Voting and Judgment Aggregation: An Axiomatic and Algorithmic Analysis

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Ann-Kathrin Selker
aus Georgsmarienhütte

Düsseldorf, Juli 2021

aus dem Institut für Informatik
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1.

2.

Tag der mündlichen Prüfung:

Abstract

This thesis deals with preference and judgment aggregation in the context of computational social choice, a subfield of multiagent systems and artificial intelligence. In preference aggregation, agents (typically called voters) have preferences over a given set of candidates and the goal is to aggregate ballots derived from these preferences to determine an output (e.g., a single candidate, a set of candidates, or a linear order over the candidates). In judgment aggregation, agents (typically called judges) have judgments over a given set of issues and the goal is to aggregate these judgments into a collective set of judgments.

In the context of multiwinner elections that elect a committee of candidates, this thesis studies a new type of ballot called ℓ-ballots. In contrast to existing ballot types, ℓ-ballots are a compromise between ordinal and cardinal ballots. This thesis focuses on the axiomatic properties of two newly defined types of multiwinner voting rules using these ℓ-ballots as input, and proposes a generalization of ℓ-ballots to fully cardinal ballots.

Furthermore, this thesis explores the computational complexity of strategic attacks on voting rules. For several prominent iterative voting rules, i.e., voting rules that proceed in rounds, it is shown that shift bribery is NP-complete in all considered cases. In the context of iterative voting, voters are encouraged to repeatedly update their ballots as a reaction to the current state of the election. This thesis uses a model where voters are connected via an underlying social network and compute their information about the current state of the election both by observing their neighbors' ballots and an opinion poll announced by a polling agency. This thesis explores the manipulation power of the polling agency for the voting rules plurality and veto and shows that manipulation is para-NP-hard even for very restricted social networks and permitted ballot deviations. In particular, the thesis focuses on distance restrictions for the polling agency in regard to the manipulated opinion poll and for the voters in regard to their deviations.

Finally, this thesis deals with control in judgment aggregation under various preference types of the chair where a chair tries to achieve a better outcome by changing the structure of the aggregation process, e.g., by adding or deleting judges. This thesis shows NP-completeness for most considered problems for the judgment aggregation procedures uniform (constant) premise-based quota rules.

Contents

Chapter 1

Introduction

Computational social choice is an interdisciplinary field that encompasses among others political science, economics, mathematics, and computer science. The aim is to aggregate, e.g, preferences, opinions, or judgments over a set of, e.g., candidates, items, or issues to reach a collective outcome. This process of collective decision making is present in everyday life, be it political elections or choosing a selection of menu items. See the book edited by Brandt et al. (2016) for an overview on the field of computational social choice, and the book edited by Endriss (2017) for an outlook on current research directions.

In *preference aggregation*, the goal is to aggregate voters' preferences over candidates to either elect a winning candidate, elect a winning committee, or create a ranking over the given candidates. Examples include electing a president, electing a parliament, or creating a ranking over politicians of a party stating in which order the party will assign them seats in parliament. Well-known contributors to the field of preference aggregation include, for example, the mathematician Jean-Charles de Borda (1733–1799), the Marquis de Condorcet (1743–1794), a philosopher and mathematician, and economist and Nobel laureate Kenneth Arrow (1921–2017). But preference aggregation is not confined to political elections. Preference aggregation provides models that can be used to describe and design the decision making process for autonomous (computer) systems, and is of particular interest to the field of artificial intelligence and multiagent systems. Computer scientists are not only concerned with combinatorial problems that result from large-scale applications, but also provide tools to analyze, among others, the suitability of an aggregation procedure for the given application. Important contributions include the design of algorithms and the study of computational complexity. The latter deals with quantifying the amount of time or space a problem takes to solve and provides, for example, barriers against strategic attacks on aggregation procedures by proving that attacks need an unfeasible span of time to execute in the worst case. The theory of parameterized complexity is able to give a more fine-grained analysis than the classical complexity setting that focuses on worst-case complexity.

In contrast to preference aggregation, the field of *judgment aggregation* only emerged quite recently. Kornhauser and Sager (1986) note that in a court of three judges that have to decide whether the defendant is guilty, the intuitive approach of aggregating consistent judgments can lead to an inconsistent collective outcome. The reformulation of this paradox by Pettit (2001) marked the starting point of a series of impossibility theorems regarding the incompatibility of desirable axioms for judgment aggregation procedures. Judgment aggregation generalizes preference aggregation and is able to express various scenarios in one framework. However, this expressiveness often comes with the toll of a high computational complexity.

This thesis is organized as follows. Chapter 2 formally introduces complexity theory, preference aggregation, and judgment aggregation, and provides a short overview of related work. The following chapters then illustrate my contribution to the field of voting and judgment aggregation. In multiwinner voting, the goal is to elect a committee of candidates. In this thesis, one focus for multiwinner voting is on the representation of voters' preferences. Chapter 3 deals with a new type of ballot that generalizes existing models for representing voters' preferences and studies the axiomatic properties of corresponding multiwinner rules. The next chapters deal with strategic attacks on aggregation procedures where an attacker tries to reach a more favorable outcome by interfering with the aggregation process. Chapter 4 extends the study of shift bribery—a type of attack where voters are influenced to change their ballots—to iterative voting rules. Chapter 5 proposes distance-based manipulation problems for iterative elections with polls, where an attacker can influence the information voters have about the election, and gives parameterized complexity results. Chapter 6 introduces the concept of control to judgment aggregation where an attacker is able to change the structure of the judgment aggregation process, for example by adding or deleting judges, and presents complexity results for various notions of what constitutes a more favorable outcome. Finally, Chapter 7 concludes the thesis and provides perspectives for future work in my field.

Chapter 2

Background and Related Work

This chapter provides the necessary background for the following chapters. Section 2.1 shortly introduces the concept of computational complexity that is used in all following chapters. Then Section 2.2 gives an overview on preference aggregation. In particular, Section 2.2.1 deals with singlewinner elections, Section 2.2.2 introduces the concept of strategic behavior in collective decision making with a focus on preference aggregation, Section 2.2.3 extends singlewinner to multiwinner preference aggregation, and Section 2.2.4 gives the background for iterative elections. Finally, Section 2.3 introduces judgment aggregation which generalizes preference aggregation.

2.1 Complexity Theory

Computer science is, among others, concerned with solving problems where problems are defined by the input they receive and the question (or task) they answer (or execute). An important factor for a given problem is the question how fast the best algorithm is able to solve it (i.e., the *computational complexity* of the problem (Hartmanis and Stearns, 1965)), and therefore how "hard" the problem is. However, if no fast algorithm for a problem is known, does that mean that a fast algorithm has just not been discovered yet, or does it not exist? While the specific running time is dependent on the algorithm, there are upper and lower bounds for a best algorithm that are due to the nature of the problem. For example, there are problems that can provably be solved in at most linear time (because there is an algorithm with that running time), or that are impossible to solve without using an exponential length of time (and therefore no algorithm can exist that solves the problem faster), where "linear" and "exponential" are seen in relation to the size of the input. This thesis mostly uses worst-case (computational) complexity, where the given upper bound is for the input that takes the most time to solve by a best (i.e., fastest) algorithm for the given problem. Problems can be grouped in so-called *complexity classes* depending on the upper bound for the worst-case input.

This thesis focuses on *decision problems*, i.e., problems that ask a yes/no question. Formally, a decision problem is a set S of words over the alphabet $\{0,1\}$,[1] where S is considered to be the set of inputs for which the answer is yes. An algorithm solves S if—given an input $x \in \{0,1\}^*$—it always correctly decides whether $x \in S$. An algorithm is *deterministic* if for each input, each step the algorithm takes is predetermined by the previous step, which among others implies that executing the algorithm several times for the same input always results in the same outcome. In contrast to a deterministic algorithm, a *nondeterministic* algorithm can take different steps in executions of the same input. In particular, a nondeterministic algorithm solves a decision problem S if for a given input x, at least one series of steps leads to the output 'yes' if $x \in S$, and the algorithm does not output 'yes' for any series of steps if $x \notin S$. The complexity class P contains all decision problems that can be solved in deterministic polynomial time (i.e., can be solved by a deterministic algorithm in polynomial time), whereas NP contains all decision problems that can be solved in nondeterministic polynomial time. Alternatively, one can define NP as the class of decision problems where a solution can be verified in deterministic polynomial time. Obviously, P is a subset of NP. One of the most prominent open problems in the field of complexity theory is the question whether P is equal to NP.

Since complexity classes only give upper bounds, but not lower bounds for the problems contained in them, the actual complexity between the problems in one class can vary. One idea to compare two problems in terms of complexity is to show that one problem is at least as hard as the other problem. This thesis uses many-one reducibility for the comparison of problems.

Definition 2.1 (polynomial-time many-one reducibility)**.** Let A and B be sets of words over $\{0,1\}$. A is *polynomial-time many-one reducible* to a problem B—denoted by $A \leq_{\mathrm{m}}^{\mathrm{P}} B$—if there exists a polynomial-time computable function $f : \{0,1\}^* \to \{0,1\}^*$ so that $x \in A \Leftrightarrow f(x) \in B$ for all $x \in \{0,1\}^*$.

If $A \leq_{\mathrm{m}}^{\mathrm{P}} B$, then A can be solved given the input $x \in \{0,1\}^*$ by solving B with the input $f(x)$. That means that problem B is at least as hard as problem A. Note that $\leq_{\mathrm{m}}^{\mathrm{P}}$ is transitive, so that $A \leq_{\mathrm{m}}^{\mathrm{P}} B$ and $B \leq_{\mathrm{m}}^{\mathrm{P}} C$ implies $A \leq_{\mathrm{m}}^{\mathrm{P}} C$ for all $A, B, C \subseteq \{0,1\}^*$.

One goal of complexity theory is to find problems that can be seen as representatives of

[1]Note that it is always possible to encode words over other, more complex alphabets into words in $\{0,1\}^*$.

their complexity class because they are a member of this complexity class and are at least as hard to solve as every other member of this class.

Definition 2.2 (hardness and completeness)**.** Let B be a set of words over $\{0,1\}^*$ and let $\mathcal{C}$ be a complexity class. B is called $\mathcal{C}$-*hard* if for all $A \in \mathcal{C}$, it holds that $A \leq_{\mathrm{m}}^{\mathrm{P}} B$, and $\mathcal{C}$-*complete* if B is $\mathcal{C}$-hard and $B \in \mathcal{C}$.

Problems in P are called *tractable*, whereas NP-hard problems are called *intractable* under the assumption that P and NP are not equal. Since NP is closed under many-one reducibility, i.e., $A \leq_{\mathrm{m}}^{\mathrm{P}} B$ and $B \in \mathrm{NP}$ implies $A \in \mathrm{NP}$, the existence of a deterministic polynomial-time algorithm for an NP-complete problem would imply deterministic polynomial-time algorithms for all problems in NP. Building on the work by Cook (1971) who proved for the first time that a decision problem—namely, deciding whether a given Boolean formula is satisfiable (SAT for short)—is NP-complete, Karp (1972) proved the NP-completeness for various other natural decision problems. While intractability is a huge disadvantage for many problems arising in natural applications, in this thesis it is mostly seen as a positive aspect in the context of problems that deal with strategic behavior (see Section 2.2.2 for a short introduction). Therefore, the following chapters explore the complexity of the considered problems in detail. See the book by Garey and Johnson (1979) for an introduction to NP-completeness and the book by Arora and Barak (2009) for an extensive introduction to complexity theory in general.

A disadvantage of classical complexity theory is the sole focus on the worst case. A brute-force search approach does not have to be the best way to find a solution for an intractable problem. In practice, the worst case might appear rarely so that algorithms might still solve the problem efficiently in the average case, or there might be efficient heuristics. Furthermore, in many applications the considered instances are bounded in size, for example because there are only a few data sets or the considered graphs are sparse. If such a bound is guaranteed to be small or even fixed, an algorithm might exploit it to find solutions efficiently. The study of problems that include a parameter corresponding to a bound on the input is called *parameterized complexity theory*.

A part of this thesis focuses on parameterized decision problems. In contrast to the classical decision problems, these problems also include a (numerical) parameter that is assumed to be small, e.g., the number of clauses in a Boolean formula for a parameterized version of SAT. Formally, a parameterized decision problem is a set $S \subseteq \{0,1\}^*$ of inputs for which the answer is yes and a parameterization κ that assigns each instance $x \in \{0,1\}^*$ a numerical value k as a parameter. The parameterized complexity class FPT (fixed-parameter tractable) is the parameterized equivalent of P (Downey and Fellows, 1995a), whereas para-NP is the parameterized equivalent of NP (Flum and Grohe, 2003). Note that the membership in a parameterized complexity class is obviously dependent on the parameter, i.e., the same classical decision problem can be a member of different parameterized complexity classes for different choices of parameterization.

Definition 2.3 (FPT, fpt-algorithm, and para-NP)**.** Let $S \subseteq \{0,1\}^*$, let κ be a parameterization $\kappa : \{0,1\}^* \to \mathbb{N}$, let $f : \mathbb{N} \to \mathbb{N}$ be a computable function, and let $g : \{0,1\}^* \to \mathbb{N}$ be a polynomial-time computable function.

1. (S,κ) is in FPT in regard to parameterization κ if there exists a deterministic algorithm that solves S in time $f(\kappa(x)) \cdot g(x)$ for each $x \in \{0,1\}^*$. Such an algorithm is called an *fpt-algorithm*.
2. (S,κ) is in para-NP in regard to parameterization κ if there is a nondeterministic algorithm that solves S in time $f(\kappa(x)) \cdot g(x)$ for each $x \in \{0,1\}^*$.

Analogously to classical complexity theory, a parameterized (polynomial-time many-one) reduction is a tool to compare the hardness of two parameterized problems.

Definition 2.4 (parameterized reduction)**.** Let A,B be sets of words over $\{0,1\}$ and let $\kappa,\kappa' : \{0,1\}^* \to \mathbb{N}$ be parameterizations. (A,κ) is *parameterized polynomial-time many-one reducible* to (B,κ') if there exists an fpt-algorithm $f : \{0,1\}^* \to \{0,1\}^*$ so that $x \in A \Leftrightarrow f(x) \in B$ for all $x \in \{0,1\}^*$, and there exists a computable function g so that $\kappa'(f(x)) \leq g(\kappa(x))$ for all $x \in \{0,1\}^*$.

Note that in contrast to the classical setting, the algorithm that transforms an instance of a problem A into an instance of the problem B is not a polynomial-time, but an fpt-algorithm. That means that the existence of a parameterized reduction does not imply the existence of a polynomial-time reduction and cannot be used to show NP-hardness. The definitions for hardness and completeness as in Definition 2.2 carry over to the parameterized context.

Downey and Fellows (1995a,b) introduce the W-hierarchy including the complexity classes W[1] and W[2]. A parameterized decision problem is in the complexity class W[t] if there exists a parameterized reduction to a certain weighted circuit satisfiability problem.

Note that FPT $= W[0] \subseteq \mathrm{W}[1] \subseteq \mathrm{W}[2] \subseteq \cdots \subseteq$ para-NP and that it is an open question whether FPT is equal to W[1]. Parameterized problems in FPT are called tractable, whereas problems that are W[t]-hard, $t > 0$, or para-NP-hard are called intractable. See the books by Flum and Grohe (2006) and Cygan et al. (2015) for an introduction to parameterized complexity theory.

In contrast to the yes/no questions by decision problems, optimization problems ask for a solution with certain "optimal" properties, e.g., a satisfying truth assignment to a Boolean formula with the maxinum number of 1's possible. Finding an optimal solution for a problem can often be intractable. However, for most applications it is sufficient if the solution is always provably close to the optimal solution. Johnson (1974) starts the comprehensive study of *approximation algorithms* and, among others, gives examples of algorithms that guarantee that the size of a solution is within a factor c the size of the optimal solution while still computing the solution in deterministic polynomial time. See the book by Vazirani (2013) for an overview on approximation.

2.2 Preference Aggregation

In the field of preference aggregation, the goal is to aggregate preferences of agents called *voters* over a set of alternatives (or *candidates*) to elect a winner. One focus in this thesis is the computational complexity of aggregation and of strategic behavior in preference aggregation.

Formally, let $C = \{c_1, \dots, c_m\}$ be a set of candidates and let $N = \{1, \dots, n\}$ be a set of voters, where each voter j has a (private, not further specified) preference over the candidates and submits a ballot v_j. In many applications, these ballots are seen as "sincere", meaning that they reflect the voters' preferences over the candidates in some way. In these contexts, the notions ballot and preference are often used interchangeably in the literature. See Section 2.2.2 and Section 2.2.4 for examples of "insincere" or "strategic" ballots. A list of ballots $P = (v_1, \dots, v_n)$ is called a *profile*.

In this thesis, the following types of ballots will be used. Note that these types can also be used to model preferences (see, e.g., Section 2.2.4), but in general, preferences are assumed to be far more complex. Using *ordinal* ballots, each voter ranks the candidates from best to worst, e.g., each ballot v_j is a linear order $>_j$ over the candidates where $c_i >_j c_k$ can be interpreted as v_j strictly preferring candidate c_i to candidate c_k. Note that a linear order $>$ is

- complete ($a > b$ or $b > a$ holds for all $a, b \in C$, $a \neq b$),
- transitive ($a > b$ and $b > c$ implies $a > c$ for all $a, b, c \in C$), and
- asymmetric (if $a > b$, then $b > a$ does not hold).

There are also models where the ballot is a weak order $\succsim_j$, where $c_i \succsim_j c_k$ denotes that voter j either prefers c_i over c_k, or does not distinguish between them. Weak orders do not have to be asymmetric. Let $pos_j(c)$ denote the rank of candidate c in ballot v_j. For example, for the ballot $c_3 >_2 c_1 >_2 c_2$, it holds that $pos_2(c_3) = 1$.

For *approval ballots*, voters only distinguish between approved and disapproved candidates, and the ballot $v_j \subseteq C$ is a set denoting the approved candidates of voter j. Note that this dichotomous model can also be generalized to allow for more groups of candidates (see Chapter 3 for my contribution to this generalization). Furthermore, for some applications, the number of candidates in an approval ballot might be fixed. The model studied in Chapter 5 uses so-called plurality ballots, where each voter can only approve of exactly one candidate, and veto ballots, where each voter can only disapprove of exactly one candidate and v_j denotes the *disapproved* candidate of voter j.

Finally, voters can assign a numerical value to each candidate. For these *cardinal ballots*, the ballot is a set $v_j = \{(c_i, k_i) \mid c_i \in C, k_i \in \mathbb{Q}, 1 \leq i \leq m\}$. Depending on the application, there might be upper or lower bounds for the values of k_i, or the ballot might be normalized so that the values for the most preferred candidate and for the most disliked candidate are fixed for all voters. See Section 3.4 for voting rules designed for cardinal ballots. In computational social choice, cardinal ballots (or *cardinal utilities*) are mostly used in the field fair division of indivisible goods, where agents (the equivalent to voters) assign values to items (the equivalent to candidates) and the goal is to find an assignment of items to agents that maximizes the utility in some way.

Example 2.5 (ballot types)**.** Let $C = \{a, b, c, d\}$ be a candidate set. Assume that a voter is a big fan of a, likes both b and c equally, and slightly dislikes d. A possible (sincere) ordinal ballot using linear orders is then $v = a > b > c > d$, a weak order $v = a \succ b \sim c \succ d$,[2] a

[2]Here, $a \succ b$ denotes that $a \succsim b$ and not $b \succsim a$, and $b \sim c$ denotes that $b \succsim c$ and $c \succsim b$.

possible approval ballot $v = \{a,b,c\}$, a plurality ballot $v = \{a\}$, a veto ballot $v = \{d\}$, and a possible cardinal ballot $v = \{(a,10),(b,3),(c,3),(d,1)\}$.

2.2.1 Singlewinner Elections

This section deals with singlewinner elections where the goal is to elect a winner among the set of candidates. To determine the outcome of the election, an *aggregation function* is needed to aggregate the voters' ballots (and therefore their underlying preferences). In this chapter, the following two types of functions are used.

Definition 2.6 (voting rule, social welfare function)**.** Let C be a set of candidates, let $N = \{1,\dots,n\}$ be the set of voters, let $\mathcal{B}(C)$ be the set of all possible ballots over C,[3] and let $\mathcal{L}(C)$ be the set of all possible linear orders over C.

- A (singlewinner) *voting rule* $\mathcal{R}$ maps a profile to a subset of candidates called the winners of the election. Formally, $\mathcal{R} : \mathcal{B}(C)^n \to 2^C$.
- A *social welfare function* $\mathcal{R}$ maps a profile to a set of linear orders over the candidates. Formally, $\mathcal{R} : \mathcal{B}(C)^n \to 2^{\mathcal{L}(C)}$.

Note that the above definition allows for an empty set of winners. However, most aggregation functions are designed to always output a winner. If the set of winners of a voting rule is a singleton, the respective candidate is called a unique winner of the election, and else a nonunique winner. In the case where the output of the aggregation function is always a singleton, the function is said to be resolute, and irresolute otherwise. A *tiebreaking scheme* is a function that maps a set to a single member from this set and can be used to make an irresolute aggregation function resolute. See also the book chapter by Baumeister and Rothe (2015) for a detailed introduction to singlewinner preference aggregation.

In this thesis, one important family of voting rules with ordinal ballots are the (positional) scoring rules. Each scoring rule is associated with a family of scoring vectors of the form $\alpha = (\alpha_1,\dots,\alpha_m)$ for each number m of candidates, where $\alpha_i \in \mathbb{N}$, $\alpha_1 \geq \cdots \geq \alpha_m$, and $\alpha_1 > \alpha_m$. For example, the scoring rule *plurality* uses scoring vectors of the form $(1,0,\dots,0)$, *veto* (also called anti-plurality) has scoring vectors of the form $(1,\dots,1,0)$, and *Borda Count* uses scoring vectors of the form $(m-1,m-2,\dots,1,0)$ for m candidates

[3] Note that most aggregation functions are only defined for a single type of ballot.

(see Example 2.9). For each voter, a candidate receives the points as indicated in the scoring vector according to the rank in the respective voter's ordinal ballot. For example, the highest-ranked candidate in a ballot receives α_1 points. These points are added to compute the resulting score for each candidate, and the winners of the election are then the candidates with the highest score (also called plurality winners, veto winners, ..., respectively). Let r_α denote the scoring rule with scoring vector α, let C be a set of candidates, and let P be a profile over C, then

$$r_\alpha(P) = \operatorname*{argmax}_{c \in C} \sum_{v_i \in P} \alpha_{pos_i(c)}.$$

In the context of approval ballots, the rule *approval voting* elects the candidates that appear in the most ballots, i.e., have the highest approval score. See the book by Laslier and Sanver (2010) for a detailed analysis of approval voting. Note that approval voting in the context of plurality ballots (respectively, veto ballots) is also called plurality (respectively, veto). Here, a plurality winner (respectively, veto winner) is the candidate with the most approvals (respectively, least disapprovals). Scoring rules can also be modified to allow for several consecutive rounds. These *iterative (positional) scoring rules* proceed in a (variable or fixed) number of rounds, where the candidates with the lowest scores are eliminated each round. After elimination, the voters' ballots are reduced to only include the candidates that are still participating. The election ends when all remaining candidates have the same score (and are therefore considered the winners of the election) or—in the case of a fixed number of rounds—after the last round (where the candidates with the highest score in the reduced profile after the final elimination are considered the winners). My contribution in Chapter 4 deals with the iterative scoring rules iterated veto and veto with runoff. For *iterated veto* (see Example 2.12), all but the candidates with the highest veto score are eliminated. *Veto with runoff* (see Example 2.10) proceeds in two fixed rounds. In the first round, all candidates that do not have the highest veto score are eliminated, unless this leaves only one candidate, then this candidate and the candidate(s) with the second-highest veto score proceed to the next round. Further iterative scoring rules in this thesis include the Hare rule, iterated plurality, and plurality with runoff (all based on plurality), the Baldwin and the Nanson rule (both based on Borda), and another veto-based rule called Coombs rule, see Chapter 4.

There are various aggregation functions using ordinal ballots that rely on the concept of pairwise comparisons. Given a profile of ordinal ballots, a candidate c wins a pairwise comparison against a candidate d if c is ranked higher than d in more ballots than the other way round. A candidate that wins the pairwise comparisons against all other candidates is called a *Condorcet winner*. Note that a Condorcet winner does not always have to exist. The voting rule *Condorcet* (Condorcet, 1785) elects the Condorcet winner, if such a candidate exists. The voting rule *Copeland* (see Example 2.8) based on a social welfare function proposed by Copeland (1951) elects the candidates with the most wins in pairwise comparisons against other candidates where a tie with another candidate counts as half a win. However, there are several variants depending on how tied candidates are treated. Other examples include counting a tied comparison as a win or as a loss. Define the Kendall tau distance d_τ (Kendall, 1938) between two linear orders $>_i$ and $>_j$ as the number of candidates where both orders disagree, i.e.,

$$d_\tau(>_i, >_j) = |\{(a,b) \mid a,b \in C \wedge a >_i b \wedge b >_j a\}|.$$

The social welfare function *Kemeny rule* (Kemeny, 1959) returns the linear orders that minimize the Kendall tau distance to the profile. Note that the rule Kemeny originally proposed uses weak orders. However, the version with linear orders is more prevalent.

For large-scale applications, an extremely important aspect of a voting rule is the computational complexity of the winner determination. In these cases, it is essential to be able to decide in polynomial time whether a candidate is a winner of the election. Formally, the winner determination decision problem is defined as follows.

$\mathcal{R}$-WINNER-DETERMINATION	
Given:	An election (C,P), where $C = \{c_1, \dots, c_m\}$ is a set of candidates and $P = (v_1, \dots, v_n)$ is a profile, and a designated candidate $w \in C$.
Question:	Is w a winner of the election using voting rule $\mathcal{R}$, i.e., is $w \in \mathcal{R}(P)$?

Note that it is possible in polynomial time to compute the set of winning candidates for all scoring rules including iterative scoring rules, so the respective winner determination problem is in P. However, the winner determination problem for the Kemeny rule is intractable (Bartholdi III et al., 1989b; Hemaspaandra et al., 2005).

Another important aspect of aggregation functions is the properties they satisfy. For example, for the voting rule that always returns the highest-ranked candidate of the first voter as the winner, the winner determination is tractable, but it definitely does a bad job at aggregating the ballots and therefore the underlying preferences of the voters. The aforementioned voting rule is dictatorial, so a *non-dictatorial* aggregation function is certainly desirable. A voting rule $\mathcal{R}$ is *non-imposed* . if for each candidate $c \in C$, there exists at least one profile P in the domain of $\mathcal{R}$ so that $\{c\} = \mathcal{R}(P)$. Based on the aforementioned notion of a Condorcet winner, a voting rule is *Condorcet consistent* if it always returns the Condorcet winner as the unique winner if such a candidate exists. The Condorcet rule is obviously Condorcet consistent, whereas the positional scoring rules are not (Fishburn, 1974). For ordinal ballots, a voting rule $\mathcal{R}$ is *monotone* (Arrow, 1950) if a winner $w \in \mathcal{R}$ remains a winner whenever a voter improves the rank of w in her ballot (and does not change any other relative rankings), and *positive responsive* (May, 1952) if this improvement leads to w being a unique winner. See Chapter 3 for an adaption of these axioms to multiwinner elections as introduced in Section 2.2.3.

Some (intuitively desirable) properties for social welfare functions for ordinal ballots include the following axioms: The axiom *universal domain* demands that the domain of the social welfare function has to consist of all possible profiles over the given candidate set. A social welfare function $\mathcal{R}$ is *Pareto optimal* (see, e.g., the book by Arrow (1963)) if for all candidates $a, b \in C$, $a >_i b$ for all $v_i \in P$ implies that $a > b$ for all $> \in \mathcal{R}(P)$. This axiom ensures that $\mathcal{R}$ actually considers the voters' ballots when returning a winning linear order, by following the voters' lead when they all agree on a relative ranking between two candidates. Furthermore, $\mathcal{R}$ is *independent of irrelevant alternatives* (see, e.g., the book by Arrow (1963)) if for each pair of candidates $a, b \in C$ and each pair of profiles $P = (p_1, \ldots, p_n)$ and $Q = (q_1, \ldots, q_n)$ that agree on the relative ranking of a and b for all voters $i \in N$,[4] all linear orders in $\mathcal{R}(P)$ and $\mathcal{R}(Q)$ also agree on the relative ranking of a and b. That means that $\mathcal{R}$ should decide on the relative ranking of a and b in the winning linear order solely by considering the relative rankings of them in the voters' ballots and not rankings of the other candidates.

However, not all aforementioned properties are compatible with each other in the sense that they can be satisfied by a single aggregation function. In the context of ordinal ballots, the famous impossibility result by Arrow (1963) states that for at least three candidates, there exists no social welfare function using ordinal ballots that has a universal domain, is Pareto optimal, independent of irrelevant alternatives, and is non-dictatorial, i.e., the result does

[4]This means that either $a >_i b$ in both p_i and q_i, or $b >_i a$ in both p_i and q_i, for all voters $i \in N$.

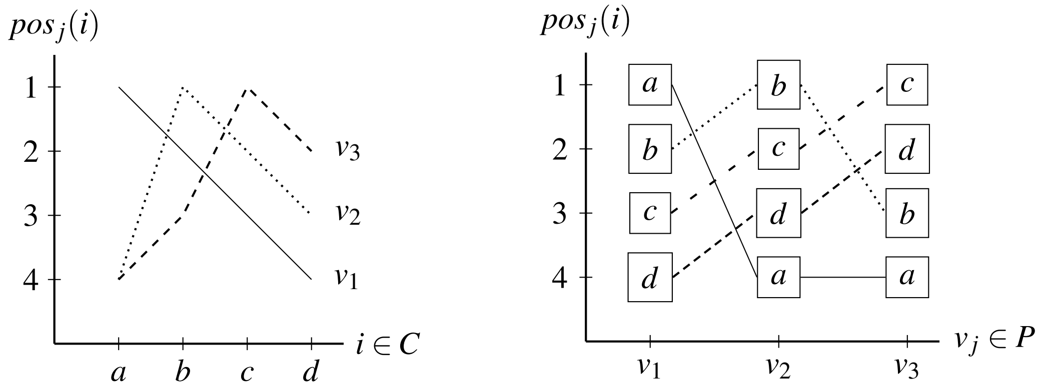
not correspond to a fixed voter's ballot for each given profile. Note that Arrow's theorem can also be restated using voting rules and adapted notions of the relevant axioms (Taylor, 2005). One way to circumvent Arrow's theorem is to restrict the domain of the considered aggregation function. Black (1958) introduces the concept of single-peaked domains for ordinal ballots. Given an ordering of the candidates π, the corresponding single-peaked domain contains all linear orders $>_i$ where $top(i)\ \pi\ c_j\ \pi\ c_k$ or $c_k\ \ \pi\ c_j\ \pi\ top(i)$ implies that $c_j >_i c_k$ for all $c_j, c_k \in C$, where $top(i)$ denotes the highest ranked candidate in $>_i$. This ordering corresponds to arranging the candidates on an axis, e.g., left to right in political elections. To illustrate this property, each single-peaked order only has one peak when the positions of the candidates in the ballots are plotted on axis π (see Example 2.7). Bartholdi and Trick (1986) show that it is possible in polynomial time to decide whether a given profile is single-peaked, i.e., whether each ballot in the given profile is single-peaked with respect to an ordering.

Another well-studied domain restriction are the single-crossing domains as introduced by Mirrlees (1971) in the context of income taxation, which contain the profiles where, given an ordering π of the voters, there are no voters v_i, v_j, v_k and candidates c, d so that $v_i\ \pi\ v_j\ \pi\ v_k$, but $c >_i d$, $d >_j c$, and $c >_k d$. This corresponds to the idea that the voters are ordered in a way so that for each pair of candidates $c, d \in C$, the voters ranking c higher than d form an interval in the profile. To illustrate this property, draw a line between each occurence of a candidate in the profile. Then for each pair of candidates, the respective lines between them are allowed to only cross once (see Example 2.7). Elkind et al. (2012) and Bredereck et al. (2013) provide polynomial algorithms to decide whether a given profile is single-crossing, and Bredereck et al. also characterize the single-crossing domain by identifying two forbidden substructures of the profiles. Faliszewski et al. (2011) adapt the notion of single-peaked and Elkind and Lackner (2015) adapt the notion of single-crossing to approval ballots, and Elkind et al. (2020b) give a characterization for domains that are single-peaked and single-crossing at the same time.

Example 2.7 (single-peaked and single-crossing profile)**.** Let $C = \{a, b, c, d\}$ be a candidate set and let P be the profile in Figure 2.1a of ordinal ballots over C. Figure 2.1b plots the candidates against their rank in the respective ballots for the candidate ordering π where $a\ \pi\ b\ \pi\ c\ \pi\ d$. Since each plot has exactly one peak, p is single-peaked. Furthermore, Figure 2.1c plots the ballots for the voter ordering π with $v_1\ \pi\ v_2\ \pi\ v_3$ and shows that the lines connecting the candidates cross each other line at most once. Therefore, P is also

$$
\begin{array}{lll}
P\colon & v_1 = & a > b > c > d \\
 & v_2 = & b > c > d > a \\
 & v_3 = & c > d > b > a
\end{array}
$$

(a) Profile P



(b) Profile P is single-peaked.

(c) Profile P is single-crossing.

Figure 2.1: The profile P in Example 2.7 is single-peaked and single-crossing.

single-crossing.

Restricted domains not only allow for non-dictatorial aggregation functions that satisfy the axioms in Arrow's theorem (and allow for non-manipulable aggregation functions that satisfy the axioms in the Gibbard-Satterthwaite Theorem, see Section 2.2.2), but can also have an impact on the computational complexity of the winner determination of other aggregation functions. For example, Brandt et al. (2015) show that the highly intractable winner determination of the Dodgson rule[5] and the Kemeny rule become tractable for single-peaked profiles. See the book chapter by Elkind et al. (2017b) for an overview on restricted domains.

[5]Given a profile of linear orders, the *Dodgson rule* returns the candidate(s) for which the number of swaps between adjacent candidates in the ballots needed to make this candidate the Condorcet winner is minimal (Dodgson, 1876). The complexity of the winner determination for the Dodgson rule was studied by Bartholdi III et al. (1989b) and Hemaspaandra et al. (1997).

2.2.2 Strategic Behavior

Unfortunately, not all actors in an election are honest. This section explores strategic behavior in elections where an attacker changes some aspects of the preference aggregation in order to reach a more desirable outcome.

There are several ways for strategic attacks on preference aggregation procedures.

- In *manipulation* (see, e.g., the book chapter by Zwicker (2016)), an agent that participates in the aggregation procedure reports an insincere ballot, i.e., a ballot that does not correspond to her underlying preferences over the candidates, in order to reach a more desired outcome of the election (see Example 2.8). In some problem variants, there can be a group of manipulators that work together to reach a joint objective, or there are (distance) restrictions on which insincere ballots may be reported. This thesis studies a slightly different notion of manipulation in the context of iterative elections as introduced in Section 2.2.4 where an attacker does not report an insincere ballot, but an insincere poll. See Chapter 5 for my contribution to the topic manipulation in iterative elections.

- In *bribery*, an external agent is able to bribe a given number of voters to submit insincere ballots in order to change the election result in his or her favor (see Example 2.9). Variations of the problem include prices for the voters, i.e., the briber has a budget and each voter (or even each change in a ballot) has a (possibly different) price attached. This thesis studies a model called shift bribery as introduced on page 23. See Chapter 4 for my contribution to the topic bribery for iterative scoring rules.

- In *control*, an external agent called the chair does not have any influence on the agents' ballots, but on the structure of the aggregation procedure (see Example 2.10). This can include adding or deleting candidates or voters, or partitioning the voters or candidates in groups where the aggregation takes place for each group separately. This thesis studies control for the related setting of judgment aggregation as introduced in Section 2.3. See the book chapter by Faliszewski and Rothe (2016) for an overview of bribery and control in voting, and see Chapter 6 for my contribution to the study of judge control problems in judgment aggregation.

Note that it is necessary to define beforehand what constitutes a more desirable outcome for an attacker. One way is via a preference (see, e.g., the book by Taylor (2005)), where for a resolute voting rule the new outcome is preferred to the old one if and only if the attacker prefers the new winner to the old winner. In the case of irresolute voting rules, the notion of a "better outcome" is more complicated and requires assumptions about the attacker's preferences over sets of candidates. For an overview over these so-called set extensions see, e.g., the book chapter by Barberà et al. (2004). In the models used in this section, there is a target candidate given. In a *constructive* attack, the attacker wants the target candidate to win the election, whereas in a *destructive* attack, the attacker tries to prevent the target candidate from winning.

Unfortunately, in most cases, voting rules are susceptible to strategic attacks. Gibbard (1973) and Satterthwaite (1975) independently proved the so-called Gibbard-Satterthwaite Theorem stating that there are no resolute, non-dictatorial and non-imposing voting rules for at least three candidates that cannot be successfully manipulated, assuming there are no restrictions on the voters' ballots.[6] This result started research on how manipulation can nevertheless be prevented. One important aspect in this research is the computational complexity of deciding whether a successful strategic attack is possible in the given election. If it turns out that the underlying problem is computationally hard for a voting system, then this might discourage an attacker.

Bartholdi III et al. (1989a) pioneered the approach of studying the computational complexity of manipulation by studying the following decision problem for voting rules using ordinal ballots.

$\mathcal{R}$-MANIPULATION	
Given:	A set of candidates C, a profile $P = (v_1, \ldots, v_n)$ over C, and a designated candidate $c \in C$.
Question:	Is there a ballot v_{n+1} so that $c \in \mathcal{R}((v_1, \ldots, v_n, v_{n+1}))$?

They show that for all voting rules that satisfy certain properties and for a single manipulator, there exists a polynomial-time greedy algorithm that computes a ballot for the manipulator to successfully make a designated candidate the winner of the election (or

[6] Black (1958) (respectively, Saporiti and Tohmé (2006)) show the existence of a non-manipulable rule when the input profile is single-peaked (respectively, single-crossing).

P:	$v_1 =$	$a > b > c > d$
	$v_2 =$	$a > b > c > d$
	$v_3 =$	$b > c > d > a$

(a) Original profile P

P':	$v_1 =$	$a > b > c > d$
	$v_2 =$	$a > b > c > d$
	$v_3 =$	$b > c > d > a$
	$v_4 =$	$b > a > c > d$

(b) Manipulated profile P'

Table 2.1: Original and manipulated profile in Example 2.8.

determines that such a ballot does not exist). Examples of such voting rules include the family of positional scoring rules. In contrast to this result, they show that the computational complexity of manipulating a voting rule called second-order Copeland (a variant of the Copeland rule) is NP-hard, although the winner can be computed in polynomial time. Bartholdi III and Orlin (1991) show that is computationally hard to manipulate the widely used iterative voting rule called STV (Single Transferable Vote, in the singlewinner version also called the Hare rule, see Chapter 4). Conitzer and Sandholm (2003) add a preround to several well-known voting rules to increase the complexity of manipulation. Conitzer et al. (2007) introduce the concept of coalitional manipulation where a group of manipulators cast ballots to achieve a common goal. A slightly different formulation of the manipulation decision problem includes the manipulator's sincere ballot in the given profile and asks whether the manipulator can successfully change her ballot to achieve a better outcome. Instead of allowing a complete change in the manipulator's ballot, one can restrict the new ballot to have at most a certain distance to the original ballot (Obraztsova and Elkind, 2012). See the survey by Faliszewski and Procaccia (2010) and the book chapter by Conitzer and Walsh (2016) for an overview of results for the computational complexity of manipulating voting rules.

