(R''_i)$.*

Proof. We take such R' and R''. If $X \in top_k(R'')$, then $R^* = R''$ and we are done. So, suppose $X \notin top_k(R'')$. This implies that there must be some $Y \in top_k(R'')$ such that $Y \notin R'$, since $|top_k(R')| = |top_k(R'')| = k$. Now, we let R^* be as R'' but where we replace Y with X and X with Y. Clearly we have that $R^* \in \mathcal{L}(\mathcal{X})$.

What is left to show is that $R^* \supseteq R$. Suppose there is some $A \succ_R B$. Now, if A and B are not X or Z, we know $A \succ_{R''} B$ so $A \succ_{R^*} B$. So, suppose A is X. Now, since $X \succ_{R'} Y$, $Y \succ_{R''} X$ we know $X \sim_R Y$. Since R is IP-transitive, we know that $Y \succ_R B$. Thus we have that $Y \succ_{R''} B$ and thus per definition $X \succ_{R^*} B$. If A is Y this is analogue. Otherwise, assume B is X. Now, since R is PI-transitive, from $B \succ_R X$ and $X \sim_R Y$ we know that $B \succ_R Y$. So again we have that $B \succ_{R''} Y$ so $B \succ_{R^*} X$. When Y is B this is analogue. So, $R^* \subseteq R$. □

Maximal solutions

The next property we will discuss is slightly harder to explain intuitively. Suppose we have some partial profile $\boldsymbol{R}$ and we know there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that alternative Y wins when we add a vote R, but that X wins when we add a vote R' instead. In most cases, this $\boldsymbol{R}'$ will not be the only profile in $E_{\boldsymbol{R}}$ with this property. In particular, if such an $\boldsymbol{R}'$ exists in the first place, there will also be a profile in which either X or Y has as many points as possible, given $\boldsymbol{R}$. Before we introduce the notion of $\boldsymbol{R}$*-maximal*, we will first introduce some new notation.

Given some profile $\boldsymbol{R}$, we will use $pnts(\boldsymbol{R}, Z)$ to denote the amount of points an alternative Z has in $\boldsymbol{R}$ e.g. how often $Z \in top_k(R_j)$ for some $R_j \in \boldsymbol{R}$. In particular, if $\boldsymbol{R}$ is a partial profile we only count a point for Z if its place in the top k is for certain, e.g. that for every full preference $R'_j \supseteq R_j$ we have $Z \in top_k(R'_j)$. In the case that $\boldsymbol{R}$ is a partial profile, we will use $pp(\boldsymbol{R}, Z)$ to denote how many points an alternative Z could have if it is placed in top_k of every vote R_j of the other voters j whenever this is possible: e.g. there exists some full preference $R'_j \supseteq R_j$ and $Z \in top_k(R'_j)$.

Definition 5.1. *If we have some alternative $X \in \mathcal{X}$, some partial profile $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^n$ and some full profile $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^n$ such that $\boldsymbol{R}' \supseteq \boldsymbol{R}$ we say that X is* $\boldsymbol{R}$**-maximal in** $\boldsymbol{R}'$ *whenever*

$$pnts(\boldsymbol{R}', X) = pp(\boldsymbol{R}, X)$$

However, there will be situations in which there are $\boldsymbol{R}'$ such that Y wins when we add a vote R and X wins when we add a vote R' instead, but none in which X or Y are $\boldsymbol{R}$-maximal. The reason is that there could be a partial preference in $\boldsymbol{R}$ in which both X and Y could get a point, but we need to make a choice which alternative actually receives it. This is when there is some preference $R_i \in \boldsymbol{R}$ in which there is only one position in the top k that is not fixed, but both X and Y can be placed in the top k. Below we will introduce the notion of a *choice set* that formalizes this idea. Then, we will prove the intuition from above and show that when choices need to be made, we can still give a lower bound on the number of X and Y receive.

Definition 5.2. *Let $X, Y \in \mathcal{X}$ be 2 alternatives and $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^n$ be a partial profile. Now, we can define the* **choice set** *$C_{\boldsymbol{R}}(X, Y)$ of $\boldsymbol{R}$ given X and Y as:*

$$\begin{aligned} C_{\boldsymbol{R}}(X, Y) = \{R_i \mid \exists R'_i, R''_i \supseteq R_i \textit{ where } X \in top_k(R'_i) \textit{ and } Y \in top_k(R''_i) \\ \forall R^*_i \supseteq R_i \textit{ it does not hold that } X, Y \in top_k(R^*_i)\} \end{aligned}$$

Lemma 5.3. *Let F denote the k-approval for any $0 < k < m$ and $\boldsymbol{R} = (R_1, ..., R_{n-1}) \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\mathbf{R}} \in$ IPPI. Now, given two $R, R' \in \mathcal{L}(\mathcal{X})$, if there exists some $\boldsymbol{R}' = (R'_1, ..., R'_{n-1}) \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R') = X$ and $F(\boldsymbol{R}', R) = Y$, then there exists some $\boldsymbol{R}'' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}'', R') = X$, $F(\boldsymbol{R}'', R) = Y$ and either:*

(1) *X is $\boldsymbol{R}$-maximal in $\boldsymbol{R}''$,*
(2) *Y is $\boldsymbol{R}$-maximal in $\boldsymbol{R}''$, or*
(3) *$pnts(\boldsymbol{R}'', X) + pnts(\boldsymbol{R}'', Y) \geq pp(\boldsymbol{R}, X) + pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)| - 1$*

Proof. Suppose we have such $\boldsymbol{R}$, R and R', such that (1),(2), and (3) do not hold for $\boldsymbol{R}'$. First, let α denote the number of points Y has more than X in $\boldsymbol{R}'$:

$$\alpha = pnts(\boldsymbol{R}', Y) - pnts(\boldsymbol{R}', X)$$

Second, note that k-approval is completely determined by the amount of points every alternative has. So, given $\boldsymbol{R}'$ we can obtain an $\boldsymbol{R}''$ with $F(\boldsymbol{R}'', R') = X$, $F(\boldsymbol{R}'', R) = Y$ by increasing the amount of points of X and Y, as long as we make sure that the difference between X and Y remains equal and there is no other alternative that receives more points in $\boldsymbol{R}''$ than in $\boldsymbol{R}'$.

Using this observation, we will show that we can always increase the points of X and Y in such a way that we obtain an $\boldsymbol{R}''$ for which either (1),(2), or (3) holds. To do so, we distinguish 3 cases:

Case 1: Suppose $pp(\boldsymbol{R}, X) \leq pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)| - \alpha$. The intuition here is that for every $R_i \in C_{\boldsymbol{R}}(X, Y)$ we place X in the top k, and then there are enough preferences $R_j \neq C_{\boldsymbol{R}}(X, Y)$ left where Y can be placed in the top k for $\boldsymbol{R}''$ to have the properties we want. Formally this goes as follows. First, we can construct an $\boldsymbol{R}^*$ such that $pp(\boldsymbol{R}, X) = pnts(\boldsymbol{R}^*, X)$ and for all $Z \in \mathcal{X} \setminus \{X\}$ it holds that $pnts(\boldsymbol{R}^*, Z) \leq pnts(\boldsymbol{R}', Z)$. From Lemma 5.2 we know such $\boldsymbol{R}^*$ always exists, since we can replace every $R'_j \supseteq R_j$ such that there exists a $R^*_j \supseteq R_j$ that has $X \in top_k(R^*_j)$, with such a R^*_j for which we have that for all $Z \in top_k(R^*_j)$, either $Z = X$ or $Z \in top_k(R'_j)$.

By our assumption that $pp(\boldsymbol{R}, X) \leq pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)| - \alpha$, it follows that we can construct some $\boldsymbol{R}''$ from $\boldsymbol{R}^*$ such that $pnts(\boldsymbol{R}'', Y) = pnts(\boldsymbol{R}'', X) + \alpha$ and for all $Z \in \mathcal{X} \setminus \{X, Y\}$ it holds that $pnts(\boldsymbol{R}'', Z) \leq pnts(\boldsymbol{R'}, Z)$. Now, by the reasoning above we know that $F(\boldsymbol{R}'', R') = X$, $F(\boldsymbol{R}'', R) = Y$, and by construction we know X is $\boldsymbol{R}$-maximal in $\boldsymbol{R}''$.

Case 2: Suppose $pp(\boldsymbol{R}, Y) - \alpha \leq pp(\boldsymbol{R}, X) - |C_{\boldsymbol{R}}(X, Y)|$. Now, with a similar argument as in case 1 it holds that there is some $\boldsymbol{R}''$ such that $F(\boldsymbol{R}'', R') = X$, $F(\boldsymbol{R}'', R) = Y$, and Y is $\boldsymbol{R}$-maximal in $\boldsymbol{R}''$.

Case 3: Suppose $pp(\boldsymbol{R}, X) > pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)| - \alpha$ and $pp(\boldsymbol{R}, Y) - \alpha > pp(\boldsymbol{R}, X) - |C_{\boldsymbol{R}}(X, Y)|$. In this case we actually have to make a choice whether we put X or Y on top in every $R_i \in C_{\boldsymbol{R}}(X, Y)$.

First, if we have some $R_i \in R$ with $R_i \neq C_{\boldsymbol{R}}(X, Y)$ but such that there exist a $R^*_i \supseteq R_i$ with either $X \in top_k(R^*_i)$ or $Y \in top_k(R^*_i)$ we know from applying Lemma 5.2, there must be some specific R^*_i such that for all $Z \in top_k(R^*_i)$ we either have $Z \in top_k(R'_i)$, $Z = X$ or $Z = Y$. Now we obtain $\boldsymbol{R}^*$ from $\boldsymbol{R}'$ by replacing all such R'_i with R^*_i.

Now, suppose $pp(\boldsymbol{R}, X) + pp(\boldsymbol{R}, Y) - C_{\boldsymbol{R}}(X, Y) - |\alpha|$ is even, we can clearly construct $\boldsymbol{R}''$ from $\boldsymbol{R}^*$ such that:

- $pnts(\boldsymbol{R}', X) + pnts(\boldsymbol{R}'', Y) = pp(\boldsymbol{R}, X) + pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)|$
- $\forall Z \in \mathcal{X} \setminus \{X, Y\}$ it holds that $pnts(\boldsymbol{R}'', Z) \geq pnts(\boldsymbol{R}', Z)$
- $pnts(\boldsymbol{R}'', Y) - pnts(\boldsymbol{R}'', X) = \alpha$

Otherwise, since we assumed (3) does not hold in $\boldsymbol{R}'$ we know there must be some $R'_j \in \boldsymbol{R}'$ such that $X, Y \notin top_k(R'_j)$ but there exists some $R^* \supseteq R_j$ with either $X \in top_k(R^*)$ or

$Y \in top_k(R^*)$. Now, we can clearly construct $\boldsymbol{R}''$ from $\boldsymbol{R}^*$ such that:

- $R''_j = R'_j$
- $pnts(\boldsymbol{R}', X) + pnts(\boldsymbol{R}'', Y) = pp(\boldsymbol{R}, X) + pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)| - 1$
- $\forall Z \in \mathcal{X} \setminus \{X, Y\}$ it holds that $pnts(\boldsymbol{R}'', Z) \geq pnts(\boldsymbol{R}', Z)$
- $pnts(\boldsymbol{R}'', Y) - pnts(\boldsymbol{R}'', X) = \alpha$

In both cases we have found some $\boldsymbol{R}'' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}'', R') = X$, $F(\boldsymbol{R}'', R) = Y$ and (3) holds.

□

The beats function

The last piece that we will need in the proof below, is the fact that given some partial profile $\boldsymbol{R}$, two preference orders R and R', and two alternatives X and Y, we can check in polynomial time if there exists a $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that voting R' will make sure X wins, while voting R will make Y the winner. To show this, we will use that we do not have to check every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, but only the ones with the properties described in Lemma 5.3. Moreover, we will use the FLOW WITH LOWER BOUNDS problem introduced in 5.3.1. Finally, we need to introduce the following notation:

Definition 5.3. *Let B some boolean expression, then we define $[\![B]\!] = 1$ when B = TRUE and $[\![B]\!] = 0$ whenever B = FALSE.*

Lemma 5.4. *Let F denote the k-approval rule for any $0 < k < m$. Now, given some partial profile $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ with $E_{\boldsymbol{R}} \in IPPI$, two preference orders $R, R' \in \mathcal{L}(\mathcal{X})$, and two alternatives $X, Y \in \mathcal{X}$, then there exists a polynomial time algorithm for the function* BEATS *that given $\boldsymbol{R}$, R', R, X and Y returns* TRUE *if and only if there exists some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R') = X$ and $F(\boldsymbol{R}', R) = Y$.*

Proof. We take some arbitrary $\boldsymbol{R}$, R, R', X and Y and will show that the Lemma holds by using the FLOW WITH LOWER BOUNDS problem described above. From Lemma 5.3 we know that we only have to search for an $\boldsymbol{R}'$, such that either (1) X is $\boldsymbol{R}$-maximal, (2) Y is $\boldsymbol{R}$-maximal, or (3) $pnts(\boldsymbol{R}', X) + pnts(\boldsymbol{R}', Y) \geq pp(\boldsymbol{R}, X) + pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)| - 1$. To show that BEATS runs in polynomial time, we will create 3 flow instances: I'_1, I'_2, and I'_3. Each I'_i will be a yes-instance if and only if there exists a $\boldsymbol{R}'$ with the property (i). Now, BEATS will return TRUE if and only if at least one if I'_i is a yes-instance.[5] Since there exists a polynomial time algorithm for the FLOW WITH LOWER BOUNDS problem and we reduce our BEATS problem to a constant number of instances I', we know that if this reduction is polynomial then BEATS can be solved in polynomial time.

[5] We could greatly improve the running time if we would use a case distinction similar to the proof of Lemma 5.3, but this results into a long and technical proof. Since our aim is to show membership of P, we choose to use this easier but less efficient method.

The reduction from I to I'_1, I'_2, and I'_3 will be similar. In all these cases the number of points that X and Y will receive in $\boldsymbol{R}'$ is fixed. The question we then need to answer is whether we can place the other alternatives in such a way that there is no $Z \in \mathcal{X} \setminus \{X, Y\}$ that beats X or Y when the manipulator votes R' or R respectively. This question can be expressed in a FLOW WITH LOWER BOUNDS problem. For now, assume that for every $Z \in \mathcal{X}$, h_Z is the maximum and e_Z the minimum amount of points we want alternative Z to receive in $\boldsymbol{R}'$ (we will determine the values of h_Z and e_Z below for all I').

Moreover, for every voter i we let v_i be the amount of alternatives that can be in the top k but are not necessarily. Formally that is:

$$v_i = |\{\, X \mid \exists R'_i, R''_i \supseteq R_i \text{ with } X \in top_k(R'_i) \text{ and } X \notin top_k(R''_i)\}|$$

For instance, if $E_{\boldsymbol{R}} \in$ TOPBOT, $v_i = k - 1$ for all i and if $E_{\boldsymbol{R}} \in$ COMPL(t), v_i is either 0 or k.

Now, we can reduce the BEATS instance $I = (\boldsymbol{R}, R', R, X, Y)$ to an instance $I' = ((V, E), c, d, q)$ of FLOW WITH LOWER BOUNDS where:

$$\begin{aligned}
V &= \{s, t\} \cup \mathcal{N} \setminus \{n\} \cup \mathcal{X} \qquad\qquad (\text{where } s, t \notin \mathcal{N} \setminus \{n\} \cup \mathcal{X})\\
E &= \{(s, i) \mid i \in \mathcal{N} \setminus \{n\}\} \cup \{(x, t) \mid x \in \mathcal{X}\}\\
&\quad \cup \{(i, x) \mid i \in \mathcal{N} \setminus \{n\}, x \in \mathcal{X}, \exists R'_i \supseteq R_i \text{ with } x \in top_k(R'_i)\}\\
c(a, b) &= \begin{cases} v_b & \text{if } a = s\\ 1 & \text{if } a \in \mathcal{N} \setminus \{n\}\\ h_a & \text{if } a \in \mathcal{X} \end{cases}\\
d(a, b) &= \begin{cases} e_a & \text{if } a \in \mathcal{X}\\ 0 & \text{otherwise} \end{cases}\\
q &= \sum_{i \in \mathcal{N} \setminus \{n\}} v_i
\end{aligned}$$

For a visualization of this reduction, see Figure 5.3.2.

No matter whether we are in (1),(2) or (3), if there always exists an $\boldsymbol{R}'$ such that $F(\boldsymbol{R}', R') = X$ and $F(\boldsymbol{R}', R) = Y$ depends heavily on R and R'. For example, if it holds that $X, Y \in top_k(R')$ and $X, Y \notin top_k(R)$, such an $\boldsymbol{R}'$ can clearly not exist. It is easy to check that we can eliminate all cases except the following:

- $X \notin top_k(R)$, $Y \in top_k(R)$ and either $X, Y \in top_k(R')$ or $X, Y \notin top_k(R')$ $(*)$.
- $X \in top_k(R')$, $Y \notin top_k(R')$, and either $X, Y \in top_k(R)$ or $X, Y \notin top_k(R)$ $(**)$.
- $X \in top_k(R')$, $Y \notin top_k(R')$, $X \notin top_k(R)$, and $Y \in top_k(R)$ $(***)$.

If R and R' are not described by one of these 5 cases, we let I'_1, I'_2 and I'_3 be trivial no-instances of FLOW WITH LOWER BOUNDS.

Moreover, the relative distance between X and Y needs to be in such a way that the manipulator can be pivotal for X and Y. We let α be the amount of points that Y has more than X in $\boldsymbol{R'}$. The value of α differs per situation, but it can be checked that if there exists a $\boldsymbol{R'}$ such that $F(\boldsymbol{R'}, R') = X$ and $F(\boldsymbol{R'}, R) = Y$, α must have the following values:

	$X \triangleright Y$	$Y \triangleright X$
$(*)$	$\alpha = 0$	$\alpha = -1$
$(**)$	$\alpha = 1$	$\alpha = 0$
$(***)$	$\begin{cases} \alpha = 1 & \text{if } \triangle_X \\ \alpha = 0 & \text{otherwise} \end{cases}$	$\begin{cases} \alpha = -1 & \text{if } \triangle_Y \\ \alpha = 0 & \text{otherwise} \end{cases}$

Where $\triangle_X := pp(\boldsymbol{R}, X) < pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X, Y)|$ and $\triangle_Y := pp(\boldsymbol{R}, Y) < pp(\boldsymbol{R}, X) - |C_{\boldsymbol{R}}(X, Y)|$.

What is left is to determine h_Z and e_Z for every alternative Z. Since these values differ based on which case we are in, we are creating these three instances. Recall that in I_1', we want X to have as many points as possible, in I_2' Y as many points as possible, and in I_3' divide all (or all but one) points given the difference remains α. Given α, we can define these as follows:

$$\text{in } I_A' : \begin{cases} e_X = h_X = pp(\boldsymbol{R}, X) - pnts(\boldsymbol{R}, X) \\ e_Y = h_Y = pp(\boldsymbol{R}, X) + \alpha - pnts(\boldsymbol{R}, Y) \end{cases}$$

$$\text{in } I_B' : \begin{cases} e_X = h_X = pp(\boldsymbol{R}, Y) - \alpha - pnts(\boldsymbol{R}, X) \\ e_Y = h_Y = pp(\boldsymbol{R}, Y) - pnts(\boldsymbol{R}, Y) \end{cases}$$

$$\text{in } I_C' : \begin{cases} e_X = h_X = \left\lfloor \frac{pp(\boldsymbol{R}, X) + pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X,Y)| - \alpha}{2} \right\rfloor - pnts(\boldsymbol{R}, X) \\ e_Y = h_Y = \left\lfloor \frac{pp(\boldsymbol{R}, X) + pp(\boldsymbol{R}, Y) - |C_{\boldsymbol{R}}(X,Y)| - \alpha}{2} \right\rfloor + \alpha - pnts(\boldsymbol{R}, Y) \end{cases}$$

If $Z \in \mathcal{X} \setminus \{X, Y\}$, we do not want Z to beat X or Y when the manipulator votes R' or R respectively. We can formally define h_Z as follows:

$$\begin{aligned} h_Z = min(&h_x + pnts(\boldsymbol{R}, X) - [\![Z \triangleright X]\!] - [\![Z \in top_k(R')]\!] + [\![X \in top_k(R')]\!], \\ &h_y + pnts(\boldsymbol{R}, Y) - [\![Z \triangleright Y]\!] - [\![Z \in top_k(R)]\!] + [\![Y \in top_k(R)]\!]) - pnts(\boldsymbol{R}, Z) \end{aligned}$$

This reduction from I to I_1', clearly runs in polynomial time. So, what is left to show is that I is a yes-instance if and only if at least one of I_1', I_2' or I_3' is a yes-instance.

($\Rightarrow$) Suppose I is a yes-instance. Then by Lemma 5.3 we know there must exist an $\boldsymbol{R'}$ such that $F(\boldsymbol{R'}, R') = X$, $F(\boldsymbol{R'}, R) = Y$ and either (1), (2) or (3) holds. Suppose (1) holds, then we know $pnts(\boldsymbol{R'}, X)$ is equal to h_X in I_1'. Now, we can define a flow f such that $f(s, i) = v_i$ for all $i \in \mathcal{N} \setminus \{n\}$, $f(i, z) = 1$ if and only if $Z \in top_k(R_i')$ for all Z such

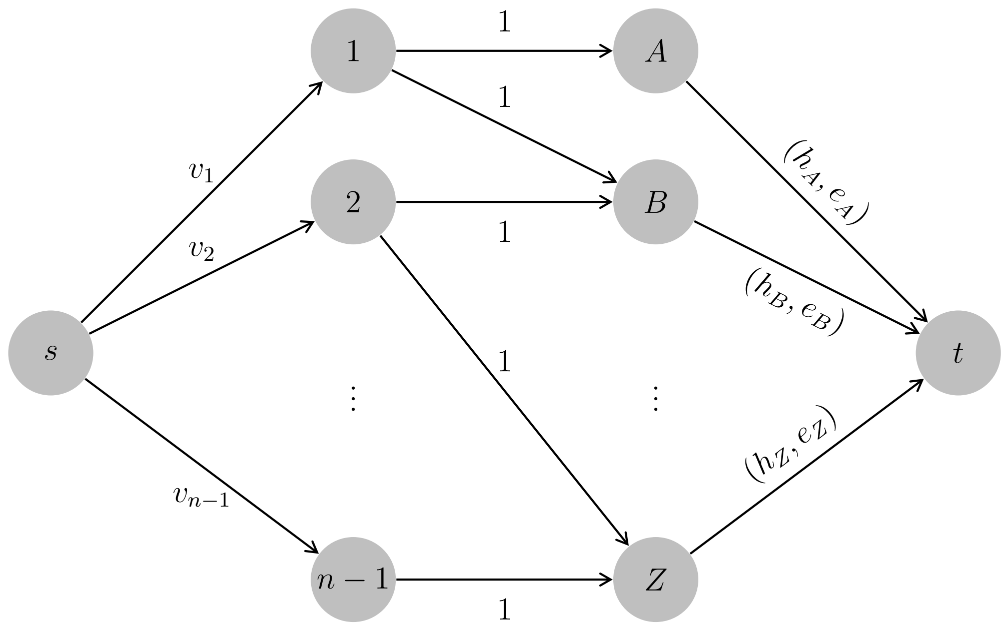

Figure 5.1: Visualization of the reduction of an BEATS instance $I = (\boldsymbol{R}, R', R, X, Y)$ to an instance $I' = ((V, E), c, d, q)$ of FLOW WITH LOWER BOUNDS as described in the proof of Lemma 5.4. The capacity and lower bounds are written next to the edges, all lower bounds that are equal to 0 are left out of the picture.

that there exists some $R''_i \supseteq R_i$ with $Z \in top_k(R''_i)$, and $f(z,t) = \sum_{i \in \mathcal{N} \setminus \{n\}} f(z,i)$. By construction we must have that f is a flow for I'_1 with $|f| \geq q$, so I'_1 is a yes-instance. We (2) or (3) hold, we similarly get that I'_2 or I'_3 is a yes-instance.

($\Leftarrow$) Suppose one of I'_1, I'_2, or I'_3 is a yes-instance. In each case there exists a flow f such that for all $(u,v) \in E$ we have $d(u,v) < f(u,v) < c(u,v)$. Now we can translate this flow to $\boldsymbol{R}'$ with the desired properties by extending $\boldsymbol{R} = (R_1, ..., R_n)$ to $\boldsymbol{R}' = (R'_1, ...R'_n)$ such that for all Z such that there exists some $R''_i \supseteq R_i$ with $Z \in top_k(R''_i)$ it holds that $Z \in top_k(R'_j)$ if and only if $f(j,z) = 1$. Now by constructions we have found an $\boldsymbol{R}'$ with the desired properties.

□

5.3.3 Complexity Results for k-approval

Theorem 5.1. MANIPULATION*(k-approval, IPPI) for fixed k can be solved in polynomial time.*

Proof. We take some arbitrary partial profile $\boldsymbol{R}$ such that $E_{\boldsymbol{R}} \in$ IPPI and a true preference $R \in \mathcal{L}(\mathcal{X})$ for the manipulator. Suppose we want to check whether some $R' \neq R$ would be a possible preference that makes this a yes-instance. By definition we know this is the case if and only if:

(1) R' is at least one time better, e.g. there exists $X, Y \in \mathcal{X}$ with $X \succ_R Y$ and some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R) = X$ and $F(\boldsymbol{R}^*, R) = Y$, and

(2) by voting R' the manipulator is never worse off, e.g. for all $X, Y \in \mathcal{X}$ with $X \succ_R Y$ there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = X$ and $F(\boldsymbol{R}', R) = Y$

From Lemma 5.4, we know that that given such R' we can check whether (1) and (2) hold in polynomial time. We can do this as follows: for every pair $X, Y \in \mathcal{X}$ with $X \succ_R Y$ we both run BEATS($\boldsymbol{R}, R', R, X, Y$) and BEATS($\boldsymbol{R}, R, R', X, Y$). If the first call returns TRUE at least once, then (1) holds and if the second returns FALSE for every pair X, Y with $X \succ_R Y$ then (2) holds. Since we know that BEATS runs in $O(p(|x|))$ time for some polynomial p and where $|x|$ denotes the input size, and there are $O(m^2)$ pairs with $X \succ_R Y$, checking whether some R' is a preference order we are looking for can be done in $O(m^2 \cdot p(|x|))$ time.

Now, from Lemma 5.1 it follows that instead of checking every $R' \neq R$ it suffices to check one arbitrary R' per equivalence class $[R'] \neq [R]$. Since there are $\binom{m}{k}$ ways to choose a subset of k out of m alternatives, there are $\binom{m}{k}$ different preferences that need to be checked. Since we above showed that checking such preference can be done in $O(m^2 \cdot p)$ time, we can solve MANIPULATION in $O(\binom{m}{k} \cdot m^2 \cdot p(|x|))$ time, for some polynomial function p. Since the parameter k is not part of the input, we have shown that MANIPULATION(k-approval, IPPI) runs in polynomial-time. □

5.4 Discussion

The conjecture, *does the NP-hardness of manipulation and the possible winner problem for all natural voting rules only hold for classes of instances that are not PI- and IP-transitive?*, poses both a theoretical and more philosophical question. The theoretical question is whether it is true. This is something that we believe would be interesting to investigate in future research. The reason why we also think it is difficult to do so is that if it is not true, a very different NP-hardness proof than the literature needs to be given. The reason for this is, as we have shown in Section 5.2, there is no straightforward way of translating the already existing results in the literature to instances that are both IP- and PI-transitive. On the other hand, if the conjecture is true, we need to show that manipulation can be solved in polynomial time. The proof of k-approval is already quite complicated, but the one of Borda and Copeland is probably going to be more complex since we cannot exploit the idea of a k-equivalence class and some smart non-trivial solution for this needs be found.