Example 2.8 (manipulation)**.** Let $C = \{a,b,c,d\}$ be a candidate set and let $P = (v_1, v_2, v_3)$ be a profile over C defined in Table 2.1a. Consider the voting rule *Copeland* (r_{Cope}) where candidates acquire one point for each win in a pairwise comparison and half a point for each tie, and the candidates with the highest score win the election. Assume that a manipulator wants candidate b to win the election in profile P. Then she can submit the ballot $v_4 = b > a > c > d$. The election with profile $P' = (v_1, v_2, v_3, v_4)$ (see Table 2.1b) proceeds as follows: Candidate a wins the pairwise comparison against c and d and ties with b, so that a has a score of 2.5. Candidate b has a score of 2.5, candidate c a score of 1, and candidate d a score of 0. Therefore, $\{a,b\} = r_{Cope}(P')$ and the manipulation was successful.

Faliszewski et al. (2009a) were the first to study the computational complexity of bribery in elections, including the following decision problem.

	$\mathcal{R}$-BRIBERY
Given:	A set of candidates C, a profile P over C, a designated candidate $c \in C$, and an integer $k \in \mathbb{N}$.
Question:	Is there a profile P' where at most k voters change their ballots compared to P, so that $c \in \mathcal{R}(P')$?

In the priced variant, voters do not have an equal price like in the above scenario, so the problem includes a price function for voters and asks whether in the profile P' only voters changed their ballots whose total price does not exceed the budget k. Hemaspaandra et al. (2007) focus on destructive bribery where the briber wants to prevent the victory of a designated candidate. Dey et al. (2017) introduce a mix of manipulation and bribery where a briber can only bribe voters that profit from this action, i.e., voters who prefer the outcome after the bribery to the original one. As for manipulation, it is possible to restrict the new ballots of the bribed voters, for example by a distance restriction to the sincere ballot. This models scenarios where voters are not willing to deviate too much from their sincere ballot, be it for ethical reasons or the fear of negative consequences for casting an insincere ballot if there is a risk of having their ballot exposed. Dey (2021) studies this problem for bribery. Baumeister et al. (2019) introduce a generalized version of distance bribery where the resulting bribed profile (and not the individual votes) is subject to a distance restriction.

Example 2.9 (bribery)**.** Let $C = \{a,b,c,d\}$ be a candidate set and let P be a profile over C defined in Table 2.2a. Consider the scoring rule Borda Count (r_{Borda}). Recall that under Borda, each candidate receives a score based on the positions in the voters' ballots, and the candidates with the highest score win. Assume that a briber wants candidate c to win in profile P and is able to bribe at most one voter. In P, candidate a receives a score of $|C|-1+|C|-1+|C|-4 = 3+3+0 = 6$, whereas candidate b has a score of 7, candidate c has a score of 4, and candidate d has a score of 1, therefore $\{b\} = r_{Borda}(P)$. If the briber

P:	$v_1 =$	$a > b > c > d$
	$v_2 =$	$a > b > c > d$
	$v_3 =$	$b > c > d > a$

(a) Original profile P

P' :	$v_1' =$	$c > d > a > b$
	$v_2 =$	$a > b > c > d$
	$v_3 =$	$b > c > d > a$

(b) Bribed profile P'

Table 2.2: The original and the bribed profile in Example 2.9.

convinces voter 1 to change v_1 into $v_1' = c > d > a > b$, then c wins in the resulting profile P' (see Table 2.2b) with a score of 6 in contrast to the scores of 4, 5, and 3 of candidates a, b, and d, respectively. Therefore, the bribery was successful.

The computational complexity of control by adding, deleting, and partitioning candidates or voters was first studied by Bartholdi III et al. (1992) for plurality and the Condorcet rule. For example, the following decision problem shows the notion of control by adding candidates. Let P_S denote the restricted profile of P where each ballot in P is restricted to the candidates in S.

$\mathcal{R}$-CONTROL-BY-ADDING-CANDIDATES	
Given:	A set of candidates C, a set of additional candidates D, a profile P over $C \cup D$, a designated candidate $c \in C$, and an integer $k \in \mathbb{N}$.
Question:	Is there a set $D' \subseteq D$ of size at most k so that $c \in \mathcal{R}(P_{C \cup D'})$?

Faliszewski et al. (2009b) study constructive control for these control types for the Copeland rule for all variants of ties in the pairwise comparisons. Furthermore, Elkind et al. (2011) provide, among others, results for the constructive control of Borda by adding candidates. Analogous to bribery, Miasko and Faliszewski (2016) introduce prices to control actions.

Example 2.10 (control by adding candidates)**.** Let $C = \{a,b,c\}$ and $D = \{d,e\}$ be candidate sets, let P be the profile over $C \cup D$ defined in Table 2.3a and let P_S be the restricted profile of P where each ballot is restricted to candidates in S.

Consider the iterative voting rule veto with runoff (r_{VRO}). Recall that veto with runoff proceeds in two fixed rounds. After the first round, all candidates that do not have the highest veto score are eliminated, unless that leaves just one candidate, then the candidates

P:	$v_1 =$	$a > e > b > c > d$
	$v_2 =$	$e > a > b > c > d$
	$v_3 =$	$e > b > c > d > a$

(a) The profile P over $C \cup D$.

P_C:	$v_1 =$	$a > b > c$
	$v_2 =$	$a > b > c$
	$v_3 =$	$b > c > a$
$P_{C \setminus \{c\}}$:	$v_1 =$	$a > b$
	$v_2 =$	$a > b$
	$v_3 =$	$b > a$

(b) Election without adding candidates.

$P_{C \cup \{d\}}$:	$v_1 =$	$a > b > c > d$
	$v_2 =$	$a > b > c > d$
	$v_3 =$	$b > c > d > a$
$P_{(C \cup \{d\}) \setminus \{a,d\}}$:	$v_1 =$	$b > c$
	$v_2 =$	$b > c$
	$v_3 =$	$b > c$

(c) Election after adding candidate d.

Table 2.3: Profiles in Example 2.10.

that do not have the highest or second-highest veto score are eliminated. The election winners are the candidates with the highest veto score in the second round. Assume that the chair wants candidate b to win and is able to add at most one candidate from D to the candidate set. In the election using profile P_C (see Table 2.3b), b is the unique candidate with the highest veto score, namely a score of 3, whereas a has the second-highest score of 2. Therefore c is eliminated after the first round, resulting in the profile $P_{C \setminus \{c\}}$. Here, a has a veto score of 2 in contrast to the veto score of 1 of b, therefore $b \notin \{a\} = r_{VRO}(P_C)$.

Next, consider the profile $P_{C \cup \{d\}}$ where the chair added candidate d (see Table 2.3c). Both b and c have the highest veto score of 3, so that a and d are eliminated after the first round, resulting in the profile $P_{(C \cup \{d\}) \setminus \{a,d\}}$. Here, b has a veto score of 3, whereas c has a veto score of 0, so $b \in \{b\} = r_{VRO}(P_{C \cup \{d\}})$. Therefore, the control action was successful.

Note that for all aforementioned results, an attacker is assumed to have complete information about the election, notably about the exact ballots of all voters. This is a highly unlikely assumption, but widely accepted since it cannot be easier to favorably change the election result in the presence of incomplete information than in an election with complete information. Therefore, the hardness results carry over to the case where the information of the attacker is limited. However, limiting the information an attacker has can make influencing an election difficult even for voting rules where favorably changing the election result is a tractable problem in the complete information model. Conitzer et al. (2011) show for many common voting rules that, given the manipulator only knows partial ordinal ballots of the other voters, it is NP-hard to decide whether a dominating manipulation strategy exists (i.e., a vote that will not be detrimental to the manipulator in comparison to his truthful vote regardless of how the full preferences of the other voters look like). Dey et al. (2018) extend this model and introduce the concept of weak (respectively, strong) manipulation where the manipulator has to be successful in at least one (respectively, in all) extensions of the partial ballots. Inspired by typical limited information a manipulator has in real life, Endriss et al. (2016b) introduce different models for incomplete information (e.g., knowledge about the candidate scores, but not about ballots) and study the effect of limited information on the computational complexity of manipulation for several voting rules including scoring rules.

The aforementioned results assume that strategic attacks are infeasible when the corresponding decision problems are intractable. However, such a computational barrier as protection against strategic attacks might still not be effective in practice. Recall that the notion of NP-hardness relates to the worst-case complexity of a decision problem. Typical real-world scenarios might involve certain restricted preferences in contrast to the unrestricted preferences in theory, or might have certain small parameters such as the number of candidates.

Faliszewski et al. (2011) show that many manipulation and control problems become tractable when the given profile is single-peaked, whereas Brandt et al. (2015) show the same for bribery. Magiera and Faliszewski (2017) show that several control problems become tractable for restricting the input to single-crossing profiles. Procaccia and Rosenschein (2007) show that for most profiles, it suffices to compute the fraction of untruthful and truthful votes to make a decision about the success of manipulation. Conitzer et al. (2007) introduce the problem of weighted coalitional manipulation (i.e., manipulation by a coalition of voters in the presence of weighted voters[7]) and determine for a variety of common voting rules the number of candidates for which weighted coalitional manipula-

[7] In an election with weighted voters, there is a given weight function $\omega : N \to \mathbb{N}$ that maps each voter to a positive integer called the *weight* w_i of the voter i. The aggregation function then uses the modified profile P' as input, where P' denotes the profile where the ballot of each voter i is duplicated $w_i - 1$ times.

tion becomes intractable. Along these lines, Chen et al. (2017) study the parameterized complexity of control by adding or deleting candidates for common voting rules when parameterized by the number of voters, and Bredereck et al. (2015) and Knop et al. (2020) study the complexity of bribery when the number of candidates is small, the latter for a form of bribery called multi-bribery (where voters can be bribed to abstain, swap adjacent candidates in their preference list, or change the number of approved candidates). Bredereck et al. (2014a) identify research challenges in voting regarding parameterized complexity. See the survey by Rothe and Schend (2013) for an overview of typical-case, parameterized, and approximation results for manipulation and control.

There are other, more positive interpretations of the aforementioned strategic attacks in voting. Manipulation can increase the satisfaction of voters with the outcome as described in Section 2.2.4. The concept of destructive bribery is closely related to the *margin of victory*, first defined by Xia (2012). Here, the goal is to compute how robust the election result is by computing the minimal number of voters that would have to report a different preference in order to change the election outcome. Reisch et al. (2014) show that the corresponding decision problem is intractable for several tournament voting rules, and Dey and Narahari (2015) provide sampling algorithms to estimate the margin of victory. Another interpretation of bribery is *campaign management*, where the bribing action can be modeled by real-life election campaigns, e.g., spending money to promote a designated candidate in the ballots. In the constructive case, campaign managers want to promote their candidate, for example by running advertisements and therefore positively influencing the voter's preference about the candidate, which leads to changes in their ballot. Furthermore, campaigns are bound by a budget and therefore need to carefully evaluate which actions yield the desired results. In this context, it makes sense to define a price per change in a voter's ballot rather than a price for each voter. Schlotter et al. (2017) focus on approval-based voting rules and give classical and parameterized complexity results for campaign management, and Elkind and Faliszewski (2010) give approximation algorithms for several voting rules including scoring rules.

One important topic for this thesis in the context of campaign management is the notion of *shift bribery* for ordinal ballots. Shift bribery is a special case of *swap bribery* (Elkind et al. (2009); introduced in another context by Faliszewski et al. (2009b) as microbribery), where a price function $\rho_j : C \times C \to \mathbb{N}$ denotes the price for swapping two adjacent candidates c_i and c_k in the ballot of voter j, and a swap bribery is successful if a given candidate wins the election after a series of swaps in the voters' ballots whose total price do not exceed the given budget. The price functions for shift bribery are defined in Definition 2.11.

Definition 2.11 (shift bribery price function)**.** Let C be a set of candidates, let $c \in C$ be a target candidate, and let v_j be a linear order over C. A *shift bribery price function* is a function $\rho_j : \mathbb{N} \to \mathbb{N}$ for a voter j where $\rho_j(0) = 0$, $\rho_j(x) \geq \rho_j(y)$ for $x > y$, and $\rho_j(x) = \rho_j(x-1)$ for all $x \geq pos_j(c)$.

Note that $\rho(i)$ indicates the price for shifting the target candidate c forward by i positions in the ballot v_j and that it is not possible to shift the candidate further than the first position in the voter's ballot. For example, $\rho_j(1) = 2$ indicates that changing voter j's ballot from $\ldots > c_k > c_i > \ldots$ to $\ldots > c_i > c_k > \ldots$ costs two units. The shift bribery problem in the constructive case is then defined as follows.

$\mathcal{R}$-SHIFT-BRIBERY	
Given:	A set of candidates C, a profile P of n linear orders over C, a target candidate $c \in C$, a budget $B \in \mathbb{N}$, and a list of price functions $\rho = (\rho_1, \ldots, \rho_n)$.
Question:	Is it possible to shift c in the ballots in P such that the total price does not exceed the budget b and $c \in \mathcal{R}(P')$ for the new profile P'?

In particular, it is only possible to shift the target candidate rather than all candidates as in swap bribery, and the shift can only be forwards in the constructive case. This model assumes that campaign managers do not run so-called smear campaigns where they negatively influence the voters' opinions about opposing candidates. To define the destructive case where the goal is to prevent a target candidate from winning, the shift bribery price functions are interpreted as the price for shifting the designated candidate *backwards*.[8]

Example 2.12 (shift bribery)**.** Let $C = \{a, b, c, d\}$ be the candidate set and let P be the profile over C defined in Table 2.4a. Note that P_S denotes the restricted profile of P where each ballot is restricted to candidates in S. Consider the iterative scoring rule iterated veto (r_{IV}). Recall that each round, the candidates that do not have the highest veto score are

[8] Note that in the destructive case, the condition $\rho_j(x) = \rho_j(x-1)$ for all $x \geq pos_j(c)$ changes to $\rho_j(x) = \rho_j(x-1)$ for all $x \geq m - pos_j(c)$ in order to prevent shifting the candidate too far backwards.

P:	$v_1 =$	$a > b > c > d$
	$v_2 =$	$a > b > c > d$
	$v_3 =$	$b > c > d > a$

$P_{C\setminus\{a,d\}}$:	$v_1 =$	$b > c$
	$v_2 =$	$b > c$
	$v_3 =$	$b > c$

(a) Election with the original profile P.

P':	$v_1 =$	$a > b > c > d$
	$v_2 =$	$a > b > c > d$
	$v_3' =$	$b > c > a > d$

$P'_{C\setminus\{d\}}$:	$v_1 =$	$a > b > c$
	$v_2 =$	$a > b > c$
	$v_3' =$	$b > c > a$

(b) Election with the bribed profile P'.

Table 2.4: The election with the original and the bribed profile in Example 2.12.

eliminated. Assume that a briber wants candidate a to win the election and that he has a budget of 2. Further assume that the price function ρ_3 for voter 3 is defined as $\rho_3(0) = 0$, $\rho_3(1) = 1$, and $\rho_3(2) = \rho_3(3) = 3$. In P, the election proceeds as follows. In the first round, candidates b and c have the highest veto score with a score of 3, so that candidates a and d are eliminated, resulting in the profile $P_{C\setminus\{a,d\}}$. Here, candidate b has the highest veto score of 3 whereas c only has a veto score of 0. Therefore, $a \notin \{b\} = r_{IV}(P)$, so the briber has to employ a bribing action.

Recall that the briber can only shift the designated candidate a forward. In the ballots of voters 1 and 2, a is already in the top position and cannot be shifted forward. Shifting a at least two positions forward in ballot v_3 costs 3 units and therefore exceeds the budget. Consider the profile P' in Table 2.4b where a was shifted forward in v_3 by one position. The election then proceeds as follows. In the first round, d has the lowest veto score and is eliminated, resulting in the profile $P'_{C\setminus\{d\}}$. In the second round, a, b, and c have a veto score of 2, 3, and 1, respectively, therefore a and c are eliminated. It holds that $a \notin \{b\} = r_{IV}(P')$, so there does not exist a successful shift bribing action.

Elkind et al. (2020a) consider swap and shift bribery in single-peaked and single-crossing domains, whereas Dorn and Schlotter (2012) analyze the parameterized complexity for swap bribery. Bredereck et al. (2016a) show that different classes of price functions lead to different complexity results for shift bribery when parameterized by the budget for a number of voting rules. They also study the parameterized complexity regarding other parameters. Faliszewski et al. (2021) investigate the approximability of shift bribery for positional scoring rules and Copeland. For the destructive variant of the problem, Kaczmarczyk and Faliszewski (2019) give efficient algorithms for solving the destructive shift bribery problems and identify cases where the complexity between the constructive and destructive problem variant differ. Similar to the margin of victory, Shiryaev et al. (2013) study the robustness of elections by viewing the destructive swap bribery problem as a mean to measure the maximal number of errors that voters can make in their ballot before the election result changes. Here, an error in a voter's ballot is defined as a pair of adjacent candidates that are swapped in the ballot in contrast to the true preference of the voter. In another variant of shift bribery called *combinatorial shift bribery*, it is possible to affect several votes at once with just a single bribery action (Bredereck et al., 2016c). See Chapter 4 for my contribution to the study of the computational complexity of shift bribery.

2.2.3 Multiwinner Elections

In contrast to singlewinner elections as presented in Section 2.2.1, in the context of multiwinner elections, the goal is to elect a group of candidates called a *committee* as the winner of the election. That means that a multiwinner voting rule maps a profile and a desired committee size k to a set that consists of the winning committees, i.e., subsets of candidates of size exactly k.

Definition 2.13 (multiwinner voting rule)**.** Let C be a set of candidates, let P be a profile over C, let $k < |C|$ be a positive integer, and let $W_k = \{S \mid S \subseteq C \wedge |S| = k\}$ be the set of k-sized committees. A *multiwinner voting rule* $\mathcal{F}$ maps a tuple (P,k) to sets in W_k.

In the above definition of a multiwinner voting rule, the committee size is part of the input. However, there are a few scenarios where it makes sense to not fix the size of a winning committee, called the *variable number of winners* setting. For example, this setting can be motivated in the context of a Hall of Fame where each year, only the best of the best candidates (e.g., athletes) are supposed to be elected into the Hall of Fame. However, this number can vary: Some years, there might be several outstanding candidates, whereas in other years, there might even be none that are deemed exceptional enough to be honored. Kilgour (2016) provides an in-depth study of multi-winner voting rules with a variable number of winners in the approval-based setting and explores the axiomatic properties of corresponding voting rules. Faliszewski et al. (2020) study the complexity of these approval-based rules. The rest of this thesis only considers the fixed number of winners model as defined in Definition 2.13, where the desired committee size is predetermined. For an overview on multiwinner elections, see the book chapter by Faliszewski et al. (2017b).

There are several possible directions for the design of multiwinner voting rules. One prominent design concept is individual excellence where the winning committee consists of the best candidates according to certain criteria, e.g., approval scores. One application for these rules is shortlisting, i.e., preselecting a short list of candidates for the next stage (e.g., of a hiring process). An example of an excellence-based multiwinner voting rule is the rule by Debord (1992) that extends Borda Count to multiwinner voting with a fixed committee size. Lackner and Maly (2021) study shortlisting procedures using approval ballots.

A second design concept is diversity, where the goal is to elect candidates in a way that preferably each voter has a candidate she approves of. A possible application is the facility location problem, where several facilities (e.g., hospitals) have to be placed (e.g., in a city) according to certain criteria. If candidates are possible locations and voters approve of locations that are (easily) accessible to them, multiwinner voting rules based on diversity help in finding optimal locations so that each voter has hospital access. See the article by Skowron et al. (2016) for other applications for diversity. A voting rule aiming at diversity was proposed by Chamberlin and Courant (1983). The Chamberlin-Courant rules are a family of multiwinner rules that elect the committee that—given a misrepresentation function—minimizes the misrepresentation of voters. Here, each voter is assigned their most preferred candidate from the committee as a representative.

Definition 2.14 (family of Chamberlin-Courant rules)**.** Let C be a set of candidates, let N be the set of voters, let $\mathcal{B}(C)$ be the set of all possible ballots over C, let $P \in \mathcal{B}(C)$ be a profile, let $k > 0$ be the desired committee size, and let W_k be the set of all k-sized committees over C.

1. A *misrepresentation function* $\mu : N \times C \to \mathbb{N}$ computes how much a candidate misrepresents a voter.

2. For a committee $S \subseteq C$, an *assignment function* $\phi_S : N \to S$ assigns each voter a representative from S. For the Chamberlin-Courant rules, $\phi_S(i) = \text{argmin}_{c \in S}\, \mu(i,c)$, i.e., voters are represented by the best candidate in a given committee.

3. A *committee misrepresentation function* $d : \mathcal{B}(C) \times W_k \to \mathbb{N}$ computes for a given profile P and k-sized committee S how misrepresented the voters in P are when S is elected.

4. A *Chamberlin-Courant rule* is a multiwinner voting rule f_{CC} so that

$$f_{CC}(P,k) = \underset{S \in W_k}{\operatorname{argmin}}\ d(P,S).$$

Note that the Chamberlin-Courant rules are used for both ordinal and approval ballots and that they can also be stated in terms of satisfaction instead of dissatisfaction (or misrepresentation) of the voters. A widely used misrepresentation function μ for ordinal ballots outputs the inverse Borda score, i.e., $\mu(i,c) = pos_i(c) - 1$ for a voter v_i and a candidate c, whereas the variant for approval ballots returns 0 if the voter approves of the candidate, and 1 otherwise. The choice for the committee misrepresentation function d depends on the application. The utilitarian approach that minimizes the total misrepresentation of each of the n voters with the winning committee (i.e., that uses the committee misrepresentation function $d_{sum}(P,S) = \sum_{1 \le i \le n} \mu(i, \phi_S(i)))$ is called the *minisum* principle, whereas the egalitarian approach that minimizes the maximum misrepresentation for a voter (i.e., that uses the committee misrepresentation function $d_{max}(P,S) = \max_{1 \le i \le n} \mu(i, \phi_S(i)))$ is called the *minimax* principle (see, e.g., the book chapter by Kilgour et al. (2006)).

Example 2.15 (Chamberlin-Courant rule)**.** Let $C = \{a,b,c\}$ be the candidate set, let $k = 2$ be the committee size, and let P be the profile in Table 2.5a of ordinal ballots over C.

For each committee $W \subset C$, the assignment function Φ_W assigns the respective best-ranked

P:	$v_1 =$	$a > b > c$
	$v_2 =$	$b > a > c$
	$v_3 =$	$b > a > c$
	$v_4 =$	$c > b > a$

(a) Profile P

	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$
v_1 :	$\mu(1,a) = 0$	$\mu(1,a) = 0$	$\mu(1,b) = 1$
v_2 :	$\mu(2,b) = 0$	$\mu(2,a) = 1$	$\mu(2,b) = 0$
v_3 :	$\mu(3,b) = 0$	$\mu(3,a) = 1$	$\mu(3,b) = 0$
v_4 :	$\mu(4,b) = 1$	$\mu(4,c) = 0$	$\mu(4,c) = 0$
Σ :	1	2	1

(b) Misrepresentation of each voter for committees of size 2.

Table 2.5: Profile and misrepresentation values in Example 2.15.

candidate from W to each voter. For example, $\Phi_{\{a,b\}}$ assigns candidate a to voter 1, and candidate b to voters 2, 3, and 4. Table 2.5b states the misrepresentation for each voter with a given committee and computes $d(P,W)$ for each committee W of size 2 using the minisum principle. It follows that $f_{CC}(P,2) = \{\{a,b\},\{b,c\}\}$.

Finally, the third important design concept in multiwinner elections is the concept of proportional representation. In many settings, e.g., in the election of a parliament, it is important to ensure that voters' opinions are adequately represented. Skowron (2018) studies the proportionality of common multiwinner rules and ranks them in a hierarchy. Lackner and Skowron (2019) classify multiwinner voting rules by how close they are to the design concepts individual excellence and proportionality. The voting rules by Monroe (1995) aim exactly at fully representing voters. The Monroe rules use the same underlying idea as the Chamberlin-Courant rules, but additionally require that each candidate in a committee represents the same number of voters (barring rounding differences), therefore the assignment function does not always assign a voter's most preferred committee member as their representative.

However, even determining whether there exists a committee with a misrepresentation of at most a given integer is already intractable for both the Chamberlin-Courant and the Monroe rules (Procaccia et al., 2007). Deciding whether a given committee is a winning committee for these voting rules (which is one variant of the *winner determination problem*) is at least as hard as the aforementioned problem. In contrast to the winner determination of the Monroe rules, the winner determination for the Chamberlin-Courant rules become tractable in the presence of profiles from restricted domains (see the article by Betzler et al. (2013) for parameterized complexity results and single-peaked profiles, and see the article by Skowron et al. (2015b) for single-crossing profiles). Peters (2018b) introduces a new technique to compute winning committees in polynomial time for single-peaked profiles for several voting rules whose winner determination is NP-hard, including the Chamberlin-Courant rules. Skowron and Faliszewski (2017) give approximation algorithms for Chamberlin-Courant under approval ballots, whereas Skowron et al. (2015a) prove that the winner determination of both the Chamberlin-Courant and the Monroe rules is inapproximable under linear orders.

To be able to properly decide on a multiwinner rule in a given context, it is important to study the properties of multiwinner rules. Felsenthal and Maoz (1992) adapt four singlewinner rules to elect more than one winner and study how this change affected their properties. Elkind et al. (2017a) introduce a variety of axioms—partly inspired by axioms for singlewinner rules, see Section 2.2.1—and show which of several well-known multiwinner rules satisfy them. Lackner and Skowron (2021) study the properties of several approval-based multiwinner rules. There are also a variety of axioms that deal with the *stability* of multiwinner rules. In the context of approval ballots, one axiom called *justified representation* (Aziz et al., 2017a) and variants thereof have received a lot of attention. A multiwinner rule satisfies justified representation if it always elects a committee so that there does not exist a large enough group of unrepresented voters that have a common approved candidate.

Definition 2.16 (justified representation (Aziz et al., 2017a))**.** Let C be the set of candidates, let $N = \{1,\dots,n\}$ be the set of voters, let $P = (v_1,\dots,v_n)$ be a profile over C using approval ballots, and let $k > 0$ be a committee size.

1. A committee S of size k satisfies *justified representation under P and k* if there does not exist a subset of voters $N' \subseteq N$ with size at least n/k, so that $\bigcap_{i\in N'} v_i \neq \emptyset$ and $v_i \cap S = \emptyset$ for each $i \in N'$.
2. A multiwinner voting rule $\mathcal{F}$ using approval ballots satisfies *justified representation* if for each profile P and each $k > 0$, each winning committee $W \in \mathcal{F}(P,k)$ satisfies justified representation under P and k.

Aziz et al. (2017a) also define a more demanding axiom named extended justified representation (EJR), whereas Sánchez-Fernández et al. (2017) introduce an in-between axiom called proportional justified representation (PJR). Aziz et al. (2018) give complexity results for EJR and PJR. An option to measure the stability of an elected committee for multiwinner rules based on ordinal ballots relies in principle on the notion of a Condorcet winner (see Section 2.2.1). A committee is *Gehrlein-stable* (Gehrlein, 1985; Ratliff, 2003) if there are no candidates outside of the committee that a strict majority of voters prefers to any candidate in the committee. Based on the this concept, Barberà and Coelho (2008) study which multiwinner rules used for shortlisting are stable in the sense that the elected committee contains a (weak) Condorcet winner if one exists. Aziz et al. (2017b) study Gehrlein stability and local stability (the latter based on the Condorcet winning sets introduced by Elkind et al. (2015)) and determine, among other things, the computational complexity of mutiwinner rules that satisfy these types of stability. Gupta et al. (2019) complement these results by studying the parameterized complexity of finding a Gehrlein stable committee. See Chapter 3 for my contribution to the study of properties for multiwinner voting rules.

Bredereck et al. (2018) introduce a model where not only the voters' preferences, but also certain attributes of candidates such as gender or skill level play a part in finding the best committee. In the model by Kagita et al. (2021), voters do not approve candidates, but only such attributes. Izsak et al. (2018) study the case where there are certain relations between candidates that have to be taken into account when finding a good committee. For example, there might be candidates that work exceptionally well (respectively, exceptionally poorly) with each other. Further, Faliszewski et al. (2017a) introduce voting rules where the winning committee is a good compromise between several extremes, for example individual excellency and proportionality. Brill et al. (2019) give approximation algorithms for these type of balanced rules, and Kocot et al. (2019) adapt these rules to guarantee specific score results of committee scoring rules. Note that these *committee scoring rules* defined by Elkind et al. (2017a) are the multiwinner analogues of the (singlewinner) scoring rules introduced in Section 2.2.1 (see the articles by Skowron et al. (2019) and Faliszewski et al. (2019) for an axiomatic characterization of committee scoring rules).

There are a multitude of papers about strategic attacks on multiwinner rules. Lackner and Skowron (2018) define the axioms independence of irrelevant alternatives, monotonicity, and SD-strategyproofness to study the susceptibility of approval-based multiwinner rules. Continuing this work, Peters (2018a) proves an impossibility result that says that no multiwinner rule can satisfy a form of proportionality and a form of strategy-proofness at the same time. Meir et al. (2008) study the computational complexity of manipulation and control for SNTV (Single Non-Transferable Vote; a multiwinner equivalent of Plurality), Bloc (a multiwinner equivalent of k-approval, where k is the desired committee size), approval voting, and cumulative voting (a voting rule that uses cardinal ballots). Yang (2019) also focuses on manipulation and control and, among other things, studies the parameterized complexity of these attacks for approval-based rules. Obraztsova et al. (2013) study the complexity of manipulating scoring rules and give several polynomial-time algorithms. Further, they focus on the role of tie-breaking rules for the success of manipulation. In the context of bribery for approval-based multiwinner rules, Faliszewski et al. (2017c) study the computational complexity and approximability of bribery actions where only a single approval might be added, deleted, or moved within a vote, and Yang (2020) focuses on destructive bribery and gives classical and parameterized complexity results. Bredereck et al. (2016b) study the computational complexity of shift bribery in multiwinner elections for linear orders. Related to this line of research is the question of the robustness of a multiwinner voting rule. Bredereck et al. (2021) study what impact a swap of neighboring candidates has in the linear order of a voter for several rules. Further, they show that it is NP-hard to decide whether the election result can be changed by a given number of swaps.

Baumeister and Dennisen (2015) generalize the concept of the multiwinner rules minisum and minimax from dichotomous ballots to trichotomous ballots[9] and to complete and incomplete linear orders. Then Baumeister et al. (2015a) study the complexity of winner determination and manipulation for these rules, whereas Liu and Guo (2016) focus on the parameterized complexity of winner determination for the generalized minimax rules. Cygan et al. (2018) extend the results for winner determination for the minimax approval rule including approximation and parameterized complexity results. See Chapter 3 for my contribution on the research on minisum and minimax based rules.

2.2.4 Iterative Elections

As seen in Section 2.2.2, all "reasonable" voting rules can be manipulated. This section deals with iterative elections where voters may change their reported ballot in each round, i.e., repeatedly manipulate the election. This corresponds to the idea of *white manipulation*, where manipulation by voters is encouraged as it leads to a better outcome for the voters. Iterative elections can also model the case where voters submit their ballot sequentially, for example in Doodle polls.

Formally, let $N = \{1,\ldots,n\}$ be the set of voters, let C be the set of candidates, let $P = (p_1,\ldots,p_n)$ be a *preference profile*[10] (usually ordinal preferences) over C, let $P^0 = (a_1,\ldots,a_n)$ be an (optional) profile of ballots over C, and let $\mathcal{R}$ be a voting rule. Iterative elections can be modeled as a game (see, e.g., the book chapter by Faliszewski et al. (2016) for an introduction to noncooperative game theory), where the voters are the players and the set of *actions* A_i (also called *strategies*) that each player i can take corresponds to the set of possible ballots over C of a type applicable to $\mathcal{R}$, and are therefore equal

[9] In a trichotomous ballot, the set of candidates is partitioned into three sets corresponding to approved, indifferent, and disapproved candidates.

[10] A *preference profile* is a list of preferences over a candidate set.

for all voters. The game proceeds in turns, where—depending on the model—voters either all act simultaneously or each turn t, a voter j is singled out to take an action so that $a_i^{t-1} = a_i^t$ for all voters $i \in N \setminus \{j\}$. A strategy profile for turn t is then a profile $P^t = (a_1^t, \ldots, a_n^t) \in A_1 \times \cdots \times A_n = \mathcal{A}$, where in most applications, P^0 is a profile of initial ballots that are assumed to be truthful, i.e., where the ballots $a_1^0, \ldots, a_n^0$ correspond to the preferences $p_1, \ldots, p_n$. Each voter further has a utility function $u_j : \mathcal{A} \to \mathbb{R}$ where $u_j(P^t)$ denotes the utility voter j gains from the the strategy profile P^t in turn t. In the context of iterative elections, $u_j(P^t) > u_j(P^s)$ if voter j prefers $\mathcal{R}(P^t)$ to $\mathcal{R}(P^s)$ according to preference p_j. Note that in this context, $\mathcal{R}$ is generally assumed to be resolute, for example by applying a tie-breaking scheme to the outcome. Therefore, preferences over election outcomes can be determined straightforward. Assume that voters only use best-response dynamics to change their ballots i.e., they submit a ballot that will result in the best outcome for them.

Definition 2.17 (best response)**.** Let $N = \{1, \ldots, n\}$ be the set of players, let $\mathcal{A} = A_1 \times \cdots \times A_n$ be the set of strategy profiles, and let $u_j : \mathcal{A} \to \mathbb{R}$ be the utility function of player $j \in N$. An action $a_j \in A_j$ is a *best response* to the strategy profile $(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n) \in A_1 \times \cdots \times A_{j-1} \times A_{j+1} \times \cdots \times A_n$ if for all $a_j' \in A_j$,

$$u_j(a_1, \ldots, a_{j-1}, a_j, a_{j+1}, \ldots, a_n) \geq u_j(a_1, \ldots, a_{j-1}, a_j', a_{j+1}, \ldots, a_n).$$

Note that—depending on the information model—a voter's computed best response does not have to coincide with the actual best response in that situation since voters may not have access to the current strategy profile. See Chapter 5 for a model where voters can only see the submitted ballots of their neighbors in a social network. Further, voters are assumed to have no memory, i.e., their best response does only depend on the current state of the election, not past states, and they are assumed to be myopic, i.e., they do not have access to the preference profile and do therefore not predict any future deviations by a voter.

In the basic model of Meir et al. (2010), a voter only deviates if he is *pivotal*, i.e., if his best response changes the election outcome. In contrast to this, Obraztsova et al. (2016) study the model where a voter not only changes her ballot when she knows that this change will impact the outcome positively, but also deviates optimistically, i.e., when she believes that other voters will also change their ballot in her favor. The winner is announced as

P:	$p_1 =$	$a > b > c$
	$p_2 =$	$b > c > a$
	$p_3 =$	$c > b > a$

(a) Preference profile P

P^0:	$a_1^0 =$	$\{a\}$
	$a_2^0 =$	$\{b\}$
	$a_3^0 =$	$\{c\}$

(b) Ballot profile P^0

Table 2.6: The preference profile and ballot profile in Example 2.18.

P_a^1:	$a_1^1 =$	$\{a\}$
	$a_2^1 =$	$\{b\}$
	$a_3^1 =$	$\{a\}$

P_b^1:	$a_1^1 =$	$\{a\}$
	$a_2^1 =$	$\{b\}$
	$a_3^1 =$	$\{b\}$

P_c^1:	$a_1^1 =$	$\{a\}$
	$a_2^1 =$	$\{b\}$
	$a_3^1 =$	$\{c\}$

Table 2.7: Possible profiles in the first turn in Example 2.18.

soon as the profile does not change anymore. Iterative elections where voters take turns are path-dependent, i.e., the election winner depends on the order in which voters are allowed to deviate. See the book chapter by Meir (2017) for an overview of iterative elections.

Example 2.18 (iterative elections)**.** Let r_{Plu} be the plurality rule, let $C = \{a,b,c\}$ be a candidate set, let P in Table 2.6a be the preference profile over C, and let P^0 in Table 2.6b be the profile of sincere plurality ballots over C. Assume that ties are broken lexicographically and that voters change their ballot sequentially, starting with voter 3. The plurality winner of P^0 is $r_{Plu}(P^0) = \{a\}$. It holds that $A_1 = A_2 = A_3 = \{\{a\},\{b\},\{c\}\}$, so consider the profiles P_a^1, P_b^1, and P_c^1 in Table 2.7, where voter 3 submits the ballot $\{a\}$, $\{b\}$, and $\{c\}$, respectively, and the other voters do not change their ballots. It holds that $r_{Plu}(P_a^1) = r_{Plu}(P_c^1) = \{a\}$ and $r_{Plu}(P_b^1) = \{b\}$, so that

$$u_3(P_b^1) > u_3(P_a^1) = u_3(P_c^1) \; ,$$

therefore only the action $\{b\}$ is a best response for strategy profile $(\{a\},\{b\})$. In the profile P_b^1, the best-response of voter 2 and 3 is to not change ballots, whereas voter 1 does not change ballots because she is not pivotal. Therefore the winner of the iterative election is candidate b. Now assume that given profile P^0, it is the turn of voter 2. The only best response for voter 2 is to submit the ballot $\{c\}$, after which no other voter changes ballots.

P^0:	$a_1^0 =$	$\{a\}$
	$a_2^0 =$	$\{b\}$
	$a_3^0 =$	$\{c\}$

(a) Initial profile P^0.

P^1:	$a_1^1 =$	$\{a\}$
	$a_2^1 =$	$\{c\}$
	$a_3^1 =$	$\{b\}$

(b) Profile P^1 after the first turn.

Table 2.8: Profiles after turn 0 and 1 in Example 2.19.

Therefore, in this case the winner of the iterative election is candidate c, illustrating the path-dependency of iterative elections when voting sequentially.

An important topic for iterative elections is convergence. Since voters may change their ballots repeatedly, the resulting profiles can cycle, i.e., there may be a sequence of profiles that repeat over and over again.

Example 2.19 (convergence)**.** Consider the setting from Example 2.18, but this time assume that voters submit their ballots simultaneously. The best response for voter 2 and 3 in profile P^0 in Table 2.8a is to submit the ballots $\{c\}$ and $\{b\}$, respectively, whereas voter 1 does not change ballots. In the resulting profile P^1 in Table 2.8b, the best response for voter 2 and 3 is to submit the ballots $\{b\}$ and $\{c\}$, respectively, whereas voter 1 again does not change ballots, resulting again in profile P^0. Recall that voters do not have memories, therefore this cycle repeats over and over again and the election does not converge.

Meir et al. (2010) initiate the study of conditions and voting rules for which elections are guaranteed to converge, i.e., result in a profile where no voter has any incentive to change their ballot. Such a profile is also called an *equilibrum*. Note that—depending on the deviations allowed for the voters and their ballots—there are different kinds of equilibria, e.g., a Nash equilibrium defined as follows.