The more philosophical question is that if this conjecture is true, do we believe this property is reasonable or not? We believe this question relates to the general criticism of using NP-hardness as a barrier against manipulation. NP-hardness is a worst-case result, so it tells us that there is at least one instance that is difficult to manipulate, but it does not say something about how difficult it *usually* is to manipulate (Walsh, 2010). For example, a problem could be NP-hard because of a small group of instances that make it very difficult, but it easy to solve on most of the other instances. This makes that in practice manipulation is often still possible (Conitzer and Sandholm, 2006). If the conjecture is true, this also means that manipulation can only be difficult when the partial profile is not IP- or PI-transitive. The question is whether we can then still argue that the complexity is a barrier to manipulation is practice.

5.5 Conclusion

In this chapter, we have shown that MANIPULATION is polynomial for k-approval for every class of instances that is both IP- and PI-transitive and argued why we believe this might also be the case for Borda and Copeland. At the beginning of this chapter, we formally defined the MANIPULATION PROBLEM, which corresponds to the notion of manipulating as defined in the Gibbard-Satterthwaite Theorem. Then, we compared this notion to related complexity problems and existing complexity results of these problems. We observed that most NP-hardness proofs reduce to instances that are not IP- and PI-transitive and posed the conjecture that the hardness results for natural voting rules do not translate to classes of instances that are IP- and PI-transitive, which is the case for every $\Sigma \in$ STRUCT. In the final part of the chapter, we showed the polynomial-time result for k-approval, using a reduction from MANIPULATION to the polynomial-time solvable problem FLOW WITH LOWER BOUNDS.

Chapter 6

Other Notions of Disadvantage

In the previous two chapters, we have used two ways of measuring the disadvantage of the manipulator: the manipulator either never has an incentive to manipulate (when t is too small), or manipulating could be more difficult or easier. In both cases, no hard evidence was found which supports a difference between the full and partial information case from the perspective of the manipulator. Intuitively, however, one would think that when a manipulator possess less information about the other voters' preferences, it poses a disadvantage of some kind compared to the manipulator having full information.

In this chapter, we focus on other ways this disadvantage could be investigated. In contrast with the chapters above, this chapter will mainly focus on how to investigate this question, rather than finding a lot of results. The main reason for this is that the intuition that having less information should disadvantage manipulators is very vague and informal. So, before precise statements about this can be made, this intuition should first be formalized. We believe this approach is relatively new; here we would like to demonstrate how existing theory can be used to look at manipulation from a different perspective since we believe this can advance research in this direction.

Below, we will take some first steps in the direction of this formalization and show how such a question could be investigated. Before we can say things about whether having less information about the preferences of voters should disadvantage them in some way, we believe that we have to answer the following three questions:

1. How does the partial information setting compare to the full information setting?
2. What does it mean to be disadvantaged, or what does a manipulator lose when having less information?
3. What does 'less information' mean for the manipulator, and how can we measure information?

We will consider every question in a separate section below.

6.1 Full versus Partial Information

If we want to reason about the disadvantage of having less information, we need to know how the setting of partial information relates to full information. In other models of partial information, this connection is explicitly made. In the model of Meir et al. (2014), for example, the information set is obtained by all profiles which have a certain distance to the actual profile. In the model of Reijngoud and Endriss (2012), the information set is defined as all profiles which have the same piece of information as the actual profile. In our framework, however, we have only defined which profiles are *possible situations*, but we have not defined which profile is the *actual situation*.

We think that the most straightforward way of doing this is to define a *partial information function* G_Σ for every $\Sigma \in \text{STRUCT}$ that given some full profile $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ returns some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}} \in \Sigma$. The G_Σ is comparable to the poll information functions π defined by Reijngoud and Endriss (2012), since they provide partial information to the manipulator, mathematically speaking we could even call G_Σ a PIF. However, we believe there is a conceptual difference between the two. A PIF returns polling information, which is often a summary of the result of the election, while a G_Σ returns partial information about the preference orders of the voters. Thus, the one provides partial information about the outcome while the other partial information about the preferences.

We believe there are two properties which this function G_Σ must posses to be a reasonable connection between the partial and full information setting:

- For every $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ it holds that if $G_\Sigma(\boldsymbol{R}') = \boldsymbol{R}$ then $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. In other words, the *actual situation* should be a *possible situation*.
- G_Σ must be surjective. This means that every that every partial profile $\boldsymbol{R}$ with $E_{\boldsymbol{R}} \in \Sigma$ is considered as potential partial information setting.

In Lemma 3.1, we have shown that every full profile is part of some information set. This means that for every Σ we are able to define a G_Σ with these two properties. Moreover, when Σ is TOPBOT, BOT(t), or TOP(t), these two restrictions even uniquely define G_Σ, since we have also shown that these structures are partitions of $\mathcal{L}(\mathcal{X})^{n-1}$ which implies that we can uniquely define G_Σ as follows:

$$G_\Sigma(\boldsymbol{R}') = E_{\boldsymbol{R}} \Leftrightarrow \boldsymbol{R}' \in E_{\boldsymbol{R}}$$

For COMPL(t), multiple such G_Σ exist.

6.2 Disadvantage

In this thesis we have already considered two ways of how this idea could be formalized: either the manipulator cannot manipulate at all (Chapter 4), or finding the preference to manipulate with takes more computational effort (Chapter 5). One of the alternative ways we can measure disadvantage is by using manipulation as a scale, rather than a binary concept. In this context, disadvantage means that in one situation (e.g. partial information) there are fewer profiles in which the manipulator has an incentive to manipulate than in the other situation (e.g. full information). We will go into this into more detail in Chapter 7.

Instead of comparing partial information with full information in general, we could also compare some particular full information profile to the corresponding partial information setting. If we suppose that the function G_Σ is known, we can first analyze how it can happen that a manipulator has a different response when knowing $\boldsymbol{R}'$ compared to knowing $G_\Sigma(\boldsymbol{R}')$. Roughly the following two things can happen:

1. The manipulator does not have an incentive to manipulate in $\boldsymbol{R}'$, but does have an incentive to manipulate in $G(\boldsymbol{R}')$
2. The manipulator has an incentive to manipulate in $\boldsymbol{R}'$, but does not have an incentive to manipulate in $G(\boldsymbol{R}')$

In the first case, the manipulator thinks they can manipulate, but in the actual situation, the result is the same as when the voters would choose their true preference. In the outcome of the election, the manipulator does not really lose something. However, it can be argued that they sacrifice computational power or credibility without gaining something. This is an interesting perspective to look at, especially from a more political science, ethical, or behavioral science point of view.

In the second case, if the manipulator manipulates, the outcome becomes worse for the manipulator when having less information than full information. We will make this idea more formal below. However, we will not only include the situations in which the manipulator completely loses all incentive to manipulate with less information, but also the situation in which they still have an incentive to manipulate, although the outcome is less favored than when the manipulator has full information.

To formalize this, we use the notion of a *best response*, a concept used in iterative voting (see e.g. Meir et al. (2010); Reijngoud and Endriss (2012)). It could be possible, for example, that a manipulator can manipulate with several $R' \neq R''$ and end up with a outcome that is better than with the truthful preference R. Now, R' is called a best response (R'_b) whenever there is no $R'' \in \mathcal{L}(\mathcal{X})$ such that $F(\boldsymbol{R}, R'') \succ_R F(\boldsymbol{R}, R'_b)$. The best response does not need to be unique. Now, we use this notion to formalize a slightly broader definition of case 2 as follows:

Definition 6.1. *Given some actual situation $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ and a partial information function G_Σ, we say that a manipulator with a true preference R* **has a disadvantage** *with partial information compared to full information whenever:*

- *The manipulator has an incentive to manipulate in $\boldsymbol{R}'$ with a best response $R'_b \in \mathcal{L}(\mathcal{X})^{n-1}$, and*
- *There is no $R'' \in \mathcal{L}(\mathcal{X})$ such the manipulator has an incentive to manipulate with R'' and $F(\boldsymbol{R}', R'') \succeq_R F(\boldsymbol{R}', R'_b)$.*

6.3 Less Information

The final aspect that we will consider is the concept of *less information*. Having enough information does not equal having full information. For example, when the voting rule is plurality and there is a manipulator who knows the most favorite alternative of every voter. In this case, the manipulator does not have full information, but we will show below that they do not have a disadvantage either. There is one distinction to be made when considering information. On the one hand, we can look at the *amount* of information, wondering if the manipulator has enough. On the other hand, there is the *kind* of information and looking whether the manipulator has the right kind of information.

6.3.1 Enough Information

When we want to measure the amount of information, we believe it is straightforward to do this in terms of bits, a unit that is often used in *information theory* in which among others the quantification of information is studied. Bits are based on the binary logarithm. For example, if we want to communicate three out of ten different candidates, we need to be able to distinguish between them, so we need at least $\log(10)$ bits per candidate and thus $3\log(10)$ bits to communicate three (see e.g. MacKay, 2003). So, how do we show how much bits of the full profile the manipulator needs to know to never have a disadvantage?

Communication Complexity

Here, we believe that the field of *communication complexity* is very suitable to look at this problem. Communication complexity tries to quantify how much information is needed for distributed computing. Traditionally, there are two agents, Alice and Bob, who both have their own input string x and y. Now one of them has to compute $f(x, y)$ with the least amount of information exchanged between them. In the context of voting, the function which is computed is the voting rule F and every voter i can be seen as an agent with their input ballot R_i. The communication complexity of F now corresponds to how much bits all voters in total need to communicate so that every voter has enough information to compute who is winning. The communication can proceed in multiple rounds - formally, this is called a *deterministic protocol*, in which all players announce some information to the others voters based on their input R_i and all bits announced so far. When all players know $F(R_1, ..., R_n)$ the protocol stops. The total minimum amount of bits that need to be sent in the worse case is the communication complexity of F (see e.g. Kushilevitz and Nisan, 1997; Conitzer and Sandholm, 2005).

Conitzer and Sandholm (2005) determined the communication complexity of common voting rules. They showed, for example, that the deterministic communication complexity of the plurality rule is $O(n\log(m))$, since communicating one (in this case the top) candidate requires $O(\log(m))$ bits for one player, so if every player announces their top favourite candidate takes $O(n\log(m))$ bits, and then clearly every player can calculate who wins the election.

By giving an explicit communication protocol, however, we are only able to prove upper bounds of the communication complexity since it is always possible that some more efficient protocol exists. Especially in this context, we are also very interested in the lower bound of the communication complexity of F. Several mathematical tools exist to do this. We will not go into depth into this here, but a nice overview can be found in (Kushilevitz and Nisan, 1997). Conitzer and Sandholm also showed lower bounds of the communication complexity of these voting rules. In particular, they showed that for large m and n, the lower and upper bounds converge for most common voting rules, including the plurality rule, Condorcet, and Borda. The only exception they found was STV.

Connection to STRUCT

Similarly to the upper bound on the communication complexity, we can determine how much bits are needed to communicate the partial profiles $\boldsymbol{R}$. For example to communicate some $\boldsymbol{R}$ with $E_{\boldsymbol{R}} \in \text{TOP}(t)$, we need at most $O(n \cdot t \cdot log(m))$ bits, since for every voter we need to communicate t of m alternatives. We can do this for all $\Sigma \in$ STRUCT:

Σ	Amount of bits needed to communicate
TOPBOT	$O(n \log(m))$
TOP(t)	$O(nt \log(m))$
BOT(t)	$O(nt \log(m))$
COMPL(t)	$O(tm \log(m))$

From the results above and the lower bounds on the communication complexity, we can now formulate the following concrete question:

> *When the lower bound on the communication complexity is larger than the upper bound on the number of bits of the profile for some Σ, does there exist some m and n such that there is some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ with $E_{\boldsymbol{R}} \in \Sigma$ and a true preference R in which the manipulator has a disadvantage?*

The result of this question gives us more knowledge on how the partial information compares to the full information setting. This will tell us more about what kind and extent of the that influence partial information has on manipulation.

We can answer this question in multiple ways. For example, Conitzer and Sandholm showed that the lower bound of Borda is $\Omega(nmlog(m))$ and above we showed that we need at most $O(nlog(m))$ bits to communicate some partial profile with a TOPBOT structure. Then we can wonder whether there is some situation in which the manipulator has a disadvantage when only knowing the TOPBOT. As we will show below, this is true for every $m \geq 5$ and $n \geq 4$. The proof can be found in the Appendix.

Theorem 6.1. *For F being the Borda rule, for any $n \geq 4$ and $m \geq 5$ there exists some $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ such that $G_{TOPBOT}(\boldsymbol{R}') = \boldsymbol{R}$ and $E_{\boldsymbol{R}} \in$ TOPBOT and a trutful profile R such that the manipulator has a disadvantage in the partial information setting compared to the full information setting.*

One way of extending this in future research is to do this for all structures and voting rules. We believe, however, that a more fruitful approach would be to see if the research question above can be proved directly.

6.3.2 The Right Kind of Information

Even if a manipulator has enough information in terms of bits, they can still not have the right kind of information. A notion we believe is relevant to use here is the notion of *being strongly computable from π-images*, introduced by Reijngoud and Endriss (2012, p. 4), who introduced this concept to use the Gibbard-Satterthwaite Theorem, not for this purpose. We think, however, that this notion also fits our purposes.