Definition 2.20 (Nash equlibrium)**.** Let $N = \{1,\dots,n\}$ be the set of players and let $\mathcal{A} = A_1 \times \cdots \times A_n$ be the set of strategy profiles. A strategy profile $P = (a_1,\dots,a_n) \in \mathcal{A}$ is in a *Nash equilibrium* if action a_j is a best response to the strategy profile $(a_1,\dots,a_{j-1}, a_{j+1},\dots,a_n) \in A_1 \times \cdots \times A_{j-1} \times A_{j+1} \times \cdots \times A_n$ for each player $j \in N$.

For example, when voters deviate sequentially, Meir et al. (2010) prove that pluality always converges to a Nash equilibrium when the inital ballot profile is sincere. Reyhani and

Wilson (2012) and Lev and Rosenschein (2016) show that no scoring rule except plurality and veto converges in iterative elections. The latter also study tie-breaking schemes and show that the result holds for all tie-breaking schemes and that even restricting the tie-breaking to a certain scheme does not guarantee convergence. Since iterative elections do not converge for many common voting rules, Grandi et al. (2013) and Obraztsova et al. (2015b) restrict the allowed deviations even more than the best-response dynamics in order to achieve guaranteed convergence. Gourvès et al. (2016) consider a model where voters are embedded in a social network. They introduce *considerate equilibria* to iterative voting, where voters do not selfishly update their ballot, but consider their neighbors in the network. A profile is then a considerate equilibrium when a *coalition* of voters consisting of a clique in a graph cannot deviate without harming themselves or their neighbors.

Iterative voting can improve the quality of the election outcome, i.e., lead to a higher utility of the voter. Brânzei et al. (2013) analyze the price of anarchy in iterative elections, which is defined as the ratio between the quality of the outcome in a Nash equilibrium that can be reached when starting a sequence of best responses from the original profile, and the quality of the outcome in the original profile. Meir et al. (2014) also analyze the quality of the outcome in iterative elections. They conduct experiments in the situation where voters are uncertain about the current state of the election. However, in real-world elections with underlying social networks, this effect of a better outcome is not as pronounced. Tsang and Larson (2016) explain this phenomenon with the fact that many voters are only connected to voters with a similar view to their own which lessens the possibilities for pivotal deviations.

Reijngoud and Endriss (2012) consider opinion polls as a source of information for voters. Fairstein et al. (2019) compare different models for strategic voting in the presence of opinion polls and introduce a heuristic to predict how voters will behave in this scenario. Sina et al. (2015) consider a social network that the voters are embedded in. Voters obtain their information by an opinion poll and by their neighbors in the social network and can act accordingly to this information. They introduce the concept of network control where the chair can introduce new edges to the social network to obtain a desired outcome. Wilczynski (2019) introduces the concept of manipulation by the agency that publishes the opinion poll. Here, the agency publishes an incorrect poll to influence the voters to change their ballots so that a designated candidate wins the election. See Chapter 5 for my contribution to the study of the computational complexity of poll manipulation.

2.3 Judgment Aggregation

This section generalizes the notion of preference aggregation. In the field of judgment aggregation, agents called *judges* have to come to collective decisions in the form of yes/no answers concerning several possibly logically related issues. Formally, let $\Phi = \{\varphi_1, \neg\varphi_1, \dots, \varphi_m, \neg\varphi_m\}$ be the agenda consisting of the φ_i (called *issues*, $1 \leq i \leq m$) to be decided over. Note that Φ is closed under complementation, consists of propositional formulas φ_i over a set of propositional variables built by using the standard boolean connectives, does not contain any doubly-negated formulas, and is assumed to be finite by a large part of the existing literature. Let $N = \{1, \dots, n\}$ be the set of judges. Each judge $i \in N$ has an individual *judgment set* $J_i \subseteq \Phi$ that is required to be complete (i.e., for each $\varphi \in \Phi$, J_i contains φ or its complement),[11] and J_i is required to be consistent (i.e., there exists a truth assignment for the underlying set of propositional variables so that each issue in J_i evaluates to true). Then $\mathcal{J}(\Phi)$ denotes the set of all possible individual judgment sets and $\mathbf{J} = (J_1, \dots, J_n) \in \mathcal{J}(\Phi)^n$ is called a *profile*.

Definition 2.21 (judgment aggregation procedure)**.** Let Φ be an agenda and let $N = \{1, \dots, n\}$ be the set of judges. A *judgment aggregation procedure* $\mathcal{P}$ is a function

$$\mathcal{P} : \mathcal{J}(\Phi)^n \to 2^{2^{\Phi}}$$

that maps a profile $\mathbf{J} = (J_1, \dots, J_n) \in \mathcal{J}(\Phi)^n$ to a set of (possibly incomplete and inconsistent) judgment sets over Φ.

Note that most of the prevalent procedures are resolute, i.e., the collective outcome is always a singleton. The formula-based framework based on Boolean algebra dominates the literature, but Dietrich (2007) shows that the concepts of judgment aggregation can also be employed with different types of logic. Ågotnes et al. (2011) also introduce a new framework for judgment aggregation based on modal logic. Endriss et al. (2016a) compare the formula-based framework with a framework where issues are propositional variables whose logical relations are expressed in an external integrity constraint. This model is a special case of the field of *binary aggregation*, see, e.g., the article by Dokow and Holzman (2010). For a detailed introduction to judgment aggregation, see the book chapters by Baumeister et al. (2015c), List and Puppe (2009), and Endriss (2016).

[11] See Terzopoulou et al. (2018) for a model where the judges' judgment sets are not required to be complete.

	c	b	ℓ
J_1	1	1	1
J_2	0	1	0
J_3	1	0	0
Maj	1	1	0

Table 2.9: Illustration of the doctrinal paradox/discursive dilemma

The famous *doctrinal paradox*—first presented by Kornhauser and Sager (1986) and later generalized to the *discursive dilemma* by Pettit (2001) and List and Pettit (2002)—shows that when deciding each issue majority-wise, the resulting majority outcome might be inconsistent even if all individual judgment sets are consistent. The corresponding rule is called *majority rule* and always outputs a single judgment set where a formula $\varphi \in \Phi$ is contained if and only if more than half of the judges accept it. To illustrate the discursive dilemma, consider the following example by List and Pettit (2002).

Example 2.22 (discursive dilemma)**.** In the discursive dilemma, three judges have to decide whether a contract was valid (c), whether there was a breach of contract (b), and whether the defendant is liable (ℓ), which is only the case if the contract was valid and it was breached. The agenda is $\Phi = \{c, \neg c, b, \neg b, \ell, \neg \ell\}$ where $\ell = c \wedge b$, and the profile $\mathbf{J} \in \mathcal{J}(\Phi)^3$ can be seen in Table 2.9. Note that a 1 means that the corresponding issue was accepted (i.e., a 1 for an issue φ in J_i means that $\varphi \in J_i$ and $\neg\varphi \notin J_i$), whereas a 0 indicates that this issue was rejected (i.e., a 0 for an issue φ in J_i means that $\varphi \notin J_i$ and $\neg\varphi \in J_i$). Even though all judges have consistent judgment sets (as required), taking the majority decision for each issue leads to an inconsistent outcome: A majority of judges accept that the contract was valid, that it was breached, but that the defendant is not liable.

Therefore, the majority rule—arguably the most simple and natural judgment aggregation procedure—has a major drawback, which started the research on judgment aggregation. See the article by Mongin (2012) for a detailed comparison between the doctrinal paradox and the discursive dilemma.

As a way out of this dilemma, Dietrich and Mongin (2010) separate the issues into premises and conclusions and only aggregate the judgments on the premises in the premise-based approach (respectively, the conclusions in the conclusion-based approach). The premise-based procedure then infers the conclusions from the premises. Dietrich and List (2007b) generalize the majority rule by introducing quota rules where an issue is contained in the collective outcome if the number of judges that accept this issue is at least as high as a given respective quota for this issue. However, as with the majority rule, this approach does not guarantee to produce complete and at the same time consistent outcomes, unless it is paired with the premise-based approach (resulting in the premise-based quota rules) and certain assumptions about the agenda are made. Note that the term “uniform” denotes that the quota is equal for all issues. In this thesis, the following definition for uniform premise-based quota rules is used (see also Example 2.27).

Definition 2.23 (uniform premise-based quota rules)**.** Let $\Phi = \Phi_p \cup \Phi_c$ be the agenda partioned into a set of premises Φ_p and a set of conclusions Φ_c, and let $\Phi_p = \Phi_1 \cup \Phi_2$ be partioned into sets Φ_1 and Φ_2 where Φ_2 contains the complements of all $\varphi \in \Phi_1$. The *uniform premise-based quota rule with quota q* (denoted by $UPQR_q$) for a quota $0 \leq q < 1$ maps each profile $\mathbf{J} = (J_1, \ldots, J_n) \in \mathfrak{J}(\Phi)^n$ to the collective outcome $UPQR_q(\mathbf{J})$ containing

- the premises $\varphi \in \Phi_1$ that are contained in more than $n \cdot q$ judgment sets $J \in \mathbf{J}$,
- the premises $\varphi \in \Phi_2$ that are contained in at least $n \cdot (1-q)$ judgment sets $J \in \mathbf{J}$,
- and all conclusions $\varphi \in \Phi_c$ that can be derived from the premises in the collective outcome.

In contrast to this approach, List (2004) uses a slightly modified majority rule in a sequential context where—following a given ordering of the issues in the agenda—the majority quota for a formula only comes into play when the acceptance or rejection of an issue cannot already be inferred by the already determined part of the collective outcome, thus ensuring a complete and consistent outcome. Following the concept of scoring rules in voting (see Section 2.2.1), Dietrich (2014) introduces scoring rules in judgment aggregation that include an equivalent of Borda Count. Other counterparts of voting rules include rules based on minimization studied by Lang et al. (2011). Furthermore, Lang and Slavkovik (2013) investigate how the different judgment aggregation procedures relate to the established voting rules.

The family of distance-based judgment aggregation procedures as introduced by Pigozzi (2006) and Miller and Osherson (2009) consists of procedures that try to minimize the distance between the collective outcome and the individual judgment sets. In a similar approach, Botan et al. (2021) design an egalitarian rule where the difference in satisfaction with the collective outcome between the judges is minimized.

Another direction to circumvent the discursive dilemma is to analyze the agenda. An agenda is called safe for a judgment aggregation procedure $\mathcal{P}$ if $\mathcal{P}$ produces a consistent outcome regardless of the input profile. Several axioms for the agenda such as the median property were introduced to characterize safe agendas. See the article by Endriss et al. (2012) for an overview and a multitude of axioms regarding the safety of agendas. However, Endriss et al. (2012) also show that for all considered axioms and procedures, it is intractable to check whether a given agenda is safe. Endriss et al. (2015) strengthen these results for the majority rule by showing intractability for agenda safety when parameterized by, among others, the maximum formula size or the maximum variable degree, but also show fixed-parameter tractability when parameterizing by the size of the agenda.

There are quite a few properties that judgment aggregation procedures can satisfy. See, e.g., the article by Lang et al. (2017) for a list of judgment aggregation procedures and their properties. Corresponding to Arrow's theorem in preference aggregation (see Section 2.2.1), Dietrich and List (2007a) show that an analogue to this theorem holds for the field of judgment aggregation stating that a judgment aggregation procedure can only fulfill certain (deemed reasonable) properties if it is a dictatorship. This theorem strengthens earlier impossibility results, for example by List and Pettit (2002, 2004). Dietrich and List (2010) further weaken the requirements posed on the judgment aggregation procedure to be dictatorial. Following these results, there are several papers that investigate whether the aforementioned "reasonable properties" are in fact reasonable for a judgment aggregation procedure. A judgment aggregation procedure $\mathcal{P}$ satisfies *neutrality* if for each agenda Φ, each profile $\mathbf{J}$, and each pair of formulas $\varphi, \psi \in \Phi$ where $\varphi \in J \Leftrightarrow \psi \in J$ for all $J \in \mathbf{J}$, it holds that $\varphi \in \mathcal{P}(\mathbf{J})$ if and only if $\psi \in \mathcal{P}(\mathbf{J})$. Intuitively, neutrality requires that any two issues have to be treated equally. Slavkovik (2014) and Terzopoulou and Endriss (2019a, 2020) argue that there are several cases where the axiom of neutrality is too strong, and the latter propose weaker versions of neutrality. List (2003) suggests to implement domain restrictions to circumvent impossibility theorems and introduces the *unidimensional alignment* domain that is the judgment aggregation equivalent to the single-crossing domain in voting.

As in voting, it is important to take the computational complexity of winner determination into account when choosing the best judgment aggregation procedure for the given application. Endriss et al. (2012) define the first problem for winner determination as follows.

$\mathcal{P}$-WINNER-DETERMINATION	
Given:	An agenda Φ, a profile $\mathbf{J} \in \mathcal{J}(\Phi)^n$, and a subset of formulas $S \subseteq \Phi$.
Question:	Is there a $J \in \mathcal{P}(\mathbf{J})$ so that $S \subseteq J$?

Note that using the winner determination problem defined above, a winning judgment set can be computed with a polynomial number of queries: Start with the empty set S. For each $\varphi \in \Phi$, ask whether $S \cup \{\varphi\}$ is part of a judgment set in $\mathcal{P}(\mathbf{J})$, and set $S := S \cup \{\varphi\}$ if the answer is yes. Endriss et al. (2012) then show that the winner determination for quota rules and the premise-based rule is tractable, while the winner determination for the distance-based procedure that returns the judgment sets from $\mathcal{J}(\Phi)$ minimizing the sum of Hamming distances to the profile is intractable. Here, the *Hamming distance* $HD(S,T)$ between two complete and consistent judgment sets S and T is defined as the number of issues $\varphi \in \Phi$ on which both sets differ. De Haan and Slavkovik (2017) expand the aforementioned intractability result by considering several more procedures from the family of the distance-based procedures. Further, they show that winner determination for scoring rules in judgment aggregation is also intractable. Lang and Slavkovik (2014) study the computational complexity of winner determination for several majority-preserving rules, and Endriss and de Haan (2015) show the intractability of winner determination for several analogues of voting rules for judgment aggregation, e.g., analogues of the Kemeny rule defined on page 11 and the Slater rule (see the article by Endriss et al. (2020) for an overview). De Haan (2016) further studies the parameterized complexity of winner determination for the Kemeny procedure and shows that the computational complexity does not coincide in the respective frameworks. For the family of sequential quota rules, Baumeister et al. (2021a) show that winner determination is intractable.

Based on the concept of iterative voting in preference aggregation (see Section 2.2.4 and Chapter 5), Terzopoulou and Endriss (2018) introduce a model for iterative judgment aggregation where judges can update their individual judgment sets. Further, they study the convergence to equilibria.

Analogous to Arrow's theorem, the Gibbard-Satterthwaite Theorem introduced in Section 2.2.2 for voting can also be transferred to judgment aggregation. To this end, Dietrich and List (2007c) introduce the following notion of non-manipulability.

Definition 2.24 (non-manipulability)**.** A resolute judgment aggregation procedure $\mathcal{P}$ is *non-manipulable* if for each agenda Φ, each set of judges $N = \{1,\ldots,n\}$, each profile $\mathbf{J} = (J_1,\ldots,J_n) \in \mathcal{J}(\Phi)^n$, each formula $\varphi \in \Phi$, and each judge $i \in N$, it holds that if $\varphi \in J_i$, but $\varphi \notin \mathcal{P}(\mathbf{J})$, then $\varphi \notin \mathcal{P}(\mathbf{J}')$ for each $\mathbf{J}' = (J_1,\ldots,J_{i-1},J_i',J_{i+1},\ldots,J_n)$ where $J_i' \in \mathcal{J}(\Phi)$.

They show that, given a certain agenda type, each resolute, non-imposing[12], non-manipulable judgment aggregation procedure that has no further restrictions on the input profile and always returns a complete and consistent collective outcome is a dictatorship of some judge i in the sense that $\mathcal{P}$ always returns the individual judgment set J_i. Furthermore, they define the following types of preferences over judgment sets to study whether a manipulator prefers a new outcome to the original collective outcome and therefore has an incentive to change her judgment set. Note that these preference types are only defined over complete and consistent judgment sets to simplify the presentation.

Definition 2.25 (preference types)**.** Let Φ be an agenda, and let $J \in \mathcal{J}(\Phi)$ be a judgment set.

1. The set U_J of *unrestricted preferences* contains all possible weak orders $\succsim$ over $\mathcal{J}(\Phi)$.
2. Define the set $TR_J \subseteq U_J$ of *top-respecting J-induced preferences* by
$$TR_J = \{\succsim \in U_J \mid J \succ Y \ \forall Y \in \mathcal{J}(\Phi) \setminus \{J\}\}.$$
3. Define the set $CR_J \subseteq TR_J$ of *closeness-respecting J-induced preferences* by
$$CR_J = \{\succsim \in U_J \mid X \succsim Y \ \forall\, X,Y \in \mathcal{J}(\Phi) \text{ where } X \cap J \supseteq Y \cap J\}.$$
4. Finally, define the Hamming-distance J-induced preference $HD_J \subseteq CR_J$ by
$$HD_J = \{\succsim \in U_J \mid (X \succsim Y \Leftrightarrow HD(X,J) \leq HD(Y,J)) \ \forall X,Y \in \mathcal{J}(\Phi)\}.$$

[12] A judgment aggregation procedure $\mathcal{P}$ is non-imposing if for each agenda Φ and each $\varphi \in \Phi$, there exist two profiles $\mathbf{J_1},\mathbf{J_2}$ in the domain of $\mathcal{P}$ so that $\varphi \in \mathcal{P}(\mathbf{J_1})$ and $\varphi \notin \mathcal{P}(\mathbf{J_2})$.

		b	c	$b \wedge c$
$\mathbf{J}$:	J_1	1	1	1
	J_2	1	0	0
	J_3	0	1	0
	$UPQR_{1/2}$	1	1	1

(a) The original profile $\mathbf{J} = (J_1, J_2, J_3)$.

		b	c	$b \wedge c$
$\mathbf{J}'$:	J_1	1	1	1
	J_2	1	0	0
	J_3'	0	0	0
	$UPQR_{1/2}$	1	0	0

(b) The manipulated profile $\mathbf{J}' = (J_1, J_2, J_3')$.

Table 2.10: The original and the manipulated profile in Example 2.27.

Define strategy-proofness as follows.

Definition 2.26 (strategy-proofness)**.** A resolute judgment aggregation procedure $\mathcal{P}$ is strategy-proof for a preference type $\mathcal{C}$ if for each agenda Φ, each profile $\mathbf{J} = (J_1, \dots, J_n) \in \mathcal{J}(\Phi)^n$, each judge i, each profile $\mathbf{J}' = (J_1, \dots, J_{i-1}, J_i', J_{i+1}, \dots, J_n) \in \mathcal{J}(\Phi)^n$, and each $\succsim_i \in \mathcal{C}_{J_i}$, it holds that $\mathcal{P}(\mathbf{J}) \succsim_i \mathcal{P}(\mathbf{J}')$.

Analogous to the impossibility theorem for non-manipulability, Dietrich and List (2007c) further prove that, given a certain agenda type, each resolute and non-imposing judgment aggregation procedure that has no further restrictions on the input profile, always returns a complete and consistent collective outcome, and is strategy-proof for preference type $\mathcal{C}$ is a dictatorship of some judge i. However, as Terzopoulou and Endriss (2019b) show, this impossibility relies on the full information model, i.e., the model where a potential manipulator knows the complete profile of judgment sets. When the judgment sets of some judges are unknown, most impossibility results concerning strategy-proofness do not hold anymore.

Example 2.27 (non-manipulability and strategy-proofness)**.** Consider the uniform premise-based quota rule with quota $q = 1/2$ and the agenda and profile from Example 2.22, i.e., $\Phi = \{b, \neg b, c, \neg c, b \wedge c, \neg(b \wedge c)\}$ and the profile $\mathbf{J} \in \mathcal{J}(\Phi)^3$ in Table 2.10a. Define $\Phi_p = \{b, \neg b, c, \neg c\}$ as the premises with $\Phi_1 = \{b, c\}$ and define $\Phi_c = \{b \wedge c, \neg(b \wedge c)\}$ as the conclusions of the agenda. A respective majority accepts b and c, so that $b \wedge c$ is also part of the collective outcome ($UPQR_{1/2}(\mathbf{J}) = \{b, c, b \wedge c\}$). In particular, judge 3 disagrees with the collective outcome on issue $b \wedge c$. However, if she changes her judgment to $J_3' = \{\neg b, \neg c, \neg(b \wedge c)\}$ resulting in profile $\mathbf{J}'$ as seen in Table 2.10b, then the collective outcome $UPQR_{1/2}(\mathbf{J}') = \{b, \neg c, \neg(b \wedge c)\}$ agrees with J_3 on issue $b \wedge c$. Therefore, the uniform premise-based quota rule with quota $q = 1/2$ is manipulable.

Now consider the same scenario and assume that judge 3 has closeness-respecting (J_3-induced) preferences. There are four complete and consistent judgment sets for Φ, namely J_1, J_2, J_3, and J_3'. According to Definition 2.25, she prefers J_3 to all other sets. Furthermore, she prefers J_3' to J_2 since $J_3' \cap J_3 = \{b, \neg(b \wedge c)\} \supseteq \{\neg(b \wedge c)\} = J_2 \cap J_3$. The relation between J_1 and both J_2 and J_3' is not fixed. Therefore,

$$\begin{aligned} CR_{J_3} = \{ & J_3 \succ J_1 \succ J_3' \succ J_2, \\ & J_3 \succ J_3' \succ J_1 \succ J_2, \\ & J_3 \succ J_3' \succ J_2 \succ J_1 \} \end{aligned}$$

Since there is an $\succsim \in CR_{J_3}$ so that $UPQR_{1/2}(\mathbf{J}') \succ UPQR_{1/2}(\mathbf{J})$, it follows that the uniform premise-based quota rule with quota $q = 1/2$ is not strategyproof under closeness-respecting preferences.

Again, as with preference aggregation, there exist a multitude of papers investigating the complexity of strategic attacks in judgment aggregation. Manipulation includes manipulation by a single judge as defined above as well as manipulation by a coalition of judges. Endriss et al. (2012) study the complexity of manipulation for the premise-based procedure where the manipulator has Hamming-distance-induced preferences, whereas Baumeister et al. (2015b) consider the (uniform) premise based (quota) rules for the preference types in Definition 2.25 and for the case where the manipulator wants to include a set D as a subset of the manipulated collective outcome. Coalitional manipulation was first studied by Botan et al. (2016). The concept of bribery is closely related to lobbying (see, e.g., the papers by Christian et al. (2007) and Bredereck et al. (2014b) for lobbying in judgment aggregation) and was introduced to judgment aggregation by Baumeister et al. (2015b). Baumeister et al. (2015b,d) study the complexity of bribery for the (uniform) premise based (quota) rules and the preference types defined above. De Haan (2017) studies manipulation and bribery for the Kemeny procedure. Similar to control in preference aggregation, the concept of control in judgment aggregation can include adding or deleting judges (Baumeister et al., 2012, 2015d), bundling issues (Alon et al., 2013) and bundling judges (Baumeister et al., 2013), adding or deleting issues (Dietrich, 2016), or changing the order of issues in sequential rules (Bredereck et al., 2017). See the book chapter by Baumeister et al. (2017) for an overview of strategic behavior in judgment aggregation and see Chapter 6 for my contribution to the investigation of the complexity of strategic attacks in judgment aggregation.

De Haan (2018) argues that the intractability results for judgment aggregation are due to the used framework being overly expressive and proposes to use more limited languages that yield tractability results and are still able to model certain applications. De Haan and Slavkovik (2019) give an encoding for several procedures and problems in judgment aggregation into answer set programming.

Chapter 3

Minisum and Minimax Committee Election Rules for General Preference Types

This chapter deals with new types of ballots and new corresponding types of multiwinner voting rules, also called committee election rules. The corresponding publication is as follows.

Baumeister, D., Böhnlein, T., Rey, L., Schaudt, O., and Selker, A.-K. (2016). Minisum and minimax committee election rules for general preference types. In *Proceedings of the 22nd European Conference on Artificial Intelligence*, pages 1656–1657. IOS Press. Extended Abstract

Brams et al. (2007) introduce the minimax procedure for electing a committee. The goal is to select a committee that minimizes the maximum Hamming distance to a voter in the given approval-based profile. They also define a corresponding minisum procedure that aims to minimize the sum of Hamming distances to the voters' ballots. Baumeister and Dennisen (2015) modify these procedures to allow trichotomous ballots, complete linear orders, and incomplete linear orders. Alcantud and Laruelle (2014) characterize a new voting rule based on trichotomous ballots.

This chapter further extends these articles by introducing a type of ballot called ℓ-ballot, applying these ballots to minisum and minimax rules, and studying correspondingly modified axiomatic properties. In an ℓ-ballot, voters partition the candidates into ℓ (possibly empty) groups and then rank these groups. This can be seen as a compromise between dichotomous and trichotomous ballots on the one hand where voters lose the ability to express more fine-grained relations between the candidates, and linear orders on the other hand where voters are forced to express a strict preference between each pair of the (possibly large) candidate set. Note that a similar type of ballot was previously introduced by Obraztsova et al. (2015a, 2017). However, they study the ballots in a game-theoretic approach, whereas this chapter deals with the axiomatic properties of minisum and minimax voting rules using these kind of ballots. Furthermore, Balinski and Laraki (2011) obtain

ℓ-ballots in an intermediate stage of their majority judgment procedure where each judge (who correspond to voters in this context) first assigns an integer grade in a fixed interval to each contestant (or candidate) via a grade function, and the procedure then returns the contestant with the best median grade of all judges.

This chapter is organized as follows. Section 3.1 introduces the concept of ℓ-ballots and the corresponding multiwinner rules called ℓ-group rules, and illustrates these concepts in a detailed example. Section 3.2 then modifies axiomatic properties for rules using ordinal ballots to ones for rules using ℓ-ballots, and shows which of these properties are fulfilled by the ℓ-group rules. Next, Section 3.3 studies the complexity of winner determination for the new rules. Section 3.4 generalizes ℓ-ballots to cardinal ballots that allow for two independent values a and b for the dissatisfaction with a candidate being a member or not a member of a committee. Finally, Section 3.5 details my contribution to the findings of this chapter.

3.1 Minisum and Minimax ℓ-group rules

This section introduces a new type of ballot and corresponding multiwinner rules. For a more detailed overview on ballot types see Section 2.2, and for the basics of multiwinner elections see Section 2.2.3. Recall that $N = \{1,\dots,n\}$ is a set of voters, $C = \{c_1,\dots,c_m\}$ is a set of candidates, $k \in \mathbb{N}$ is a committee size, and $\mathcal{F}$ is a multiwinner voting rule.

Definition 3.1 (ℓ-ballots)**.** An *ℓ-ballot* v for an integer $\ell \geq 2$ over candidate set C is a partition of C into ℓ pairwise disjoint, possibly empty sets (called *groups*). Formally, $v = (G_1,\dots,G_\ell)$ so that $G_i \cap G_j = \emptyset$ for $1 \leq i < j \leq \ell$ and $\bigcup_{i=1}^{\ell} G_i = C$.

Note that 2-ballots correspond to approval ballots, whereas n-ballots do not correspond to linear orders or weak orders since groups might be empty. For each group G_j, j is called the *group number* of G_j. Intuitively, a voter prefers candidates with a low group number to ones with a higher one, and is indifferent between candidates with the same group number. Furthermore, let the candidates in C have a fixed ordering $c_1,\dots,c_m$ and let $v(k)$ denote the group number of a candidate c_k in ballot v. Here, $P_\ell = (v_1,\dots,v_n)$ denotes the profile of ℓ-ballots for a fixed $\ell \in \mathbb{N}$ where v_i is the ballot submitted by voter $i \in N$. A *committee election* is a tuple $\mathcal{E} = (C, P_\ell, k)$.

Definition 3.2 (Confirmed and potential committee members)**.** Let $\mathcal{E} = (C, P_\ell, k)$ be a committee election and let $\mathcal{F}$ be a multiwinner rule. A candidate $c \in C$ is a *confirmed committee member for* $\mathcal{E}$ (under $\mathcal{F}$) if $c \in W$ for all $W \in \mathcal{F}(P_\ell, k)$, and a *potential committee member for* $\mathcal{E}$ (under $\mathcal{F}$) if there exists a $W \in \mathcal{F}(P_\ell, k)$ so that $c \in W$.

In this chapter, the focus is on minimizing the dissatisfaction voters have with an elected committee W. Let $F_k(C) = \{S \mid S \subseteq C \wedge |S| = k\}$ be the set of all committees of size k over C. The dissatisfaction of a voter j—associated with ℓ-ballot v_j—with a committee $W \in F_k(C)$ is measured as

$$\delta_\ell(v_j, W) = \sum_{i=1}^{m} |v_j(i) - W(i)|,$$

where $W(i) = 1$ if $c_i \in W$, and $W(i) = \ell$ else. Note that $\delta_\ell(v_j, W)$ can be interpreted as the distance between ℓ-ballot v_j and committee W and generalizes the Hamming distance between two vectors.

Define the following two families of multiwinner rules tailored to ℓ-ballots and minimizing voters' dissatisfaction.

Definition 3.3 (minisum and minimax ℓ-group rules)**.** For each $\ell > 0$, let $\mathcal{E} = (C, P_\ell, k)$ be a committee election and let $F_k(C) = \{S \mid S \subseteq C \wedge |S| = k\}$.

- The *minisum ℓ-group rules* are functions f^ℓ_{sum} that return the committees minimizing the total dissatisfaction of the voters with the elected committees, i.e.,

$$f^\ell_{sum}(P_\ell, k) = \operatorname*{argmin}_{W \in F_k(C)} \sum_{v \in P_\ell} \delta_\ell(v, W).$$

- The *minimax ℓ-group rules* are functions f^ℓ_{max} that return the committees minimizing the dissatisfaction of the respective least satisfied voter with the elected committees, i.e.,

$$f^\ell_{max}(P_\ell, k) = \operatorname*{argmin}_{W \in F_k(C)} \max_{v \in P_\ell} \delta_\ell(v, W).$$

Define $\sum_{v \in P_\ell} v(i)$ as the *minisum score* of a candidate c_i for profile P_ℓ.

Claim 3.4. *For each committee election $\mathcal{E} = (C, P_\ell, k)$, an alternative way to compute the result of $f^\ell_{sum}(P_\ell, k)$ is to calculate the minisum score of each candidate for P_ℓ and then return the committees $W \in F_k(C)$ that contain all candidates with a minisum score lower than s, as well as contain only candidates with a minisum score lower or equal than s, where s denotes the k-lowest minisum score of a candidate.*

Proof. Assume that the claim is not true, i.e., there exists a committee election $\mathcal{E} = (C, P_\ell, k)$ so that either (1) for a winning committee W, there exists a candidate $c \notin W$ that has a lower minisum score than a member of W, or (2) a committee satisfying the requirements in the claim does not win.

Case (1): Let $W \in f^\ell_{sum}(P_\ell, k)$ be a winning committee, so that there exists a $c_i \in W$ and a $c_j \in C \setminus W$ so that $\sum_{v \in P_\ell} v(j) < \sum_{v \in P_\ell} v(i)$. Consider the committee $W' = (W \cup \{c_j\}) \setminus \{c_i\}$.

$$\begin{aligned}
&\sum_{v \in P_\ell} \delta_\ell(v, W) - \sum_{v \in P_\ell} \delta_\ell(v, W') \\
=&\sum_{v \in P_\ell} \big(v(i) - 1\big) + \sum_{v \in P_\ell} \big(\ell - v(j)\big) - \sum_{v \in P_\ell} \big(v(j) - 1\big) - \sum_{v \in P_\ell} \big(\ell - v(i)\big) \\
=&2 \cdot \sum_{v \in P_\ell} v(i) - 2 \cdot \sum_{v \in P_\ell} v(j) > 0 \,,
\end{aligned}$$

i.e., the total dissatisfaction of the voters with W is greater than with W', a contradiction to the fact that W is a winning committee.

Case (2): It follows from case (1) that all winning committees satisfy the requirements in the claim statement. Assume a committee $W \notin f^\ell_{sum}(P_\ell, k)$ satisfying the requirements does not win, whereas a committee $W' \in f^\ell_{sum}(P_\ell, k)$ does. It holds that

$$\sum_{v \in P_\ell} \Big(\sum_{\substack{c_i \in W \\ \wedge c_i \notin W'}} v(i) + \sum_{\substack{c_i \notin W \\ \wedge c_i \in W'}} v(i) \Big) = 2 \cdot |\{c \mid c \in W \wedge c \notin W'\}| \cdot s = \sum_{v \in P_\ell} \Big(\sum_{\substack{c_i \in W' \\ \wedge c_i \notin W}} v(i) + \sum_{\substack{c_i \notin W' \\ \wedge c_i \in W}} v(i) \Big),$$

and therefore

$$
\begin{aligned}
&\sum_{v \in P_\ell} \delta_\ell(v,W) \\
&= \sum_{v \in P_\ell} \Big(\sum_{\substack{c_i \in W \\ \wedge c_i \in W'}} \big(v(i)-1\big) + \sum_{\substack{c_i \in W \\ \wedge c_i \notin W'}} \big(v(i)-1\big) + \sum_{\substack{c_i \notin W \\ \wedge c_i \in W'}} \big(\ell - v(i)\big) + \sum_{\substack{c_i \notin W \\ \wedge c_i \notin W'}} \big(\ell - v(i)\big) \Big) \\
&= \sum_{v \in P_\ell} \Big(\sum_{\substack{c_i \in W \\ \wedge c_i \in W'}} \big(v(i)-1\big) + \sum_{\substack{c_i \in W' \\ \wedge c_i \notin W}} \big(v(i)-1\big) + \sum_{\substack{c_i \notin W' \\ \wedge c_i \in W}} \big(\ell - v(i)\big) + \sum_{\substack{c_i \notin W \\ \wedge c_i \notin W'}} \big(\ell - v(i)\big) \Big) \\
&= \sum_{v \in P_\ell} \delta_\ell(v,W')
\end{aligned}
$$

This is a contradiction to the fact that W' is a winning committee and W is not.

Since both cases end in a contradiction, the claim is true. □

The following example illustrates the minisum and minimax ℓ-group rules.

Example 3.5 (minisum and minimax ℓ-group rules)**.** Let $\mathcal{E} = (C, P_3, 3)$ be a committee election where $C = \{a,b,c,d\}$ and P_3 is the profile of 3-ballots over C in Table 3.1.

		G_1	G_2	G_3	
P_3:	$v_1 = ($	$\{a\}$	$\{b,c\}$	$\{d\}$	$)$
	$v_2 = ($	$\{a,b,d\}$	$\{\}$	$\{c\}$	$)$
	$v_3 = ($	$\{c\}$	$\{a,b,d\}$	$\{\}$	$)$

Table 3.1: The profile P_3 in Example 3.5.

Start with the minisum 3-group rule f^3_{sum}. The minisum scores of the candidates are the total of the respective group numbers, see Table 3.2a. For example, candidate a is in G_1 for both voter 1 and 2, and in G_2 for voter 3, adding up to a score of $1+1+2=4$. For $k=3$, the third lowest score is 6. According to Claim 3.4, each winning committee has to contain the candidates with a score lower than 6, namely a and b, and can only contain candidates with a minisum score of at most 6. It follows that $f^3_{sum}(P_3,3) = \{\{a,b,c\},\{a,b,d\}\}$.

Next, turn to the minimax 3-group rule f^3_{max}. Recall that the dissatisfaction of a voter j with a given committee W is the total of $v_j(i)-1$ for each $c_i \in W$ and $\ell - v_j(i)$ for each

	a	b	c	d
v_1 :	1	2	2	3
v_2 :	1	1	3	1
v_3 :	2	2	1	2
$\sum$:	4	5	6	6

(a) Minisum scores

	$\{a,b,c\}$	$\{a,b,d\}$	$\{a,c,d\}$	$\{b,c,d\}$
v_1 :	2	4	4	6
v_2 :	4	0	4	4
v_3 :	3	5	3	3
max :	4	5	4	6

(b) Voters' dissatisfaction with each committee of size 3.

Table 3.2: The minisum scores of each candidate and the dissatisfaction of each voter for the minimax 3-group in Example 3.5.

$c_i \notin W$. See Table 3.2b for the dissatisfaction of each voter for each possible committee of size 3.

For example, voter 1 has a dissatisfaction with committee $\{b,c,d\}$ of $(3-1)+(2-1)+(2-1)+(3-1)=6$. The winning committees are then the committees with the lowest maximum of dissatisfaction. It follows that $f^3_{max}(P_3,3) = \big\{\{a,b,c\},\{a,c,d\}\big\}$ with a maximum dissatisfaction of 4.

3.2 Axiomatic Properties

As pointed out in Sections 2.2.1 and 2.2.3, the choice of using a specific (multiwinner) voting rule depends heavily on the axiomatic properties the rules satisfy. This section studies whether the minisum and minimax ℓ-group rules satisfy some well-known properties for singlewinner and/or multiwinner voting rules. Note that the considered properties were originally defined for linear or weak orders, so some properties have to be specifically adapted to ℓ-ballots. In this section, all proofs that I did not contribute are omitted.

The first (fundamental) property is non-imposition, that–similar to the singlewinner version defined on page 12–demands that each committee can win in some election (which also implies that no candidates are incompatible with each other and therefore cannot be part of the same committee). The property homogeneity asks that multiplying the given profile does not change the election result. The multiwinner adaptations of both properties are due to Elkind et al. (2017a).

Definition 3.6. A multiwinner voting rule $\mathcal{F}$ satisfies

- *non-imposition*, if for each set of candidates C and and each committee $W \in F_k(C)$ of size k there is a profile P so that $\mathcal{F}(P,k) = \{W\}$.
- *homogeneity*, if for each committee election $\mathcal{E} = (C,P,k)$ and each $t \in \mathbb{N}$ it holds that $\mathcal{F}(tP,k) = \mathcal{F}(P,k)$, where tP denotes the concatenation of t copies of P.

The families of minisum and minimax ℓ group rules satisfy both properties.

Theorem 3.7. *For each $\ell \in \mathbb{N}$, the minisum and minimax ℓ-group rules satisfy non-imposition and homogeneity.*

The property consistency (also called reinforcement) states that when a profile can be split in two so that the two parts agree on some winning committees, these committees should also win the election for the original profile. Again, this property was adapted from the singlewinner context to multiwinner rules by Elkind et al. (2017a).

Definition 3.8. Let $P_1 + P_2$ denote the concatenation of profiles P_1 and P_2. A multiwinner voting rule $\mathcal{F}$ satisfies *consistency*, if for each pair of profiles P_1, P_2 over the same candidate set C and for each committee size $k \leq |C|$, it holds that $\mathcal{F}(P_1 + P_2, k) = \mathcal{F}(P_1, k) \cap \mathcal{F}(P_2, k)$ whenever $\mathcal{F}(P_1, k) \cap \mathcal{F}(P_2, k) \neq \emptyset$.

Theorem 3.9. *For each $\ell \in \mathbb{N}$, the minisum ℓ-group rule satisfies consistency, while the minimax ℓ-group rule does not.*

A *clone* of a candidate c is another candidate c' that is very similar to c in the voter's eyes and will therefore be ranked close to the original candidate in the ballots. For example, a clone of a job applicant might be another applicant with the same qualifications and age. Ideally, adding a clone of candidate c to the election should not be detrimental to c, but for many contexts, the cloning might lead to a split vote. Consider for example plurality ballots, i.e., approval ballots where exactly one candidate can be approved. Some of the voters approving c might switch to approve the clone c' and thereby costing c votes to the benefit of another candidate. The idea of cloning was first introduced by Tideman (1987) who considers weak orders as ballots where clones of a candidate c are tied to or ranked directly above or below c. In the setting of ℓ-ballots, cloning a candidate c means adding an additional candidate c' to the set of candidates and placing c' in the same group as c for each ballot.

Definition 3.10. A multiwinner voting rule $\mathcal{F}$ is *independent of clones* if for each committee election $\mathcal{E} = (C, P, k)$ and each candidate $c \in C$, a candidate $d \in C$ that is not a potential committee member under $\mathcal{E}$ cannot be a potential committee member in a committee election when cloning c.

Note that the candidates c and d in the above definition do not have to be distinct, i.e., the cloning of candidate c should also not be beneficial to c.

Theorem 3.11. *For each ℓ, the minisum ℓ-group rule satisfies independence of clones, while the minimax ℓ-group rule does not.*

Next, consider variants of monotonicity starting with *committee monotonicity*. This property demands that members of winning committees should remain in winning committees even if the committee size is increased. For example, it is not reasonable to replace candidates on an interview shortlist just because an additional interview slot opened up. The following definition of *committee monotonicity* is due to Elkind et al. (2017a).

Definition 3.12. A multiwinner voting rule $\mathcal{F}$ satisfies *committee monotonicity* if for each committee election $\mathcal{E} = (C, P, k)$, and for each elected committee $W \in \mathcal{F}(P, k)$, the following holds:

- If $k < |C|$ then there exists a $W' \in \mathcal{F}(P, k+1)$ such that $W \subseteq W'$, and
- if $k > 1$ there exists a $W' \in \mathcal{F}(P, k-1)$ such that $W \supseteq W'$.

Theorem 3.13. *For each $\ell \in \mathbb{N}$, the minisum ℓ-group rule satisfies committee monotonicity, whereas the minimax ℓ-group rule does not for each $\ell \geq 2$.*

Proof. For an arbitrary positive integer ℓ, let $\mathcal{E} = (C, P_\ell, k)$ be a committee election. The minisum scores $\sum_{v \in P_\ell} v(j)$ of each candidate $c_j \in C$ are independent of the committee size k and therefore do not change by increasing or decreasing k. Since each minisum ℓ-group rule picks the committees that contain the candidates with the lowest minisum score, it satisfies committee monotonicity.

However, the minimax ℓ-group rule does not satisfy committee monotonicity for each $\ell \geq 2$. To prove this, consider $C = \{c_1, c_2, c_3\}$ and the following profile P_ℓ:

P_ℓ:	$v_1 =$	$(\{c_1\},$	$\{\},$	$\dots,$	$\{\},$	$\{c_2, c_3\})$
	$v_2 =$	$(\{c_2\},$	$\{\},$	$\dots,$	$\{\},$	$\{c_1, c_3\})$

It holds that $f^\ell_{max}(P_\ell, 1) = \big\{\{c_1\}, \{c_2\}, \{c_3\}\big\}$, whereas $f^\ell_{max}(P_\ell, 2) = \big\{\{c_1, c_2\}\big\}$. This violates the first condition of committee monotonicity since for $k = 1$, the winning committee $\{c_3\}$ is not a subset of any winning committee for $k = 2$. □

The following notions of monotonicity aim at ensuring that additional support for a candidate c does not hurt the prospect of winning for c. In the example of the shortlist for a hiring process, a candidate should not lose an already secured interview spot just because an additional member of the hiring committee announces support for this candidate. The original multiwinner definition for *candidate monotonicity* and *monotonicity* for linear orders is again due to Elkind et al. (2017a) and states that an improvement of a potential committee member c should not be detrimental to c (candidate monotonicity) respectively to the winning committee of which c is a member (monotonicity). In the original (singlewinner) definition of *positive responsiveness* for linear orders that goes back to May (1952), an improvement of a winning candidate c in a ballot should even lead to c being a unique winner of the election (see Definition 3.14 for a restatement for multiwinner voting rules). However, their notion of *improvement*[1] of a candidate is tailored to linear orders and is therefore not suitable for ℓ-ballots. Definition 3.14 below therefore uses the following notion of improvement.

Given a profile $P_\ell = (v_1, \dots, v_n)$ of ℓ-ballots over candidate set C, let P'_ℓ denote the modified profile obtained by improving candidate c_i in some ballot v_j, i.e., shifting c_i into a better group in v_j while the rest of the ballot remains unchanged. Formally,

$$P'_\ell = (v_1, \dots, v_{j-1}, v'_j, v_{j+1}, \dots, v_n) \text{ where } v'_j(i) < v_j(i) \text{ and } v'_j(g) = v_j(g) \text{ for all } c_g \neq c_i.$$

Note that in contrast to the case of linear orders, no candidates are shifted backwards by improving c_j. This is to preserve the spirit of the original notion where the relation between candidates other than c_j remain unchanged. For example, let $C = \{c_1, c_2, c_3, c_4\}$ be the set of candidates and consider the 2-ballots $v = (\{c_3\}, \{c_1, c_2\})$ and $v' = (\{c_1\}, \{c_2, c_3\})$. The

[1] Improving a candidate c by one position in the linear order of voter i means swapping c and the candidate in position $pos_i(c) - 1$. An improvement of $t > 1$ steps can be carried out by improving c by one position t times.

ballot v' can be obtained by swapping the candidates c_1 and c_3, but this has consequences for the relation between c_2 and c_3.

Furthermore, it does not suffice to just require that all relative candidate relations barring the ones with c_i remain the same, but it is necessary to fix the group numbers of all candidates but c_i. In the 6-ballots

$$\begin{array}{rllllll} v = & (\{\}, & \{c_4\}, & \{c_2\}, & \{c_1\}, & \{c_3\}, & \{\}) \quad \text{and} \\ v' = & (\{c_4\}, & \{c_1, c_2\}, & \{\}, & \{\}, & \{\}, & \{c_3\}), \end{array}$$

the candidate c_1 is improved in v' in relation to v, but c_2 and c_4 improved as well which is counterintuitive to an improvement of c_1. Additionally, the candidate c_3 is even shifted a group back without changing the relation to any other candidate including c_1, which is detrimental to c_3 and again counterintuitive. For a recent adaptation of the notion of monotonicity to approval ballots (respectively, to ballots using weak orders) see the article by Sánchez-Fernández and Fisteus (2019) (respectively, Aziz and Lee (2020)).

Definition 3.14. A multiwinner voting rule $\mathcal{F}$ satisfies

- *candidate monotonicity (resp., positive responsiveness)*, if for each committee election $\mathcal{E} = (C, P, k)$ and each $c \in C$, if c is a potential committee member for $\mathcal{E}$ under $\mathcal{F}$, then it holds that c is a potential (respectively, confirmed) committee member for (C, P', k), where P' is obtained from P by improving c in some ballot v.

- *monotonicity*, if for each committee election $\mathcal{E} = (C, P, k)$, each $W \in \mathcal{F}(P, k)$, and each $c \in W$, it holds that $W \in \mathcal{F}(P', k)$ for all P' that are obtained from P by improving c in some ballot v.

Note that monotonicity as well as positive responsiveness imply candidate monotonicity.

The following theorem shows that the minisum ℓ-group rules satisfy the notions of monotonicity defined above.

Theorem 3.15. *For each $\ell \in \mathbb{N}$, the minisum ℓ-group rule satisfies monotonicity, candidate monotonicity, and positive responsiveness.*

Proof. Start with the the property of monotonicity. Let $\ell \geq 2$ be an arbitrary integer and assume that the minisum ℓ-group rule does not satisfy monotonicity, i.e., there exists a committee election $\mathcal{E} = (C, P_\ell, k)$, a committee $W \in f^{\ell}_{sum}(P_\ell, k)$, and a candidate $c_i \in W$ so that $W \notin f^{\ell}_{sum}(P'_\ell, k)$, where P'_ℓ is obtained from P_ℓ by improving c_i in ballot v_j. In particular, that means that in the new committee election, there exists a committee W' that has a lower sum of dissatisfaction of the voters with W' than with the original winning committee W. Formally,

$$\sum_{v \in P_\ell} \delta_\ell(v, W') \geq \sum_{v \in P_\ell} \delta_\ell(v, W) \text{ and} \tag{3.1}$$

$$\sum_{v \in P'_\ell} \delta_\ell(v, W') < \sum_{v \in P'_\ell} \delta_\ell(v, W). \tag{3.2}$$

where Equation (3.1) holds due to $W \in f^{\ell}_{sum}(P_\ell, k)$ and Equation (3.2) due to $W \notin f^{\ell}_{sum}(P'_\ell, k)$. Recall that $c_i \in W$ and that only the group number of c_i differs in v_j and v'_j while the placement of all other candidates remains unchanged, and all other ballots in P_ℓ are identical in P'_ℓ. Consequently, for each committee S it holds that

$$\delta_\ell(v'_j, S) - \delta_\ell(v_j, S) = v'_j(i) - v_j(i). \tag{3.3}$$

Equation (3.2) can be transformed as follows.

$$\begin{aligned}
& \sum_{v \in P'_\ell} \delta_\ell(v, W') < \sum_{v \in P'_\ell} \delta_\ell(v, W) \\
\iff\ & \Big(\sum_{v \in P'_\ell \setminus \{v'_j\}} \delta_\ell(v, W')\Big) + \delta_\ell(v'_j, W') < \Big(\sum_{v \in P'_\ell \setminus \{v'_j\}} \delta_\ell(v, W)\Big) + \delta_\ell(v'_j, W) \\
\iff\ & \Big(\sum_{v \in P_\ell \setminus \{v_j\}} \delta_\ell(v, W')\Big) + \delta_\ell(v'_j, W') < \Big(\sum_{v \in P_\ell \setminus \{v_j\}} \delta_\ell(v, W)\Big) + \delta_\ell(v'_j, W) \\
\iff\ & \Big(\sum_{v \in P_\ell} \delta_\ell(v, W')\Big) + \delta_\ell(v'_j, W') - \delta_\ell(v_j, W') < \Big(\sum_{v \in P_\ell} \delta_\ell(v, W)\Big) + \delta_\ell(v'_j, W) - \delta_\ell(v_j, W) \\
\iff\ & \sum_{v \in P_\ell} \delta_\ell(v, W') - \sum_{v \in P_\ell} \delta_\ell(v, W) < \delta_\ell(v_j, W') - \delta_\ell(v'_j, W') + \underbrace{v'_j(i) - v_j(i)}_{\text{due to Eq. (3.3)}} \\
\iff\ & \underbrace{\sum_{v \in P_\ell} \delta_\ell(v, W') - \sum_{v \in P_\ell} \delta_\ell(v, W)}_{\geq 0 \text{ due to Eq. (3.1)}} < \underbrace{v_j(i) - v'_j(i)}_{\text{due to Eq. (3.3)}} + v'_j(i) - v_j(i) = 0 \ \lightning
\end{aligned}$$

Due to the contradiction, it follows that the minisum ℓ-group rule satisfies monotonicity for each $\ell > 0$, which also implies that the minisum ℓ-group rule satisfies candidate monotonicity for each $\ell > 0$.

Next, assume that the minisum ℓ-group rule does not satisfy positive responsiveness, i.e., there exists a committee election $\mathcal{E} = (C, P_\ell, k)$, a potential committee member $c \in W$ for some $W \in f^\ell_{sum}(P_\ell, k)$, and a committee $W' \in f^\ell_{sum}(P'_\ell, k)$ so that $c \notin W'$, where P'_ℓ is obtained from P_ℓ by improving c in ballot v_j. Then $W \in f^\ell_{sum}(P'_\ell, k)$ since f^ℓ_{sum} satisfies monotonicity as proved above. The inequality

$$\sum_{v \in P_\ell} \delta_\ell(v, W') \overset{(1)}{<} \sum_{v \in P'_\ell} \delta_\ell(v, W') \overset{(2)}{=} \sum_{v \in P'_\ell} \delta_\ell(v, W) \overset{(3)}{<} \sum_{v \in P_\ell} \delta_\ell(v, W)$$

is a contradiction to the fact that W is a winning committee in the original election. Note that (1) is due to the fact that c is not a member of W' and an improvement of c therefore leads to a higher dissatisfaction of voter j with W'. Furthermore, (2) holds since both W and W' are winning committees in the new election, and (3) holds because an improvement of c leads to a lower dissatisfaction of voter j with W. Therefore, the minisum ℓ-group rule satisfies positive responsiveness for all $\ell > 0$. □

Next, turn to the minimax ℓ-group rules.

Theorem 3.16. *For each $\ell \in \mathbb{N}$, the minimax ℓ-group rule satisfies candidate monotonicity. However, for each $\ell \geq 2$, the minimax ℓ-group rule does not satisfy positive responsiveness and monotonicity.*

Proof. First, show that for each $\ell > 0$, the minimax ℓ-group rule satisfies candidate monotonicity. For each committee election $\mathcal{E} = (C, P_\ell, k)$, and each profile P'_ℓ that is obtained by improving a potential committee member $c \in W$, $W \in f^\ell_{max}(P_\ell, k)$, in a ballot in P_ℓ, the following holds:

The dissatisfaction of a least satisfied voter does not change or even grows for all committees $W' \in F_k(C)$ where $c \notin W'$, so that the following holds.

$$\max_{v \in P'_\ell} \delta_\ell(v, W) \leq \max_{v \in P_\ell} \delta_\ell(v, W) \leq \max_{v \in P_\ell} \delta_\ell(v, W') \leq \max_{v \in P'_\ell} \delta_\ell(v, W')$$

$P_\ell :$	$v_1 =$	$(\{c_1\},$	$\{c_2,c_3\},$	$\{\},$	$\dots)$
	$v_2 =$	$(\{c_3\},$	$\{c_1,c_2\},$	$\{\},$	$\dots)$

$P'_\ell :$	$v'_1 =$	$(\{c_1,c_2\},$	$\{c_3\},$	$\{\},$	$\dots)$
	$v_2 =$	$(\{c_3\},$	$\{c_1,c_2\},$	$\{\},$	$\dots)$

Table 3.3: Profiles P_ℓ and P'_ℓ in the positive responsiveness proof of Theorem 3.16.

$P_\ell :$	$v_1 =$	$(\{c_1\},$	$\{c_2,c_3\},$	$\{\},$	$\dots)$
	$v_2 =$	$(\{c_2\},$	$\{c_1,c_3\},$	$\{\},$	$\dots)$
	$v_3 =$	$(\{c_3\},$	$\{c_1,c_2\},$	$\{\},$	$\dots)$

$P'_\ell :$	$v'_1 =$	$(\{c_1,c_2\},$	$\{c_3\},$	$\{\},$	$\dots)$
	$v_2 =$	$(\{c_2\},$	$\{c_1,c_3\},$	$\{\},$	$\dots)$
	$v_3 =$	$(\{c_3\},$	$\{c_1,c_2\},$	$\{\},$	$\dots)$

Table 3.4: Profiles P_ℓ and P'_ℓ in the monotonicity proof of Theorem 3.16.

Since $c \in W$, this means that there does not exist a committee W' not containing c that has a strictly lower dissatisfaction of the least satisfied voter in the new election than all committees containing c. It follows that c has to be a potential committee member in the new election.

However, such a candidate c does not have to be a confirmed committee member after an improvement, which the following example shows. Consider $C = \{c_1, c_2, c_3\}$ and the profiles P_ℓ and P'_ℓ in Table 3.3, where the latter is obtained by improving c_2 in the ballot v_1. It holds that $f^\ell_{max}(P_\ell, 1) = f^\ell_{max}(P'_\ell, 1) = \big\{\{c_1\},\{c_2\},\{c_3\}\big\}$, so c_2 is not a confirmed committee member in the new election despite being a potential committee member in the original election and being improved. This proves that the minimax ℓ-group rule does not satisfy positive responsiveness for each $\ell \geq 2$ because the improvement of a candidate c in a vote that is not the single vote with the minimal maximal dissatisfaction for any committee has no effect on the election result.

Furthermore, the minimax ℓ-group rule does not satisfy monotonicity for each $\ell \geq 2$. Consider for example $C = \{c_1, c_2, c_3\}$ and the profiles P_ℓ and P'_ℓ in Table 3.4, where the latter is obtained by improving c_2 in the ballot v_1. It holds that

$$f^{\ell}_{max}(P_\ell, 2) = \big\{\{c_1, c_2\}, \{c_2, c_3\}, \{c_1, c_3\}\big\}.$$

But the improvement of c_2 in ballot v_1 leads to a lower dissatisfaction of the single most dissatisfied voter—namely the first voter—with committee $\{c_2, c_3\}$, while the dissatisfaction of the most dissatisfied voter for the other winning committees does not change. Therefore, $\{c_1, c_2\} \notin f^{\ell}_{max}(P'_\ell, 2) = \big\{\{c_2, c_3\}\big\}$, although $c_2 \in \{c_1, c_2\}$ was improved in the original election. $\square$

Elkind et al. (2017a) also consider the weaker notion of *non-crossing monotonicity* where only improvements for $c \in W$ are allowed that do not change the respective relations between all members of W. The results presented above still hold for this variant.

Next, consider Condorcet consistency. In a profile P_ℓ of ℓ-ballots, a candidate c_i beats a candidate c_j in a pairwise comparison if more voters prefer c_i to c_j than the other way round, i.e., $|\{v \in P_\ell \mid v(i) < v(j)\}| > |\{v \in P_\ell \mid v(j) < v(i)\}|$. Recall that a Condorcet winner $c \in C$ beats all other candidate in pairwise comparisons, and that a (singlewinner) voting rule is Condorcet consistent when the Condorcet winner is the unique winner of the election whenever it exists (see the definitions on pages 11 and 12). As in the case of monotonicity, one can define two notions of Condorcet consistency in the multiwinner context, namely one focusing on the candidates and one focusing on the committees. The candidate variant is due to Felsenthal and Maoz (1992) who adapted the singlewinner property of Condorcet consistency to resolute multiwinner voting rules by requiring that a Condorcet winner has to be a member of the winning committee. See Definition 3.17 for a restatement of this definition for irresolute multiwinner voting rules. The committee variant states that the committee Condorcet winner has to be the unique winning committee of the election if it exists, where the *committee Condorcet winner* of size k is a committee $W \in F_k(C)$ for a set of candidates C so that each $c \in W$ beats each $d \in C \setminus W$ in a pairwise comparison (Gehrlein, 1985). Note that a committee Condorcet winner is unique, and that it contains the Condorcet winner if such a candidate exists.

Definition 3.17. A multiwinner voting rule $\mathcal{F}$ satisfies

- *Condorcet consistency*, if for each committee election $\mathcal{E} = (C, P, k)$ where a Condorcet winner $c \in C$ exists, c is a confirmed committee member under $\mathcal{E}$.

- *committee Condorcet consistency*, if for each committee election $\mathcal{E} = (C, P, k)$ where a committee Condorcet winner W exists, it holds that $\{W\} = \mathcal{F}(P, k)$.

Note that committee Condorcet consistency implies Condorcet consistency.

Theorem 3.18. *Neither the minisum nor the minimax ℓ-group rules are Condorcet consistent or committee Condorcet consistent.*

Similar to the property defined for social welfare functions on page 12, a singlewinner voting rule is Pareto optimal (or satisfies the *Pareto criterion*) if only candidates that are not dominated by another candidate can win. A candidate c dominates a candidate d if c is preferred to d by all voters. However, this definition has to be slightly modified for the multiwinner context. For example, a dominated candidate might have to be part of the winning committee when there are not enough not dominated candidates left to fill the committee (as can be the case, e.g., in a *unanimous profile* where all voters prefer a candidate c to every other candidate). See, e.g., the work by Felsenthal and Maoz (1992).

Definition 3.19. Let $\mathcal{E} = (C, P, k)$ be a committee election. A multiwinner voting rule $\mathcal{F}$ satisfies the *Pareto criterion* if the following holds: If a candidate $c_i \in C$ is preferred to a candidate $c_j \in C$ by all voters in P, i.e., $v(i) < v(j)$ for all $v \in P$ in the context of ℓ-ballots, it holds that $c_j \in W$ for some $W \in \mathcal{F}(P, k)$ implies that $c_i \in W$.

Theorem 3.20. *For each $\ell \in \mathbb{N}$, both the minisum and the minimax ℓ-group rules satisfy the Pareto criterion.*

Proof. Le ℓ be an arbitrary positive integer. Assume that the minisum ℓ-group rule (respectively, the minimax ℓ-group rule) does not satisfy the Pareto criterion. Then there exists a committee election $\mathcal{E} = (C, P_\ell, k)$ and candidates $c_i, c_j \in C$, so that $v(i) < v(j)\ \forall v \in P_\ell$, but $c_j \in W$ and $c_i \notin W$ for some $W \in f^\ell_{sum}(P_\ell, k)$ (respectively, $W \in f^\ell_{max}(P_\ell, k)$). A voter's dissatisfaction with a committee only depends on the individual members. The dissatisfaction of each voter v with the committee $W' = (W \cup \{c_i\}) \setminus \{c_j\}$ is strictly less than the dissatisfaction with W:

$$\delta_\ell(v, W') - \delta_\ell(v, W) = v(i) - 1 + \ell - v(j) - \big(v(j) - 1 + \ell - v(i)\big) = 2v(i) - 2v(j) < 0,$$

and therefore

$$\sum_{v \in P_\ell} \delta_\ell(v, W') < \sum_{v \in P_\ell} \delta_\ell(v, W)$$

$$\Big(\text{respectively,} \qquad \max_{v \in P_\ell} \delta_\ell(v, W') < \max_{v \in P_\ell} \delta_\ell(v, W)\Big)$$

This is a contradiction to the fact that W is a winning committee under f^ℓ_{sum} (respectively, f^ℓ_{max}). It follows that both the minisum and the minimax ℓ-group rules satisfy the Pareto criterion for all $\ell > 0$. □

The following properties aim at ensuring proportionality by demanding that a committee W wins the election when a large enough group of voters prefers the members in W to the candidates outside W. They were introduced by Elkind et al. (2017a) and weaken an axiom proposed by Dummett (1984). Here, they are stated in the context of ℓ-ballots. See the axioms of justified representation and Gehrlein stability on page 29 for similar ideas.

Definition 3.21. A multiwinner voting rule $\mathcal{F}$ satisfies

- *solid coalitions* if for each committee election $\mathcal{E} = (C, P_\ell, k)$ with n voters where

$$|\{v \in P_\ell \mid v(i) < v(j)\ \forall c_j \in C \setminus \{c_i\}\}| \geq n/k$$

  for a candidate $c_i \in C$ implies that c_i is a confirmed committee member,

- *consensus committee* if for each committee election $\mathcal{E} = (C, P_\ell, k)$ with n voters and each $W \in F_k(C)$ so that each voter ranks some member of W higher than all other candidates and each member of W is ranked higher than all other candidates by either $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$ voters, it holds that $\mathcal{F}(P_\ell, k) = \{W\}$, and

- *strong (respectively, weak) unanimity* if for each committee election $\mathcal{E} = (C, P_\ell, k)$, where $v(i) < v(j)$ for all $v \in P_\ell$, $c_i \in W$, and $c_j \in C \setminus W$, it holds that $\{W\} = \mathcal{F}(P_\ell, k)$ (respectively, $W \in \mathcal{F}(P_\ell, k)$).

Note that committee Condorcet consistency implies strong unanimity, which implies weak unanimity.

Theorem 3.22. *For each $\ell \geq 2$, neither the minisum nor the minimax ℓ-group rules satisfy solid coalitions and consensus committee, but they satisfy strong unanimity.*

Table 3.5 summarizes the axiomatic property results in this section. Note that the notion of (candidate) monotonicity in the attached paper (Baumeister et al., 2016) corresponds to candidate monotonicity as defined in Definition 3.14. The notion of (crossing and non-crossing) monotonicity was not previously considered for ℓ-group rules.

Property	ℓ-group rules			
	minisum		minimax	
Non-imposition	✓	Thm. 3.7	✓	Thm. 3.7
Homogeneity	✓	Thm. 3.7	✓	Thm. 3.7
Consistency	✓	Thm. 3.9	×	Thm. 3.9
Independence of clones	✓	Thm. 3.11	×	Thm. 3.11
Committee monotonicity	✓	Thm. 3.13	×	Thm. 3.13
Candidate monotonicity	✓	Thm. 3.15	✓	Thm. 3.16
Monotonicity	✓	Thm. 3.15	×	Thm. 3.16
Positive responsiveness	✓	Thm. 3.15	×	Thm. 3.16
Condorcet consistency	×	Thm. 3.18	×	Thm. 3.18
Committee Condorcet consistency	×	Thm. 3.18	×	Thm. 3.18
Pareto criterion	✓	Thm. 3.20	✓	Thm. 3.20
Solid coalitions	×	Thm. 3.22	×	Thm. 3.22
Consensus committee	×	Thm. 3.22	×	Thm. 3.22
Unanimity	strong	Thm. 3.22	strong	Thm. 3.22

Table 3.5: Property results for the minisum and minimax ℓ-group rules. The respective results hold for each applicable ℓ.

3.3 Computational Complexity of Winner Determination

This section studies the computational complexity of winner determination for the minisum and minimax ℓ-group rules, starting with the minisum ℓ-group rules. As stated in Claim 3.4, the winning committees for f_{sum}^{ℓ} can be determined by computing the minisum scores of all candidates, which is obviously possible in polynomial time.

Theorem 3.23. *For each $\ell \in \mathbb{N}$, all winning committees for the minisum ℓ-group rule can be computed in polynomial time.*

However, this is not the case for the minimax ℓ-group rules. Recall that 2-ballots correspond to approval ballots. Therefore, f_{max}^{2} corresponds to the original minimax rule introduced by Brams et al. (2007). LeGrand (2004) shows the NP-hardness of the following problem for $\ell = 2$.

MINIMAX-ℓ-SCORE	
Given:	A committee election $\mathcal{E} = (C, P_\ell, k)$, and a nonnegative integer d.
Question:	Is there a committee $W \in F_k(C)$ such that $\max_{v \in P_\ell} \delta(v, W) \leq d$?

Since this problem would be tractable if it were possible to compute a winning committee for f^2_{max} in polynomial time, it follows that the winner determination for f^2_{max} is intractable as well. Furthermore, Misra et al. (2015) show that MINIMAX-2-SCORE is W[2]-hard when parameterized by the size of the committee k. These results can be generalized to all values of ℓ. However, Misra et al. (2015) also show that there exists an fpt-algorithm for MINIMAX-2-SCORE when parameterized by d. The following theorem generalizes the result for each ℓ.

Theorem 3.24. *For each ℓ,* MINIMAX-ℓ-SCORE *is in* FPT *when parameterized by d.*

3.4 (a,b)-rules

The ℓ-ballots introduced in Section 3.1 are a compromise between ordinal and cardinal ballots in the way that voters can assign a dissatisfaction value to candidates where the values are bound by an underlying ranking of the candidates. There are some articles exploring cardinal preferences,[2] but cardinal ballots and corresponding voting rules are rarely considered in preference aggregation. One example is range voting for singlewinner elections introduced by Smith (2000). Voters assign each candidate a real number from the interval $[-1, 1]$ where a higher number implies a higher satisfaction with the respective candidate, and range voting then elects the candidate(s) with the highest sum of satisfaction.

[2] Ballester and Rey-Biel (2006) study a model where voters have cardinal preferences and have to translate these to the ballots allowed for the respective voting rules. In particular, they focus on approval voting which uses approval ballots and study whether the optimal strategy for voters with cardinal preferences is to submit sincere ballots, i.e., ballots correspond to their underlying preferences in a certain way. Procaccia and Rosenschein (2006) also focus on cardinal preferences and define a notion of "distortion" that occurs when translating cardinal preferences into linear orders.

This section introduces (a,b)-ballots that generalize the concept of ℓ-ballots to fully cardinal ballots. In contrast to range voting, voters assign not only a dissatisfaction value a for the case that a candidate is part of a committe, but also a separate, independent value b of dissatisfaction with a candidate not being a member of a committee. This models for example the case when candidates have different attributes and voters have a ranking for each attribute. A possible attribute for a candidate is, e.g., the party membership. Two independent dissatisfaction values are able to model cases where the conflicting preferences over candidates' attributes do not correspond to an underlying weak order.

Definition 3.25 ((a,b)-ballots)**.** An (a,b)-ballot v_j over candidate set C is a set

$$v_j = \{(c_i, a_i^j, b_i^j) \mid c_i \in C,\ a_i^j, b_i^j \in \mathbb{Q},\ 1 \le i \le |C|\}.$$

Then voter j is said to strictly prefer candidate c_1 to candidate c_2 if and only if

$$(a_1^j < a_2^j \text{ and } b_1^j \ge b_2^j) \text{ or } (a_1^j \le a_2^j \text{ and } b_1^j > b_2^j),$$

and voter j is said to be indifferent between c_1 and c_2 if and only if $a_1^j = a_2^j$ and $b_1^j = b_2^j$. In the remaining cases, it remains unknown which candidate the voter prefers. Note that an ℓ-ballot is a special case of an (a,b)-ballot over the same candidate set C where $a_i^j = v_j(i) - 1$ and $b_i^j = \ell - v_j(i)$ for each candidate $c_i \in C$, but any (a,b)-ballot where the sum of both values is a constant can be interpreted as an ℓ-ballot for a corresponding ℓ.

Recall that $F_k(C) = \{S \mid S \subseteq C \wedge |S| = k\}$ is the set of all committees of size k over C. The dissatisfaction of a voter j—associated with (a,b)-ballot v_j—with a committee $W \in F_k(C)$ is measured as

$$\delta_{(a,b)}(v_j, W) = \sum_{c_i \in W} a_i^j + \sum_{c_i \notin W} b_i^j.$$

Analogous to the minisum and minimax ℓ-group rules, define the following multiwinner rules using (a,b)-ballots:

Definition 3.26 (minisum and minimax (a,b)-rules)**.** Let $\mathcal{E} = (C, P_{(a,b)}, k)$ be a committee election with a profile $P_{(a,b)}$ of (a,b)-ballots over C and let $F_k(C) = \{S \mid S \subseteq C \wedge |S| = k\}$ be the set of committees of size k.

- The *minisum* (a,b)*-rule* is a function $f_{sum}^{(a,b)}$ that returns the committees minimizing the total dissatisfaction of the voters with the elected committees, i.e.,

$$f_{sum}^{(a,b)}(P_{(a,b)},k) = \operatorname*{argmin}_{W \in F_k(C)} \sum_{v \in P_{(a,b)}} \delta_{(a,b)}(v,W).$$

- The *minimax* (a,b)*-rule* is a function $f_{max}^{(a,b)}$ that returns the committees minimizing the dissatisfaction of the respective least satisfied voter with the elected committees, i.e.,

$$f_{max}^{(a,b)}(P_{(a,b)},k) = \operatorname*{argmin}_{W \in F_k(C)} \max_{v \in P_{(a,b)}} \delta_{(a,b)}(v,W).$$

The following example illustrates (a,b)-ballots and the corresponding multiwinner rules.

Example 3.27. (a,b)-rules Let $\mathcal{E} = (C, P_{(a,b)}, 2)$ be a committee election where $C = \{c_1, c_2, c_3\}$ and $P_{(a,b)} = (v_1, v_2)$ is a profile of (a,b)-ballots over C with

$$\begin{aligned} v_1 =& \{(c_1,3,0),(c_2,2,2),(c_3,1,2)\}, \\ v_2 =& \{(c_1,1,2),(c_2,4,3),(c_3,0,4)\}. \end{aligned}$$

Voter 1 strictly prefers c_3 to both c_1 and c_2, since $(a_3^1 < a_1^1 \wedge b_3^1 \geq b_1^1)$ and $(a_3^1 < a_2^1 \wedge b_3^1 \geq b_2^1)$, and also strictly prefers c_2 to c_1. Voter 2 strictly prefers c_3 to c_2 and strictly prefers c_1 to c_2, but the ballot does not allow to draw conclusions about the voter's preference over c_1 and c_3.

Now turn to the minisum and the minimax (a,b)-rules. Recall that the dissatisfaction of a voter j with a committee W is the total of a_i^j for all $c_i \in W$ and b_i^j for all $c_i \notin W$. Table 3.6 shows the dissatisfaction of the voters with each committee of size 2. Therefore, $f_{sum}^{(a,b)}(P_{(a,b)},2) = \{\{c_2,c_3\}\}$ and $f_{max}^{(a,b)}(P_{(a,b)},2) = \{\{c_1,c_3\},\{c_2,c_3\}\}$.

The properties presented in Section 3.2 are defined for ballots that correspond to ordinal ballots. Note that in the restricted model where (a,b)-ballots have to correspond to an underlying weak order, the results in Section 3.2 also hold for the minisum and minimax (a,b)-rules. However, the results for winner determination differ slightly from the results presented in Section 3.3. My coauthors Böhnlein and Schaudt (personal communication, 2016) prove that while a winning committee for the (a,b)-minisum-rule can be computed

	$\{c_1, c_2\}$	$\{c_1, c_3\}$	$\{c_2, c_3\}$
v_1 :	$3+2+2=7$	$3+2+1=6$	$0+2+1=3$
v_2 :	$1+4+4=9$	$1+3+0=4$	$2+4+0=6$
$\sum$:	16	10	9
max :	9	6	6

Table 3.6: Dissatisfaction of voters with each committee of size 2 in Example 3.27.

in polynomial time, the auxiliary problem MINIMAX (a,b)-SCORE—the (a,b)-version of the problem defined in Section 3.3—is W[2]-hard when parameterized by d and k. However, when the (a,b)-ballots are restricted in a way that they correspond to ℓ-ballots after some normalization (where the value of ℓ may be different for each ballot), they show that Theorem 3.24 also holds for the (a,b)-minimax rule.

Unfortunately, the expressiveness of the presented model demands a great deal of the voters since the necessity to assign two values to each candidate becomes more and more infeasible as the number of candidates grows. Depending on the application, it might therefore be reasonable to restrict the possible values of a and b by, e.g., providing lower and upper bounds, fixing the total sum of a voter's a-values and of a voter's b-values, or by using ℓ-ballots.

3.5 My Contribution

In joint work with my coauthors, I developed the models of ℓ-ballots and (a,b)-ballots and the corresponding rules and modified the properties in Section 3.2—when necessary—for the context of ℓ-ballots. Furthermore, I contributed the proofs to Claim 3.4 and Theorems 3.13, 3.15, 3.16, 3.20, and 3.23. The writing of the attached article was done jointly with my coauthors.

Minisum and Minimax Committee Election Rules for General Preference Types

Dorothea Baumeister[1] and **Toni Böhnlein**[2] and **Lisa Rey**[1] and **Oliver Schaudt**[2] and **Ann-Kathrin Selker**[1]

Abstract. In committee elections it is often assumed that voters only (dis)approve of each candidate or that they rank all candidates, as it is common for single-winner elections. We suggest an intermediate approach, where the voters rank the candidates into a fixed number of groups. This allows more diverse votes than approval votes, but leaves more freedom than in a linear order. A committee is then elected by applying the minisum or minimax approach to minimize the voters' dissatisfaction. We study the axiomatic properties of these committee election rules as well as the complexity of winner determination and show fixed-parameter tractability for our minimax rules.

1 Introduction

A central point in computational social choice is the analysis of voting systems, see for example the book chapter by Zwicker [10]. Whereas the initial focus was mainly on single-winner elections, the study of committee elections recently received considerable attention. In a committee election a winner is a subset of candidates of a predefined size.

Most voting rules require the voters to either rank all candidates in a strict linear order, which might be impossible given a large set of candidates, or to divide them into two groups, i. e., approval ballots, which might be too rough to fully express the voters' preferences. As an intermediate approach, we propose ℓ-ballots. Voters group the candidates into a fixed number of groups, where all candidates in one group are tied. We use this type of ballot – a slight variant of the model proposed by Obraztsova et al. [9]– to define committee election rules that minimize the voters' dissatisfaction and study computational and axiomatic properties of these rules. To that end, we apply the well-known minisum method where the sum of the distances to the individual votes is minimized, and the minimax method where the maximal distance to an individual vote is minimized. Originally, the minisum and minimax methods have been applied to approval votes by Brams et al. [3]. The most relevant papers for our study are those by Baumeister et al. [1, 2] who extended this approach to determine winning committees for different forms of votes, namely trichotomous votes as well as complete and incomplete linear orders. Elkind et al. [5] studied axiomatic properties such as consistency, monotonicity, and solid coalitions for different multiwinner voting rules, including STV, Bloc, k-Borda and different variants of the Chamberlin-Courant and Monroe's rule. We adapt some of these properties to our setting and study them for the class of ℓ-group rules. The parameterized complexity of minimax voting rules has been studied by Misra et al. [8] for approval votes as well as by Liu and Guo [7] for trichotomous votes and linear and partial orders. In both papers it is shown that, for their respective voting rules, computing a winning committee is W[2]-hard when parameterized by the size of the committee and that computing a winning committee is fixed-parameter tractable with respect to a distance parameter.

[1] Heinrich-Heine-Universität Düsseldorf, Germany, email: {baumeister, lrey, selker}@cs.uni-duesseldorf.de

[2] Universität zu Köln, Germany, email: boehnlein@zpr.uni-koeln.de, schaudto@uni-koeln.de

2 Definitions

Let $C = \{c_1, \dots, c_m\}$ be a set of candidates and $V = (v_1, \dots, v_n)$ a profile, i. e., a list of voters represented by their vote. In an *ℓ-ballot* over C, a vote is given as a list of ℓ pairwise disjoint sets of candidates, which may also be empty: $v = (G_1, \dots, G_\ell)$ where $G_i \cap G_j = \emptyset$ for $1 \le i, j \le \ell$ and $i \ne j$, and $\bigcup_{1 \le i \le \ell} G_i = C$. Considering a set of candidates $C = \{c_1, c_2, c_3, c_4\}$, a possible 3-ballot is $(\{c_3, c_4\}, \{\}, \{c_1, c_2\})$ which means that candidates c_3 and c_4 are preferred to all other candidates, and candidates c_1 and c_2 are the most disliked ones.

A very similar ballot model has been introduced by Obraztsova et al. [9]. The predefined ℓ groups correspond to their preference levels. In contrast to our model, they assume that the first and last group are never empty and that at least one voter specifies no empty group. However, these are only technical requirements that are not crucial for our results.

A *committee* is a subset of C. Let $F_k(C)$ denote the set of all committees of size k. A *committee election* is a triple $E = (C, V, k)$, where C is the set of candidates, V is a list of voters, represented by ℓ-ballots for some fixed constant ℓ over C, and $k \in \mathbb{N}$ denotes the committee size. A *committee election rule* $\mathcal{R}$ is a function that, given a committee election, returns a set of tied winning committees.

Now we introduce the ℓ-group voting rules discussed in this paper. For this sake we define $\delta_\ell(v, W) = \sum_{c \in C} |v(c) - W(c)|$ as the *dissatisfaction* (or *distance*) between a ballot v and a committee $W \in F_k(C)$ where $W(c) = 1$ for a candidate $c \in W$, and $W(c) = \ell$ for a candidate $c \notin W$, and where $v(c)$ denotes the group number of a candidate c. For the case of $\ell = 2$ this distance corresponds to the Hamming distance between the vote and the committee. The following two rules elect the winning committee(s) for profiles consisting of ℓ-ballots.