Recall that Reijngoudt and Endriss model uncertainty by a poll information function (PIF) $\pi : \mathcal{L}(\mathcal{X})^n \to \mathcal{I}$, a function from a profile to an information set. In this context, a voting rule F is strongly computable from a π-image (for some specific PIF π) whenever there exists a function $H : \mathcal{I} \to \wp(\mathcal{X}) \setminus \{\emptyset\}$ such that $F = H \circ \pi$, and for any two full profiles $\boldsymbol{R} = (R_1, \ldots, R_n), \boldsymbol{R}' = (R'_1, \ldots, R'_n) \in \mathcal{L}(\mathcal{X})^n$, if for any voter i we have that $\pi(\boldsymbol{R}) = \pi(\boldsymbol{R}'_{-i}, R_i)$ then it must also hold that $F(\boldsymbol{R}_{-i}, R'_i) = F(\boldsymbol{R}')$. This means that if a voter knows $\pi(\boldsymbol{R})$, she can compute the winner for any way of voting, not only when voting R_i.

Intuitively, we think that being computable from a structure means that no matter what G_Σ is, we can always compute who is winning. Therefore, we have formalized 'being computable' as follows and can immediately show that this implies that every possible situation should have the same winner.

Definition 6.2. *A voting rule F is* **computable from** Σ *when there exists a function H such that for any partial information function G_Σ it holds that $F = H \circ G_\Sigma$*

Lemma 6.1. *If a voting rule F is* **computable from some structure** Σ *for any $\Sigma \in$ STRUCT, then for every partial profile $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^n$ such that $E_{\boldsymbol{R}} \in \Sigma$, we have that for all $\boldsymbol{R}', \boldsymbol{R}'' \in E_{\boldsymbol{R}}$ it hold that $F(\boldsymbol{R}') = F(\boldsymbol{R}'')$.*

The proof can be found in the Appendix.

To illustrate why we want to say that this H should work for any arbitrary G_Σ, and not for some specific G_Σ, consider the following example.

Example 6.1. *Suppose we have 3 voters and 3 alternatives $\mathcal{X} = \{A, B, C\}$. Now we let Σ be COMPL(1) and let F denote the plurality rule. Note that every $\boldsymbol{R}$ such that $E_{\boldsymbol{R}} \in$ COMPL(1) consists of one full vote and two empty votes. Moreover, since F denotes plurality, we know that if $F(\boldsymbol{R}') = X$ for any alternative X there must be at least one voter* *that has X as the top position. This means that we can construct a G_Σ such that we map every $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^3$ to the $\boldsymbol{R}$ in which the vote with the winner is on top of the vote is still complete. If there are multiple votes with the winner on top, we can pick the corresponding $\boldsymbol{R}$ such that G_Σ becomes surjective.*

Now, note that G_Σ is a partial information function, since it is surjective and $G_\Sigma(\mathbf{R}') = \boldsymbol{R}$ implies that $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. Moreover, we can now define a H such that $F = H \circ G_\Sigma$ by letting H elect the alternative that is on the top of the only defined vote. Now by construction we have that $H(G_\Sigma(\boldsymbol{R}')) = F(\boldsymbol{R}')$. However, we clearly do not have that for any $\boldsymbol{R}', \boldsymbol{R}'' \in E_{\boldsymbol{R}}$ that $F(\boldsymbol{R}') = F(\boldsymbol{R}'')$.

Finally, we will define the notion of being strongly computable similar to Reijngoud and Endriss and show that we believe this is a notion that incorporates having enough information:

Definition 6.3. *A voting rule F is* **strongly computable from** Σ *if F is computable from Σ and for any $E_{\boldsymbol{R}} \in \Sigma$ it holds that for all $\boldsymbol{R}', \boldsymbol{R}'' \in E_{\boldsymbol{R}}$ and $R \in \mathcal{L}(\mathcal{X})$ we have $F(\boldsymbol{R}', R) = F(\boldsymbol{R}'', R)$.*

Theorem 6.2. *If F is is strongly computable from a structure Σ, the manipulator cannot have a disadvantage in the partial information setting compared to the full information setting.*

Proof. Suppose we have some full profile $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ and some truthful preference R of the manipulator such that they have an incentive to manipulate in bmR' with a best response is R'_b. Now we show that the manipulator also has an incentive to manipulate with R'_b in $G_\Sigma(\boldsymbol{R}')$.

Towards a contradiction, suppose this is not the case. Since $\boldsymbol{R}' \in E_{G_\Sigma(\boldsymbol{R}')}$ by the definition of G_Σ, this can only hold if there is some $\boldsymbol{R}'' \in E_{G_\Sigma(\boldsymbol{R}')}$ such that $F(\boldsymbol{R}'', R) \succ_R F(\boldsymbol{R}'', R'_b)$. However, since F is strongly computable from Σ, we have that $F(\boldsymbol{R}', R) = F(\boldsymbol{R}'', R)$ and $F(\boldsymbol{R}', R'_b) = F(\boldsymbol{R}'', R'_b)$ and since we know by assumption that $F(\boldsymbol{R}', R'_b) \succ_R F(\boldsymbol{R}', R)$, this can never be the case. □

6.4 Conclusion

In the previous chapters, we have seen that partial information does not always imply that a manipulator does not have an incentive to manipulate anymore and we have not been able to show manipulation becomes more significantly more difficult. In this chapter, we have investigated other ways the manipulator could be worse off. First, we looked how the partial information setting compared to the full information setting. We suggested to define a partial information function G_Σ that given some full profile $\mathbf{R}'$ determines which partial profile is the corresponding partial situation. For TOPBOT, BOT(t) and TOP(t), the definition of G_Σ is straightforward, for COMPL(t) it is not. Second, we looked at how we could formalize the notion of disadvantage. Here we suggested to compare concrete situations in which a manipulator can manipulate in the full information setting, but no longer with partial information setting. Finally, we looked at what the notion of less information could mean formally. Here, we considered both communication complexity and the strong computability.

Chapter 7

Discussion and Outlook

In the previous chapters, we have investigated the influence of the structure of partial information on the manipulability of different voting rules. The main conclusion of this investigation was that the available knowledge about the preferences of others impacts the manipulability of several voting rules. In other words: what you about people's preferences matters. We have provided evidence for this claim from two perspectives: showing that there are situations in which the manipulator has an incentive to manipulate and investigating the difficulty of manipulation. We are aware that there are other ways and methods which can be applied to investigate the research question. Due to limits in time, this thesis mainly focused on the two above mentioned approaches, because we believe they are the most straightforward way of answering this question and correspond to the approaches used in the existing literature.

We believe, however, that a critical discussion of previous literature is necessary to go forward. In Chapter 6, we have already investigated different notions of disadvantage for the manipulator. To provide an outlook for future research, this chapter shortly considers and discusses alternative and interesting ways of investigation which may advance future research on how partial information influences manipulability. Three aspects will be central in discussing these alternative methods. First, we look at alternatives to the set-up of this thesis. Second, we consider an alternative measure of manipulability where we see it as a scale. Finally, we investigate another paradigm to measure the difficulty of manipulation.

7.1 Alternative Set-Ups

Throughout this study, we compared four structures with three different voting rules. In this section, we will focus on which other voting rules and settings would be interesting to consider for future work.

Structures

We already considered some possible limitations of our structures in the discussion of Chapter 3. For example, since the amount of alternatives that the manipulator knows from the preference of the other voters is fixed by the parameter t, we believe it would be interesting to also look at structures in which the amount of alternatives a manipulator knows the location of per structure differs. In particular, we believe it would be interesting to compare our work with the more general notions defined by Erdélyi and Reger (2016).[1]

Second, we found that all three voting rules showed a difference between the structures that contain some information about all of the voters (TOPBOT, TOP(t), and BOT(t)) and the structure that contained all information about the preference of some of the voters (COMPL(t)). In the first case, we found a constant value for t for which it was possible to manipulate, in the second case the value of t depended on the input size n. However, there is a difference between these two groups of structures. In the case of TOP(t) and BOT(t), the amount of alternatives that the manipulator knows in total grows when n grows, even t remains constant. In the case of COMPL(t), this is not the case. Here it does hold, however, that the number of alternatives that the manipulator knows grows when the total number of alternatives increases, even when t is constant. We believe this difference is consistent with the assumptions we started with: knowing something about all the voters implies you have more information when there are more voters, knowing everything about some of the voters implies you have more information when there are more alternatives. One of the potential future research directions would be to investigate this phenomenon in more detail, for example by defining t in such a way that it measures the amount of information the manipulator knows in total.

Tournament Rules

The present study has limited its investigation to k-approval, Borda, and Copeland. By looking at k-approval, we have included one of the most used and well-known voting rules: plurality. By comparing Borda and Copeland, we attempted to see whether the fundamental difference between Condorcet extensions and positional scoring rules influences the manipulability with partial information. By focusing on these three rules, we covered two out of the three types of rules as explained in Chapter 2. For future research, therefore, it would be interesting to consider a tournament rule, such as single transferable vote, and see if our findings are also consistent with this type of rule.

Ties

We only considered resolute voting rules, which are rules that always elect a unique winner. To break ties between alternatives, we use lexicographic tie-breaking, which is a relatively simple way of breaking ties. This is also the procedure used in comparable work (see e.g. Reijngoud and Endriss, 2012; Endriss et al., 2016). The tie-breaking order does, however, influence both axiomatic and computational properties of voting rules (see e.g. Freeman et al., 2015). For example, all three rules are no longer neutral.[2] We believe that for future

[1] See Chapter 3 for more information about these structures

[2] Recall from Chapter 2 that a voting rule is neutral when swapping alternatives in the input corresponds to swapping the alternatives in a similar way in the output. When an alternative is tied, however, this does

work we could investigate whether different tie-breaking procedures lead to different results.

Another way to deal with tie-breaking is to simply allow them in the outcome. We have already shown that there are exist impossibility results for irresolute voting rules as the Gibbard-Satterthwaite Theorem (see e.g. Duggan and Schwartz, 2000). For this reason, we believe it would be very relevant to investigate whether similar results can be obtained by investigating irresolute rules to see whether differences between these structures also result in differences for these type of rules.

Higher order reasoning

Throughout this study, we only considered situations in which all voters vote honestly, or at least in which the manipulators assume that they do. In practice, however, all voters can manipulate and manipulators can also take this information into account. To model this, more complex models can be applied in which manipulators also take the strategies of other voters into account. A popular approach to study this is to see voting as a strategic game and apply several game-theoretic notions to predict who will win. Several models of strategic reasoning are also applied to the context of voting (see e.g. Obraztsova and Elkind, 2012) and social choice in general.[3] We think it would be interesting to investigate how different structures of uncertainty influence different game-theoretic properties of voting.

7.2 Manipulation as a Scale

In the previous chapters, we have viewed susceptibility to manipulation as a binary concept: a voting rule is either susceptible or immune to manipulation. Alternatively, we could also consider *how many* profiles there exist that can be manipulated. Kelly (1993), for example, argued that a voting rule F_1 is more manipulable than a voting rule F_2 when there are more profiles in which there is a voter who has an incentive to manipulate in F_1 rather than F_2. The idea behind this is to not define manipulability as some property that a rule has or does not have, but to consider manipulability as a scale and look at how often manipulation occurs. Most of the existing research concerning counting and simulating the number of manipulable profiles focus on the setting in which voters have full information. Veselova (2016), however, runs experiments in which she assumes that voters do not have full information about the preferences of others. She used the model of Reijngoud and Endriss (2012) to define uncertainty.

This approach could be used to compare structures in a different way than we did in Chapter 4. Instead of looking whether it is possible to manipulate, we could investigate how often manipulation occurs. For example, although manipulation for two structure is

not hold. Suppose alternative A and B receive the same amount of points and the tie-breaking decides that A wins. Now, if we swap A and B in the input, the two alternatives still tie so A still wins. However, from neutrality it follows that we should apply the same permutation to the output so B should have won.

[3]Interesting related work is done by Terzopoulou (2017), who combines higher-order reasoning with partial information, but in the context of judgment aggregation, a generalization of voting that studies the aggregation of individual judgments regarding the truth-value of logically related propositions.

possible for the same value t, it can be the case that one structure has more profiles in which one can manipulate than in the other. A similar approach can be used to compare different voting rules. Since the amount of profiles grows exponentially with m, we think these type of results can provide us with an interesting alternative perspective that might tell us more about the probability that manipulation occurs in practice.

However, an important side note needs to be addressed here. This research is all based on the *impartial culture assumption* which means that every profile is considered to be equally likely. This assumption is often not found empirically plausible, which means this assumption will probably not hold in practice (Regenwetter et al., 2006). For this reason, this approach might be more suited to apply when a specific domain is known in which we know more about the actual probability distribution of the profiles.

7.3 Another Difficulty Measure

In this thesis, we have investigated the difficulty of manipulation using classical complexity theory. While this is an approach that is often used, especially in the context of voting (see e.g. Brandt et al., 2016b), other paradigms exist. One interesting alternative is *parameterized complexity*. While in classical complexity theory the complexity is expressed *only* in the terms of the input, here a problem is studied both in terms of the input size and in terms of several parameters. Such results provide a more detailed analysis and can show whether the computational intractability of a problem depends on some specific aspect of the input. For example, while VERTEX COVER is one of Karp's NP-complete problems (Karp, 1972), it is fixed-parameter tractable with respect to the size of the vertex cover that needs to be found. This tells us that even though vertex cover is hard to solve in general, when the size of the vertex cover is small, efficient algorithms can probably be found in practice (Downey and Fellows, 1999). Moreover, parameterized complexity results may help to justify computational barriers against manipulation. As said before, NP-hardness might not prevent manipulation in practice. If we, however, find that a problem belongs to a hardness class with respect to a parameter k, it is unlikely that an efficient algorithm can be found to solve it, even when k is small (Dorn and Schlotter, 2017). We, therefore, believe that a parameterized complexity analysis of the MANIPULATION problem could be an interesting addition to the classical approach.

In the context of collective decision making, several natural parameters exist, such as the number of voters, the number of alternatives, or the 'distance' from an instance to an instance with a desirable property such as single-peakedness. In the case of partial information, we can define several parameters which 'measure' the uncertainty of the preferences of other voters, for example the number of undetermined alternatives. Moreover, some parameterized complexity results of strategic voting behavior are already found in the literature. Betzler et al. (2009) show, for example, that for all positional scoring rules, Copeland, and some other natural rules the POSSIBLE WINNER PROBLEM is fixed-parameter tractable with respect to the number of candidates, but cannot be fixed-parameter tractable with respect to the number of voters for Copeland and k-approval. For future work, we believe it would be interesting to investigate if such results can be generalized for Manipulation.[4]

7.4 Conclusion

In final chapter of this thesis investigated alternative ways we could consider the difference between restricted partial information and the full or general partial information setting. This chapter showed therefore some important and interesting directions for future research, which further build on the result of this thesis. We focused on three aspects. First, we looked at which other voting rules and settings would be interesting to investigate. We mainly believe that considering a tournament rule, such as STV, more elaborate ways of tie-breaking, and resolute rules would be interesting future research that can show whether our findings are consistent among a broader part of computational social choice. Second, we explained how treating manipulation as a scale rather than some binary concept can provide us with some extra more fine-grained notion of the probability of manipulation in practice. Finally, Moreover, we explained how parameterized complexity can help us to understand the computational properties of manipulation into more depth.

[4]Note that this does not automatically follow from the observation of Section 5.2, in which we argued that there is a polynomial algorithm when the number of candidates is fixed. A parameterized problem is fixed-parameter tractable for some parameter k when it runs in $O(f(k) \cdot p(|x|))$ time, where $|x|$ is the size of the input, p some polynomial function and f any function. The brute-force algorithm described in section 5.2 runs in $O(m! \cdot n^m)$ time, so this does not suffice to show the problem is fixed-parameter tractable.

Chapter 8

Conclusion

In this thesis, we have investigated how partial knowledge about the preferences of other voters influences the manipulability of several voting rules. We did this research in the context of three voting rules: k-approval, Borda, and Copeland. Moreover, we have modeled partial information by restricted partial orders of four types: TOPBOT in which the voters knows top and bottom alternative of each voter, TOP(t) in which the manipulator knows the top t alternatives of each voter, BOT(t) in which the manipulator knows the bottom t alternatives of each voter, and COMPL(t) in which the manipulator knows the complete preference of t voters and nothing about the other voters.

We have considered manipulability from two perspectives. In Chapter 4, we have investigated in which cases manipulation was and was not possible. For all structures and voting rules we have found values of t in which the manipulator had the incentive to manipulate. This means that partial information, in general, does not prevent manipulators to misreport their preferences. For some specific values of t, however, this is the case. Particular noteworthy was the fact that all three voting rules showed a difference between the three structures in which the manipulator has some information about all of the voters (TOPBOT, TOP(t), and BOT(t)) and the case where the manipulator knows the complete preferences of some of the voters and nothing about the other voters (COMPL(t)). In the first case, we found some constant value of t for which this is possible, in the second case t depended on the number of voters.

In Chapter 5, we have investigated the computational complexity of the MANIPULATION problem. We have showed that for k-approval this problem could be solved in polynomial time, for every structure that was both IP- and PI transitive, which includes all structures Σ that we defined initially. Moreover, we have observed that most NP-hardness proofs of related problems reduce to instances that are not IP- and PI-transitive. For this reason, we have posed the conjecture that for most natural voting rules, these results to do not translate to classes of instances that only contain structures that are IP- and PI-transitive.

In Chapter 6, we have investigated other notions of disadvantage the manipulator can have when having partial information compared to full information. In this chapter, we mainly provided some formalizations of how this question could be investigated in the future using existing theories. Finally, in Chapter 7, we have investigated alternatives ways in which our research question could have been answered. This showed some important and interesting directions of future work, which can help to strengthen the results found in this thesis.

To sum up, we have shown that partial information influences the manipulability of voting rules: it matters *what* you know about the other voters.

List of Symbols

$\wp$ A function that returns all subsets of a set (e.g. the powerset)

N The set of all n voters ($N \subseteq \mathbb{N}$)

$\mathcal{X}$ The set of all m alternatives or candidates

$\mathcal{L}(\mathcal{X})$ The set of all linear orders over the set of alternatives, containing all full preferences R

$\succ_R$ The preference relation R, if $X \succ_R Y$ then $(X, Y) \in R$, e.g. X is preferred over Y with respect to R. Moreover, if $\mathcal{X}' \subseteq \mathcal{X}$ and $A \in \mathcal{X}$ for some set of alternatives $\mathcal{X}$, we write $\mathcal{X}' \succ_R A$, when $B \succ_R A$ for every $B \in \mathcal{X}'$ and $A \succ_R \mathcal{X}'$ when $A \succ_R B$ for every $B \in \mathcal{X}'$

$\rhd$ The tie-breaking order. If $A \rhd B$, A is preferred over B with respect to $\rhd$

top A function that returns the top alternative of some preference order R

top_k A function that returns the set of top k alternatives of some preference order R

$\mathcal{P}(\mathcal{X})$ The set of all partial orders over the set of alternatives, containing all partial preference R

$\sim_R$ The incomparable relation, defined for partial preferences R. If $A \sim_R B$ then both $(A, B) \notin R$ and $(B, A) \notin R$

$(\boldsymbol{R}, R)$ The profile in which preference R is added to $\boldsymbol{R}$

$(\boldsymbol{R}_{-i}, R')$ The profile that corresponds to $\boldsymbol{R}$ but in which R_i is replaced by R'

$E_{\boldsymbol{R}}$ The information set induced by partial order $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})$

STRUCT The set that contains all structures Σ we consider in this thesis, see Definition 3.2

$[R]_k$ The k-equivalence class of a preference $R \in \mathcal{L}(\mathcal{X})$, containing all preference orders that have the same top k alternatives. See Definition 4.2

$[\boldsymbol{R}]_k$ The k-equivalence class of a profile contains all profiles from which the preference orders of similar voters are in the same k-equivalence class. See Definition 4.3

IPPI A structure that contains all partial orders that are both IP- and PI-transitive

pnts The function $pnts(\boldsymbol{R}, Z)$ denotes the amount of points an alternative Z has (for certain) in $\boldsymbol{R}$

pp If $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})$, the function $pp(\boldsymbol{R}, Z)$ denotes how much points an alternative Z can have it is places in the top k whenever possible

$C_{\boldsymbol{R}}(X, Y)$ The choice set, see Definition 5.2

$[\![B]\!]$ IF B some boolean expression, then we define $[\![B]\!] = 1$ as the numerical truth-value of B, see Definition 5.3

G_Σ The partial information function that given any $\Sigma \in$ STRUCT partial and some full profile $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ returns some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ such that $E_{\boldsymbol{R}} \in \Sigma$

R'_b The best response of a manipulator with respect to some full profile $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^n$ and true preference R, which means there is no $R'' \in \mathcal{L}(\mathcal{X})$ with $F(\boldsymbol{R}, R'') \succ_R F(\boldsymbol{R}, R'_b)$

Bibliography

Ahuja, R., Magnanti, T., and Orlin, J. (1993). *Network Flows: Theory, algorithms, and applications*. Pearson.

Arora, S. and Barak, B. (2007). *Computational Complexity: a Modern Approach*. Cambridge University Press.

Arrow, K. (1963). *Social Choice and Individual Values*. John Wiley and Sons.

Bartholdi, J. and Orlin, J. (1991). Single Transferable Vote Resists Strategic Voting. *Social Choice and Welfare*, 8(4):341–354.

Bartholdi, J., Tovey, C., and Trick, M. (1989). The Computational Difficulty of Manipulating an Election. *Social Choice and Welfare*, 6(2):227–241.

Baumeister, D., Faliszewski, P., Lang, J., and Rothe, J. (2012). Campaigns for Lazy Voters: Truncated ballots. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 577–584.

Benoît, J.-P. (2000). The Gibbard – Satterthwaite Theorem: A Simple Proof. *Economics Letter*, 69(1):319–322.

Betzler, N. and Dorn, B. (2010). Towards a Dichotomy for the Possible Winner Problem in Elections Based on Scoring Rules. *Journal of Computer and System Sciences*, 76(8):812–836.

Betzler, N., Hemmann, S., and Niedermeier, R. (2009). A Multivariate Complexity Analysis of Determining Possible Winners Given Incomplete Votes. *Proceedings of 21st International Joint Conference of Artificial Intelligence (IJCAI)*, pages 53–58.

Black, D. (1948). On the Rationale of Group Decision-making. *Journal of Political Economy*, 56(1):23–34.

Brandt, F., Brill, M., and Harrenstein, P. (2016a). Tournament Solutions. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A., editors, *Handbook of Computational Social Choice*, chapter 3, pages 57–84. Cambridge University Press.

Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. (2016b). *Handbook of Computational Social Choice*. Cambridge University Press, New York.

Briskorn, D., Erdélyi, G., and Reger, C. (2016). Bribery in k-Approval and k-Veto Under Partial Information (Extended Abstract). In *Proceedings of the 15th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1299–1301.

Conitzer, V. and Sandholm, T. (2005). Communication Complexity of Common Voting Rules. In *Proceedings of the 6th ACM Conference on Electronic Commerce (ACM-EC)*, pages 78–87.

Conitzer, V. and Sandholm, T. (2006). Nonexistence of Voting Rules That Are Usually Hard to Manipulate. In *Proceedings of the American Association for Artificial Intelligence (AAAI)*, volume 6, pages 627–634.

Conitzer, V., Sandholm, T., and Lang, J. (2007). When Are Elections with Few Candidates Hard to Manipulate? *Journal of the ACM*, 54(3):14–47.

Conitzer, V. and Walsh, T. (2016). Barriers to Manipulation in Voting. In F. Brandt, V. Conitzer, U. Endriss, J. Lang and Procaccia, A., editors, *Handbook of Computational Social Choice*, chapter 6, pages 127–145. Cambridge University Press.

Conitzer, V., Walsh, T., and Xia, L. (2011). Dominating Manipulations in Voting with Partial Information. In *Proceedings of the 25th American Association for Artificial Intelligence (AAAI)*, volume 11, pages 638–643.

Damanik, R., Ionescu, S., Seip, P., and van Oostveen, L. (2017). Strategic Manipulation for Voting under Distance-Based Uncertainty. Technical report, University of Amsterdam.

Dey, P., Misra, N., and Narahari, Y. (2016). Complexity of Manipulation with Partial Information in Voting. *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 229–235.

Dorn, B. and Schlotter, I. (2017). Having a Hard Time? Explore Parameterized Complexity! In Endriss, U., editor, *Trends in Computational Social Choice*, pages 209–239. AI Access.

Downey, R. G. and Fellows, M. R. (1999). *Parameterized Complexity.* Springer.

Duggan, J. and Schwartz, T. (2000). Strategic Manipulability without Resoluteness or Shared Beliefs: Gibbard-Satterthwaite Generalized. *Social Choice and Welfare*, 17(1):85–93.

Dummett, M. and Farquharson, R. (1961). Stability in Voting. *Econometrica*, 29(1):33–43.

Edmonds, J. and Karp, R. (1972). Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems. *Journal of the ACM*, 19(2):248–264.

Endriss, U. (2011). Logic and Social Choice Theory. *Logic and Philosophy Today*, 2:333–377.

Endriss, U. (2017). Computational Social Choice: Voting Rules. (Lecture notes), Available from: `https://staff.science.uva.nl/u.endriss/teaching/comsoc/2017/slides/comsoc-voting-rules-2017.pdf`.

Endriss, U., Obraztsova, S., Polukarov, M., and Rosenschein, J. (2016). Strategic Voting with Incomplete Information. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*.

Erdélyi, G. and Reger, C. (2016). Possible Bribery in k-Approval and k-Veto Under Partial Information (Long Version). In *Proceedings of the 17th International Conference on Artificial Intelligence: Methodology, Systems, and Applications (AIMSA)*, pages 299–309.

Fischer, F., Hudry, O., and Niedermeier, R. (2016). Weighted Tournament Solutions. In Brandt, F., Conitzer, V., Endriss, U., and Lang, J., editors, *Handbook of Computational Social Choice*, chapter 4, pages 85–105. Cambridge University Press.

Freeman, R., Brill, M., and Conitzer, V. (2015). General Tiebreaking Schemes for Computational Social Choice. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1401–1409.

Gibbard, A. (1973). Manipulation of Voting Schemes: A General Result. *Econometrica*, 41(4):587–601.

Hansson, S. (2001). *The Structure of Values and Norms*. Cambridge University Press.

Harris, T. and Ross, F. (1955). *Fundamentals of a Method for Evaluating Rail Net Capacities*. Rand Corp.

Karp, R. (1972). Reducibility Among Combinatorial Problems. In Miller, R., Thatcher, W., and Bohlinger, J., editors, *Complexity of Computer Computations*, pages 85–103. Springer, Boston, MA.

Kelly, J. (1993). Almost All Social Choice Rules Are Highly Manipulable, But a Few Aren't. *Social Choice and Welfare*, 10(2):161–175.

Kushilevitz, E. and Nisan, N. (1997). *Communication Complexity*. Cambridge University Press.

MacKay, D. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.

Marquis de Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la probabilité des voix*.