Definition 1 (minisum/minimax ℓ-group rule)

- *Minisum ℓ-group rules are functions f^ℓ_{sum} so that $f^\ell_{sum}((C,V,k)) = \text{argmin}_{W \in F_k(C)} \sum_{v \in V} \delta_\ell(v, W)$, i. e., f^ℓ_{sum} minimizes the sum of the voters' dissatisfaction to the winning committees.*
- *Minimax ℓ-group rules are functions f^ℓ_{max} so that $f^\ell_{max}((C,V,k)) = \text{argmin}_{W \in F_k(C)} \max_{v \in V} \delta_\ell(v, W)$, i. e., f^ℓ_{max} minimizes the dissatisfaction of the least satisfied voter with the winning committees.*

Note that the minisum/minimax voting rules defined by Baumeister and Dennisen [1] correspond to our minisum/minimax ℓ-group rules for $\ell = 2, 3$, and m, and without allowing empty groups.

3 Results

Due to space restrictions we present only the results of our work. For the axiomatic study, we first adapt the existing definitions for some properties to handle the more general input type of ℓ-ballots. Then we can show that the minisum ℓ-group rules satisfy nearly all properties at hand, whereas the minimax ℓ-group rules violate some of them. An overview of our results is given in Table 1.

Properties	ℓ-group rules minisum	ℓ-group rules minimax
Non-imposition, Homogeneity	✓	✓
Consistency	✓	×
Independence of clones	✓	×
Committee monotonicity	✓	×
(Candidate) monotonicity	✓	✓
Positive responsiveness	✓	×
Pareto criterion	✓	✓
(Committee) Condorcet consistency	×	×
Solid coalitions, Consensus committee	×	×
Unanimity	strong	strong

Table 1: Properties for minisum and minimax ℓ-group rules

Next, we study the complexity of computing a winning committee for minisum and minimax ℓ-group rules. For the minisum rule the problem can be solved in polynomial time, as it can be shown that the candidates c with the lowest score $\sum_{v \in V} v(c)$ form a winning committee.

For the study of minimax rules we need the following auxiliary decision problem.

MINIMAX ℓ-SCORE	
Given:	A committee election $E = (C, V, k)$, and a nonnegative integer d.
Question:	Is there a committee $W \in F_k(C)$ such that $\max_{v \in V} \delta_\ell(v, W) \leq d$?

LeGrand et al. [6] show that a problem corresponding to our MINIMAX 2-SCORE is NP-hard, a result that can be generalized to every greater value of ℓ. On these grounds we resort to the study of parameterized complexity. Thus, our goal is to formulate an efficient algorithm when certain parameters of the problem are small, i. e., can be treated as a constant.[3] For approval voting Misra et al. [8] show that the problem is W[2]-hard, when parameterized by the size of the committee. This hardness result also applies to MINIMAX ℓ-SCORE. Hence, an attempt to tune an algorithm with respect to the size of the committee is most likely going to result in failure.

As a positive result we give an algorithm that efficiently solves the MINIMAX ℓ-SCORE problem when the parameter d is treated as a constant. Which proofs the following theorem.

Theorem 1 *There is an algorithm solving* MINIMAX ℓ-SCORE *whose running time is in* $O\left((mn + m\log m)\left(\frac{\sqrt{33}}{2}d\right)^d\right)$. *In particular,* MINIMAX ℓ-SCORE *is fixed-parameter tractable when parameterized by d.*

[3] For formal definitions and background regarding parameterized complexity we refer to the book of Downey and Fellows [4].

4 Conclusion

We have introduced different ways of expressing the voters' preferences in committee elections, namely ℓ-ballots, an intermediate between approval votes and linear orders. In addition to axiomatic properties, we have studied the computational complexity of winner determination. While in the minisum case computing a winning committee under the ℓ-group rule can be done efficiently, the case of minimax is NP-hard, however there exists a fixed-parameter tractable algorithm that determines a winning committee.

Note that the input type of ℓ-ballots is only one form of a more general vote. In our setting the differences in scores between two groups are always equivalent and there may be situations where for example the first two groups are of greater importance than the other ones. So as a very general framework one could consider that each voter reports two dissatisfaction values (a, b) to each candidate, one for the case that the candidate is in the committee, the other one for the case where the candidate is not in the committee.[4] We call the resulting voting rules minisum/minimax-(a, b)-rules. Obviously our ℓ-group rules are obtained as a special case of such (a, b)-rules, when we restrict the input to $a + b = l$ for each voter. More interestingly, we can show that under some mild restrictions the results obtained in this paper even hold for the very general class of (a, b)-rules.

As a task for future work we propose to identify other interesting special cases of (a, b)-rules and provide a characterization for them. Furthermore, we want to consider different rules for these types of input and identify which of the properties from Table 1 are satisfied, and especially find rules that fulfill Condorcet consistency and committee Condorcet consistency. Closely related to the setting of minisum and minimax elections are the systems of proportional representation, which itself is related to the interesting concept of justified representation so a task for future research is to redefine and study these concepts for more general types of votes.

REFERENCES

[1] D. Baumeister and S. Dennisen. Voter Dissatisfaction in Committee Elections. In *Proc. AAMAS'15*. IFAAMAS, 2015. Extended abstract.

[2] D. Baumeister, S. Dennisen, and L. Rey. Winner determination and manipulation in minisum and minimax committee elections. In *Proc. ADT'15*, pages 469–485. Springer-Verlag *LNCS #9346*, 2015.

[3] S. Brams, D. Kilgour, and R. Sanver. A minimax procedure for electing committees. *Public Choice*, 132:401–420, 2007.

[4] R. Downey and M. Fellows. *Fundamentals of Parameterized Complexity*. Springer-Verlag, 2013.

[5] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. Properties of multiwinner voting rules. In *Proc. AAMAS'14*, pages 53–60. IFAAMAS, 2014.

[6] R. LeGrand. Analysis of the Minimax Procedure. Technical Report WUCSE-2004-67, Department of Computer Science and Engineering, Washington University, St. Louis, Missouri, 2004.

[7] H. Liu and J. Guo. Parameterized Complexity of Winner Determination in Minimax Committee Elections. In *Proc. AAMAS'16*. IFAAMAS, 2016.

[8] N. Misra, A. Nabeel, and H. Singh. On the Parameterized Complexity of Minimax Approval Voting. In *Proc. AAMAS'15*, pages 97–105. IFAAMAS, 2015.

[9] S. Obraztsova, E. Elkind, M. Polukarov, and Z. Rabinovich. Doodle poll games. In *Proc. of the 1st IJCAI-Workshop on Algorithmic Game Theory*, 2015.

[10] W. Zwicker. Introduction to the Theory of Voting. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 2. Cambridge University Press, 2016.

[4] Note that we consider dissatisfaction values, as the goal of our minisum and minimax rules is to minimize the voters' dissatisfaction.

Chapter 4

Complexity of Shift Bribery for Iterative Voting Rules

This chapter deals with shift bribery for iterative (positional) scoring rules as defined on page 23 and on page 10, respectively. The attached article (Maushagen et al., 2021) was submitted to the *Journal of Artificial Intelligence Research* and is based on a preliminary conference version (Maushagen et al., 2018b). An earlier version was also presented at ISAIM'18 (Maushagen et al., 2018a).

Maushagen, C., Neveling, M., Rothe, J., and Selker, A.-K. (2021). Complexity of shift bribery for iterative voting rules. Submitted to *Journal of Artificial Intelligence Research*

Summary

My coauthors and I study the computational complexity of shift bribery for the following iterative scoring rules.

Hare (defined, e.g., in the textbook by Taylor (2005) eliminates the candidates with the lowest plurality score.

Coombs (defined, e.g., by Levin and Nalebuff (1995)) eliminates the candidates with the lowest veto score.

Baldwin eliminates the candidates with the lowest Borda score (Baldwin, 1926).

Nanson eliminates all candidates with lower than average Borda score (Nanson, 1882).

Iterated plurality (defined, e.g., in the textbook by Taylor (2005)) eliminates the candidates that do not have the highest plurality score.

Iterated veto eliminates the candidates that do not have the highest veto score (see Example 2.12 on page 23).

Plurality with runoff (again defined, e.g., by Taylor (2005)) proceeds in two rounds. In the first round, all candidates that do not have the highest plurality score are eliminated, unless there is a unique plurality winner, then all candidates that do not have the highest or second-highest score are eliminated.

Veto with runoff is the veto variant of plurality with runoff (see Example 2.10 on page 19).

We show that shift bribery is NP-complete for all considered voting rules for both the constructive and the destructive variants. Our results hold for both the unique and the nonunique winner model: In the constructive case of the *unique winner model*, the attacker wants to make the target candidate the unique winner of the election, whereas in the *nonunique winner model*, the attack is successful when the target candidate is a member of the set of winning candidates. Further, as an example we state modified proofs for constructive shift bribery for the rules Hare and plurality with runoff that show that the computational complexity does not change when we allow the attacker to exploit nonmonotonicity. We conjecture that shift bribery for all our considered nonmonotonic rules does not become tractable in this setting.

My Contribution

The writing was done jointly with my co-authors. I was responsible for Section 6 (the complexity results for iterated veto and veto with run-off including the proofs) and Example 2.

Journal of Artificial Intelligence Research () Submitted ; published

Complexity of Shift Bribery for Iterative Voting Rules*

Cynthia Maushagen CYNTHIA.MAUSHAGEN@HHU.DE
Marc Neveling MARC.NEVELING@HHU.DE
Jörg Rothe ROTHE@HHU.DE
Ann-Kathrin Selker ANN-KATHRIN.SELKER@HHU.DE
Heinrich-Heine-Universität Düsseldorf
Universitätsstr. 1
40225 Düsseldorf, Germany

Abstract

In iterative voting systems, candidates are eliminated in consecutive rounds until either a fixed number of rounds is reached or the set of remaining candidates does not change anymore. We focus on iterative voting systems based on the positional scoring rules plurality, veto, and Borda and study their resistance against shift bribery attacks due to Elkind, Faliszewski, and Slinko (2009) and Kaczmarczyk and Faliszewski (2016). In constructive shift bribery (Elkind et al., 2009), an attacker seeks to make a designated candidate win the election by bribing voters to shift this candidate in their preferences; in destructive shift bribery (Kaczmarczyk & Faliszewski, 2016), the briber's goal is to prevent this candidate's victory. We show that many iterative voting systems are resistant to these types of attack, i.e., the corresponding decision problems are NP-hard. These iterative voting systems include iterated plurality as well as the voting rules due to Hare (see, e.g., the book by Taylor, 2005), Coombs (see, e.g., the article by Levin & Nalebuff, 1995), Baldwin (1926), and Nanson (1882); variants of Hare voting are also known as single transferable vote, instant-runoff voting, and alternative vote.

1. Introduction

One of the main themes in computational social choice (Brandt, Conitzer, Endriss, Lang, & Procaccia, 2016; Rothe, 2015) is to study the complexity of manipulative attacks on voting systems, in the hope that proving computational hardness of such attacks may provide some sort of protection against them. Besides manipulation (Bartholdi, Tovey, & Trick, 1989; Conitzer, Sandholm, & Lang, 2007)—also referred to as strategic voting—and electoral control (Bartholdi, Tovey, & Trick, 1992; Hemaspaandra, Hemaspaandra, & Rothe, 2007), much work has been done to study bribery attacks. For a comprehensive overview of the formal models and the related complexity results, we refer to the book chapters by Conitzer and Walsh (2016) for manipulation, by Faliszewski and Rothe (2016) for control and bribery, and by Baumeister and Rothe (2015) for all three topics.

Bribery in voting was introduced by Faliszewski, Hemaspaandra, and Hemaspaandra (2009a, see also the article by Faliszewski, Hemaspaandra, Hemaspaandra, & Rothe, 2009b). In their model, a briber intends to change the outcome of an election to his or her own advantage by bribing certain voters without

*. This paper extends the preliminary conference versions that appear in the proceedings of the *17th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS'18, see Maushagen, Neveling, Rothe, & Selker, 2018) and in the nonarchival website proceedings of the *International Symposium on Artificial Intelligence and Mathematics* (ISAIM'18) by presenting all proofs (some of which were omitted in the conference versions due to space limitations) in full detail, by adding new results on iterated veto and veto with runoff in Theorems 11 and 12, and by adding more illustrating examples and discussion (such as the discussion in Section 7 with new Theorems 13 and 14).

exceeding a given budget. Bribery shares some features with manipulation, as the briber (just like a strategic voter) has to find the right preference orders that the bribed voters are then requested to change their votes to. Bribery also shares some features with electoral control, as the briber (just like an election chair) has to pick the right voters to bribe so as to make the cost of bribing them as inexpensive as possible and to stay within the allowed budget.

We will focus on *shift bribery*, which was introduced by Faliszewski et al. (2009b) in the context of so-called irrational voters for Copeland elections and was then studied in detail by Elkind et al. (2009) for the constructive variant (where the briber's goal is to make a favorite candidate win the election) and was later studied by Kaczmarczyk and Faliszewski (2016) in the destructive variant (where the briber's goal is to make sure that a despised candidate does not win the election). In *swap bribery*, which generalizes shift bribery, the briber has to pay for each swap of any two candidates in the votes. Shift bribery additionally requires that swaps always involve the designated candidate that the briber wants to see win (in the constructive case) or not win (in the destructive case).

A natural interpretation of swap bribery—and thus in particular of shift bribery—regards *campaign management*: A campaign manager organizing a political campaign for some candidate seeks to influence the public opinion about this candidate by legal activities such as, e.g., running targeted television ads. Those ads might influence voters to change their opinion (and consequently their vote) of the targeted candidate positively or negatively. Campaign managers are restricted by a budget and need to choose the right ads to run in order to increase their candidates' chances of winning. Shift bribery can be seen to model campaign management in a more ethical way than general swap bribery, as campaign managers then always target their own candidates only and thus cannot change the voters' opinions over pairs of other candidates.

Another natural interpretation of swap bribery regards election fraud detection: If the winner of an election can be dethroned by only a few changes (by swapping candidates) to the votes then the election might have been tampered with or, from a more optimistic viewpoint, small errors in the counting of the votes might have influenced the election result. In that situation, a recounting would be required since for a close election result only few errors in the counting are needed to elect a candidate that is not the "true" winner of the election. This has been studied as the *margin of victory* (Xia, 2012; Reisch, Rothe, & Schend, 2014), which is closely related to destructive bribery. In this context, shift bribery models a more fine-grained search for election fraud which targets only a specific candidate.

Swap bribery generalizes the possible winner problem (Konczak & Lang, 2005; Xia & Conitzer, 2011), which itself is a generalization of unweighted coalitional manipulation. Therefore, each of the many hardness results known for the possible winner problem is directly inherited by the swap bribery problem. This was the motivation for Elkind et al. (2009) to look at restricted variants of swap bribery such as shift bribery.

Even though shift bribery possesses a number of hardness results (Elkind et al., 2009), it has also been shown to allow exact and approximate polynomial-time algorithms in a number of cases (Elkind et al., 2009; Elkind & Faliszewski, 2010; Schlotter, Faliszewski, & Elkind, 2017). For example, Elkind et al. (2009) provided a 2-approximation algorithm for shift bribery when using Borda voting.[1] This result was

1. In *Borda* with m candidates, each vote is a linear order of the candidates, the ith candidate in a vote scores $m-i$ points, and whoever has the most points wins. Borda is a very prominent positional scoring rule and can be described by the scoring vector $(m-1, m-2, \dots, 0)$. Other prominent positional scoring rules are *plurality*, where only the top candidates in the votes score a point and no one else (i.e., plurality has the scoring vector $(1,0,\dots,0)$), and *veto* (a.k.a. *antiplurality*), where all except the bottom candidates in the votes score a point (i.e., veto has the scoring vector $(1,\dots,1,0)$); again, whoever has the most points wins in these rules.

extended by Elkind and Faliszewski (2010) to all positional scoring rules; they also obtained somewhat weaker approximations for Copeland and maximin voting. Very recently Faliszewski, Manurangsi, and Sornat (2019) further extended this result to a polynomial-time approximation scheme. For Bucklin and fallback voting, the shift bribery problem is even exactly solvable in polynomial time (Schlotter et al., 2017).[2] In addition, Bredereck, Chen, Faliszewski, Nichterlein, and Niedermeier (2014b) were the first to analyze shift bribery in terms of *parameterized* complexity, and only recently a long-standing open problem regarding the parameterized complexity of bribery (including shift bribery) with the number of candidates as the parameter (see the survey by Bredereck, Chen, Faliszewski, Guo, Niedermeier, and Woeginger (2014a) for a deeper discussion on this problem) was solved by Knop, Koutecký, and Mnich (2017) for a multitude of voting rules. Furthermore, Bredereck, Faliszewski, Niedermeier, and Talmon (2016b) introduced combinatorial shift bribery in which a single shift bribery action affects multiple voters and Bredereck, Faliszewski, Niedermeier, and Talmon (2016a) studied shift bribery in the context of multiwinner elections for various committee selection rules.

While the complexity of shift bribery has been comprehensively investigated for many standard voting rules, it has not been considered yet for *iterative* voting systems. To close this glaring gap, we study shift bribery for eight iterative voting systems that are based on any one of the Borda, plurality, and veto rules (see Footnote 1 for their definitions) and that each proceed in rounds, eliminating after each except the last round the candidates performing worst in a certain sense:

- The *system of Baldwin (1926)* eliminates the candidates with the lowest Borda score and
- the *system of Nanson (1882)* eliminates the candidates whose scores are lower than the average Borda score, while
- the *system of Hare* (see, e.g., the book by Taylor, 2005) eliminates the candidates with the lowest plurality score,
- the system called *iterated plurality* (again see, e.g., the book by Taylor, 2005) eliminates the candidates that do not have the highest plurality score,
- the system called *iterated veto* is defined analogously to iterated plurality, except based on the veto rather than the plurality score, and
- the *system of Coombs* (defined, e.g., in the paper by Levin & Nalebuff, 1995) eliminates the candidates with the lowest veto score.

The last two systems that we consider differ from the above iterative voting systems because they always use exactly two rounds:

- *Plurality with runoff* (as defined, e.g., in the book by Taylor, 2005) eliminates the candidates that do not have the highest plurality score, except in the case where there is a unique plurality winner—it then eliminates all candidates that do not have the highest or second-highest plurality score; in the second round, all remaining candidates with the highest plurality score then win.

2. Faliszewski, Reisch, Rothe, and Schend (2015) have complemented these results on Bucklin and fallback voting. In particular, they studied a number of bribery problems for these rules, including a variant called "extension bribery," which was previously introduced by Baumeister, Faliszewski, Lang, and Rothe (2012) in the context of campaign management when the voters' ballots are truncated.

	Hare	Coombs	Baldwin	Nanson
Constructive	NP-c (Thm. 1)	NP-c (Thm. 3)	NP-c (Thm. 5)	NP-c (Thm. 7)
Destructive	NP-c (Thm. 2)	NP-c (Thm. 4)	NP-c (Thm. 6)	NP-c (Thm. 8)
	Iterated Plurality	Plurality with Runoff	Iterated Veto	Veto with Runoff
Constructive	NP-c (Thm. 9)	NP-c (Thm. 9)	NP-c (Thm. 11)	NP-c (Thm. 11)
Destructive	NP-c (Thm. 10)	NP-c (Thm. 10)	NP-c (Thm. 12)	NP-c (Thm. 12)

Table 1: Summary of complexity results for shift bribery problems

- *Veto with runoff* is defined analogously, except that veto scores instead of plurality scores and veto winners instead of plurality winners are considered.

These voting systems have been thoroughly studied and are also used in the real world. Among the systems we consider, Hare voting and variants thereof (some of which are called single transferable vote, instant-runoff voting, or alternative vote) are most widely used, for example in Australia, India, Ireland, New Zealand, Pakistan, the UK, and the USA.

Table 1 gives an overview of our complexity results for constructive and destructive shift bribery in our eight voting systems,[3] where the shorthand NP-c stands for "NP-complete." Our results complement results by Davies, Katsirelos, Narodytska, Walsh, and Xia (2014) who have shown unweighted coalitional manipulation to be NP-complete for Baldwin and Nanson voting (even with just a single manipulator)—and also for the underlying Borda system (with two manipulators; for the latter result, see also the paper by Betzler, Niedermeier, and Woeginger (2011)). Davies et al. (2014) also list various appealing features of the systems by Baldwin and Nanson, including that they have been applied in practice (namely, in the State of Michigan in the 1920s, in the University of Melbourne from 1926 through 1982, and in the University of Adelaide since 1968) and that (unlike Borda itself) they both are Condorcet-consistent.[4] Axiomatic properties of iterative voting systems were also studied by Freeman, Brill, and Conitzer (2014) who showed, in particular, that Hare is the only iterative voting system based on scoring rules that satisfies the independence-of-clones property. Further, it was shown by Bartholdi and Orlin (1991) that Hare (which is called STV in their work) is NP-hard to manipulate even with only one manipulator. This result was complemented by Davies, Narodytska, and Walsh (2012) who showed the same result for Coombs and a general class of iterative versions of scoring rules. For plurality with runoff, it was shown by Conitzer et al. (2007) that unweighted coalitional manipulation is NP-hard. Finally, plurality with runoff and veto with runoff were also studied by Erdélyi, Neveling, Reger, Rothe, Yang, and Zorn (2021) with respect to electoral control.

This paper is organized as follows. In Section 2, we will provide the needed definitions regarding elections and voting systems (in particular, iterative voting systems), define the shift bribery problem, and give some background on computational complexity. We will then study the complexity of shift bribery for Hare and Coombs elections in Section 3, for Baldwin and Nanson elections in Section 4, for iterated plurality and plurality with runoff in Section 5, and for iterated veto and veto with runoff in Section 6.

3. As shown by Xia (2012), destructive bribery is closely related to the *margin of victory*, a critical robustness measure for voting systems. Reisch et al. (2014) add to this connection by showing that the former problem can be easy while the latter is hard.

4. A *Condorcet winner* is a candidate who defeats every other candidate in a pairwise comparison. Such a candidate does not always exist. A voting rule is *Condorcet-consistent* if it chooses only the Condorcet winner whenever there exists one.

Further, in Section 7 we will discuss how the nonmonotonicity property of our iterative voting systems can be exploited in our reductions showing NP-hardness, exemplified for Hare voting and plurality with runoff. Finally, we will conclude in Section 8 by presenting some open problems related to our work.

2. Preliminaries

Below, we provide the needed notions and notation.

Elections and voting systems. An *election* is specified as a pair (C,V) with C being a set of candidates and V a profile of the voters' preferences over C, typically given by a list of linear orders of the candidates. A *voting system* is a function that maps each election (C,V) to a subset of C, the *winner(s) of the election*. An important class of voting systems is the family of positional scoring rules whose most prominent members are plurality, veto, and Borda count, see, e.g., the book chapters by Zwicker (2016) and Baumeister and Rothe (2015) and also the survey by Rothe (2019) on using Borda in collective decision making.

Recall from Footnote 1 in Section 1 that, in *plurality*, each voter gives her top-ranked candidate one point; in *veto* (a.k.a. *antiplurality*), each voter gives all except the bottom-ranked candidate one point; in *Borda* with m candidates, each candidate in position i of the voters' rankings scores $m-i$ points; and the winners in each case are those candidates scoring the most points.

Iterative voting systems. The iterative voting systems we will study are based on plurality, veto, and Borda but, unlike those, their election winner(s) are determined in consecutive rounds. For all iterative voting systems considered here except for plurality with runoff and veto with runoff (which will be defined shortly afterwards), if in some round all remaining candidates have the same score (for instance, there may be only one candidate left), then all those candidates are proclaimed winners of the election. In each preceding round, however, all candidates with the lowest score are eliminated.[5]

Recall from Section 1 that the eight scoring methods we will use work as follows: The iterative voting systems due to *Hare*, *Coombs*, and *Baldwin* use, respectively, plurality, veto, and Borda scores in order to decide which candidates are the weakest and thus to be removed. The *Nanson* system eliminates in every (except the last) round all candidates that have less than the average Borda score. *Iterated plurality* eliminates all candidates that do not have the highest plurality score, and *iterated veto* eliminates all candidates that do not have the highest veto score.

Unlike the above multiple-round iterative voting systems, *plurality with runoff* (respectively, *veto with runoff*) always proceeds in two rounds: In the first round, it eliminates all candidates that do not have the highest plurality score (respectively, veto score), unless there is a unique plurality winner (respectively, veto winner) in which case all candidates are eliminated except those with the highest or second-highest plurality score (respectively, veto score); in the second round, all candidates with the highest plurality score (respectively, veto score) win.

Shift bribery. For any given voting system $\mathscr{E}$, we now define the problem $\mathscr{E}$-SHIFT-BRIBERY, which is a special case of $\mathscr{E}$-SWAP-BRIBERY, introduced by Faliszewski et al. (2009b) in the context of so-called irrational voters for Copeland and then comprehensively studied by Elkind et al. (2009). Informally stated, given a profile of votes, a swap-bribery price function exacts a price for each swap of any two candidates in the votes, and in shift bribery only swaps involving the designated candidate are allowed.

5. In the original sources defining these iterative voting systems as stated in the Introduction, certain tie-breaking schemes are used whenever more than one candidate has the lowest score in some round. For the sake of convenience and uniformity, however, we prefer eliminating them all and will therefore disregard tie-breaking issues in such a case.

$\mathscr{E}$-CONSTRUCTIVE-SHIFT-BRIBERY	
Given:	An election (C,V) with n votes, a designated candidate $p \in C$, a budget B, and a list of price functions $\rho = (\rho_1, \dots, \rho_n)$.
Question:	Is it possible to make p the unique $\mathscr{E}$ winner of the election by shifting p in the votes such that the total price does not exceed B?

In the corresponding problem $\mathscr{E}$-DESTRUCTIVE-SHIFT-BRIBERY, given the same input, we ask whether it is possible to prevent p from being a unique winner.

These problems are here defined in the unique-winner model where a constructive (respectively, destructive) bribery action is considered successful only if the designated candidate can be made (respectively, can be prevented from being) the only winner of the election. We also consider these problems in the nonunique-winner model where for a constructive (respectively, destructive) bribery action to be considered successful it is required that the designated candidate is merely one among possibly several winners (respectively, does not win at all). Note that a yes-instance of $\mathscr{E}$-CONSTRUCTIVE-SHIFT-BRIBERY in the unique-winner model is also a yes-instance of the same problem in the nonunique-winner model, whereas a yes-instance of $\mathscr{E}$-DESTRUCTIVE-SHIFT-BRIBERY in the nonunique-winner model is also a yes-instance of the same problem in the unique-winner model; analogous statements apply to the no-instances of these problems by swapping the unique-winner model with the nonunique-winner model. We will make use of these facts in our proofs that all work in both winner models.

Membership in NP is obvious for all considered problems, so it will be enough to show only NP-hardness so as to prove in fact NP-completeness.

Regarding the list of price functions $\rho = (\rho_1, \dots, \rho_n)$ with $\rho_i : \mathbb{N} \to \mathbb{N}$, in the constructive case $\rho_i(k)$ indicates the price the briber has to pay in order to move p in vote i by k positions to the top (respectively, to the bottom in the destructive case). For all i, we require that ρ_i is nondecreasing ($\rho_i(\ell) \leq \rho_i(\ell+1)$), $\rho_i(0) = 0$, and if p is at position r in vote i then $\rho_i(\ell) = \rho_i(\ell-1)$ whenever $\ell \geq r$ in the constructive case (respectively, whenever $\ell \geq |C| - r + 1$ in the destructive case). The latter condition ensures that p can be shifted upward no farther than to the top (respectively, the bottom).[6] When the voter i in ρ_i is clear from the context, we omit the subscript and simply write ρ.

Our proofs use the following notation: A vote of the form $a\ b\ c$ indicates that the voter ranks candidate a on top position, then candidate b, and last candidate c. If a set $S \subseteq C$ of candidates appears in a vote as $\overrightarrow{S}$, its candidates are placed in this position in lexicographical order. By $\overleftarrow{S}$ we mean the reverse of the lexicographical order of the candidates in S. If S occurs in a vote without an arrow on top, the order in which the candidates from S are placed here does not matter for our argument. We use $\cdots$ in a vote to indicate that the remaining candidates may occur in any order.

Computational complexity. We assume familiarity with the standard concepts of complexity theory, including the classes P and NP, polynomial-time many-one reducibility, and NP-hardness and -completeness. We will use the following NP-complete problem:

EXACT-COVER-BY-3-SETS (X3C)	
Given:	A set $X = \{x_1, \dots, x_{3m}\}$ and a family of sets $\mathscr{S} = \{S_1, \dots, S_n\}$ such that $S_i \subseteq X$ and $\|S_i\| = 3$ for all $S_i \in \mathscr{S}$.
Question:	Does there exist an exact cover of X, i.e., a subset $\mathscr{S}' \subseteq \mathscr{S}$ such that $\|\mathscr{S}'\| = m$ and $\bigcup_{S_i \in \mathscr{S}'} S_i = X$?

6. If p is in the first (respectively, the last) position of a vote, this voter cannot be bribed and we tacitly assume a price function of $\rho(t) = 0$ for each $t \geq 0$. We will disregard these voters when setting price functions for the other voters in our proofs.

In instances of X3C, we assume that each $x_j \in X$ is contained in exactly three sets $S_i \in \mathscr{S}$; thus $|X| = |\mathscr{S}|$. Gonzalez (1985) shows that X3C under this restriction remains NP-hard. Note that if not stated otherwise, we will use $(X, \mathscr{S})$ to denote an X3C instance, where $X = \{x_1, \ldots, x_{3m}\}$, $\mathscr{S} = \{S_1, \ldots, S_{3m}\}$, and $S_i = \{x_{i,1}, x_{i,2}, x_{i,3}\}$. Also note that we assume $x_{i,1}$ to be the $x_j \in S_i$ with the smallest subscript and $x_{i,3}$ to be the $x_j \in S_i$ with the largest subscript.

ONE-IN-THREE-POSITIVE-3SAT	
Given:	A set X of boolean variables, a set S of clauses over X, each containing exactly three unnegated literals.
Question:	Does there exist a truth assignment to the variables in X such that exactly one literal is set to true for each clause in S?

In instances of ONE-IN-THREE-POSITIVE-3SAT, we assume that each $x_j \in X$ is contained in exactly three clauses. Porschen, Schmidt, Speckenmeyer, and Wotzlaw (2014) show that this restricted problems remains NP-complete.

For more background on computational complexity, the reader is referred to, for instance, the textbooks by Garey and Johnson (1979), Papadimitriou (1995), and Rothe (2005).

3. Hare and Coombs

We start by showing NP-hardness of shift bribery for Hare elections.

Theorem 1. *In both the unique-winner and the nonunique-winner model, Hare-*CONSTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof. NP-hardness follows by a reduction from X3C. Given an X3C instance $(X, \mathscr{S})$, construct an instance $((C,V), p, B, \rho)$ of Hare-CONSTRUCTIVE-SHIFT-BRIBERY with candidate set $C = X \cup \mathscr{S} \cup \{p\}$, designated candidate p, and the following list V of votes, with # denoting their number:

#	vote	for
1	$S_i \; x_{i,1} \; \overrightarrow{X \setminus \{x_{i,1}\}} \; \cdots$	$1 \leq i \leq 3m$
1	$S_i \; x_{i,2} \; \overrightarrow{X \setminus \{x_{i,2}\}} \; \cdots$	$1 \leq i \leq 3m$
1	$S_i \; x_{i,3} \; \overrightarrow{X \setminus \{x_{i,3}\}} \; \cdots$	$1 \leq i \leq 3m$
4	$x_i \; \overrightarrow{X \setminus \{x_i\}} \cdots$	$1 \leq i \leq 3m$
1	$S_i \; p \; \cdots$	$1 \leq i \leq 3m$
3	$p \; \cdots$	

For votes of the form $S_i \; p \; \cdots$, we use the price function $\rho(1) = 1$, and $\rho(t) = m+1$ for all $t \geq 2$; and for every other vote, we use the price function ρ with $\rho(t) = m+1$ for all $t \geq 1$. Finally, set the budget $B = m$. Without loss of generality, we assume that $m > 1$.

Note that p scores three points while the rest of the candidates score four points each, so p is eliminated in the first round and does not win the election without bribing voters.

We claim that $(X,\mathscr{S})$ is in X3C if and only if $((C,V),p,B,\rho)$ is in Hare-CONSTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

($\Rightarrow$) Suppose that $(X,\mathscr{S})$ is a yes-instance of X3C. Then there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of size m. We now show that it is possible for p to become a unique Hare winner of an election obtained by shifting p in the votes without exceeding the budget B. For every $S_i \in \mathscr{S}'$, we bribe the voter with the vote of the form $S_i\ p\ \cdots$ by shifting p once, so her new vote is of the form $p\ S_i\ \cdots$; each such bribe action costs us only 1 from our budget, so the budget will not be exceeded. In the first round, p now has $m+3$ points, every candidate from $\mathscr{S}'$ has 3 points, and every other candidate has 4 points. Therefore, all candidates in $\mathscr{S}'$ are eliminated. In the second round, all candidates in X now gain one point from the elimination of $\mathscr{S}'$, since it is an exact cover. Therefore, p and all candidates in X proceed to the next round and the remaining candidates $\mathscr{S} \setminus \mathscr{S}'$ are eliminated. In the next round with only p and the candidates from X remaining, p has $3m+3$ points, while every candidate in X scores 7 points (recall that every $x_i \in X$ is contained in exactly three members of $\mathscr{S}$). Since all candidates from X have been eliminated now, p is the only remaining candidate and thus the unique Hare winner.

($\Leftarrow$) Suppose that $(X,\mathscr{S})$ is a no-instance of X3C. Then no subset $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| \leq m$ covers X. We now show that we cannot make p become a Hare winner of an election obtained by bribing voters without exceeding budget B. Note that we can only bribe at most m voters with votes of the form $S_i\ p\ \cdots$ without exceeding the budget. Let $\mathscr{S}' \subseteq \mathscr{S}$ be such that $S_i \in \mathscr{S}'$ exactly if the voter with the vote $S_i\ p\ \cdots$ has been bribed. Clearly, $|\mathscr{S}'| \leq m$ and in all those votes p has been shifted once to the left, so p is now ranked first in these votes. Therefore, p now has $3+|\mathscr{S}'|$ points and every $S_i \in \mathscr{S}'$ scores 3 points. Since every other candidate scores as many points as before the bribery (namely, 4 points), the candidates in $\mathscr{S}'$ are eliminated in the first round. Let $X' = \{x_i \in X \mid x_i \notin \bigcup_{S_j \in \mathscr{S}'} S_j\}$ be the subset of candidates $x_i \in X$ that are not covered by $\mathscr{S}'$. We have $X' \neq \emptyset$ (otherwise, $\mathscr{S}'$ would have been an exact cover of X). In the second round, unlike the candidates from $X \setminus X'$, the candidates in X' will not gain additional points from eliminating the candidates in $\mathscr{S}'$. Thus, in the current situation, the candidates from X' and $\mathscr{S} \setminus \mathscr{S}'$ are trailing behind with 4 points each and are eliminated in this round.[7] Therefore, in the next round, only p and the candidates from $X \setminus X'$ are remaining in the election. Let $x_\ell \in X \setminus X'$ be the candidate from $X \setminus X'$ with the smallest subscript. Since all candidates from $\mathscr{S}$ are eliminated, p has $3m+3$ points and every candidate from $X \setminus X'$ except x_ℓ has 7 points. On the other hand, x_ℓ gains additional points from eliminating the candidates from X'; therefore, x_ℓ survives this round by scoring more than 7 points. In the final round with only p and x_ℓ remaining, p is eliminated, since $3m \cdot 7 > 3m+3$. $\square$

Example 1. *Let $(X,\mathscr{S})$ be a yes-instance of* X3C *defined by*

$$\begin{aligned} X &= \{x_1,\ldots,x_6\} \text{ and} \\ \mathscr{S} &= \{\{1,2,3\},\{4,5,6\},\{2,3,6\},\{2,4,5\},\{1,3,4\},\{1,5,6\}\}. \end{aligned}$$

Construct $((C,V),p,B,\rho)$ from $(X,\mathscr{S})$ as in the proof of Theorem 1; in particular, the budget is $B=2$. If we bribe the voters with $S_1\ p\ \cdots$ and $S_2\ p\ \cdots$ so as to shift p to the top of their votes, p will be the unique winner of the election that proceeds as follows (where the numbers in the columns below candidates give their scores):

7. Note that in the case that $|\mathscr{S}'| = 1$, i.e., only one voter was bribed, p also gets eliminated in this round and is consequently not a Hare winner, which is what we want to show. Therefore, we will now assume that at least two voters were bribed.

Round	p	$x \in X$	S_1, S_2	S_3, S_4, S_5, S_6
1	5	4	3	4
2	5	5	*out*	4
3	9	7	*out*	*out*

Now consider a no-instance $(X, \mathscr{S})$ *of* X3C *with*

$$\begin{aligned} X &= \{x_1, \ldots, x_6\} \text{ and} \\ \mathscr{S} &= \{\{1,2,4\}, \{4,5,6\}, \{2,3,6\}, \{2,3,5\}, \{1,3,4\}, \{1,5,6\}\}. \end{aligned}$$

If we bribe no voter, p gets eliminated in the first round and so does not win. If we bribe one voter, say the one with vote $S_1\ p\ \cdots$*, then p gets eliminated in the second round:*

Round	p	x_1	$x_2,\ x_4$	$x_3,\ x_5,\ x_6$	S_1	$S_i \in \mathscr{S} \setminus \{S_1\}$
1	4	4	4	4	3	4
2	4	5	5	4	*out*	4
3	*out*	≥ 28	≥ 7	*out*	*out*	*out*

Since $(X, \mathscr{S})$ *is a no-instance of* X3C*, no matter which two subsets* $S_i, S_j \in \mathscr{S}$ *we choose, at least one* x_k *is in both subsets, so p loses the direct comparison in the last round. For example, if we bribe the voters with* $S_1\ p\ \cdots$ *and* $S_2\ p\ \cdots$*, the election proceeds as follows:*

Round	p	x_1	x_3	x_4	x_2, x_5, x_6	$S_1,\ S_2$	S_3, S_4, S_5, S_6
1	5	4	4	4	4	3	4
2	5	5	4	6	5	*out*	4
3	9	14	*out*	7	7	*out*	*out*
4	9	42	*out*	*out*	*out*	*out*	*out*

This completes Example 1.

Next, we show that shift bribery is NP-hard for Hare also in the destructive case.

Theorem 2. *In both the unique-winner and the nonunique-winner model, Hare-*DESTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof. Again, we use a reduction from X3C. Construct from a given X3C instance $(X, \mathscr{S})$ a Hare-DESTRUCTIVE-SHIFT-BRIBERY instance $((C,V), p, B, \rho)$ as follows. Let $D = \{d_1, \ldots, d_{3m}\}$ be a set of $3m$ dummy candidates. The candidate set is $C = X \cup \mathscr{S} \cup D \cup \{p, w\}$ with designated candidate p. The list V of votes is constructed as follows:

#	vote	for
2	$S_i\ x_{i,1}\ \overrightarrow{X \setminus \{x_{i,1}\}}\ w\ p\ \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,2}\ \overrightarrow{X \setminus \{x_{i,2}\}}\ w\ p\ \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,3}\ \overrightarrow{X \setminus \{x_{i,3}\}}\ w\ p\ \cdots$	$1 \leq i \leq 3m$
7	$x_i\ \overrightarrow{X \setminus \{x_i\}}\ w\ p\ \cdots$	$1 \leq i \leq 3m$
1	$p\ S_i\ \cdots$	$1 \leq i \leq 3m$
12	$w\ p\ \cdots$	
$18m$	$p\ \cdots$	
6	$d_i\ S_i\ p\ \cdots$	$1 \leq i \leq 3m$

For votes of the form $p\ S_i\ \cdots$, we use the price function $\rho(1) = 1$, and $\rho(t) = m+1$ for all $t \geq 2$; and for every other vote, we use the price function ρ with $\rho(t) = m+1$ for all $t \geq 1$. Finally, set the budget $B = m$.