Meir, R., Lev, O., and Rosenschein, J. (2014). A Local-Dominance Theory of Voting Equilibria. In *Proceedings of the 15th ACM Conference on Economics and Computation*, pages 313–330.

Meir, R., Polukarov, M., Rosenschein, J., and Jennings, N. (2010). Convergence to Equilibria in Plurality Voting. In *Proceedings of 24th Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence (AAAI)*, pages 823–828.

Moulin, H. (1980). On Strategy-Proofness and Single Peakedness. *Public Choice*, 35(4):437–455.

Obraztsova, S. and Elkind, E. (2012). Optimal Manipulation of Voting Rules. In *Proceedings of the 26th Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence (AAAI)*, pages 2141–2147.

Regenwetter, M., Grofman, B., Marley, A., and Tsetlin, I. (2006). *Behavioral Social Choice: Probabilistic Models, Statistical Inference, and Applications.* Cambridge University Press.

Reijngoud, A. and Endriss, U. (2012). Voter Response to Iterated Poll Information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS).*

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

Terzopoulou, Z. (2017). *Manipulating the Manipulators: Richer Models of Strategic Behavior in Judgment Aggregation.* Master thesis in logic, University of Amsterdam.

Veselova, Y. (2016). Does Incomplete Information Reduce Manipulability? Technical report, National Research University Higher School of Economics (HSE).

Walsh, T. (2010). Is Computational Complexity a Barrier to Manipulation? In *Proceedings of the 11th international Workshop on Computational Logic in Multi Agent Systems (CLIMA)*, volume 6245, pages 1–7.

Xia, L. and Conitzer, V. (2011). Determining Possible and Necessary Winners under Common Voting Rules given Partial Orders. *Journal of Artificial Intelligence Research*, 41:25–67.

Young, H. (1988). Condorcet's Theory of Voting. *The American Political Science Review*, 82(4):1231–1244.

Zwicker, W. (2016). Introduction to the Theory of Voting. In F. Brandt, V. Conitzer, U. Endriss, J. Lang and Procaccia, A., editors, *Handbook of Computational Social Choice*, chapter 2, pages 23–56. Cambridge University Press.

Appendix

Part of the proof of Lemma 4.2

Proof. What is left to show is that the theorem also holds when $k = 1$. So, suppose $k = 1$ and take some arbitrary $n, m \geq 3$, let $\mathcal{X} = \{A, B, ..., Z\}$ with $|\mathcal{X}| = m$ and the ties be broken by $A \rhd B \rhd \cdots \rhd Z$. Now, suppose that for any $E_{\boldsymbol{R}} \in \Sigma$ we know that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ it holds that $E_{\boldsymbol{R}} \subseteq [\boldsymbol{R}']_k$. This again implies that for any $\boldsymbol{R}$ with $E_{\boldsymbol{R}} \in \Sigma$, all $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ are in the same profile equivalence class.

First, suppose n is odd and the manipulator has the true preference $R : Z \succ B \succ A \succ \cdots$. Now, we can consider the $\boldsymbol{R}$ in which for every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that for half of the other voters $top(R_i) = A$ and the other half $top(R_i) = B$. If the manipulator votes R, it holds that $F(\boldsymbol{R}', R) = A$. If the manipulator votes $R' : B \succ Z \succ A \succ \cdots$ instead, we know that $F(\boldsymbol{R}', R') = B$ holds for all $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. So, the manipulator has an incentive to manipulate.

Finally, suppose n is even. Now, we suppose the manipulator has a true preference $Z \succ A \succ B \succ \cdots$. We can consider the $\boldsymbol{R}$ such that t for half of the other voters $top(R_i) = A$ and the other half $top(R_i) = B$ and there is one voter more who has B on top than A. Now, if the manipulator votes R, it holds that $F(\boldsymbol{R}', R) = B$ for all $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. If the manipulator votes $R' : A \succ Z \succ B \succ \cdots$ instead, we have that $F(\boldsymbol{R}', R') = A$, since $A \rhd B$. Again, the manipulator has an incentive to manipulate. $\square$

Proof of Theorem 4.2

Proof. Take some arbitrary $m \geq 3$ and $n \geq 3$ and let $\mathcal{X} = \{A, B, ..., Z\}$ with $|\mathcal{X}| = m$, where $A \rhd B \rhd \ldots \rhd Z$. Let F denote k-approval with this tie-breaking scheme. Similar as in in Theorem 4.1 we let the manipulator have the true preference $R : B \succ \cdots \succ A \succ \cdots \succ Z$ such that for all $C \succeq_R A$ it holds that $C \in top_k(R)$ and for all Y such that $A \succ_R Y$ we have $Y \notin top_k(R)$ and the untruthful profile R' in which is similar to R except that Z receives one point instead of A.

Suppose $t \geq k$. Now we know that for every $E_{\boldsymbol{R}} \in \mathrm{TOP}(t)$ it must hold that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $E_{\boldsymbol{R}} \subseteq [\boldsymbol{R}]_k$. So, from Lemma 4.2 we know that k-approval with lexicographic tie-breaking is susceptible to $\mathrm{TOP}(t)$-manipulation.

Now, suppose $t < k$. Note that this implies $k > 1$, since we assumed $t > 0$. Now, let the partial profile $\boldsymbol{R}$ consists of only voters with the partial preference R_1 such that $B \in top_k(R_1)$, which other alternatives are fixed does not matter for the proof. Now, we will show that (1) there exists a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$ and (2) that there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

(1) We can extend $\boldsymbol{R}$ to $\boldsymbol{R}^*$ as follows: we extend R_1 to R_1' such that $A \in top_k(R_1')$. Since $t < k$, this is always possible. Now with a similar argument as in Theorem 4.1, we have found an $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}^*, R') = B$ and $F(\boldsymbol{R}^*, R) = A$, and since $B \succ_R A$, (1) holds.

(2) Note that for every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ we know that if the manipulator votes R', B has n points and A has at most $n-1$ points. All other $X \in \mathcal{X} \setminus \{A, B\}$ have at most n points, but it also holds that $B \rhd X$. So, it must hold that $F(\boldsymbol{R}', R') = B$. Since there is not $Y \in \mathcal{X}$ such that $X \succ_R B$, (2) must hold.

□

Proof of Lemma 4.3

Proof. We take an arbitrary $m, n \geq 3$, let $\mathcal{X} = \{A, B, ..., Z\}$ be the set of m alternatives and the usual tie-breaking order $A \rhd \cdots \rhd Z$. Now, we let the manipulator have the opposite preference order as the $\rhd$, that is $R = Z \succ \cdots \succ A$. Moreover, we let X denote the highest ranked alternative that is not in the top k of R. If $t \geq m - k$, then we have that $E_{\boldsymbol{R}} \subseteq [\boldsymbol{R}']_k$ for any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ and we know from Lemma 4.2 that the Theorem holds. So, we assume that $t < m - k$.

Now, we construct $\boldsymbol{R}$ by placing all alternatives in $top_k(R)$ in the bottom t of every R_i in $\boldsymbol{R}$ and if $t > k$, dividing $\mathcal{X} \setminus top_k(R) \cup \{X, A\}$ as equally among the remaining places in the bottom t of every R_i in $\boldsymbol{R}$. We know this is always possible because from $m \geq 3$, $t \geq k$, and $t > m - k - t$, it follows that $t \geq 2$ thus $|\mathcal{X} \setminus top_k(R) \cup \{X, A\}| \geq m - k - t$. If $k = 1$ and n is even, make sure there is one R_i in $\boldsymbol{R}$ such that some C other than Z, X or A is not fixed in the bottom t. This is also always possible since we know from $t \geq 2$ and the assumption that $t < m - k$ that $m \geq 4$.

Note that all $D \in top_k(R)$ receive zero points in every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, so it cannot be the case that $F(\boldsymbol{R}', R) = D$ for any $D \in top_k(R)$, since there will always be some $C \notin top_k(R)$ that receives at least one point in $\boldsymbol{R}'$ and by construction $C \rhd D$. Now, let R' be the ballot such that Z and X are swapped in R, such that X receives one point instead of Z.

Towards a contradiction, suppose that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = E$ and $F(\boldsymbol{R}', R') = E'$ with $E \succ_R E'$. Now, by the properties of k-approval it must hold that either $E \in top_k(R)$ and $E \notin top_k(R')$, or $E' \in top_k(R')$ and $E' \notin top_k(R)$. We have just shown that the first case cannot be possible, since $E \in top_k(R)$ implies it cannot be that $F(\boldsymbol{R}', R) = E$. The second case, however, implies that $E' = X$, but since X denotes the highest ranked alternative not in top k, $E \succ_R X$ implies that $E \in top_k(R)$ and we have just argued this cannot be the case. Contradiction! So, such $\boldsymbol{R}'$ does not exists.

What is left to show is that there exists some $\boldsymbol{R}^*$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. If $k \geq 2$, let X and A be in the top_k of every R_i^* in $\boldsymbol{R}^*$, then since $n \geq 3$ and $A \rhd Y$ for any other $Y \in \mathcal{X}$ we get $F(\boldsymbol{R}^*, R') = X$ and $F(\boldsymbol{R}^*, R) = A$. If $k = 1$, we alternate between putting X and A on the top of R_i^*. If n is even, by construction there is one R_i that we can extend to R_i^* such that some C other than Z, X or A is on top. So we have again have that $F(\boldsymbol{R}^*, R') = X$ and $F(\boldsymbol{R}^*, R) = A$. □

Proof of Theorem 4.5

Proof. Suppose $t = 0$, and take an arbitrary n with both $n \geq 3$ and $n \geq \lceil \frac{m-2}{m-k} \rceil$. Note that if $E_{\boldsymbol{R}} \in \text{TOP}(0)$, we know that $E_{\boldsymbol{R}} = \mathcal{L}(\mathcal{X})^{n-1}$, since $t = 0$ implies that $\boldsymbol{R} = (\emptyset, ..., \emptyset)$.

Now, towards a contradiction suppose that there is some manipulator with a true preference R that has an incentive to manipulate with a preference $R' \neq R$. This means there is some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. Now, we know that there must be some X, Y such that $X \in top_k(R')$ and $X \notin top_k(R)$ and $Y \notin top_k(R')$ and $Y \in top_k(R)$. Hence, it holds that $Y \succ_R X$. Now, we claim that there exists some $\boldsymbol{R}' \in E_{\boldsymbol{R}} = \mathcal{L}(\mathcal{X})^{n-1}$ such that $F(\boldsymbol{R}', R') = X$ and $F(\boldsymbol{R}', R) = Y$.

If $k \geq 2$, we can construct $\boldsymbol{R}'$ by creating $n - 1$ profiles such that X and Y are both in the top k, and every $Z \in \mathcal{X} \setminus \{X, Y\}$ is not in the top k at least one time. This is always possible since we assumed $n \geq \lceil \frac{m-2}{m-k} \rceil$. Now, if the manipulator votes R', X is the only alternative with n points, so $F(\boldsymbol{R}', R') = X$, and similarly $F(\boldsymbol{R}', R) = Y$.

If $k = 1$, we can create the profile $\boldsymbol{R}'$ in which X and Y are both $\lfloor \frac{n-1}{2} \rfloor$ times on top. If n is even, (so $n-1$ is odd) we have one preference left to extend, in which we put Y on top when $X \rhd Y$ and X on top otherwise. Now, all $Z \in \mathcal{L}(\mathcal{X})^n$ have 0 points, while X and Y have at least 2 points when the manipulator votes R' or R respectively. So, we similarly found an $\boldsymbol{R}'$ such that $F(bfR', R') = X$ and $F(\boldsymbol{R}', R) = Y$.

Contradiction! So, k-approval is immune to TOP(0)-manipulation. □

Proof of Theorem 4.6

Proof. Towards a contradiction, suppose there are some $\boldsymbol{R}$ $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ and R, R' such that the manipulator has an incentive to manipulate. This means that $F(\boldsymbol{R}^*, R') = E'$ and $F(\boldsymbol{R}^*, R) = E$ with $E' \succ_R E$. From the structure of k-approval, this means that either (1) $E' \in top_k(R')$ and $E' \notin top_k(R)$ or (2) $E \in top_k(R)$ and $E \notin top_k(R')$.

(1) Suppose that $E' \in top_k(R')$ and $E' \notin top_k(R)$. We will show that this means that there must be some $C \in top_k(R)$ and some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = C$ and $F(\boldsymbol{R}', R') = E'$. This is a contradiction with the assumption that the manipulator has an incentive to manipulate with R'.

First, in every $\boldsymbol{R}$ it must the case that there is some $C \in top_k(R)$ that is fixed in the bottom at most

$$\left\lfloor \frac{(n-1) \cdot t}{k} \right\rfloor$$

times, since this is the case in which all k alternatives are divided among the t bottom positions of every $n - 1$ voters. Let q denote the amount of times that C is not fixed in the bot t of some R_i in $\boldsymbol{R}$. Moreover, we let q' denote the amount of times E' is not fixed in the bot t of $\boldsymbol{R}$. Now, if $q' > q$, we construct some $\boldsymbol{R}'$ such that C is on top as often as possible, and E' as often such that if all other alternatives receive less points than E', $F(\boldsymbol{R}', R) = C$ and $F(\boldsymbol{R}', R') = E'$. This depends on $\triangleright$ and whether $C \in top_k(R')$, but is clearly always possible since $q' > q$

What is left to show is that there is no $D \notin top_k(R)$ that has more points than C or E' in $\boldsymbol{R}'$. However, note that if we distribute the lower $m-k$ alternatives equally over $m-k-t$ unfixed positions of $\boldsymbol{R}'$, that the most points any of such alternative can get is:

$$(n-1) - \lfloor \frac{(n-1)\cdot(m-k-t)}{(m-k)} \rfloor$$

But, from our assumption this is less than $(n-1)-q$, so $F(\boldsymbol{R}', R) = C$ and $F(\boldsymbol{R}', R') = E'$ holds.