Without bribing, the election (C,V) proceeds as follows:

Round	p	w	$x_i \in X$	$S_i \in \mathscr{S}$	$d_i \in D$
1	$21m$	12	7	6	6
2	$39m$	12	13	out	out
3	$39m+12$	out	13	out	out

It follows that p has won the election after three rounds.

We claim that $(X,\mathscr{S})$ is in X3C if and only if $((C,V),p,B,\rho)$ is in Hare-DESTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

$(\Rightarrow)$ Suppose that $(X,\mathscr{S})$ is a yes-instance of X3C. Then there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of size m. We now show that it is possible to eliminate p from an election obtained by shifting p in the votes without exceeding the budget B. For every $S_i \in \mathscr{S}'$, we bribe the voter with the vote of the form $p\ S_i\ \cdots$ by shifting p once, so her new vote is of the form $S_i\ p\ \ldots$; each such bribe action costs us only 1 from our budget, so the budget will not be exceeded. Now the election proceeds as follows:

Round	p	w	$x_i \in X$	$S_i \in \mathscr{S}'$	$S_i \in \mathscr{S} \setminus \mathscr{S}'$	$d_i \in D$
1	$20m$	12	7	7	6	6
2	$32m$	12	11	13	out	out
3	$32m$	$33m+12$	out	13	out	out
4	$39m$	$39m+12$	out	out	out	out

We see that p is eliminated in the fourth round and w wins.

$(\Leftarrow)$ Suppose that $(X,\mathscr{S})$ is a no-instance of X3C. Then no subset $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| \leq m$ covers X. We now show that p will not be eliminated in any election obtained by bribing voters without exceeding budget B but will in fact become the only winner. Note that we can only bribe at most m voters with votes of the form $p\ S_i\ \cdots$ without exceeding the budget. Let $\mathscr{S}' \subseteq \mathscr{S}$ be such that for every $S_i \in \mathscr{S}'$ we have bribed the voter whose vote is $p\ S_i\ \cdots$. We can assume that $|\mathscr{S}'| > 0$. Every candidate in $\mathscr{S}'$ will gain an additional point and therefore survives the first round. All candidates from D and $\mathscr{S} \setminus \mathscr{S}'$ will be eliminated, since p only loses at most m points.

In the second round, the remaining candidates from $\mathscr{S}$ will additionally gain six points from the elimination of candidates in D and will score 13 points in this round (and in all subsequent rounds with p still standing). If a candidate $S_i \in \mathscr{S}$ was eliminated in the previous round, every $x_i \in S_i$ gains two additional points in this round. Partition X into sets X_0, X_1, X_2, and X_3 so that $x_i \in X_k \Leftrightarrow |\{S_j \in \mathscr{S}' \,|\, x_i \in S_j\}| = k$ for $k \in \{0,1,2,3\}$. Note that X_0, X_1, X_2, and X_3 are disjoint and $|X_0| > 0$, but one or two of X_1, X_2, and X_3 may be empty. Then $x_i \in X_j$ scores $7+(6-2j) \in \{7,9,11,13\}$ points depending on how many times x_i is covered by $\mathscr{S}'$. Therefore, every $x_i \in X_0$ scores more points than w who has 12 points. Thus there are candidates from X that survive this round and other candidates from X (more precisely, candidates from X_1, X_2, or X_3) who are eliminated.

In the third round, the candidate $x_\ell \in X$ with the smallest subscript who is still standing gains at least seven points from the eliminated candidates, so that x_ℓ scores at least 16 points.[8] All other candidates still score the same number of points as in the last round. Therefore, p scores at least $20m$ points, w scores still 12 points, every $S_i \in \mathscr{S}'$ scores 13 points, and every still standing candidate from X except x_ℓ scores at most 13 points. Since w can only gain additional points when all candidates from X are eliminated and only x_ℓ gains points from the elimination of candidates from $X \setminus \{x_\ell\}$ in the subsequent rounds, all candidates $X \setminus (\{x_\ell\} \cup X_0)$ and w are eliminated. Then all still standing candidates from $X_0 \setminus \{x_\ell\}$ and candidates from $\mathscr{S}'$ who each score 13 points are eliminated, which leaves p and x_ℓ in the last round. In this round, p scores $39m+12$ points and x_ℓ scores $39m$ points, so p solely wins the election, no matter how we bribe voters within the budget, i.e., we have a no-instance of Hare-DESTRUCTIVE-SHIFT-BRIBERY in both winner models. □

Next, we turn to shift bribery for Coombs elections. While the idea of the reduction is similar, and perhaps even simpler than in the previous two proofs, the proof of correctness is way more involved.

Theorem 3. *In both the unique-winner and the nonunique-winner model, Coombs-*CONSTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof. To prove NP-hardness, we now describe a reduction from X3C to Coombs-CONSTRUCTIVE-SHIFT-BRIBERY. Given an X3C instance $(X, \mathscr{S})$, construct an election (C,V) with the set $C = \{p, w, d_1, d_2, d_3\} \cup X \cup Y$ of candidates, where p is the designated candidate and $Y = \{y_i \,|\, x_i \in X\}$. Construct the following list V of votes:

#	vote	for
1	$\cdots\ x_{i,1}\ x_{i,2}\ x_{i,3}\ p$	$1 \le i \le 3m$
$2m$	$\cdots\ p\ \overrightarrow{Y \setminus \{y_i\}}\ y_i\ x_i$	$1 \le i \le 3m$
$2m$	$\cdots\ p\ \overrightarrow{Y}\ w\ d_1\ d_2\ d_3$	
1	$\cdots\ p\ \overrightarrow{Y}\ w\ X\ d_1\ d_2\ d_3$	
m	$\cdots\ p\ \overrightarrow{Y}\ w$	

For votes of the form $\cdots\ x_{i,1}\ x_{i,2}\ x_{i,3}\ p$, we use the price function $\rho(1) = \rho(2) = \rho(3) = 1$, and $\rho(t) = m+1$ for all $t \ge 4$; and for all the remaining votes, we use the price function $\rho(t) = m+1$ for all $t \ge 1$. Furthermore, our budget is $B = m$.

8. Since this candidate x_ℓ is still in the election, x_ℓ cannot have been in X_3 and thus must have had at least nine points.

The candidates have the following veto counts: p has $3m$ vetoes, each $x_i \in X$ has $2m$ vetoes, w has m vetoes, d_3 has $2m+1$ vetoes, and the remaining candidates each have 0 vetoes. Therefore, p will be eliminated in the first round and thus does not win the election.

We claim that $(X, \mathscr{S})$ is in X3C if and only if $((C,V), p, B, \rho)$ is in Coombs-CONSTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

($\Rightarrow$) Assume that $(X, \mathscr{S})$ is in X3C. This means that there exists a subset $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| = m$ and $\bigcup_{S_i \in \mathscr{S}'} S_i = X$. So we have a partition of X into three sets, $X = X_1 \cup X_2 \cup X_3$, such that:

$$
\begin{aligned}
X_1 &= \{x_i \in S_i \mid x_i \text{ has the lowest subscript in } S_i \in \mathscr{S}' \}, \\
X_3 &= \{x_i \in S_i \mid x_i \text{ has the highest subscript in } S_i \in \mathscr{S}' \}, \text{ and} \\
X_2 &= X \setminus (X_1 \cup X_3).
\end{aligned}
$$

Let $Y = Y_1 \cup Y_2 \cup Y_3$ be the corresponding partition of Y.

We bribe the voters with votes of the form $\cdots\ x_{i,1}\ x_{i,2}\ x_{i,3}\ p$ for $S_i \in \mathscr{S}'$ so that they change their votes to $\cdots\ p\ x_{i,1}\ x_{i,2}\ x_{i,3}$. Since $\mathscr{S}'$ is an exact cover of X, it follows that p now has a total of $2m$ vetoes, whereas each $x \in X_3$ receives an additional veto for a total of $2m+1$. The number of vetoes for the remaining candidates remain unchanged. If a candidate has the highest number of vetoes then she has the fewest number of points and cannot proceed to the next round (unless all candidates have the same score). Here, the candidates in X_3 and d_3 have the fewest number of points (and fewer than the other candidates) and therefore are eliminated in the first round.

Without the candidates in X_3, each candidate in X_2 gets an additional veto and the candidates in Y_3 each take all but one of the vetoes of the eliminated candidates in X_3. Furthermore, d_2 receives the vetoes of d_3. As a consequence, in the second round the candidates in X_2 and d_2 have the fewest number of points (and fewer than the remaining candidates) and are eliminated.

Similarly to the first round, vetoes from candidates in X_2 and d_2 are passed on to candidates in X_1 and Y_1 and to d_1. Thus the candidates have the following veto counts in the third round: p and each $y \in Y_2 \cup Y_3$ receive $2m$ vetoes, w receives m vetoes, each $y \in Y_1$ receives zero vetoes, and d_1 and each $x_i \in X_1$ receive $2m+1$ vetoes. Consequently, all the candidates $x_i \in X_1$ and d_1 are eliminated in the third round, so in the next round there are no candidates from X and no d_i, $1 \leq i \leq 3$.

It follows that w receives $2m+1$ additional vetoes in the fourth round, so w has the most vetoes in this round and is eliminated. We need $3m$ further rounds until p ends up as the last remaining candidate and sole winner of the election. In each of these rounds, the candidate in Y that is still alive and has the highest subscript has at least $2m+2m+1+m = 5m+1$ vetoes, while p always has only $3m$ vetoes.

($\Leftarrow$) Suppose that $(X, \mathscr{S})$ is a no-instance of X3C. We will show that $((C,V), p, B, \rho)$ then is a no-instance of Coombs-CONSTRUCTIVE-SHIFT-BRIBERY in the nonunique-winner (and thus also in the unique-winner) model. Observe that if we were going to make p a winner of the election, we would have to bribe at least m voters with a vote of the form $\cdots\ x_{i,1}\ x_{i,2}\ x_{i,3}\ p$; otherwise, p would have at least $2m+1$ vetoes and would be eliminated right away in the first round. Due to our budget, on the other hand, we can bribe no more than m (and thus would have to bribe exactly m) such voters and cannot bribe any further voters. Let $\mathscr{S}' \subseteq \mathscr{S}$ be such that $S_i \in \mathscr{S}'$ exactly if the voter with the vote of the form $\cdots\ x_{i,1}\ x_{i,2}\ x_{i,3}\ p$ has been bribed. Note that $|\mathscr{S}'| = m$ and $\mathscr{S}'$ does not cover X because we have a no-instance of X3C. Now p has only $2m$ vetoes and will not be eliminated in the first round.

Let X_1 be the set of candidates $x_i \in S_i$ for $S_i \in \mathscr{S}'$ with the smallest subscript in S_i, let X_2 be the set of candidates $x_i \in S_i$ for $S_i \in \mathscr{S}'$ with the second-smallest subscript in S_i, and let X_3 be the set of candidates $x_i \in S_i$ for $S_i \in \mathscr{S}'$ with the highest subscript in S_i. Note that $X_1 \cup X_2 \cup X_3 \neq X$, since $\mathscr{S}'$ does not cover X.

For w to have more vetoes than p, the candidates d_1, d_2, and d_3 need to be eliminated. For that to happen, there must be three rounds in which no other candidate has more than $2m+1$ vetoes. In the round where d_i, $1 \leq i \leq 3$, is eliminated, all still standing candidates in X_i are eliminated as well. Assume there were three rounds in which $2m+1$ was the maximal number of vetoes for a candidate. Then d_1, d_2, d_3, and all candidates in $X_1 \cup X_2 \cup X_3$ are eliminated. Note that those candidates that are not covered by $\mathscr{S}'$ always have only $2m$ vetoes and are still participating in the election. Therefore, in the next round, p and w have $3m$ vetoes each, the remaining candidates from X have at most $2m+1$ vetoes, and the candidates from Y have at most $2m$ vetoes. So even if p survives the first rounds with the candidates d_1, d_2, and d_3 still present, p will then surely be eliminated in the following round. If there is at least one voter who shifts p only one or two positions upward, then p has to drop out with d_1 or even before d_1 drops out, because at the latest after two rounds (with $2m+1$ being the maximal number of vetoes for a candidate) p receives another veto and thus has at least the same number of vetoes as d_1. $\square$

Example 2. *Let $(X, \mathscr{S})$ be a yes-instance of* X3C *defined by*

$$\begin{aligned} X &= \{x_1, \ldots, x_6\} \text{ and} \\ \mathscr{S} &= \{\{1,2,3\}, \{4,5,6\}, \{2,3,6\}, \{2,4,5\}, \{1,3,4\}, \{1,5,6\}\}. \end{aligned}$$

Construct $((C,V), p, B, \rho)$ from $(X, \mathscr{S})$ as in the proof of Theorem 3; in particular, the budget is $B = 2$. If we bribe the voters that correspond to the sets in the exact cover, S_1 and S_2, to change their votes from $\cdots\ x_1\ x_2\ x_3\ p$ and $\cdots\ x_4\ x_5\ x_6\ p$ to $\cdots p\ x_1\ x_2\ x_3$ and $\cdots p\ x_4\ x_5\ x_6$, then p alone wins the election that proceeds as follows, where in order to make this example easier to follow, the numbers in the table count the candidates' vetoes, not their points, i.e., the candidates with the highest *number in a round (row) get eliminated:*

Round	p	w	x_1,x_4	x_2,x_5	x_3,x_6	y_1	y_2	y_3	y_4	y_5	y_6	d_1	d_2	d_3
1	4	2	4	4	5	0	0	0	0	0	0	0	0	5
2	4	2	4	5	*out*	0	0	4	0	0	4	0	5	*out*
3	4	2	5	*out*	*out*	0	4	4	0	4	4	5	*out*	*out*
4	6	7	*out*	*out*	*out*	4	4	4	4	4	4	*out*	*out*	*out*
5	6	*out*	*out*	*out*	*out*	4	4	4	4	4	11	*out*	*out*	*out*
6	6	*out*	*out*	*out*	*out*	4	4	4	4	15	*out*	*out*	*out*	*out*
7	6	*out*	*out*	*out*	*out*	4	4	4	19	*out*	*out*	*out*	*out*	*out*
8	6	*out*	*out*	*out*	*out*	4	4	23	*out*	*out*	*out*	*out*	*out*	*out*
9	6	*out*	*out*	*out*	*out*	4	27	*out*	*out*	*out*	*out*	*out*	*out*	*out*
10	6	*out*	*out*	*out*	*out*	31	*out*	*out*	*out*	*out*	*out*	*out*	*out*	*out*

It follows that p is the sole winner of the election.
Now consider a no-instance $(X, \mathscr{S})$ with

$$\begin{aligned} X &= \{x_1, \ldots, x_6\} \text{ and} \\ \mathscr{S} &= \{\{1,2,4\}, \{4,5,6\}, \{2,3,6\}, \{2,3,5\}, \{1,3,4\}, \{1,5,6\}\}. \end{aligned}$$

Recall that we can bribe at most two voters. If we bribe fewer than two voters, however, p will be eliminated in the first round. Since $(X, \mathscr{S})$ is a no-instance of X3C*, no matter which two subsets $S_i, S_j \in \mathscr{S}$ we choose, at least one x_k is in both S_i and S_j. For example, if we bribe the voters that correspond to the sets S_1 and S_2, changing their votes from $\cdots\ x_1\ x_2\ x_4\ p$ and $\cdots\ x_4\ x_5\ x_6\ p$ to $\cdots p\ x_1\ x_2\ x_4$ and $\cdots p\ x_4\ x_5\ x_6$, then the election proceeds as follows:*

Round	p	w	x_1	x_2,x_5	x_3	x_4,x_6	y_1	y_2,y_5	y_3	y_4,y_6	d_1	d_2	d_3
1	4	2	4	4	4	5	0	0	0	0	0	0	5
2	4	2	4	5	4	*out*	0	0	0	4	0	5	*out*
3	5	2	5	*out*	4	*out*	0	4	0	4	5	*out*	*out*
4	*out*	2	*out*	*out*	4	*out*	4	4	0	4	*out*	*out*	*out*
...	...	...	...	...	...	...	...	...	...	...	...	...	...

Since x_4 is in both S_1 and S_2, p gets an additional veto in round 3 and is subsequently eliminated. The same will happen for similar reasons in every other case.

This completes Example 2.

We now modify the previous reduction so as to work for the destructive case in Coombs elections.

Theorem 4. *In both the unique-winner and the nonunique-winner model, Coombs-*DESTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof. To prove NP-hardness, we again reduce from the NP-complete problem X3C to Coombs-DESTRUCTIVE-SHIFT-BRIBERY. Given an X3C instance $(X, \mathscr{S})$ where we may assume that $m > 2$ for $|X| = 3m$, we construct a DESTRUCTIVE-SHIFT-BRIBERY instance $((C,V), p, B, \rho)$ as follows. Let $C = X \cup \mathscr{S} \cup D \cup \{p, w, y\}$ be the candidate set with designated candidate p and a set $D = \{d_{i,j} \mid 1 \leq i \leq m-1, 1 \leq j \leq 4\}$ of dummy candidates. Let $D = D_1 \cup D_2 \cup D_3 \cup D_4$ be a partition of D with $D_j = \{d_{i,j} \mid 1 \leq i \leq m-1\}$ for $1 \leq j \leq 4$. The list V of votes is then constructed as follows:

#	vote	for
1	$\cdots\ p\ S_i$	$1 \leq i \leq 3m$
$4m$	$p\ \cdots\ w\ x_{i,1}\ x_{i,2}\ x_{i,3}\ S_i$	$1 \leq i \leq 3m$
$4m+1$	$\cdots\ p\ X\ d_{i,1}\ d_{i,2}\ d_{i,3}\ d_{i,4}$	$1 \leq i \leq m-1$
1	$p\ \cdots\ y\ x_i$	$1 \leq i \leq 3m$
3	$\cdots\ p$	
2	$p\ \cdots\ w$	

Unlike in the previous proofs, it is here necessary that the candidates that are represented by "$\cdots$" are placed in lexicographical order. For votes of the form $\cdots\ p\ S_i$, we use the price function $\rho(1) = 1$, and $\rho(t) = 2m+1$ for all $t \geq 2$; and for all the remaining voters, we use the price function $\rho(t) = 2m+1$ for all $t \geq 1$. Finally, we set the budget $B = 2m$.

Analyzing the constructed election without bribing voters, the candidates have the following veto counts: p has three vetoes, w has two vetoes, each $x \in X$ has one veto, each $S_i \in \mathscr{S}$ and each $d \in D_4$ has $4m+1$ vetoes, and the remaining candidates each have zero vetoes. It follows that all candidates from $\mathscr{S}$ and D_4 are eliminated. The candidates from D_4 transfer their vetoes to candidates in D_3 who each have $4m+1$ vetoes now; p gets $3m$ additional vetoes from the eliminated candidates in $\mathscr{S}$; and the remaining $12m^2$ vetoes (from the second group of voters) are shared among candidates from X. Since they are ordered lexicographically in those votes, there must be one candidate from X (now and in the following rounds) that obtains more than $4m+1$ vetoes leading to the elimination of all candidates from X in the following rounds. In each of these following rounds, the candidate who receives some of those $12m^2$ vetoes from a previously eliminated candidate (starting with w) will now be eliminated, eventually leaving p as the last standing candidate and sole winner.

We claim that $(X,\mathscr{S})$ is in X3C if and only if $((C,V),p,B,\rho)$ is in Coombs-DESTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

($\Rightarrow$) Assume that $(X,\mathscr{S})$ is in X3C. This means that there exists a subset $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| = m$ and $\bigcup_{S_i \in \mathscr{S}'} S_i = X$. So we have a partition of X into three sets, $X = X_1 \cup X_2 \cup X_3$, such that:

$$
\begin{aligned}
X_1 &= \{x_i \in S_i \mid x_i \text{ has the lowest subscript in } S_i \in \mathscr{S}' \}, \\
X_3 &= \{x_i \in S_i \mid x_i \text{ has the highest subscript in } S_i \in \mathscr{S}' \}, \text{ and} \\
X_2 &= X \setminus (X_1 \cup X_3).
\end{aligned}
$$

We bribe the voters with a vote of the form $\cdots\ p\ S_i$ with $S_i \in \mathscr{S} \setminus \mathscr{S}'$ such that they change their vote to $\cdots\ S_i\ p$. Now the election proceeds as follows, where we again count the vetoes and not the points:

Round	p	w	y	$\mathscr{S}'$	$\mathscr{S} \setminus \mathscr{S}'$	X_1	X_2	X_3	D_1	D_2	D_3	D_4
1	$2m+3$	2	0	$4m+1$	$4m$	1	1	1	0	0	0	$4m+1$
2	$3m+3$	2	0	out	$4m$	1	1	$4m+1$	0	0	$4m+1$	out
3	$3m+3$	2	m	out	$4m$	1	$4m+1$	out	0	$4m+1$	out	out
4	$3m+3$	2	$2m$	out	$4m$	$4m+1$	out	out	$4m+1$	out	out	out
5	$4m^2+2$	$4m^2+2$	$3m$	out	$4m$	out	out	out	out	out	out	out

We see that p is eliminated in the fifth round, whereas y and some other candidates from $\mathscr{S} \setminus \mathscr{S}'$ are still in the election. Hence, p does not win.

($\Leftarrow$) Suppose that $(X,\mathscr{S})$ is a no-instance of X3C. Then no subset $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| \leq m$ covers X. We now show that p will not be eliminated in an election obtained by bribing voters without exceeding budget B but will in fact become the only winner. Note that we can only bribe at most $2m$ voters with votes of the form $\cdots\ p\ S_i$ without exceeding the budget. Let $\mathscr{S}' \subseteq \mathscr{S}$ be such that for every $S_i \in \mathscr{S} \setminus \mathscr{S}'$ we have bribed the voter whose vote was $\cdots\ p\ S_i$ and now is $\cdots\ S_i\ p$. We can assume that $|\mathscr{S} \setminus \mathscr{S}'| > 0$.

Every candidate in $\mathscr{S} \setminus \mathscr{S}'$ will gain an additional point and therefore survives the first round. All candidates in D_4 and $\mathscr{S}'$ will be eliminated in the first round. It follows that p has $3m+3$ vetoes in the second round. At this point, p is in each voter group other than the third voter group (with votes of the form $\cdots\ p\ X\ d_{i,1}\ d_{i,2}\ d_{i,3}\ d_{i,4}$) either the most (groups 2, 4, and 6) or the least preferred (groups 1 and 5) candidate; therefore, p does not receive any further vetoes before some candidate $d \in D_1$ is eliminated.

We note that $|\mathscr{S}'| \geq m$. Since $\mathscr{S}'$ is not an exact cover of X, we have at least one $x \in X$ which is in two sets $S, S' \in \mathscr{S}'$. Let $X' = \{x \in X \mid \exists\, S,S' \in \mathscr{S}',\ S \neq S',\ x \in S \cap S'\}$. After two further rounds in which $4m+1$ is the maximum number of vetoes, the candidates $d \in D \setminus D_1$ are eliminated. If each $x \in X'$ is still in the election, it follows that each $x \in X'$ has at least $4m+2$ vetoes such that some candidates $x \in X'$ will be eliminated. It follows that in the next round w receives at least $4m+2$ vetoes such that w has the most vetoes while the candidates $d \in D_1$ still have $4m+1$ vetoes. Otherwise, if at least one candidate $x \in X'$ is eliminated, it follows that w receives at least $4m+2$ vetoes at the latest in the fourth round, while each $d \in D_1$ still has $4m+1$ vetoes. After w is eliminated, in each following round the candidate x with the highest subscript and later the candidate S with the highest subscript and y will be eliminated. It follows that only p and the candidates $d \in D_1$ are still in the election. In each following round, p has at most $4m^2-4m+1$ vetoes while the still standing candidate $d \in D_1$ with the highest subscript receives at least $12m^2+7m+3$ vetoes. Hence, eventually p alone wins the election. $\square$

4. Baldwin and Nanson

We now show NP-hardness of shift bribery for Baldwin and Nanson elections. Note that our reductions are inspired by and similar to those used by Davies et al. (2014) to show NP-hardness of the unweighted coalitional manipulation problem for these voting systems.

For a preference profile V over a set of candidates C, let $avg(V)$ be the average Borda score of the candidates in V (i.e., $avg(V) = (|C|-1)|V|/2$). To conveniently construct votes, for a set of candidates C and $c_1, c_2 \in C$, let

$$W_{(c_1,c_2)} = (c_1\ c_2\ \overrightarrow{C \setminus \{c_1,c_2\}},\ \overleftarrow{C \setminus \{c_1,c_2\}}\ c_1\ c_2).$$

Under Borda, from the two votes in $W_{(c_1,c_2)}$ candidate c_1 scores $|C|$ points, c_2 scores $|C|-2$ points, and all other candidates score $|C|-1$ points. Also, observe that if a candidate $c^* \in C$ is eliminated in some round and $c^* \notin \{c_1,c_2\}$ then all other candidates lose one point due to the votes in $W_{(c_1,c_2)}$; if $c^* = c_1$ then c_2 loses no points but all other candidates lose one point; and if $c^* = c_2$ then c_1 loses two points and all other candidates lose one point. Therefore, if c^* is eliminated, the point difference caused by this elimination with respect to the votes in $W_{(c_1,c_2)}$ remains the same for all candidates, with two exceptions: (a) If $c^* = c_1$ then c_2 gains a point with respect to every other candidate, and (b) if $c^* = c_2$ then c_1 loses a point with respect to every other candidate. Furthermore, let $score_{(C,V)}(x)$ denote the number of points candidate x obtains in a Borda election (C,V), and let $dist_{(C,V)}(x,y) = score_{(C,V)}(x) - score_{(C,V)}(y)$.

We start with the complexity of shift bribery in Baldwin elections for the constructive case.

Theorem 5. *In both the unique-winner and the nonunique-winner model, Baldwin-*CONSTRUCTIVE-SHIFT-BRIBERY *is* NP-*hard.*

Proof. To prove NP-hardness, we reduce the NP-complete problem X3C to Baldwin-CONSTRUCTIVE-SHIFT-BRIBERY. From a given X3C instance $(X, \mathscr{S})$, we construct an election (C,V) with the set of candidates $C = \{p,w,d\} \cup X \cup \mathscr{S}$ and designated candidate p and with V consisting of two lists of votes, V_1 and V_2, where V_1 contains the following votes:

#	votes	for	#	votes	for
1	$W_{(S_j,p)}$	$1 \le j \le 3m$	2	$W_{(x_{j,3},S_j)}$	$1 \le j \le 3m$
2	$W_{(x_{j,1},S_j)}$	$1 \le j \le 3m$	2	$W_{(w,x_i)}$	$1 \le i \le 3m$
2	$W_{(x_{j,2},S_j)}$	$1 \le j \le 3m$	7	$W_{(w,p)}$	

The votes in V_1 give the following scores to the candidates in C:

$$\begin{aligned}
score_{(C,V_1)}(x_i) &= avg(V_1) + 4 \text{ for every } x_i \in X,\\
score_{(C,V_1)}(S_j) &= avg(V_1) - 5 \text{ for every } S_j \in \mathscr{S},\\
score_{(C,V_1)}(p) &= avg(V_1) - 3m - 7,\\
score_{(C,V_1)}(w) &= avg(V_1) + 6m + 7,\\
score_{(C,V_1)}(d) &= avg(V_1).
\end{aligned}$$

Furthermore, V_2 contains the following votes:

#	votes	for	#	votes
$2m+1$	$W_{(d,S_j)}$	$1 \le j \le 3m$	1	$W_{(p,d)}$
$2m+9$	$W_{(d,x_i)}$	$1 \le i \le 3m$	$2m+14$	$W_{(d,w)}$

The votes in V_2 give the following scores to the candidates in C:

$$\begin{aligned}
score_{(C,V_2)}(x_i) &= avg(V_2) - (2m+9) \text{ for every } x_i \in X,\\
score_{(C,V_2)}(S_j) &= avg(V_2) - (2m+1) \text{ for every } S_j \in \mathscr{S},\\
score_{(C,V_2)}(p) &= avg(V_2) + 1,\\
score_{(C,V_2)}(w) &= avg(V_2) - (2m+14),\\
score_{(C,V_2)}(d) &= avg(V_2) + 12m^2 + 32m + 13.
\end{aligned}$$

Let $V = V_1 \cup V_2$ and $avg(V) = avg(V_1) + avg(V_2)$. Then we have the following Borda scores for the complete preference profile V over C:

$$\begin{aligned}
score_{(C,V)}(x_i) &= avg(V) - 2m - 5 \text{ for every } x_i \in X,\\
score_{(C,V)}(S_j) &= avg(V) - 2m - 6 \text{ for every } S_j \in \mathscr{S},\\
score_{(C,V)}(p) &= avg(V) - 3m - 6,\\
score_{(C,V)}(w) &= avg(V) + 4m - 7,\\
score_{(C,V)}(d) &= avg(V) + 12m^2 + 32m + 13.
\end{aligned}$$

Regarding the price function, for every first vote of $W_{(S_j,p)}$ (i.e., a vote of the form $S_j\ p\ \overrightarrow{C \setminus \{S_j,p\}}$), let $\rho(1) = 1$ and $\rho(t) = m+1$ for every $t \geq 2$. For every other vote, let $\rho(t) = m+1$ for every $t \geq 1$. Finally, we set the budget $B = m$.

It is easy to see that p is eliminated in the first round in the election (C,V) and thus does not win.

We claim that $(X, \mathscr{S})$ is in X3C if and only if $((C,V), p, B, \rho)$ is in Baldwin-CONSTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

($\Rightarrow$) Suppose there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$. Then we bribe the first votes of $W_{(S_j,p)}$ for every $S_j \in \mathscr{S}'$ by shifting p to the left once. Note that we won't exceed our budget, since shifting once costs 1 in those votes and $|\mathscr{S}'| = m$. After this bribery, for every $S_j \in \mathscr{S}'$, the two votes from $W_{(S_j,p)}$ result in two votes that are symmetric to each other (i.e., $p\ S_j\ \overrightarrow{C \setminus \{S_j,p\}}$ equals the vote $\overleftarrow{C \setminus \{S_j,p\}}\ S_j\ p$ in reverse order) and can thus be disregarded from now on, as all candidates gain the same number of points from those votes and all candidates lose the same number of points if a candidate is eliminated from the election. After those m votes have been bribed, only the scores of p and every $S_j \in \mathscr{S}'$ change. With $score_{(C,V)}(p) = avg(V) - 2m - 6$ and $score_{(C,V)}(S_j) = avg(V) - 2m - 7$, all candidates in $\mathscr{S}'$ are tied for the last place. If any $S_j \in \mathscr{S}'$ is eliminated in a round, the three candidates $x_{j,1}$, $x_{j,2}$, and $x_{j,3}$ will lose two points more than the candidates from $\mathscr{S}' \setminus \{S_j\}$ that were in the last position before S_j was eliminated. Therefore, those three candidates from X will then be in the last position in the next round. This means that all candidates $\mathscr{S}'$ and every $x_i \in X$ that is covered by $\mathscr{S}'$ will be eliminated in the subsequent rounds. Since $\mathscr{S}'$ is an exact cover, now there is no candidate from X left. Thus the point difference between p and w is 1 and between p and the remaining $S_j \in (\mathscr{S} \setminus \mathscr{S}')$ is -6. Note that p can beat d only if no candidate of $C \setminus \{p,d\}$ is still participating. So in the next round, w is eliminated. From this p gains seven points on all $S_j \in (\mathscr{S} \setminus \mathscr{S}')$, so these are tied for the last place. Therefore, the remaining candidates from $\mathscr{S}$ are eliminated, which leaves p and d for the next and final round, where d is eliminated and p wins the election alone.

($\Leftarrow$) Suppose there is no exact cover. It is obvious that at most m of the first votes of $W_{(S_j,p)}$ can be bribed without exceeding the budget. Without bribing, p is in the last place and the point difference to the second-to-last candidate(s) is $dist_{(C,V)}(p,S_j) = m$, $1 \leq j \leq 3m$. By bribing, as explained above, p gains $m+1$ points on m candidates from $\mathscr{S}$, which then will be eliminated from the election. This leads to the elimination of all $x_i \in X$ that are covered by the set $\mathscr{S}' \subseteq \mathscr{S}$ of candidates that were eliminated. Since there is no exact cover, $\mathscr{S}'$ doesn't cover X. So there are candidates $X' \subseteq X$, $|X'| \geq 1$, who were not eliminated before, as for every candidate $x_i \in X'$ all three candidates $S_j \in (\mathscr{S} \setminus \mathscr{S}')$ with $x_i \in S_j$ are still in the election. With the candidates $C_1 = \{p,w,d\} \cup (\mathscr{S} \setminus \mathscr{S}') \cup X'$ still standing, the point differences of p to the other remaining candidates are as follows:

$$\begin{aligned}
dist_{(C_1,V)}(p,d) &= -2m-5-2m(2m+1)-|X'|(2m+9)-(2m+14) < 0,\\
dist_{(C_1,V)}(p,w) &= 1-2|X'| < 0,\\
dist_{(C_1,V)}(p,x_i) &= -1 \text{ for every } x_i \in X', \text{ and}\\
dist_{(C_1,V)}(p,S_j) &\leq 0 \text{ for every } S_j \in \mathscr{S} \setminus \mathscr{S}'.
\end{aligned}$$

Therefore, p is in the last place and is eliminated and thus does not win. □

The proof of the following theorem, which handles the destructive variant for Baldwin, uses a similar idea as the proof of Theorem 5. That is why we refrain from presenting all proof details in full; a proof sketch will suffice.

Theorem 6. *In both the unique-winner and the nonunique-winner model, Baldwin-*DESTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof Sketch. To prove NP-hardness, we reduce the NP-complete problem X3C to Baldwin-DESTRUCTIVE-SHIFT-BRIBERY. From a given X3C instance $(X,\mathscr{S})$, we construct an election (C,V), where $C = \{p,w,b,d\} \cup X \cup \mathscr{S}$ is the set of candidates, p is the designated candidate, and V consists of two lists of votes, V_1 and V_2, where V_1 contains the following votes:

#	votes	for	#	votes	for
1	$W_{(p,S_j)}$	$1 \leq j \leq 3m$	2	$W_{(w,x_i)}$	$1 \leq i \leq 3m$
2	$W_{(S_j,x_{j,1})}$	$1 \leq j \leq 3m$	$3m+7$	$W_{(w,d)}$	
2	$W_{(S_j,x_{j,2})}$	$1 \leq j \leq 3m$	$m+10$	$W_{(b,S_j)}$	$1 \leq j \leq 3m$
2	$W_{(S_j,x_{j,3})}$	$1 \leq j \leq 3m$			

Furthermore, V_2 contains the following votes:

#	votes	for	#	votes
1	$W_{(d,p)}$		$6m+14$	$W_{(p,w)}$
$2m+7$	$W_{(p,S_j)}$	$1 \leq j \leq 3m$	$3m^2+33m+12$	$W_{(p,b)}$
$3m+3$	$W_{(p,x_i)}$	$1 \leq i \leq 3m$		

Let $V = V_1 \cup V_2$. Then we have the following Borda scores for the complete profile V:

$$\begin{aligned}
score_{(C,V)}(x_i) &= avg(V) - 3m - 11 \text{ for every } x_i \in X,\\
score_{(C,V)}(S_j) &= avg(V) - 3m - 12 \text{ for every } S_j \in \mathscr{S},\\
score_{(C,V)}(d) &= avg(V) - 3m - 6,\\
score_{(C,V)}(w) &= avg(V) + 3m - 7,\\
score_{(C,V)}(b) &= avg(V) - 3m - 12,\\
score_{(C,V)}(p) &= avg(V) + 18m^2 + 72m + 25.
\end{aligned}$$

Regarding the price function, for every first vote of $W_{(p,S_j)}$ (i.e., a vote of the form $p\ S_j\ C \setminus \{S_j, p\}$), let $\rho(1) = 1$ and $\rho(t) = m+1$ for every $t \geq 2$. For every other vote, let $\rho(t) = m+1$ for every $t \geq 1$. Finally, we set the budget $B = m$.

It is easy to see that p wins the election (C,V).

We claim that $(X, \mathscr{S})$ is in X3C if and only if $((C,V), p, B, \rho)$ is in Baldwin-DESTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

$(\Rightarrow)$ Suppose there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$. Then we bribe the first votes of $W_{(p,S_j)}$ for every $S_j \in \mathscr{S}'$ by shifting p to the right once. With a similar argument as in the proof of Theorem 5, d alone wins the election, i.e., p is not among the winners.

$(\Leftarrow)$ Suppose there is no exact cover. Then, for every $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| \leq m$, there is at least one $x_i \in X$ that is not covered by $\mathscr{S}'$. It is obvious that at most m of the first votes of $W_{(p,S_j)}$ can be bribed without exceeding the budget. We can then show, similarly as in the proof of Theorem 5, that d will always be eliminated before w and therefore p cannot be prevented from winning the election alone. □

Finally, we turn to Nanson elections for which we again will show that shift bribery is NP-hard. The reduction below will only use pairs of votes of the form $W_{(c_1,c_2)}$. The average Borda score for those two votes is $|C|-1$. The candidate c_1 scores one point more than the average Borda score and c_2 scores one point fewer than the average Borda score. The other candidates score exactly the average Borda score. If a candidate is eliminated in a round, the average Borda score required to survive the next round decreases by one. Regardless of which candidate is eliminated, all remaining candidates that are not c_1 or c_2 lose one point and still have exactly the average Borda score. If c_2 is eliminated, c_1 loses its advantage with respect to the average Borda score and now scores exactly the average Borda score as well. If one of the other candidates is eliminated, c_1 continues to have one point more than the average Borda score. By symmetry, this holds analogously for c_2: If c_1 is eliminated, c_2 scores the average Borda score, and if one of the other candidates is eliminated, c_2 still has one point fewer than the average Borda score.

Theorem 7. *In both the unique-winner and the nonunique-winner model, Nanson-*CONSTRUCTIVE-SHIFT-BRIBERY *is* NP-*hard.*

Proof. To prove NP-hardness, we reduce the NP-complete problem X3C to Nanson-CONSTRUCTIVE-SHIFT-BRIBERY. Again, starting from a given X3C instance $(X, \mathscr{S})$, we construct an election (C,V) with the set of candidates $C = \{p, w_1, w_2, d\} \cup X \cup \mathscr{S}$, where p is the designated candidate. Then we construct two sets of votes, V_1 and V_2, where V_1 contains the following votes:

#	votes	for	#	votes	for
1	$W_{(S_j,p)}$	$1 \leq j \leq 3m$	1	$W_{(x_{j,3},S_j)}$	$1 \leq j \leq 3m$
1	$W_{(x_i,p)}$	$1 \leq i \leq 3m$	4	$W_{(S_j,w_1)}$	$1 \leq j \leq 3m$
1	$W_{(x_{j,1},S_j)}$	$1 \leq j \leq 3m$	$15m$	$W_{(w_1,w_2)}$	
1	$W_{(x_{j,2},S_j)}$	$1 \leq j \leq 3m$	$3m$	$W_{(p,w_1)}$	

Furthermore, V_2 contains the following votes:

#	votes	for
$2m$	$W_{(p,d)}$	
2	$W_{(d,S_j)}$	$1 \leq j \leq 3m$
4	$W_{(d,x_i)}$	$1 \leq i \leq 3m$

Let $V = V_1 \cup V_2$. Then we have the following Borda scores for the complete profile V:

$$\begin{aligned}
score_{(C,V)}(x_i) &= avg(V) \text{ for every } x_i \in X,\\
score_{(C,V)}(S_j) &= avg(V) \text{ for every } S_j \in \mathscr{S},\\
score_{(C,V)}(p) &= avg(V) - m,\\
score_{(C,V)}(w_1) &= avg(V),\\
score_{(C,V)}(w_2) &= avg(V) - 15m,\\
score_{(C,V)}(d) &= avg(V) + 16m.
\end{aligned}$$

The price function is again defined as follows. For every first vote of $W_{(S_j,p)}$ (i.e., a vote of the form $S_j \; p \; C \setminus \{S_j, p\}$), let $\rho(1) = 1$ and $\rho(t) = m+1$ for every $t \geq 2$. For every other vote, let $\rho(t) = m+1$ for every $t \geq 1$. Finally, we set the budget $B = m$.

It is easy to see that p is eliminated in the first round of the election (C,V) and so does not win.

We claim that $(X, \mathscr{S})$ is in X3C if and only if $((C,V), p, B, \rho)$ is in Nanson-CONSTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

($\Rightarrow$) Suppose there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$. Then, for every $S_j \in \mathscr{S}'$, we bribe the first vote of $W_{(S_j,p)}$ by shifting p to the left once in all those votes. Note that we won't exceed our budget, since this bribe action costs 1 per vote and $|\mathscr{S}'| = m$. With the additional m points, p reaches the average Borda score and is not eliminated in the first round. However, all candidates in $\mathscr{S}'$ lose one point and are eliminated. Additionally, w_2 will be eliminated as well.

In the next round, w_1 will be eliminated, since she has $11m$ points fewer than the average Borda score required to survive this round. Since the candidates in $\mathscr{S}'$ were eliminated in the last round and $\mathscr{S}'$ is an exact cover, every candidate in X now has fewer points than the average Borda score and is eliminated.

In the third round, only p, d, and the candidates in $\mathscr{S} \setminus \mathscr{S}'$ are still standing. Therefore, the only pairs of votes that are not symmetric are $W_{(S_j,p)}$, twice $W_{(d,S_j)}$ for every $S_j \in (\mathscr{S} \setminus \mathscr{S}')$, and $2m$ pairs of $W_{(p,d)}$. Since $|\mathscr{S} \setminus \mathscr{S}'| = 2m$, we have that p scores exactly the average Borda score and survives this round, just as d. Every $S_j \in (\mathscr{S} \setminus \mathscr{S}')$ has one point fewer than the average Borda score and is eliminated. This leaves only p and d in the last round, which p alone wins.

($\Leftarrow$) Suppose there is no exact cover. Then, for every $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| = m$, there is at least one $x_i \in X$ that is not covered by $\mathscr{S}'$. Note that we can only bribe the first votes of any $W_{(S_j,p)}$ without exceeding the budget. For p to survive the first round, we need to bribe m of those votes by shifting p to the left once. Let $\mathscr{S}' \subseteq \mathscr{S}$ be such that $\mathscr{S}'$ contains S_j exactly if the first vote of $W_{(S_j,p)}$ has been bribed. Then every $S_j \in \mathscr{S}'$ has a score of $avg(V) - 1$ and p has a score of $avg(V)$. Therefore, in the first round, every candidate from $\mathscr{S}'$ and w_2 are eliminated from the election.

In the second round, w_1 will be eliminated because of the $15m$ pairs of votes $W_{(w_1,w_2)}$ and the elimination of w_2. Furthermore, a candidate $x_i \in X$ reaches the average Borda score with p and d still standing only if all three $S_j \in \mathscr{S}$ with $x_i \in S_j$ are also not yet eliminated. Since the candidates in $\mathscr{S}'$ were eliminated in the previous round, for every $S_j \in \mathscr{S}'$, all three $x_i \in S_j$ will be eliminated in this round. Since $\mathscr{S}'$ is not an exact cover, there are candidates $X' \subseteq X$ that survive this round. d also reaches the average Borda score, as there are $2m$ candidates $\mathscr{S} \setminus \mathscr{S}'$ and those candidates $\mathscr{S} \setminus \mathscr{S}'$ survive due to w_1.

In the next round, the candidates still standing are p, d, X', and $\mathscr{S} \setminus \mathscr{S}'$. Because $|X'| \geq 1$, candidate p has $|X'|$ points fewer than the average Borda score and is eliminated in this round. Thus p does not win. □

Our last result in this section shows that the destructive variant of shift bribery in Nanson elections is intractable as well.

Theorem 8. *In both the unique-winner and the nonunique-winner model, Nanson-*DESTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof. To prove NP-hardness, we reduce the NP-complete problem X3C to Nanson-DESTRUCTIVE-SHIFT-BRIBERY. Once more, given an X3C instance $(X,\mathscr{S})$, we construct an election (C,V) with the set of candidates $C=\{p,w_1,w_2,w_3,d\}\cup X\cup\mathscr{S}$, where p is the designated candidate and $(X,\mathscr{S})$ is the given X3C instance. Then we construct two sets of votes, V_1 and V_2, where V_1 contains the following votes:

#	votes	for	#	votes	for
1	$W_{(p,S_j)}$	$1\le j\le 3m$	6	$W_{(S_j,w_3)}$	$1\le j\le 3m$
1	$W_{(d,x_i)}$	$1\le i\le 3m$	$20m$	$W_{(w_1,w_2)}$	
2	$W_{(x_{j,1},S_j)}$	$1\le j\le 3m$	$19m$	$W_{(w_3,w_1)}$	
2	$W_{(x_{j,2},S_j)}$	$1\le j\le 3m$	$3m+1$	$W_{(w_3,d)}$	
2	$W_{(x_{j,3},S_j)}$	$1\le j\le 3m$			

Furthermore, V_2 contains the following votes:

#	votes	for	#	votes
1	$W_{(d,p)}$		$3m+1$	$W_{(p,w_3)}$
1	$W_{(p,x_i)}$	$1\le i\le 3m$		

Let $V=V_1\cup V_2$. Then we have the following Borda scores for the complete profile V:

$$\begin{aligned}
score_{(C,V)}(x_i) &= avg(V)+4 \text{ for every } x_i\in X,\\
score_{(C,V)}(S_j) &= avg(V)-1 \text{ for every } S_j\in\mathscr{S},\\
score_{(C,V)}(d) &= avg(V),\\
score_{(C,V)}(w_1) &= avg(V)+m,\\
score_{(C,V)}(w_2) &= avg(V)-20m,\\
score_{(C,V)}(w_3) &= avg(V)+m,\\
score_{(C,V)}(p) &= avg(V)+9m.
\end{aligned}$$

The price function is again defined as follows. For every first vote of $W_{(p,S_j)}$ (i.e., a vote of the form $p\ S_j\ C\setminus\{S_j,p\}$), let $\rho(1)=1$ and $\rho(t)=m+1$ for every $t\ge 2$. For every other vote, let $\rho(t)=m+1$ for every $t\ge 1$. Finally, we set the budget $B=m$.

It is easy to see that p will only have fewer points than the average Borda score if all candidates from $\mathscr{S}$, X, and the candidate w_3 are eliminated while d is still standing. Without bribing, d is eliminated in the third round while w_3 is still standing, and eventually p wins the election (C,V).

We claim that $(X,\mathscr{S})$ is in X3C if and only if $((C,V),p,B,\rho)$ is in Nanson-DESTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model.

($\Rightarrow$) Suppose there is an exact cover $\mathscr{S}'\subseteq\mathscr{S}$. Then, for every $S_j\in\mathscr{S}'$, we bribe the first vote of $W_{(p,S_j)}$ by shifting p to the right once in all those votes. Note that we won't exceed our budget, since this bribe action costs 1 per vote and $|S'| = m$. After those m votes have been bribed, every $S_j \in \mathscr{S}'$ gains a point and therefore survives the first round. All other candidates $\mathscr{S} \setminus \mathscr{S}'$ and w_2 are eliminated.

Let $C_1 = \{p, d, w_1, w_3\} \cup X \cup \mathscr{S}'$ be the set of candidates present in the second round. w_1 loses $19m$ points on the average Borda score from the elimination of w_2 and is eliminated. Additionally, all candidates of X lose four points on the average Borda score but still survive this round, as they now have exactly the average Borda score.

Let $C_2 = \{p, d, w_3\} \cup X \cup \mathscr{S}'$ be the candidates in the third round. In this round, only w_3 is eliminated because w_3 lost $19m$ points on the average Borda score from the elimination of w_1.

Let $C_3 = \{p, d\} \cup X \cup \mathscr{S}'$ be the candidates in the fourth round. The scores are as follows:

$$\begin{aligned}
score_{(C_3,V)}(x_i) &= avg(V) \text{ for every } x_i \in X, \\
score_{(C_3,V)}(S_j) &= avg(V) - 6 \text{ for every } S_j \in \mathscr{S}', \\
score_{(C_3,V)}(d) &= avg(V) + 3m + 1, \\
score_{(C_3,V)}(p) &= avg(V) + 3m - 1.
\end{aligned}$$

Therefore all candidates in $\mathscr{S}'$ are eliminated. In the following round, all candidates in X are eliminated. This leaves only p and d in the final round in which p is eliminated and thus cannot win.

($\Leftarrow$) Suppose there is no exact cover. Then, for every $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| \leq m$, there is at least one $x_i \in X$ that is not covered by $\mathscr{S}'$. Note that we can only bribe the first votes of any $W_{(p,S_j)}$ without exceeding the budget.

We now show that, even with optimal bribing, d will be eliminated in the third round and, therefore, p alone wins the election. Within our budget, we can prevent at most m candidates from $\mathscr{S}$, say $\mathscr{S}'$, of being eliminated in the first round by bribing the corresponding vote of $W_{(p,S_j)}$. Since $\mathscr{S}'$ cannot be an exact cover of X, there is at least one $x_i \in X$ for which all $S_j \in \mathscr{S}$ with $x_i \in S_j$ are eliminated. This x_i is eliminated in the second round, as it has lost six points on the average Borda score from the eliminations of candidates in the previous round. In the third round, w_3 is still participating since w_2 and w_1 were only eliminated in the first and second round, respectively. Therefore, the score of d minus the average Borda score of this round is at most -1, which means that d is eliminated in this round. Thus, there is no candidate left that can prevent p from winning the election. $\square$

5. Iterated Plurality and Plurality with Runoff

In this section, we show hardness of shift bribery for iterated plurality and plurality with runoff, handling both voting systems simultaneously and starting with the constructive case.

Theorem 9. *In both the unique-winner and the nonunique-winner model, for iterated plurality and plurality with runoff,* CONSTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof. To prove NP-hardness, we reduce X3C to CONSTRUCTIVE-SHIFT-BRIBERY for these two voting systems. Let $(X, \mathscr{S})$ be a given X3C instance. We construct the CONSTRUCTIVE-SHIFT-BRIBERY instance $((C,V), p, B, \rho)$ as follows. Let $C = \{p, w\} \cup X \cup \mathscr{S} \cup D$ be the set of candidates, where p is the designated candidate and $D = \{d_{i,j} \mid 1 \leq i \leq 3m \text{ and } 1 \leq j \leq m-7\}$ is a set of dummy candidates. The list V of votes is constructed as follows:

#	vote	for
1	$S_i\ p\ \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,1}\ \overrightarrow{X \setminus \{x_{i,1}\}}\ \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,2}\ \overrightarrow{X \setminus \{x_{i,2}\}}\ \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,3}\ \overrightarrow{X \setminus \{x_{i,3}\}}\ \cdots$	$1 \leq i \leq 3m$
1	$S_i\ d_{i,j}\ \overrightarrow{X \setminus \{x_i\}}\ \cdots$	$1 \leq i \leq 3m$, $1 \leq j \leq m-7$
m	$x_i\ \overrightarrow{X \setminus \{x_i\}}\ \cdots$	$1 \leq i \leq 3m$
m	$d_{i,j}\ \overrightarrow{X}\ \cdots$	$1 \leq i \leq 3m$, $1 \leq j \leq m-7$
3	$w\ p\ \cdots$	

For voters with votes of the form $S_i\ p\ \cdots$, we use the price function $\rho(1) = 1$, and $\rho(t) = m+1$ for all $t \geq 2$; and for every other voter, we use the price function $\rho(t) = m+1$ for $t \geq 1$. Finally, set the budget $B = m$.

Without bribing, p has a score of zero and is eliminated immediately in both voting systems.

We claim that $(X, \mathscr{S})$ is in X3C if and only if $((C,V), p, B, \rho)$ is in CONSTRUCTIVE-SHIFT-BRIBERY for either of the two voting systems, regardless of the winner model.

($\Rightarrow$) Suppose that $(X, \mathscr{S})$ is a yes-instance of X3C. Then there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of size m. We now show that it is possible for p to become a unique iterated-plurality (respectively, plurality-with-runoff) winner of an election obtained by shifting p in the votes without exceeding the budget. For every $S_i \in \mathscr{S}'$, we bribe the voter with the vote of the form $S_i\ p\ \cdots$, so her new vote is of the form $p\ S_i\ \cdots$. In the first round p, every $x_i \in X$, every $d_{i,j} \in D$, and every $S_i \in \mathscr{S} \setminus \mathscr{S}'$ is a plurality winner, so only these candidates participate in the next round. In the second round, p receives three further points from the three voters whose vote is $w\ p\ \cdots$. Every candidate $x_j \in X$ receives two further points from the votes of the form $S_i\ x_j\ \cdots$ with $x_j \in S_i$ and $S_i \in \mathscr{S}'$. Every $d_{i,j}$ with $S_i \in \mathscr{S}'$ and $1 \leq j \leq m-7$ receives one additional point from the voters with vote $S_i\ d_{i,j}\ \cdots$. It follows that p has the most points and therefore p is the unique iterated-plurality (respectively, plurality-with-runoff) winner.

($\Leftarrow$) Suppose that $(X, \mathscr{S})$ is a no-instance of X3C. Then, for every $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| = m$, there is at least one candidate in X that is not covered and, therefore, at least one candidate in X occurring in at least two sets from $\mathscr{S}'$. We show that it is not possible for p to become a winner of the election obtained from the original election by bribing without exceeding the budget.

To become a winner of such a bribed election, it is necessary for p to get at least m points in the first round. Due to the budget, it is also necessary to bribe m voters with a vote of the form $S_i\ p\ \cdots$ with $S_i \in \mathscr{S}'$. It follows that p, each $x \in X$, each $S_i \in \mathscr{S} \setminus \mathscr{S}'$, and each $d_{i,j} \in D$ participate in the second round. As mentioned above, at least one candidate in X receives at least four further points due to the fact that $\mathscr{S}'$ is not a cover of X. Thus p does not win. That means that $((C,V), p, B, \rho)$ is a no-instance of CONSTRUCTIVE-SHIFT-BRIBERY for either of iterated plurality and plurality with runoff regardless of the winner model. $\square$

We have the same result in the destructive case. This is the first proof where we use an NP-complete problem other than X3C to show NP-hardness, namely ONE-IN-THREE-POSITIVE-3SAT, which was also defined in Section 2.

Theorem 10. *In both the unique-winner and the nonunique-winner model, for iterated plurality and plurality with runoff,* DESTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard.*

Proof. To prove NP-hardness, we reduce the NP-complete problem ONE-IN-THREE-POSITIVE-3SAT to DESTRUCTIVE-SHIFT-BRIBERY for both voting systems. Let (X,S) be a given ONE-IN-THREE-POSITIVE-3SAT instance, where $X = \{x_1,\ldots,x_{3m}\}$ and $S = \{S_1,\ldots S_{3m}\}$ with $S_i = \{x_{i,1},x_{i,2},x_{i,3}\} \subseteq X$ for each $1 \leq i \leq 3m$. Without loss of generality, we can assume that $m > 6$. We construct the DESTRUCTIVE-SHIFT-BRIBERY instance for both voting systems as follows. Let $C = \{p,w,e,f\} \cup D \cup Y \cup X$ with $D = \{d_{i,j} \mid 1 \leq i \leq 3m$ and $1 \leq j \leq 2m-1\}$ and $Y = \{y_{i,j} \mid 1 \leq i \leq 3m$ and $1 \leq j \leq 4\}$ and where p is the designated candidate. The list V of votes is constructed as follows:

#	votes	for
1	$p\ x_i\ \cdots$	$1 \leq i \leq 3m$
1	$y_{i,1}\ x_{i,1}\ x_{i,2}\ w\ p\ \cdots$	$1 \leq i \leq 3m$
1	$y_{i,2}\ x_{i,2}\ x_{i,3}\ w\ p\ \cdots$	$1 \leq i \leq 3m$
1	$y_{i,3}\ x_{i,1}\ x_{i,3}\ w\ p\ \cdots$	$1 \leq i \leq 3m$
4	$y_{i,4}\ x_{i,1}\ x_{i,2}\ x_{i,3}\ p\ \cdots$	$1 \leq i \leq 3m$
1	$x_i\ d_{i,j}\ p\ \cdots$	$1 \leq i \leq 3m$, $1 \leq j \leq 2m-1$
$2m$	$d_{i,j}\ p\ \cdots$	$1 \leq i \leq 3m$, $1 \leq j \leq 2m-1$
$2m$	$w\ p\ \cdots$	
$2m-1$	$e\ p\ \cdots$	
m	$f\ p\ \cdots$	

For votes of the form $p\ x_i\ \cdots$ we use the price function $\rho(1) = 1$ and $p(t) = m+1$ for all $t \geq 2$. For every other vote, we use the price function $\rho(t) = m+1$ for $t \geq 1$. Finally, set the budget $B = m$.

Without bribing, the election proceeds as follows. In the first round, p scores $3m$ points, w and every $d_{i,j} \in D$ scores $2m$ points, and each of the remaining candidates scores fewer than $2m$ points. In the second round, p scores $18m-1$ points, w scores $11m$ points, and every $d_{i,j}$ scores $2m+1$ points. It follows that p is the unique winner for either of iterated plurality and plurality with runoff.

We claim that (X,S) is in ONE-IN-THREE-POSITIVE-3SAT if and only if $((C,V),p,B,\rho)$ is in DESTRUCTIVE-SHIFT-BRIBERY for either of the two voting systems, regardless of the winner model.

($\Rightarrow$) Suppose that (X,S) is a yes-instance of ONE-IN-THREE-POSITIVE-3SAT. Then there exists a subset $U \subseteq X$ such that for each clause S_j we have $|U \cap S_j| = 1$. We bribe the voters with the vote of the form $p\ x_i\ \cdots$ with $x_i \in U$ so that the new vote has the form $x_i\ p\ \cdots$. It follows that p, w, every $x_i \in U$, and every $d_{i,j} \in D$ reach the second round with $2m$ points each. In the second round, p gains $3m-1$ additional points while w gains $3m$ additional points. It follows that p is not a winner of the election, so $((C,V),p,B,\rho)$ is a yes-instance of DESTRUCTIVE-SHIFT-BRIBERY for both voting systems, regardless of the winner model.

($\Leftarrow$) Suppose that (X,S) is a no-instance of ONE-IN-THREE-POSITIVE-3SAT. We show that $((C,V),$ $p,B,\rho)$ is also a no-instance of DESTRUCTIVE-SHIFT-BRIBERY for both voting systems. To ensure that p is not the only plurality winner in the first round, it is necessary to bribe m voters with votes of the form $p\ x_i\ \cdots$ to now vote $x_i\ p\ \cdots$. Note that we can only bribe at most m such voters without exceeding the budget. Let $U \subseteq X$ be the set of candidates that benefit from the bribery action. It follows that p, every $d_{i,j} \in D$, every $x_i \in U$, and w can move forward to the next round with $2m$ points each. In this round, the designated candidate p gains $3m-1$ additional points from the votes of the form $e\ p\ \cdots$ and $f\ p\ \cdots$; every candidate $d_{i,j}$ with $x_i \notin U$ gains one additional point; every candidate $x_i \in U$ can receive at most 18 additional points; and w is discussed separately in the following paragraph.

To prevent the victory of p, it is necessary that w gains at least $3m$ points (since if w gains only $3m-1$ points, it follows that w and p move forward to the final round, where p would achieve a clear victory). For w to gain at least one point from any one of the three votes of the form $y_{i,1}\ x_{i,1}\ x_{i,2}\ w\ p\ \cdots$, $y_{i,2}\ x_{i,2}\ x_{i,3}\ w\ p\ \cdots$, and $y_{i,3}\ x_{i,1}\ x_{i,3}\ w\ p\ \cdots$, it is necessary that at most one candidate $x_{i,j}$ participates in the second round. On the other hand, if no candidate $x_{i,j}$ participates in the second round, p gains four points from the voters of the fifth line, whose vote is $y_{i,4}\ x_{i,1}\ x_{i,2}\ x_{i,3}\ p\ \cdots$, i.e., this clause harms w. Only a clause S_i with $|S_i \cap U| = 1$ helps w to reduce the point difference to p. Since (X,S) is a no-instance of ONE-IN-THREE-POSITIVE-3SAT, there are at most $3m-2$ clauses with this property.

With these clauses w can reduce the point difference to two. With the two remaining clauses the point difference is growing. This implies that p is always a unique winner of the election, i.e., $((C,V),p,B,\rho)$ is a no-instance of DESTRUCTIVE-SHIFT-BRIBERY for both voting systems, regardless of the winner model. □

6. Iterated Veto and Veto with Runoff

In this section, we show hardness of shift bribery for iterated veto and veto with runoff, again handling both voting systems simultaneously and starting with the constructive case.

Theorem 11. *In both the unique-winner and the nonunique-winner model, for veto with runoff and iterated veto,* CONSTRUCTIVE-SHIFT-BRIBERY *is* NP-*hard.*

Proof. To prove NP-hardness, we reduce X3C to CONSTRUCTIVE-SHIFT-BRIBERY for veto with runoff and iterated veto at the same time. Let $(X,\mathscr{S})$ be a given X3C instance and construct the CONSTRUCTIVE-SHIFT-BRIBERY instance $((C,V),p,B,\rho)$ as follows. Let $C=\{p,d_1,d_2\}\cup X\cup\mathscr{S}$ be the set of candidates, where p is the designated candidate, and construct the voter preferences in V as follows:

#	votes	for
1	$\cdots\ S_i\ p$	$1\le i\le 3m$
2	$\cdots\ x_{i,1}\ S_i$	$1\le i\le 3m$
2	$\cdots\ x_{i,2}\ S_i$	$1\le i\le 3m$
2	$\cdots\ x_{i,3}\ S_i$	$1\le i\le 3m$
$2m-6$	$\cdots\ d_2\ S_i$	$1\le i\le 3m$
$2m$	$\cdots\ x_i$	$1\le i\le 3m$
m	$\cdots\ d_2\ x_i\ d_1$	$1\le i\le 3m$
$m+2$	$\cdots\ d_2\ S_i\ d_1$	$1\le i\le 3m$
$2m$	$\cdots\ d_2$	
1	$\cdots\ p\ d_1$	

For votes of the form $\cdots\ S_i\ p$, we use the price function $\rho(1)=1$, and $\rho(t)=m+1$ for all $t\ge 2$; and for every other voter, we use the price function $\rho(t)=m+1$ for $t\ge 1$. Finally, set the budget $B=m$.

Note that for both voting rules, p is eliminated in the first round with $3m$ vetoes and therefore cannot be the winner without bribing voters.

We claim that $(X,\mathscr{S})$ is in X3C if and only if $((C,V),p,B,\rho)$ is in CONSTRUCTIVE-SHIFT-BRIBERY for either of iterated veto and veto with runoff, regardless of the winner model.

$(\Rightarrow)$ Suppose that $(X,\mathscr{S})$ is a yes-instance of X3C. Then there exists an exact cover $\mathscr{S}'\subseteq\mathscr{S}$ of size m. Shift p one position forward in the votes of the form $\cdots\ S_i\ p$ for each $S_i\in\mathscr{S}'$, so that the new vote has the form $\cdots\ p\ S_i$. It follows that p, each $S \in \mathscr{S} \setminus \mathscr{S}'$, each x_i for $1 \leq i \leq 3m$, and d_2 are veto winners with $2m$ vetoes each and thus proceed to the second round. Since $\mathscr{S}'$ is an exact cover, each x_i receives two additional vetoes from the voters in lines 2–4 corresponding to the sets in the exact cover and m vetoes from the voters in line 7. Furthermore, each $S \in \mathscr{S} \setminus \mathscr{S}'$ receives $m+2$ vetoes from the voters in line 8, whereas p receives m vetoes from the voters in line 1 and only one additional veto from the voter in the last line. Since d_2 gains far more than $m+1$ vetoes in this round, it follows that p is the unique veto winner of the bribed election. Thus $((C,V),p,B,\rho)$ is a yes-instance of CONSTRUCTIVE-SHIFT-BRIBERY for either of iterated veto and veto with runoff, regardless of the winner model.

($\Leftarrow$) Suppose that $(X,\mathscr{S})$ is a no-instance of X3C. This means that for every $\mathscr{S}' \subseteq \mathscr{S}$, $|\mathscr{S}'| \leq m$, there is an $x' \in X$ that is not covered by any $S \in \mathscr{S}'$.

To not be eliminated in the first round and to not exceed the budget of m, p has to lose exactly m vetoes so as to tie with the $2m$ vetoes of the x_i. This is only possible by bribing the voters in the first line. Let $\mathscr{S}' \subseteq \mathscr{S}$, $|\mathscr{S}'| = m$, be the set that corresponds to the S_i of the bribed voters. Candidates p and d_2 as well as each $S \in \mathscr{S} \setminus \mathscr{S}'$ and each x_i, $1 \leq i \leq 3m$, reach the second round with $2m$ vetoes. However, in the second round, the $x' \in X$ that was not covered by $\mathscr{S}'$ receives only m additional vetoes in contrast to p who receives $m+1$ additional vetoes. It follows that p is not winning the election for either of the two voting rules. That means that $((C,V),p,B,\rho)$ is a no-instance of CONSTRUCTIVE-SHIFT-BRIBERY for either of iterated veto and veto with runoff, regardless of the winner model. $\square$

We now turn to the destructive variant of shift bribery for iterated veto and veto with runoff.

Theorem 12. *In both the unique-winner and the nonunique-winner model, for veto with runoff and iterated veto,* DESTRUCTIVE-SHIFT-BRIBERY *is* NP-*hard.*

Proof. To prove NP-hardness, we reduce the NP-complete problem ONE-IN-THREE-POSITIVE-3SAT to DESTRUCTIVE-SHIFT-BRIBERY for veto with runoff and iterated veto simultaneously. Given an instance (X,S) of ONE-IN-THREE-POSITIVE-3SAT, where $X = \{x_1,\dots,x_{3m}\}$ and $S = \{S_1,\dots,S_{3m}\}$, with $S_i = \{x_{i,1},x_{i,2},x_{i,3}\} \subseteq X$ for each $1 \leq i \leq 3m$, we construct the election (C,V) with candidate set $C = \{p,w,d_1,d_2\} \cup X$, designated candidate p, and the following list V of votes:

#	votes	for
1	$\cdots\ p\ x_i$	$1 \leq i \leq 3m$
2	$\cdots\ p\ x_{i,1}\ x_{i,2}\ d_1$	$1 \leq i \leq 3m$
2	$\cdots\ p\ x_{i,2}\ x_{i,3}\ d_1$	$1 \leq i \leq 3m$
2	$\cdots\ p\ x_{i,1}\ x_{i,3}\ d_1$	$1 \leq i \leq 3m$
7	$\cdots\ w\ x_{i,1}\ x_{i,2}\ x_{i,3}\ d_1$	$1 \leq i \leq 3m$
$2m$	$\cdots\ d_2\ x_i$	$1 \leq i \leq 3m$
$22m$	$\cdots\ d_2\ x_i\ d_1$	$1 \leq i \leq 3m$
$2m$	$\cdots\ d_2$	
m	$\cdots\ p$	
$2m$	$\cdots\ w$	
$8m-1$	$\cdots\ w\ d_1$	

For every vote of the form $\cdots\ p\ x_i$, let the price function be $\rho(1) = 1$ and $\rho(t) = m+1$ for every $t \geq 2$. For every other vote, define $\rho(t) = m+1$ for every $t \geq 1$. Finally, we set the budget $B = m$.

It is easy to see that p is the winner of this election for both voting rules.

We claim that (X,S) is in ONE-IN-THREE-POSITIVE-3SAT if and only if $((C,V),p,B,\rho)$ is in DESTRUCTIVE-SHIFT-BRIBERY for either of veto with runoff and iterated veto, regardless of the winner model.

$(\Rightarrow)$ Assume that (X,S) is in ONE-IN-THREE-POSITIVE-3SAT. Then there is a subset $X' \subseteq X$ such that for each clause S_i we have $|X' \cap S_i| = 1$. Bribe the voters with votes of the form $\cdots\ p\ x_i$ with $x_i \in X'$ so that the new vote has the form $\cdots\ x_i\ p$. It follows that p, w, d_2, and each $x_i \in X'$ have the fewest vetoes (namely, $2m$) and therefore proceed to the second round. In the second round, p receives $2m$ vetoes from the votes in line 1 and for each of the $3m$ clauses two vetoes from the voters in lines 2–4 for a total of $8m$ additional vetoes, whereas w only receives a total of $8m-1$ vetoes. It follows that p is not a winner of the election for either of the two voting rules.

$(\Leftarrow)$ Let (X,S) be a yes-instance of DESTRUCTIVE-SHIFT-BRIBERY for veto with runoff (respectively, iterated veto), i.e., it is possible to bribe voters so that p does not win the election. Recall that it is only possible to bribe voters in line 1 without ecxeeding the budget. In the first round, p receives m vetoes, i.e., the fewest vetoes of all candidates. Due to the votes in line 7, the only candidate capable of receiving fewer vetoes than p or the same number of vetoes as p in the second round is w.[9] However, this is only possible if p receives at least $9m-1$ additional vetoes since w has $10m-1$ vetoes in the second round from the last two lines alone. p receives $3m$ of these additional vetoes from line 1—after bribing voters so that p is in the last position, or eliminating the x_i in the first round—leaving a gap of $6m-1$ vetoes. For each clause S_j such that no $x_i \in S_j$ is present in the second round, p receives six additional vetoes (lines 2–4), whereas w receives in this case seven additional vetoes from the voters in line 5, i.e., this widens the gap between p and w instead of closing it. That means that for each clause S_j, there has to be at least one $x_i \in S_j$ present in the second round, i.e., for each clause S_j, a voter with a vote of the form $\cdots\ p\ x_i$ with $x_i \in S_j$ needs to be bribed to cast a vote of the form $\cdots\ x_i\ p$ to bring the respective vetoes down to $2m$, the same as, e.g., d_2. However, if at least two literals, say x_i and x_k, in a clause S_j are present in the second round, p receives no additional veto, which does not help to close the gap between p and w. The only possibility remaining for p not to be a winner of the bribed election is that the bribed voters correspond to the variables set to true in an assignment where in each clause there is exactly one literal true, i.e., we have a yes-instance of ONE-IN-THREE-POSITIVE-3SAT. □

7. Using the Nonmonotonicity Property

Informally stated, a voting rule is said to be *monotonic* if winners can never be turned into nonwinners by improving their position in some votes, everything else remaining the same.[10] Intuitively, that is to say that only shifting a candidate forward (closer to the top) is beneficial, whereas shifting a candidate backward (closer to the bottom) is not. In shift bribery under some monotonic voting rule, it thus makes only sense for the briber to shift the designated candidate forward in the constructive case (respectively, backward in the destructive case). However, all voting rules considered here except iterated plurality and iterated veto are *not* monotonic, and in nonmonotonic voting rules, shifting the designated candidate backward in the constructive case (respectively, forward in the destructive case) could also be beneficial for the briber.

It would therefore be interesting to find out whether the complexity of our problems changes when the nonmonotonicity of voting rules is specifically allowed, or even required, to be exploited in shift bribery

9. Note that d_1 will definitely be eliminated in the first round.
10. This definition captures just one common notion of monotonicity, the one we will be using here; but note that there are also other notions of monotonicity for voting rules known in social choice theory.

actions. Indeed, with slight modifications to the proofs, we can show that Hare-CONSTRUCTIVE-SHIFT-BRIBERY and plurality-with-runoff-CONSTRUCTIVE-SHIFT-BRIBERY are still NP-hard if the designated candidate can *only* be shifted *backward*. We conjecture that all other proofs (except the proofs for the monotonic voting rules iterated plurality and iterated veto) can be adapted in such a way as well.

We start with constructive shift bribery in Hare elections where the only allowed bribery action is to shift the designated candidate *backward*.

Theorem 13. *In both the unique-winner and the nonunique-winner model, Hare-*CONSTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard even if the designated candidate can only be shifted backward.*

Proof. NP-hardness again follows by a reduction from X3C. Construct from a given X3C instance $(X,\mathscr{S})$ an instance $((C,V),p,B,\rho)$ of Hare-CONSTRUCTIVE-SHIFT-BRIBERY with candidate set $C = X \cup \mathscr{S} \cup D \cup \{p,w\}$, where $D = \{d_1,\dots,d_{3m}\}$ is a set of dummy candidates and p the designated candidate, and the following list V of votes:

#	vote	for
1	$S_i\ x_{i,1}\ \overrightarrow{X \setminus \{x_{i,1}\}}\ w\ p\ \cdots$	$1 \le i \le 3m$
1	$S_i\ x_{i,2}\ \overrightarrow{X \setminus \{x_{i,2}\}}\ w\ p\ \cdots$	$1 \le i \le 3m$
1	$S_i\ x_{i,3}\ \overrightarrow{X \setminus \{x_{i,3}\}}\ w\ p\ \cdots$	$1 \le i \le 3m$
4	$x_i\ \overrightarrow{X \setminus \{x_i\}}\ w\ p\ \cdots$	$1 \le i \le 3m$
6	$w\ \overrightarrow{X}\ p\ \cdots$	
1	$p\ S_i\ \cdots$	$1 \le i \le 3m$
6	$p\ \cdots$	
3	$d_i\ S_i\ p\ w \cdots$	$1 \le i \le 3m$

For votes of the form $p\ S_i\ \cdots$, we use the price function $\rho(1) = 1$, and $\rho(t) = m+1$ for all $t \ge 2$; and for every other vote, we use the price function ρ with $\rho(t) = m+1$ for all $t \ge 1$. Finally, set the budget $B = m$.

Without bribing the voters the election proceeds as follows:

Round	p	w	x_1	$x_i \in X \setminus \{x_1\}$	$S_i \in \mathscr{S}$	$d_i \in D$
1	$3m+6$	6	4	4	3	3
2	$12m+6$	6	7	7	out	out
3	$12m+6$	out	13	7	out	out
4	$12m+6$	out	$21m+6$	out	out	out

It follows that p is eliminated in the last round and does not win the election.

We claim that $(X,\mathscr{S})$ is in X3C if and only if $((C,V),p,B,\rho)$ is in Hare-CONSTRUCTIVE-SHIFT-BRIBERY, regardless of the winner model, even if the designated candidate can only be shifted backward.

($\Rightarrow$) Suppose that $(X,\mathscr{S})$ is a yes-instance of X3C. Then there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of size m. We now show that it is possible for p to become a unique Hare winner of an election obtained by shifting p in the votes without exceeding the budget B. For every $S_i \in \mathscr{S}'$, we bribe the voter with the vote of the form $p\ S_i\ \cdots$ by shifting p once, so her new vote is of the form $S_i\ p\ \cdots$; each such bribe action costs us only 1 from our budget, so the budget will not be exceeded. Now the election proceeds as follows:

Round	p	w	$x_i \in X$	$S_i \in \mathscr{S}'$	$S_i \in \mathscr{S} \setminus \mathscr{S}'$	$d_i \in D$
1	$2m+6$	6	4	4	3	3
2	$8m+6$	6	6	7	out	out
3	$26m+12$	out	out	7	out	out

We see that p is the only candidate still standing in the fourth round and thus the only Hare winner of the bribed election.

($\Leftarrow$) Suppose that $(X, \mathscr{S})$ is a no-instance of X3C. Then no subset $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| \leq m$ covers X. We now show that p will be eliminated in all elections obtained by bribing voters without exceeding budget B. Note that we can only bribe at most m voters with votes of the form $p\ S_i\ \cdots$ without exceeding the budget. Let $\mathscr{S}' \subseteq \mathscr{S}$ be such that for every $S_i \in \mathscr{S}'$ we have bribed the voter whose vote is $p\ S_i\ \cdots$. We can assume that $|\mathscr{S}'| > 0$.

Every candidate in $\mathscr{S}'$ will gain an additional point and therefore survives the first round. All candidates from D and $\mathscr{S} \setminus \mathscr{S}'$ will be eliminated, since p only loses at most m points.

In the second round, the remaining candidates from $\mathscr{S}$ will gain three additional points from the elimination of candidates in D and score seven points in this round (and in all subsequent rounds with p still standing). If a candidate $S_i \in \mathscr{S}$ was eliminated in the previous round, every $x_j \in S_i$ gains one additional point in this round. Partition X into sets X_0, X_1, X_2, and X_3 so that $x_i \in X_k \Leftrightarrow |\{S_j \in \mathscr{S}' \mid x_i \in S_j\}| = k$ for $k \in \{0,1,2,3\}$. Note that X_0, X_1, X_2, and X_3 are disjoint and $|X_0| > 0$, but one or two of X_1, X_2, and X_3 may be empty. Then $x_i \in X_j$ scores $4 + (3 - j) \in \{4,5,6,7\}$ points depending on how many times x_i is covered by $\mathscr{S}'$. Therefore, every $x_i \in X_0$ scores more points than w who has six points. So, there are candidates from X that survive this round and other candidates from X (i.e., candidates from X_1, X_2, or X_3), who are eliminated.

In the third round, the candidate $x_\ell \in X$ with the smallest subscript who is still standing gains at least four points from the eliminated candidates, so that she scores at least nine points now (since no candidates from X_3 are left in the election). All other candidates still score the same number of points as in the previous round. Therefore, p scores $4|\mathscr{S} \setminus \mathscr{S}'| + 6$ points, w scores six points (if w was not already eliminated along with the candidates from X_1), every $S_i \in \mathscr{S}'$ scores seven points, and every still standing candidate from X except x_ℓ scores at most seven points. Since w can only gain additional points when all candidates from X are eliminated and only x_ℓ gains points from the elimination of w or candidates from $X \setminus \{x_\ell\}$ in the subsequent rounds, all candidates $X \setminus (\{x_\ell\} \cup X_0)$ and w are eliminated. Then all still standing candidates from $X_0 \setminus \{x_\ell\}$ and candidates from $\mathscr{S}'$ who score seven points each are eliminated, which leaves p and x_ℓ in the last round. In this round, p scores $12m + 6$ points and x_ℓ scores $21m + 6$ points, so p is eliminated from the election and does not win. □

Next, we show the corresponding result for plurality with runoff.