If $q' \leq q$, we can create a $\boldsymbol{R}'$ in which E' is q' times on top and C on top such that $F(\boldsymbol{R}', R) = C$ and $F(\boldsymbol{R}', R') = E'$ holds when the other alternatives receive less points. Since $F(\boldsymbol{R}^*, R') = E'$, we know that we must be able to divide the other alternatives this way.

(2) Otherwise, suppose that $E \in top_k(R)$ and $E \notin top_k(R')$. In this case, we will show that there must be some $D \notin top_k(R)$ such that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ and $F(\boldsymbol{R}', R) = E$ and $F(\boldsymbol{R}', R') = D$. This again gives us a contradiction with the assumption that the manipulator has an incentive to manipulate and proves this Theorem.

Similarly as in the previous case, we know that there must be one $D \notin top_k(R)$ that is not fixed more than:

$$(n-1) \cdot \left\lfloor \frac{(n-1)\cdot t}{(m-k)} \right\rfloor$$

times in the bot t of some R_i in $\boldsymbol{R}$. Since we assumed that

$$\left\lfloor \frac{(n-1)\cdot t}{(m-k)} \right\rfloor < \left\lfloor \frac{(n-1)\cdot(m-k-t)}{k} \right\rfloor$$

we can use a similar construction as above to show that there must be some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = E$ and $F(\boldsymbol{R}', R') = D$.

□

Proof of Theorem 4.9

Proof. We take some arbitrary $t > 0$, $n \geq 3$ and $m \geq 4$ and let $\mathcal{X} = \{A, B, C, D...\}$ with $|\mathcal{X}| = m$ and $A \triangleright B \triangleright C \triangleright D \triangleright \cdots$. Let F denote Borda in combination with lexicographic tie-breaking. Now, if $t \geq m-1$, we have that $|E_{\boldsymbol{R}}| = 1$ for every $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^{n-1}$ with $E_{\boldsymbol{R}} \in \text{BOT}(t)$ and since F is clearly surjective, by Lemma 4.1 we can conclude that Borda in combination with lexicographic tie-breaking is susceptible to BOT(t)-manipulation.

So, we suppose that $t < m - 1$. Now, we define $\boldsymbol{R}$ such that A is fixed on the bottom everywhere, and B and C are never fixed on the bottom. Since we assumed $t < m - 1$, this is always possible. Now, we let the manipulator have the true preference $R = \rhd$ and R' be similar to R but with A and B swapped, that is $R' : B \succ A \succ C \succ \cdots$. We will show that (1) there exists some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$ and (2) there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

(1) We distinguish between two cases:

  - when is n is odd, we extend $\boldsymbol{R}$ to $\boldsymbol{R}^*$ by letting two voters vote $C \succ B \succ \cdots \succ A$, and let the rest of the $n - 3$ voters alternate between $B \succ C \succ \cdots \succ A$ and $C \succ B \succ \cdots \succ A$.
  - when n is even, we extend one voter to $C \succ D \succ B \succ \cdots \succ A$ and let the rest of the $n - 2$ voters alternate between $B \succ C \succ \cdots \succ A$ and $C \succ B \succ \cdots \succ A$.

  In both cases, C receives 2 more points than B in $\boldsymbol{R}^*$ and there is no other alternative that receives more points than C. Hence, since $B \rhd C$ we have that $F(\boldsymbol{R}^*, R') = B$ and $F(\boldsymbol{R}^*, R) = C$, so (1) holds.

(2) Towards a contradiction we suppose that there is some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$. Then, either A was winning when voting R and by moving it down, some other alternative is now winning. It can clearly not be the case that $F(\boldsymbol{R}', R) = A$ for any $\boldsymbol{R}'$ so, it must be the case that B wins with voting R' because it now receives more points than the alternative winning. But this again implies that $F(\boldsymbol{R}', R) = A$ and this cannot be the case. Hence, there is no such $\boldsymbol{R}'$.

□

Proof of Theorem 4.10:

Proof. Take an arbitrary $m \geq 4$ and $n \geq 3$ and $t > \left\lceil \frac{n-1}{2} \right\rceil$. Let $\mathcal{X} = \{A, B, C, .., Z\}$ be the set of m alternatives with the usual tie-breaking order $\rhd$. If $t = n-1$, we have that $|E_{\boldsymbol{R}}| = 1$, so from Lemma 4.1 it follows that the theorem holds. So, we assume $t < n - 1$. Now, we construct $\boldsymbol{R}$ as follows:

- $\left\lfloor \frac{n-1}{2} \right\rfloor$ times $B \succ C \succ \cdots \succ C$
- $t - \left\lfloor \frac{n-1}{2} \right\rfloor$ times $C \succ B \succ \cdots \succ A$
- and when n is even, one $C \succ Z \succ B \succ \cdots \succ A$

Clearly $E_{\boldsymbol{R}} \in \text{COMPL}(t)$. Moreover, we let $R = \rhd$ and R' be similar to R but with A and B swapped, that is $B \succ A \succ C \succ \cdots \succ Z$.

We first show that there exists some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. Since $t < n - 1$, there must be at least one $R_i \in \boldsymbol{R}$ with $R_i = \emptyset$. We extend this R_i to $C \succ Z \succ$

$B \succ \cdots \succ A$ and if there are other R_j with R_j is $\emptyset$, we extend them to $C \succ B \succ \cdots \succ A$. Now, we have that C has 2 more points than B in $\boldsymbol{R}^*$ and all other alternative have less points. Hence, we have that $F(\boldsymbol{R}^*, R') = C$ and $F(\boldsymbol{R}^*, R) = B$.

What is left to show is that there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$. First note that in every $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ if the manipulator votes R, A has at most $(n-t-1)\cdot(m-1)+(m-1)$ points and B has at least $\left\lfloor \frac{n-1}{2} \right\rfloor \cdot (m-1) + \Delta$ where

$$\Delta = \begin{cases} (t - \frac{n-1}{2} + 1) \cdot (m-2) & \text{if } n \text{ is odd} \\ (t - \left\lfloor \frac{n-1}{2} \right\rfloor + 1) \cdot (m-2) + (m-3) & \text{if } n \text{ is even} \end{cases}$$

Since we know that $t > \left\lceil \frac{n-1}{2} \right\rceil$ it also holds that $\left\lceil \frac{n-1}{2} \right\rceil > n-1-t$. So $(n-t-1)\cdot(m-1) < \left\lfloor \frac{n-1}{2} \right\rfloor \cdot (m-1)$. Moreover, since $m \geq 4$, we know that $(m-1) \leq (m-2)+(m-3)$ (this can be shown by a straightforward proof by induction) and thus also that $m-1 < 2(m-2)$. Since $t > \lfloor \lceil \frac{n-1}{2} \rceil \rfloor$ we have that $\Delta \geq (m-1)$ always holds, so B must have more points than A. This means that it can never be the case that $F(\boldsymbol{R}', R) = A$. Since difference between R and R' is that B receives more points than A, this implies that if $F(\boldsymbol{R}', R) \neq F(\boldsymbol{R}', R')$ it must hold that $F(\boldsymbol{R}', R') \succ_R F(\boldsymbol{R}', R)$. □

Proof of Theorem 4.13

Proof. We take an arbitrary $n, m \geq 3$ and $\Sigma \in \{\text{TOP}(0), \text{BOT}(0)\}$. First note that if $t = 0$, we know that $E_{\boldsymbol{R}} \in \Sigma$ implies that $E_{\boldsymbol{R}} = \mathcal{L}(\mathcal{X})^{n-1}$. Now, towards a contradiction suppose that there exists some R and R' such that the manipulator has an incentive to manipulate. This means there must be some $R^* \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. This means there must be some alternatives A and B such that $A \succ_R B$ and $B \succ_{R'} A$. We will show that this implies there is also some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) = A$ and $F(\boldsymbol{R}', R') = B$, which implies that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$ and gives us a contradiction that shows the theorem must hold.

Since $E_{\boldsymbol{R}} = \mathcal{L}(\mathcal{X}^{n-1}$, we can define any profile $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ and we know $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. Now, we define $\boldsymbol{R}'$ by having half of the preferences be $A \succ B \succ \cdots$ and the other half $B \succ A \succ \cdots$. If n is even, we have one profile left. In this case we extend to $A \succ B \cdots$ if $B \rhd A$ and $B \succ A \succ \cdots$ otherwise.

Now, A and B have more point than the rest, since

$$\frac{n-1}{2} \cdot (2m-3) > \frac{n-1}{2} + (m-1)$$

holds for all $n \geq 3$ and $m \geq 3$. Now, by construction, $F(\boldsymbol{R}', R) = A$ and $F(\boldsymbol{R}', R') = B$. □

Part of the proof of Theorem 4.14

Proof. What remains to be shown is that the Theorem also holds when n is even. So, assume n is even and take similar $\mathcal{X}$, $R = Z \succ Y \succ \cdots \succ B \succ A$, and R' as R but with Y and Z swapped. Now, we let $\boldsymbol{R}$ be defined as follows:

- $\lfloor \frac{n-1}{2} \rfloor$ voters vote $R_1 : Y \succ \mathcal{X} \setminus \{Y, B\} \succ B$
- $\lfloor \frac{n-1}{2} \rfloor$ voters vote $R_2 : A \succ \mathcal{X} \setminus \{A, B\} \succ B$
- 1 voter vote $R_3 : A \succ \mathcal{X} \setminus \{A, Z\} \succ Z$

Now, let $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ be defined by extending R_1 to $R'_1 : Y \succ Z \succ A \succ \cdots \succ B$, R_2 to $R'_2 : A \succ Z \succ Y \succ \cdots \succ B$, and R_3 to $R'_3 : A \succ Y \succ B \succ \cdots \succ Z$. Now, if the manipulator votes R note that A, Z and Y all beat all other alternatives and tie among each other. So, because of the tie-breaking we have that $F(\boldsymbol{R}', R) = A$. If the manipulator votes R', however, Y beats Z so $F(\boldsymbol{R}', R') = Y$. So, that there is a $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$.

What is left to show is that there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that the manipulator is worse of voting R'. This can only be the case when $F(\boldsymbol{R}', R) = Z$. Note, however, that the best Z can do in any $\boldsymbol{R}'$ is to tie with A and Y and beat all other alternatives. Note that Y and A always beat all other alternatives and at least tie with each other and Z. Since $Y \rhd Z$ and $A \rhd Z$, we know that $F(\boldsymbol{R}^*, R) = Z$ can never be the case. □

Proof of Theorem 4.15

Proof. Take some arbitrary $m, n \geq 3$ and some $t > 1$. If $t \geq m - 1$, we know that $|E_{\boldsymbol{R}}| = 1$ for every $E_{\boldsymbol{R}} \in \text{TOP}(t)$, so from Lemma 4.1 we know the theorem holds. So, we assume $t < m - 1$. Now let $\mathcal{X}$ be a set of m alternatives, with the usual tie-breaking order $\rhd$. We let R be the opposite of $\rhd$, that is $Z \succ Y \succ X \succ \cdots$. First, suppose that $t = 2$. Now, if n is odd, we construct $\boldsymbol{R}$ as follows:

- $\frac{n-1}{2}$ voters with $R_1 : X \succ Z \succ \mathcal{X} \setminus \{X, Z\}$
- $\frac{n-1}{2}$ voters with the partial preference $R_2 : Y \succ X \succ \mathcal{X} \setminus \{X, Y\}$

Now consider the non-truthful preference R' that is similar to R but with $Y \succ_{R'} Z$ instead of $Z \succ_R Y$.

First, we show there cannot be any $\boldsymbol{R}'$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$. Towards a contradiction, suppose such $\boldsymbol{R}'$ does exists. Note that since the only difference between R and R' is the order of Z and Y, the only way this can happen is when $F(\boldsymbol{R}', R) = Z$. However, we know that X beats everything besides Y in every $\boldsymbol{R}'$ when voting R, so it receives $m - 3$ points ($m - 2$ points for every pairwise majority contest won, and -1 for the one lost). In the best case, Z also beats everything besides Y so also receives $m - 3$ points and since $X \rhd Z$, Z can never win when the manipulator votes R. So, such $\boldsymbol{R}'$ can never exist.

What is left to show for this case is that there also is some $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. Now, we extend $\boldsymbol{R}$ such that R_1 is extended to R'_1 with $X \succ Z \succ Y \succ \cdots$ and $R'_2 Y \succ X \succ Z \succ \cdots$. Now, when the manipulator votes R X, Y, Z beat all other alternatives, and X beats Z, Z beats Y and Y beats X, so all alternatives have $m - 3$ points. So, $F(\boldsymbol{R}^*, R) = X$. However, when the manipulator votes R', Y becomes the Condercet winner, so $F(\boldsymbol{R}^*, R') = Y$.

Now, suppose n is even. We let R' be the same as when n is odd and define $\boldsymbol{R}$ as follows:

- $\lfloor \frac{n-1}{2} \rfloor$ voters with $R_1 : X \succ Z \succ \mathcal{X} \setminus \{X, Z\}$
- $\frac{n-1}{2}$ voters with the partial preference $R_4 : Y \succ Z \succ \mathcal{X} \setminus \{Y, Z\}$
- 1 voter with the partial preference $R_5 : X \succ Y \succ \mathcal{X} \setminus \{X, Y\}$

For any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, we have that X, Y, and Z all tie and beat all other alternatives when the manipulator votes R, so we have that $F(\boldsymbol{R}', R) = X$. Now when the manipulator votes R', however, Y beats X so it must hold that $F(\boldsymbol{R}', R') = Y$. So, clearly the theorem holds.

When $t > 2$, we know that there must be some $E_{\boldsymbol{R}''} \in \text{TOP}(t)$ such that $E_{\boldsymbol{R}''} \subseteq E_{\boldsymbol{R}}$ and $\boldsymbol{R}^* \in E_{\boldsymbol{R}''}$ when n is odd. From this it follows that the Theorem also holds in this case. □

Proof of Theorem 4.16

Proof. Take some arbitrary $m, n \geq 3$ and some $t > 1$. If $t \geq m - 1$, we know that $|E_{\boldsymbol{R}}| = 1$ for every $E_{\boldsymbol{R}} \in \text{BOT}(t)$, so from Lemma 4.1 we know the theorem holds. So, we assume $t < m - 1$. Now let $\mathcal{X}$ be a set of m alternatives, with the usual tie-breaking order $\rhd$. We distinguish between two cases, n is odd and n is even. First, suppose $t = 2$.

First, suppose n is odd. We let R be the opposite of $\rhd$, that is $Z \succ Y \succ X \succ \cdots$. Now, we construct $\boldsymbol{R}$ as follows:

- $\frac{n-1}{2}$ of the voters has the partial preference order $R_1 : \mathcal{X} \setminus \{X, Z\} \succ X \succ Z$
- $\frac{n-1}{2}$ of the voters has the partial preference order R_2: $\mathcal{X} \setminus \{Y, Z\} \succ Z \succ Y$

We let R' be similar to R but with Y and Z swapped.

First, we show that there cannot be any $\boldsymbol{R}'$ such that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$. Towards a contradiction, suppose there is such an $\boldsymbol{R}'$. Since the only difference between R and R' is that $Y \succ_{R'} Z$ and $Z \succ_R Y$ it must be the case that $F(\boldsymbol{R}', R) = Z$. However, since $n \geq 3$, Z loses every pairwise majority contest, except against Y, now X beats Z we know that X will always have at least as many points as Z. Since $X \rhd Z$, Z can never win when the manipulator votes R. So, such $\boldsymbol{R}'$ cannot exist.

What is left to show is that there is some $\boldsymbol{R}^*$ with $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. We extend R_1 to $R_1' : Y \succ \cdots \succ X \succ Z$ and R_2 to $R_2' : X \succ \cdots \succ Z \succ Y$. Now, when the manipulator votes R, Y beats everything besides Z and X beats everything besides Y. Since $X \rhd Y$, this means we have $F(\boldsymbol{R}^*, R) = X$. When the manipulator votes R', however, Y becomes the Condorcet winner and thus $F(\boldsymbol{R}^*, R') = Y$.

Now, suppose n is even. Since we know that $n \geq 4$ and assumed that $t > 1$ and $t < m - 1$, it must hold that $m \geq 4$. We let R be equal to $\rhd$, that is $R : A \succ B \succ C \succ \cdots \succ Z$. Now, we construct $\boldsymbol{R}$ as follows:

- $\lceil \frac{n-1}{2} \rceil$ of the voters has the partial preference order $R_3 : \mathcal{X} \setminus \{C, A\} \succ C \succ A$
- $\lfloor \frac{n-1}{2} \rfloor$ of the voters has the partial preference order R_4: $\mathcal{X} \setminus \{A, B\} \succ A \succ B$

We let R' be similar to R but with $B \succ_{R'} A$ instead of $A \succ_R B$. With a similar argument as in the case that n is odd, we know there can never be some $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ such that $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

What is left to show is that there is a $\boldsymbol{R}^*$ with $F(\boldsymbol{R}^*, R') \succ_R F(\boldsymbol{R}^*, R)$. We define $\boldsymbol{R}^*$ as follows: We extend *one* R_3 to $Z \succ B \succ \cdots \succ C \succ A$ and the rest to $B \succ Z \succ \cdots \succ C \succ A$, and all R_4 to $Z \succ \cdots \succ C \succ A \succ B$. Now, if the manipulator votes R, B beats all alternatives except it ties with both Z and A, which means it receives $m-3$ points. While Z beats all alternatives and only ties with B, so receives $m-2$ points. Any other alternative Y besides A has at most $m-3$ points, so we have that $F(\boldsymbol{R}^*, R) = Z$. However, if the manipulator votes R', B beats A and also receives $m-2$ points. Since $B \triangleright Z$, we now have that $F(\boldsymbol{R}^*, R') = B$.

With a similar argument as in Theorem 4.15, we know the theorem must also hold for all $t > 2$. □

Part of the proof of Theorem 4.17

Proof. What is left to show is that the Theorem also holds when n is even. We take $\mathcal{X}$, R and R' similar to the odd case. First, suppose $n = 4$, now from the assumption that $t > \frac{n-1}{2}$, we know $t \geq 2$, but since the Theorem holds from Lemma 4.1 when $t = n-1$, we may also assume that $t < n-1 = 3$, so $t = 2$. Now, let $\boldsymbol{R}$ contain the two votes: $X \succ Z \succ Y \succ \cdots$ and $X \succ Y \succ Z \succ \cdots$.

First, note that no matter what the third vote is, if the manipulator votes R, Z will at most tie with a pairwise majority contest against X. Moreover, X will at least tie against Z and X will beat all other alternatives. Since $X \triangleright Z$, this means we never have that $F(\boldsymbol{R}', R) = Z$, for any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$. This implies there can never be any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

Now, let $\boldsymbol{R}^*$ be the extension of $\boldsymbol{R}$ with the vote $Y \succ Z \succ X \succ \cdots$. Now, it is easy to check that since $X \triangleright Y$, we have $F(\boldsymbol{R}^*, R') = Y$ and $F(\boldsymbol{R}^*, R) = X$, which shows the theorem holds when n is 4.

Now, suppose $n > 4$. Now, we define $\boldsymbol{R}$ as follows:

- $\lfloor \frac{n-1}{2} \rfloor$ voters vote $Y \succ X \succ Z \succ \cdots$
- The rest of the voters vote $X \succ Z \succ Y \cdots$

First, note that again we know that $F(\boldsymbol{R}', R) \neq Z$ for any $\boldsymbol{R}' \in E_{\boldsymbol{R}}$, which implies there is no $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}', R) \succ_R F(\boldsymbol{R}', R')$.

Finally, consider the $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$ with:

- $\lfloor \frac{n-1}{2} \rfloor$ voters vote $Y \succ X \succ Z \succ \cdots$
- $\lfloor \frac{n-1}{2} \rfloor$ voters vote $X \succ Z \succ Y \succ \cdots$
- and 1 voter who also votes R

Since we may assume $t < n-1$, it must hold that $\boldsymbol{R}^* \in E_{\boldsymbol{R}}$. Now, since $n > 4$, we know that if the manipulator votes R, X beats every alternative besides Y, Y every alternative besides Z, and Z every alternative besides X. So, we again have that $F(\boldsymbol{R}^*, R') = Y$ and $F(\boldsymbol{R}^*, R) = X$. $\square$

Proof of Theorem 6.1

Proof. Let $\mathcal{X} = \{A, B, ...,\}$ such that $|\mathcal{X}| = m$ and let voter n be the manipulator with a true preference $RA \succ B \succ C \succ \cdots$ similar to the tie-breaking order. Now, we define the actual situation $\boldsymbol{R}' \in \mathcal{L}(\mathcal{X})^{n-1}$ as follows. When n is even, $\boldsymbol{R}'$ contains:

- $\lfloor \frac{n-1}{2} \rfloor$ voters who voter $R'_1 : B \succ C \succ ... \succ A \succ E$
- $\lfloor \frac{n-1}{2} \rfloor$ voter who vote: $R'_2 : C \succ B \succ \cdots \succ A \succ E$.
- 1 voter who votes $R'_3 : C \succ D \succ B \succ \cdots \succ A \succ E$

When n is odd, there are also $\frac{n-1}{2}$ voter who vote R'_1, but $\frac{n-1}{2} - 2$ voters who vote R'_2 and two voters who votes $R'_3 : C \succ D \succ B \succ \cdots \succ A \succ E$. This means that $G_{tb}(\boldsymbol{R}') = E_{\boldsymbol{R}}$, $\boldsymbol{R}$ contains for every n:

- $\lceil \frac{n-1}{2} \rceil$ voters with partial preference $R_1 : C \succ \mathcal{X} \setminus \{B, E\} \succ E$
- $\lfloor \frac{1}{2} \rceil$ voters with partial preference $R'_2 : B \succ \mathcal{X} \setminus B, F \succ F$.

Now, we need to show two things, (1) the manipulator can manipulate in $\boldsymbol{R}'$ and (2) the manipulator cannot get such a good result in $E_{\boldsymbol{R}}$.

(1) Let R' be $B \succ A \succ C \succ \cdots \succ E$, i.e. the preference order in which A and B are swapped. Since $B \triangleright C$, we have that $F(\boldsymbol{R}', R') = B$ while $F(\boldsymbol{R}', R) = C$. Now, to show that R' is a best response, it suffices to show that there cannot be any R'' such that $F(\boldsymbol{R}', R'') = A$. This holds since A has at most $n - 1 + m - 1$ points, i.e. A is on top in R'', while C has at least $\lceil \frac{n-1}{2} \rceil \cdot (m-1) + \lfloor \frac{n-1}{2} \rfloor \cdot (m-2)$ points, i.e. C receives zero points in R''. Since $m > 4$ and $n > 2$, C will always receive more points than A, so such R'' can never exist.

(2) Towards a contradiction, suppose there exists some R'' such that $F(\boldsymbol{R}', R'') = B$ and for every $\boldsymbol{R}'' \in E_{\boldsymbol{R}}$ we have that $F(\boldsymbol{R}'', R'') \succeq_R F(\boldsymbol{R}'', R)$. Now, since C has two more points than B in $\boldsymbol{R}'$, it must hold that either B has a higer position in R'' than in R or C has a lower position in R'' than R. If B is moved up, this means that A must be moved down. But note that it is easy to check that there is some $\boldsymbol{R}''$ such that A and B tie (if n is odd) or B has one more point than A (if n is even) (put A second and B third when possible in the case that n is odd, when n is even swop A and B once when C on top), so then it holds that $F(\boldsymbol{R}'', R) = A$ and $F(\boldsymbol{R}'', R'') = B$. So, this is not possible.

So, it must be the case that C is moved down. But this means that must exists some D such that $D \succ_{R''} C$ while $C \succ_R D$. Moreover, since $m > 4$ there must be one such that $D \neq E$. But, now again we can define a $\boldsymbol{R}''$ such that C and D tie or D has one more point than C and all other alternatives have less points. This implies that $F(\boldsymbol{R}'', R) = C$ and $F(\boldsymbol{R}'', R'') = D$. Contradiction! So, no such $\boldsymbol{R}''$ exists.

□

Proof of Lemma 6.1

Proof. If $\Sigma \neq \text{COMPL}(t)$ for any t, we know that $G_\Sigma(\boldsymbol{R}') = \boldsymbol{R} \Leftrightarrow \boldsymbol{R}' \in E_{\boldsymbol{R}}$. So, for a $\boldsymbol{R}$ with $E_{\boldsymbol{R}} \in \Sigma$, we have that for any two $\boldsymbol{R}', \boldsymbol{R}'' \in E_{\boldsymbol{R}}$ it must hold that

$$F(\boldsymbol{R}') = H(G_\Sigma(\boldsymbol{R}')) = H(\boldsymbol{R}) = H(G_\Sigma(\boldsymbol{R}'')) = F(\boldsymbol{R}'')$$

If $\Sigma = \text{COMPL}(t)$ and either $t = (n-1)$ or $m = 1$, then we again know that $G_\Sigma(\boldsymbol{R}') = \boldsymbol{R} \Leftrightarrow \boldsymbol{R}' \in E_{\boldsymbol{R}}$, since $|E_{\boldsymbol{R}}| = 1$. Again the theorem holds.

Now, if $t < n$ and $m > 1$ note that if $\boldsymbol{R}' \in E_{\boldsymbol{R}}$ we can always define a partial information function G_Σ such that $G_\Sigma(\boldsymbol{R}') = \boldsymbol{R}$. Let $\mathcal{N}'$ be the set of voters that have a linear order in $\boldsymbol{R}$, then we let $G_\Sigma(\boldsymbol{R}') = (R_1, ...R_{n-1})$ such that $R_i = R'_i$ if $i \in \mathcal{N}'$ and $R_i = \emptyset$ otherwise. This G_Σ is clearly surjective and we must have $G_\Sigma(\boldsymbol{R}') = \boldsymbol{R}$ implies $\boldsymbol{R}' \in E_{\boldsymbol{R}}$.

We will prove the final part with contraposition. Suppose we have that there is some $\boldsymbol{R} \in \mathcal{P}(\mathcal{X})^n$ such that $E_{\boldsymbol{R}} \in \Sigma$ and two $\boldsymbol{R}', \boldsymbol{R}'' \in E_{\boldsymbol{R}}$ with $F(\boldsymbol{R}') \neq F(\boldsymbol{R}'')$. Now, we have just shown we can define a G_Σ such that $G_\Sigma(\boldsymbol{R}') = R$. this means that we can define the partial information function G'_Σ that is equal to G_Σ except that $\boldsymbol{R}''$ is also mapped to $\boldsymbol{R}$. Morover, by the contruction above, we know that for any $\boldsymbol{R}$ there must be at least $t \cdot (n-1) \cdot m!$ full profiles that are mapped to any partial profile. Since $t < (n-1)$ and $m > 1$, so G'_Σ is still surjective. So, now we have that $G'_\Sigma(\boldsymbol{R}') = G'_\Sigma(\boldsymbol{R}'')$ but $F(\boldsymbol{R}') \neq F(\boldsymbol{R}'')$. So a H such that $F = H \circ G'_\Sigma$ can never exist and we shown that F cannot be computable from Σ. □