Theorem 14. *In both the unique-winner and the nonunique-winner model, plurality-with-runoff-*CONSTRUCTIVE-SHIFT-BRIBERY *is* NP*-hard even if the designated candidate can only be shifted backward.*

Proof. To prove NP-hardness, we reduce X3C to CONSTRUCTIVE-SHIFT-BRIBERY for plurality with runoff. Let $(X, \mathscr{S})$ be a given X3C instance, where $X = \{x_1, \ldots, x_{3m}\}$ and $\mathscr{S} = \{S_1, \ldots, S_{3m}\}$. Also, we require that $m > 3$. We construct the CONSTRUCTIVE-SHIFT-BRIBERY instance $((C,V), p, B, \rho)$ as follows. Let $C = \{p\} \cup X \cup \mathscr{S} \cup D \cup Y$ with sets of dummy candidates $D = \{d_{i,j} \mid 1 \leq i \leq 3m$ and $1 \leq j \leq 2m^2 - 5m - 4\}$ and $Y = \{y_i \mid 1 \leq i \leq 3m+1\}$ and designated candidate p. The list V of votes is constructed as follows:

#	vote	for
1	$p\ S_i\ \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,1}\ w\ \overrightarrow{X \setminus \{x_{i,1}\}} \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,2}\ w\ \overrightarrow{X \setminus \{x_{i,2}\}} \cdots$	$1 \leq i \leq 3m$
2	$S_i\ x_{i,3}\ w\ \overrightarrow{X \setminus \{x_{i,3}\}} \cdots$	$1 \leq i \leq 3m$
$3m$	$w\ p\ \cdots$	
1	$y_i\ p$	$1 \leq i \leq 3m+1$
$m-3$	$S_i\ w\ p$	$1 \leq i \leq 3m$
$m-4$	$S_i\ p\ w$	$1 \leq i \leq 3m$
$2m$	$x_i\ w\ p$	$1 \leq i \leq 3m$
1	$d_{i,j}\ x_i\ w\ p\ \cdots$	$1 \leq i \leq 3m$, $1 \leq j \leq 2m^2-5m-4$

For votes of the form $p\ S_i\ \cdots$, we use the price function $\rho(1) = 1$, and $\rho(t) = m+1$ for all $t \geq 2$; and for every other vote, we use the price function $\rho(t) = m+1$ for $t \geq 1$. Finally, set the budget $B = m$.

Without bribing, only p and w reach the second and final round with $3m$ points each. Clearly, w alone wins the election with only p and w present.

We claim that $(X, \mathscr{S})$ is in X3C if and only if $((C,V), p, B, \rho)$ is in CONSTRUCTIVE-SHIFT-BRIBERY for plurality with runoff, regardless of the winner model.

($\Rightarrow$) Suppose that $(X, \mathscr{S})$ is a yes-instance of X3C. Then there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of size m. We now show that it is possible for p to become a unique plurality-with-runoff winner of an election obtained by shifting p in the votes without exceeding the budget. For every $S_i \in \mathscr{S}'$, we bribe the voter with the vote of the form $p\ S_i\ \cdots$ once, so her new vote is of the form $S_i\ p\ \cdots$.

In the first round, w scores $3m$ points; p, every $x_i \in X$, and every $S_i \in \mathscr{S}'$ score $2m$ points each; every $S_i \in \mathscr{S} \setminus \mathscr{S}'$ scores $2m-1$ points; and every candidate from D and Y scores only one point. Since w is the only plurality winner, all second-place candidates (namely, p, every $x_i \in X$, and every $S_i \in \mathscr{S}'$) proceed to the second round.

In the second round, every $S_i \in \mathscr{S}'$ still scores the same number of points as in the first round, w gains $2m(m-3)$ additional points, p gains $(3m+1)+2m(m-4)$ additional points, and every $x_i \in X$ gains $(2m^2-5m-4)+4$ additional points. Therefore, p alone wins the election with $2m^2-3m+1$ points, ahead of w and every $x_i \in X$ with $2m^2-3m$ points each, and every $S_i \in \mathscr{S}'$ with $2m$ points each.

($\Leftarrow$) Suppose that $((C,V), p, B, \rho)$ is a yes-instance of Plurality-with-runoff-CONSTRUCTIVE-SHIFT-BRIBERY. Notice that if no voters are bribed, p and w are leading in the election with $3m$ points each, so they both proceed to the final round. It is easy to see that w wins against p in a one-on-one election. To prevent w and p from being the only candidates in the second round, m voters with votes of the form $p\ S_i\ \cdots$ have to be bribed. Let $\mathscr{S}' \subseteq \mathscr{S}$ be such that $S_i \in \mathscr{S}'$ if the voter with vote $p\ S_i\ \cdots$ has been bribed. Then w, p, every $x_i \in X$, and every $S_i \in \mathscr{S}'$ survive the first round. Since every other candidate is deleted in the first round, p now scores $2m^2-5m+1$ points and beats w by a margin of one point. Moreover, p beats every $S_i \in \mathscr{S}'$ since the candidates from $\mathscr{S}'$ did not gain any additional points in this round. Regarding the candidates from X, every $x_i \in X$ gains $2m^2-5m-4$ points and two additional points for every $S_j \in \mathscr{S} \setminus \mathscr{S}'$ with $x_i \in S_j$ that was eliminated in the first round. Since there are exactly three $S_j \in \mathscr{S}$ with $x_i \in S_j$, every $x_i \in X$ can gain six points if all those candidates were eliminated in the last round, which would let x_i overtake p by one point. In order for p to beat all $x_i \in X$, at least one $S_j \in \mathscr{S}$ with $x_i \in S_j$ needs to be in $\mathscr{S}'$ and is therefore still standing in the second round. Since $|\mathscr{S}'| = m$ and there are $3m$ candidates in X, p can beat every $x_i \in X$ (and subsequently win the election) only if $\mathscr{S}'$ is an exact cover of X. $\square$

8. Conclusions and Open Questions

We have shown that shift bribery is NP-complete for each of the iterative voting systems of Hare, Coombs, Baldwin, Nanson, iterated plurality, plurality with runoff, iterated veto, and veto with runoff, each for both the constructive and the destructive case and in both the unique-winner and the nonunique-winner model. This contrasts previous results due to Elkind et al. (2009), Elkind and Faliszewski (2010), and Schlotter et al. (2017) showing that shift bribery can be solved efficiently by exact or approximation algorithms for many natural voting rules that do not proceed iteratively. Indeed, the iterative nature of the voting rules we have studied seems to be responsible for the hardness of shift bribery.

While these are interesting theoretical results complementing earlier work both on shift bribery and on these voting systems, NP-hardness of course has its limitations in terms of providing protection against shift bribery attacks in practice. Therefore, it would be interesting to also study shift bribery for these voting systems in terms of approximation and parameterized complexity and to do a typical-case analysis. Based on our results in this article, Zhou and Guo (2020) already obtained first results regarding the parameterized complexity of iterative voting systems with respect to a fixed number of shifts, votes, or candidates. Further, they have shown that the hardness of shift bribery for the Hare, Coombs, Baldwin, and Nanson rules also holds for unit price cost functions. It would be particularly interesting to determine the role of the cost function for the hardness of shift bribery. Furthermore, it would be interesting future work to study in detail the effect that specific tie-breaking models (such as the "parallel universes" model (Conitzer, Rognlie, & Xia, 2009) and other models) may have on the complexity of shift bribery problems for iterative voting rules.

A feature shared by most of the iterative voting rules we have studied is that many of them are not monotonic. This has the somewhat counterintuitive effect that shifting the designated candidate forward in some votes can actually hurt this candidate's chances to win, and shifting the designated candidate backward can increase these chances. We have discussed this feature in Section 7, showing that constructive shift bribery remains NP-hard even if we are allowed to only shift the designated candidate backward in some votes for two iterative voting systems: Hare voting and plurality with runoff. We leave the analogous question open for the remaining iterative voting systems studied here (except, of course, for the monotonic rules iterated plurality and iterated veto), and conjecture that they share this property. Even more interestingly, we pose as an open question whether there is a nonmonotonic voting system—a natural one or an artificially constructed one—for which unrestricted shift bribery is NP-hard but becomes efficiently solvable when restricted to shift bribery actions specifically exploiting their nonmonotonicity (i.e., allowing to shift the designated candidate only backward in the constructive case, or forward in the destructive case).

Acknowledgments

We thank the anonymous AAMAS'18 and ISAIM'18 reviewers for helpful comments. This work was supported in part by DFG grants RO 1202/21-1, RO 1202/15-1, RO 1202/14-2, and BA 6270/1-1.

References

Baldwin, J. (1926). The technique of the Nanson preferential majority system of election. *Transactions and Proceedings of the Royal Society of Victoria*, *39*, 42–52.

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

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

Bartholdi III, J., Tovey, C., & Trick, M. (1992). How hard is it to control an election? *Mathematical and Computer Modelling*, *16*(8/9), 27–40.

Baumeister, D., Faliszewski, P., Lang, J., & Rothe, J. (2012). Campaigns for lazy voters: Truncated ballots. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 577–584. IFAAMAS.

Baumeister, D., & Rothe, J. (2015). Preference aggregation by voting. In Rothe, J. (Ed.), *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, Springer Texts in Business and Economics, chap. 4, pp. 197–325. Springer-Verlag.

Betzler, N., Niedermeier, R., & Woeginger, G. (2011). Unweighted coalitional manipulation under the Borda rule is NP-hard. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pp. 55–60. AAAI Press/IJCAI.

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

Bredereck, R., Chen, J., Faliszewski, P., Guo, J., Niedermeier, R., & Woeginger, G. (2014a). Parameterized algorithmics for computational social choice: Nine research challenges. *Tsinghua Science and Technology*, *19*(4), 358–373.

Bredereck, R., Chen, J., Faliszewski, P., Nichterlein, A., & Niedermeier, R. (2014b). Prices matter for the parameterized complexity of shift bribery. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, pp. 1398–1404. AAAI Press.

Bredereck, R., Faliszewski, P., Niedermeier, R., & Talmon, N. (2016a). Complexity of shift bribery in committee elections. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pp. 2452–2458. AAAI Press.

Bredereck, R., Faliszewski, P., Niedermeier, R., & Talmon, N. (2016b). Large-scale election campaigns: Combinatorial shift bribery. *Journal of Artificial Intelligence Research*, *55*, 603–652.

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

Conitzer, V., Sandholm, T., & Lang, J. (2007). When are elections with few candidates hard to manipulate? *Journal of the ACM*, *54*(3), Article 14.

Conitzer, V., & Walsh, T. (2016). Barriers to manipulation in voting. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. (Eds.), *Handbook of Computational Social Choice*, chap. 6, pp. 127–145. Cambridge University Press.

Davies, J., Katsirelos, G., Narodytska, N., Walsh, T., & Xia, L. (2014). Complexity of and algorithms for the manipulation of Borda, Nanson's and Baldwin's voting rules. *Artificial Intelligence*, *217*, 20–42.

Davies, J., Narodytska, N., & Walsh, T. (2012). Eliminating the weakest link: Making manipulation intractable? In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, pp. 1333–1339. AAAI Press.

Elkind, E., & Faliszewski, P. (2010). Approximation algorithms for campaign management. In *Proceedings of the 6th International Workshop on Internet & Network Economics*, pp. 473–482. Springer-Verlag *Lecture Notes in Computer Science #6484*.

Elkind, E., Faliszewski, P., & Slinko, A. (2009). Swap bribery. In *Proceedings of the 2nd International Symposium on Algorithmic Game Theory*, pp. 299–310. Springer-Verlag *Lecture Notes in Computer Science #5814*.

Erdélyi, G., Neveling, M., Reger, C., Rothe, J., Yang, Y., & Zorn, R. (2021). Towards completing the puzzle: Complexity of control by replacing, adding, and deleting candidates or voters. *Journal of Autonomous Agents and Multi-Agent Systems*, *forthcoming*. Conference versions due to Erdélyi, Reger, and Yang (2019) and Neveling, Rothe, and Zorn (2020).

Erdélyi, G., Reger, C., & Yang, Y. (2019). Towards completing the puzzle: Solving open problems for control in elections. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pp. 846–854. IFAAMAS.

Faliszewski, P., Hemaspaandra, E., & Hemaspaandra, L. (2009a). How hard is bribery in elections? *Journal of Artificial Intelligence Research*, *35*, 485–532.

Faliszewski, P., Hemaspaandra, E., Hemaspaandra, L., & Rothe, J. (2009b). Llull and Copeland voting computationally resist bribery and constructive control. *Journal of Artificial Intelligence Research*, *35*, 275–341.

Faliszewski, P., Manurangsi, P., & Sornat, K. (2019). Approximation and hardness of shift-bribery. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pp. 1901–1908. AAAI Press.

Faliszewski, P., Reisch, Y., Rothe, J., & Schend, L. (2015). Complexity of manipulation, bribery, and campaign management in Bucklin and fallback voting. *Journal of Autonomous Agents and Multi-Agent Systems*, *29*(6), 1091–1124.

Faliszewski, P., & Rothe, J. (2016). Control and bribery in voting. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. (Eds.), *Handbook of Computational Social Choice*, chap. 7, pp. 146–168. Cambridge University Press.

Freeman, R., Brill, M., & Conitzer, V. (2014). On the axiomatic characterization of runoff voting rules. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, pp. 675–681. AAAI Press.

Garey, M., & Johnson, D. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company.

Gonzalez, T. (1985). Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, *38*, 293–306.

Hemaspaandra, E., Hemaspaandra, L., & Rothe, J. (2007). Anyone but him: The complexity of precluding an alternative. *Artificial Intelligence*, *171*(5–6), 255–285.

Kaczmarczyk, A., & Faliszewski, P. (2016). Algorithms for destructive shift bribery. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pp. 305–313. IFAAMAS.

Knop, D., Koutecký, M., & Mnich, M. (2017). Voting and bribing in single-exponential time. In *Proceedings of the 34th Annual Symposium on Theoretical Aspects of Computer Science*, Vol. 66 of *LIPIcs*, pp. 46:1–46:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

Konczak, K., & Lang, J. (2005). Voting procedures with incomplete preferences. In *Proceedings of the Multidisciplinary IJCAI-05 Workshop on Advances in Preference Handling*, pp. 124–129.

Levin, J., & Nalebuff, B. (1995). An introduction to vote-counting schemes. *The Journal of Economic Perspectives*, *9*(1), 3–26.

Maushagen, C., Neveling, M., Rothe, J., & Selker, A. (2018). Complexity of shift bribery in iterative elections. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1567–1575. IFAAMAS.

Nanson, E. (1882). Methods of election. *Transactions and Proceedings of the Royal Society of Victoria*, *19*, 197–240.

Neveling, M., Rothe, J., & Zorn, R. (2020). The complexity of controlling Condorcet, fallback, and k-veto elections by replacing candidates or voters. In *Proceedings of the 15th International Computer Science Symposium in Russia*, pp. 314–327. Springer-Verlag *Lecture Notes in Computer Science #12159*.

Papadimitriou, C. (1995). *Computational Complexity* (Second edition). Addison-Wesley.

Porschen, S., Schmidt, T., Speckenmeyer, E., & Wotzlaw, A. (2014). XSAT and NAE-SAT of linear CNF classes. *Discrete Applied Mathematics*, *167*, 1–14.

Reisch, Y., Rothe, J., & Schend, L. (2014). The margin of victory in Schulze, cup, and Copeland elections: Complexity of the regular and exact variants. In *Proceedings of the 7th European Starting AI Researcher Symposium*, pp. 250–259. IOS Press.

Rothe, J. (2005). *Complexity Theory and Cryptology. An Introduction to Cryptocomplexity*. EATCS Texts in Theoretical Computer Science. Springer-Verlag.

Rothe, J. (Ed.). (2015). *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*. Springer Texts in Business and Economics. Springer-Verlag.

Rothe, J. (2019). Borda count in collective decision making: A summary of recent results. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pp. 9830–9836. AAAI Press.

Schlotter, I., Faliszewski, P., & Elkind, E. (2017). Campaign management under approval-driven voting rules. *Algorithmica*, *77*, 84–115.

Taylor, A. (2005). *Social Choice and the Mathematics of Manipulation*. Cambridge University Press.

Xia, L. (2012). Computing the margin of victory for various voting rules. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pp. 982–999. ACM Press.

Xia, L., & Conitzer, V. (2011). Determining possible and necessary winners given partial orders. *Journal of Artificial Intelligence Research*, *41*, 25–67.

Zhou, A., & Guo, J. (2020). Parameterized complexity of shift bribery in iterative elections. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1665–1673. IFAAMAS.

Zwicker, W. (2016). Introduction to the theory of voting. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. (Eds.), *Handbook of Computational Social Choice*, chap. 2, pp. 23–56. Cambridge University Press.

Chapter 5

Manipulation of Opinion Polls to Influence Iterative Elections

This chapter deals with the complexity of manipulation of opinion polls by a polling agency in the context of iterative elections. See Section 2.2.4 for a short overview on iterative elections and opinion polls. The corresponding publication (Baumeister et al., 2020b) is as follows.

Baumeister, D., Selker, A.-K., and Wilczynski, A. (2020b). Manipulation of opinion polls to influence iterative elections. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, pages 132–140. IFAAMAS

Note that a version including the omitted proofs was published in the non-archival proceedings of COMSOC'21 (Baumeister et al., 2021b).

Summary

In iterative elections, voters repeatedly update their ballots to achieve a better outcome for them. In the model we use, the necessary information to compute a best response stems from an underlying social network where voters can see the ballots (and updates) of their neighbors, and from an opinion poll announced by a polling agency. Following the work by Wilczynski (2019), we study the manipulative power of the polling agency in iterative elections. First, we introduce a best-response variant for the voting rule veto that significantly differs from the already known best-response definition for plurality. This is due to the fact that under veto, changing who to veto might directly benefit a voter's most despised candidate since that candidate loses a veto from this voter. Second, as an addition to the already known poll manipulation problem without a parameter, we introduce two distance-restricted variants of poll manipulation, i.e., a poll-restricted variant that includes an upper bound on the distance between the sincere and the manipulated poll, and a voter-restricted variant where voters do not vote for candidates that need more than a given maximum number of swaps in their preferences to become their most preferred candidates for plurality (respectively, their most despised candidate for veto). Third, we consider destructive manipulation in the context of poll manipulation.

To this end, we study the (parameterized) computational complexity of all three problem variants, for constructive and destructive manipulation, and for the voting rules plurality and veto, among others parameterized by the length of the longest path in the social network and by the allowed deviation distance. We show that all problems are NP-hard even in an acyclic social network, and for the constructive poll-restricted poll manipulation under plurality even when the social network does not contain any edges. Manipulation under veto is tractable for all considered problem variants, whereas for plurality we are only able to prove tractability in the destructive case and only for the unrestricted and the poll-restricted variant. The other three cases remain open.

Furthermore, we design efficient heuristics to compare both voting rules and come to the conclusion that—in our setting—manipulation is more successful under the veto rule and that destructive manipulation is more successful than constructive manipulation.

My Contribution

The writing of the attached article was done jointly with my coauthors. I defined the poll-restricted and voter-restricted problem variants and contributed the complete results in Section 3, Theorem 4.1, and Proposition 4.2.

Manipulation of Opinion Polls to Influence Iterative Elections

Dorothea Baumeister, Ann-Kathrin Selker
Heinrich-Heine-Universität Düsseldorf
Düsseldorf, Germany
{d.baumeister,ann-kathrin.selker}@uni-duesseldorf.de

Anaëlle Wilczynski
Technical University of Munich
Munich, Germany
anaelle.wilczynski@in.tum.de

ABSTRACT

In classical elections, voters only submit their ballot once, whereas in iterative voting, the ballots may be changed iteratively. Following the work by Wilczynski [20], we consider the case where a social network represents an underlying structure between the voters, meaning that each voter can see her neighbors' ballots. In addition, there is a polling agency, which publicly announces the result for the initial vote. This paper investigates the manipulative power of the polling agency. Previously, Wilczynski [20] studied constructive manipulation for the plurality rule. We introduce destructive manipulation and extend the study to the veto rule. Several restricted variants are considered with respect to their parameterized complexity. The theoretical results are complemented by experiments using different heuristics.

KEYWORDS

Iterative voting; Opinion polls; Manipulation; Complexity theory

ACM Reference Format:
Dorothea Baumeister, Ann-Kathrin Selker and Anaëlle Wilczynski. 2020. Manipulation of Opinion Polls to Influence Iterative Elections. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 9 pages.

1 INTRODUCTION

In the field of computational social choice there have been a lot of studies on elections, see the book by Brandt et al. [2]. The usual assumption is that voters once submit their ballot and then the winner is determined. This assumption completely neglects the reasoning about how the voters come to their individual decision. Especially in the ages of digital democracy, opinion polls may be executed efficiently and also repeatedly, which may lead the voters to strategically think about their ballot. We focus on iterative elections where voters can update (i.e., manipulate) their ballots. Following the seminal paper by Bartholdi III et al. [1], the issue of manipulation through strategic voting has been studied intensively in the computational social choice context. The most common, but often criticized, assumption is that the manipulator has complete knowledge over all ballots. There are different approaches to tackle this issue, for example the study of incomplete information settings like the possible winner problem introduced by Konczak and Lang [8]. However, we assume that voters only have partial knowledge about the other ballots. They have two sources of information. The first one is the result of some opinion poll, while the second one is the information they get from their neighbors in a social network.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

We then assume that every voter can update her ballot with respect to this information. In this process, the opinion poll is critical, and hence the polling agency has a lot of power. In this paper, we investigate the manipulative power of the polling agency with respect to different situations. In contrast to the manipulation by the voters, it is reasonable to assume complete information for the polling agency, since it collects the votes.

Wilczynski [20] recently introduced the problem of constructive manipulation, where the polling agency tries to make some distinguished candidate win by announcing some strategic opinion poll. The only condition the opinion poll has to satisfy is that no voter may directly detect the manipulation since the poll contradicts with her actual information. Formally this is modeled through a likelihood condition. The influence of opinion polls on the behavior of voters has also been studied by Reijngoud and Endriss [16] and Endriss et al. [3]. Their focus is on the strategic response of voters to different types of information communicated by the opinion polls, without an underlying social network, whereas we focus on the manipulative actions of the opinion poll itself. Our work analyzes poll manipulation in the setting of iterative voting, a widely studied topic in social choice (see Meir [10] for a recent survey), where voters can successively change their ballot in a strategic way. In this context, Sina et al. [18] have previously investigated election control in the presence of a social network. However, they focus on manipulation by an external agent who can add or remove links in the social network whereas we study manipulation of the initial scores communicated by the polling agency.

Recently, many works have investigated voting where voters are embedded in a social network. Tsang and Larson [19] analyze the consequences, in the strategic behavior of voters, of inferring the outcome of the election from the votes of neighbors in a social network. Alternatively, Gourvès et al. [7] study how voters can manipulate by coalitions which come from a social network structure. However, both works do not consider election control questions. Our work is also related to the study of opinion diffusion in graphs. Faliszewski et al. [4] study the effects of campaigning for manipulating election outcomes in an opinion diffusion process with voter clusters. In a similar context, Wilder and Vorobeychik [21, 22] investigate the game-theoretic properties of a game where an attacker tries to influence the election outcome by diffusing fake news in a social network and a defender aims to limit their impact.

We extend the study by Wilczynski [20] in several different ways. First, the current results are restricted to the use of the plurality rule, and we will also consider the veto rule. Although the rules are very similar, the results differ for restricted cases. Second, we introduce a destructive variant, where the opinion poll aims to prevent the victory of some designated candidate by announcing a strategic poll. Third, we analyze this problem for different parameters and with respect to various distance restrictions. This is particularly important when considering real-world problems. Usually, the voters have some rough idea about how the opinion poll will look like, so the announcement should not deviate too much from the original outcome. We show that the corresponding decision problems are NP-hard for acyclic networks under both plurality and veto, and in P for empty networks under veto, whereas under plurality, the exact complexity for empty networks depends on the restriction model. Furthermore, we prove that all decision problems are tractable for a small number of candidates. Additionally, we design efficient heuristics for the manipulation of both voting rules and compare the manipulation results in our experimental section.

2 POLL-CONFIDENT ITERATIVE MODEL

We first describe the poll-confident iterative voting model.

2.1 Basic notations

Let $N = \{1, \dots, n\}$ be a set of n agents (or voters) and $M = \{x_1, \dots, x_m\}$ a set of m candidates. Each voter i has strict ordinal preferences over the candidates, represented by a linear order $\succ_i$ over M. The preference profile is denoted by $\succ = (\succ_1, \dots, \succ_n)$. Let $\overrightarrow{M}$ (respectively, $\overleftarrow{M}$) be a shorthand for $x_1 \succ \dots \succ x_m$ (respectively, $x_m \succ \dots \succ x_1$), and let $N_{x \succ y}$ denote the set of voters who prefer candidate x to candidate y, i.e., $N_{x \succ y} = \{i \in N : x \succ_i y\}$. The winner of the election is determined by a voting rule $\mathcal{F}$. We focus on single-winner elections and use a deterministic tie-breaking rule based on a linear order $\rhd$ over the candidates, in case of ties.

In this article, we focus on two voting rules, namely plurality, denoted by $\mathcal{F}_P$, and veto, denoted by $\mathcal{F}_V$. Under both voting rules, each voter i is asked to submit a ballot $\mathbf{b}_i \in M$ corresponding to a single candidate, i.e., voter i approves candidate $\mathbf{b}_i$ under plurality, whereas voter i vetoes candidate $\mathbf{b}_i$ under veto. A profile of ballots is said to be *truthful* if each agent submits as a ballot her most preferred candidate under plurality and her least preferred candidate under veto. Given a profile of ballots $\mathbf{b} \in M^n$, the score $\mathbf{s_b}(x)$ of each candidate x is computed as follows: $\mathbf{s_b}(x) = |\{i \in N : \mathbf{b}_i = x\}|$. Then the winner under plurality $\mathcal{F}_P(\mathbf{b})$ maximizes the number of approvals, i.e., $\mathcal{F}_P(\mathbf{b}) \in \arg\max_{x \in M} \mathbf{s_b}(x)$, whereas the winner under veto $\mathcal{F}_V(\mathbf{b})$ minimizes the number of vetoes, i.e., $\mathcal{F}_V(\mathbf{b}) \in \arg\min_{x \in M} \mathbf{s_b}(x)$. For the sake of simplicity, we may use $\mathcal{F}(\mathbf{s})$ to denote the winner of a profile of ballots whose score function corresponds to $\mathbf{s}$.

We consider a strategic game called *iterative voting* [10] where, starting from an initial voting profile, agents can successively deviate from their current submitted ballot in order to get a better outcome at the next election. The strategy profile at step t is denoted by $\mathbf{b}^t$. We assume that the initial profile $\mathbf{b}^0$ is truthful, indeed the agents do not have any information to enable them to deviate yet. A single voter is assumed to deviate between two consecutive steps. In case of several voters having incentive to deviate at the same step, one of them is arbitrarily chosen, unless a particular *turn function* τ is specified for choosing the deviator. Note that the choice of the turn function might influence the election outcome.

In the classical iterative voting setting, there is common knowledge of the current strategy profile (or at least the associated scores). For realistic reasons we consider, following Wilczynski [20], that the voters only get partial information about the current strategy profile which is determined by a *social network* and an *opinion poll* and thus which can be biased. More precisely, we assume that the agents are embedded in a social network represented by a directed graph $G = (N, E)$ such that for each arc $(i, j) \in E$, agent i is able to observe the current ballot of agent j. The social network is said to be *empty* if $E = \emptyset$ and *acyclic* if there is no directed cycle in G. The set of agents that a given agent i can observe is denoted by $\Gamma(i) := \{j \in N : (i, j) \in E\} \cup \{i\}$. For a voting profile $\mathbf{b}$, the score of candidate x that agent i is able to observe is denoted by $\mathbf{s}^i_{\mathbf{b}}(x) = |\{j \in \Gamma(i) : \mathbf{b}_j = x\}|$. Moreover, as a prior information, the agents know the scores of the initial profile, given by a polling agency, through the vector of scores $\Delta = (\Delta_1, \dots, \Delta_m)$ where Δ_j stands for the score of candidate x_j which is announced by the polling agency. By abuse of notation, we may also use $\Delta(x)$ for the announced score of candidate x. To summarize, an instance of the poll-confident iterative voting model is a tuple $\mathcal{I} = (N, M, \succ, G, \rhd, \tau)$.

2.2 Manipulation by voters

The manipulation moves of voters are conditioned by the information they get, which is determined by the deviations that they are able to observe. Each agent i has a specific belief regarding the scores of the strategy profile at step t which is given by a believed score vector $B^t_i = (B^t_i(x_1), \dots, B^t_i(x_m))$. The voters trust the results communicated by the polling agency, and thus $B^0_i = \Delta$ for every agent $i \in N$. The believed score vector for both the plurality and veto rules is updated at each step as follows.

Definition 2.1 (Score Belief Update). At step $t+1$, after the deviation of an agent j from ballot $\mathbf{b}^t_j = x$ to ballot $\mathbf{b}^{t+1}_j = y$ at step t, the score of candidate z that agent i believes is given by

$$B^{t+1}_i(z) = \begin{cases} B^t_i(z) - 1 & \text{if } z = x \text{ and } j \in \Gamma(i) \\ B^t_i(z) + 1 & \text{if } z = y \text{ and } j \in \Gamma(i) \\ B^t_i(z) & \text{otherwise} \end{cases}$$

According to the belief of agent i, the *current believed winner* at step t is candidate $\mathcal{F}(B^t_i)$. We assume that the voters only deviate when they believe that they are pivotal, i.e., they believe that their deviation changes the winner of the election.[1] In such a context, identifying the *potential winners* which are the candidates that an agent can make win is essential. However, this mainly depends on the belief of the agents.

Definition 2.2 (Potential winner). A candidate x is a potential winner for agent i at step t, i.e., $x \in PW^t_i$, if, without considering the current ballot $\mathbf{b}^t_i$ of agent i, agent i believes that one more vote in favor of x under plurality or one more veto against another candidate under veto, will make candidate x the new winner.

Observe that the two voting rules under consideration are not symmetric with respect to the set of potential winners. Under plurality, for a given agent, there may be several potential winners other than the current believed winner and it seems rational that

[1] Introducing thresholds to relax the assumption of strict pivot, in the spirit of the works of Meir et al. [11], Obraztsova et al. [14] or Wilczynski [20], also makes sense. We do not make such an assumption for the sake of simplicity and to especially focus on the impact of the social environment of the agents (social network, opinion poll).

the agent will choose to favor the candidate that she prefers. In contrast, under veto, vetoing candidates other than the current believed winner would not produce any direct change according to the belief of an agent. Therefore, there is only one potential winner other than the believed winner, i.e., the one which becomes the new winner after one more veto for the current believed winner. This difference strongly conditions the dynamics of deviations that we consider for each voting rule. While best response deviations are considered under plurality, deviations consisting of vetoing the current believed winner are considered under veto.

Definition 2.3 (Best response deviation (plurality)). A voter i deviates from ballot $\mathbf{b}_i^t$ to ballot $\mathbf{b}_i^{t+1} := y$ at step t following a best response if $y \in PW_i^t \setminus \{\mathcal{F}_P(B_i^t)\}$ and $y \succ_i z$ for any $z \in PW_i^t \setminus \{y\}$.

Definition 2.4 (Veto-winner deviation (veto)). A voter i deviates from ballot $\mathbf{b}_i^t$ to ballot $\mathbf{b}_i^{t+1}$ at step t following a veto-winner deviation if $\mathbf{b}_i^{t+1} = \mathcal{F}_V(B_i^t)$ and $\mathcal{F}_V(B_i^{t+1}) \succ_i \mathcal{F}_V(B_i^t)$.

Both best response and veto-winner dynamics are proved to converge under plurality and veto, respectively, when the social network is complete, i.e., the scores of the current strategy profile are common knowledge [9, 12, 17]. Moreover, convergence is also satisfied when the social network is acyclic or transitive [20].

When the dynamics converges, it reaches a stable state where no voter has an incentive to deviate according to her belief. In this article, we are interested in the identity of the *iterative winner*, i.e., the winner of the stable state reached by the dynamics. We aim to analyze how the polling agency can influence the outcome of the dynamics by manipulating the scores of the initial poll which is communicated to the voters.

2.3 Manipulation by the polling agency

In order for the polling agency not to be detected manipulating the initial poll, it is important that the manipulated poll meets the following criterion, first introduced by Wilczynski [20].

Definition 2.5 (Likelihood condition). A polling vector Δ is *plausible* if $n = \sum_{j=1}^m \Delta_j$ and it gives for each candidate at least the highest score that an agent can observe, i.e., $\Delta_j \geq \max_{i \in N} \mathbf{s}^i_{\mathbf{b}^0}(x_j)$.

Note that checking whether a poll satisfies the likelihood condition is possible in polynomial time.

In this paper, we will study whether the polling agency is able to influence the outcome of the iterative election via the following decision problem for voting rule $\mathcal{F} \in$ {plurality, veto}.

$\mathcal{F}$-{Constructive / Destructive}-Election-Enforcing:
Instance: Instance $(N, M, \succ, G, \triangleright, \tau)$, target candidate p.
Question: Can the polling agency announce a plausible poll Δ so that p {is / is not} the iterative winner?

In reality, the likelihood condition as shown in Definition 2.5 might be too weak and give the polling agency too much power. Especially in cases where some organizations keep an eye on the polling agency or where there have been recent election results, the polling agency should not announce a poll that extremely differs from the correct poll. The motivation is similar to the one presented by Obraztsova and Elkind [13] for optimal manipulation in voting. They propose to bound the manipulative action by some distance, which makes manipulation possibly harder to detect in real-world instances. Therefore, we introduce the following distance-restricted problems, where the *Manhattan distance* between two m-vectors Δ and Δ' is defined as $dist(\Delta, \Delta') = \sum_{i=1}^m |\Delta_i - \Delta'_i|$.

$\mathcal{F}$-Poll-Restricted-{Constr. / Destr.}-Election-Enforcing:
Instance: $(N, M, \succ, G, \triangleright, \tau)$, target candidate p, distance d.
Question: Can the polling agency announce a plausible poll Δ so that p {is / is not} the iterative winner and $dist(\Delta, \mathbf{s}(\mathbf{b}^0)) \leq d$?

Example 2.6. Let us consider an instance with 6 voters and 4 candidates where $G = (N, \{(1, 2), (3, 4)\})$ and $x_3 \triangleright x_2 \triangleright x_1 \triangleright x_4$. The preferences are as follows.

$$\begin{aligned} 1,2,3: &\quad x_1 \succ x_2 \succ x_3 \succ x_4 \\ 4: &\quad x_3 \succ x_1 \succ x_4 \succ x_2 \\ 5,6: &\quad x_1 \succ x_4 \succ x_2 \succ x_3 \end{aligned}$$

Under veto, the truthful winner is x_1. If $\Delta = \mathbf{s}(\mathbf{b}^0)$, there is no deviation: x_1 is the top candidate of all voters except voter 4, but she cannot deviate, otherwise her worst candidate x_2 will be elected. Suppose that the polling agency aims to avoid the election of x_1. By the likelihood condition, it must hold that $\Delta(x_2) \geq 1$, $\Delta(x_3) \geq 1$ and $\Delta(x_4) \geq 2$. If $\Delta = (0, 3, 1, 2)$, then voter 4 believes that x_1 is the winner and x_3 a potential winner. She thinks that she can safely deviate without making x_2 elected, so she deviates for vetoing x_1 and makes x_2 the new winner. Voter 3 observes this deviation and then deviates to veto x_3 that she believes to be the winner. However, x_2 remains the real iterative winner. This is the only successful poll manipulation, thus if the distance to the truthful scores is limited to less than 4, there is no poll-restricted manipulation.

Voters and their current votes are visible for their neighbors. Especially when candidates can be positioned on a spectrum, voters might be inclined to vote for candidates that do not clash with their preference order, either for ideological reasons or because they are worried about what their friends might think of them. Therefore, we introduce the following problem, where the distance between a ballot submitted by agent i approving (resp., vetoing) candidate x under $\mathcal{F}_P$ (resp., $\mathcal{F}_V$) and her truthful ballot is given by the number of swaps between two consecutive candidates in ranking $\succ_i$ that are necessary to put x at the top (resp., bottom) of $\succ_i$.

$\mathcal{F}$-Voter-Restricted-{Constr./Destr.}-Election-Enforcing:
Instance: $(N, M, \succ, G, \triangleright, \tau)$, target candidate p, distance d.
Question: Can the polling agency announce a plausible poll Δ so that p {is / is not} the iterative winner when voters can only submit a ballot at distance at most d from their truthful ballot?

We assume our reader to be familiar with the complexity classes P, NP, para-NP, FPT, and the W-hierarchy, as well as the concepts of polynomial-time many-one reducibility and fpt-reducibility (see, e.g., Papadimitriou [15] and Flum and Grohe [5]).

The winner determination for the considered iterative elections might exceed polynomial time, even for converging elections and acyclic networks. However, for each of the constructed instances in our hardness proofs, the winner determination is possible in polynomial time, therefore proving the intractability of the problems does not depend on the complexity of the winner determination.

3 MANIPULATING POLL PLURALITY SCORES

In this section, we investigate the election enforcing problem under the plurality rule and best response dynamics. It turns out that most of the variants of the problem are intractable, except when the number of candidates is relatively small.

All hardness results in this section hold even when the social network is acyclic and the turn function is constructed so that each voter changes her vote at most once. We use the following NP-complete decision problem to prove our first result. Hitting Set asks—given a universe $X = \{x_1, \dots, x_m\}$, a collection $\mathcal{S} = \{S_1, \dots, S_n\}$ of subsets over X, and a nonnegative integer k—whether there exists a hitting set of size k, i.e., a set $X' \subseteq X$ of size k such that $S \cap X' \neq \emptyset$ for all $S \in \mathcal{S}$. Note that Hitting Set is also W[2]-complete when parameterized by the size of the hitting set k.

Theorem 3.1. *$\mathcal{F}_P$-Destr.-Election-Enforcing is* NP*-hard.*

Sketch of proof. Let $(X, \mathcal{S}, k)$ be an instance of Hitting Set where $X = \{x_1, \dots, x_m\}$ and $\mathcal{S} = \{S_1, \dots, S_n\}$. Without loss of generality, we assume that $k > 3$. Construct an instance of $\mathcal{F}_P$-Destructive-Election-Enforcing as follows:

Let $X \cup Y \cup \{p, z\}$ be the set of candidates, where p is the target candidate and $Y = \{y_0, y_1, \dots y_m\}$. The table below shows the preferences of the voters, partitioned into parts A to F.

Part	Name	Preference	for
A	$a_1:$	$y_0 \succ p \succ \overleftarrow{X} \succ \cdots \succ z$	
	$a_2:$	$y_0 \succ z \succ \overrightarrow{X} \succ \cdots \succ p$	
	$a_3:$	$y_0 \succ p \succ z \succ \overrightarrow{X} \succ \dots$	
	$a_4:$	$p \succ z \succ \overrightarrow{X} \succ \dots$	
B	$b_i:$	$y_i \succ x_i \succ z \succ p \succ \overrightarrow{X \setminus \{x_i\}} \succ \dots$	$1 \leq i \leq m$
C	c_j	$p \succ z \succ \overrightarrow{X} \succ \dots$	$1 \leq j \leq n$
D	$d_j:$	$z \succ p \succ \overrightarrow{X} \succ \dots$	$1 \leq j \leq n$
E	$e_j:$	$p \succ z \succ \overrightarrow{X} \succ \dots$	$1 \leq j \leq k$
F	$f_{i,j}:$	$x_i \succ z \succ p \succ \overrightarrow{X \setminus \{x_i\}} \succ \dots$	$1 \leq i \leq m, 1 \leq j \leq n$

The complete set of arcs in the social network is as follows. There is an arc from a_2 and a_3 to a_1, and an arc from each voter in B to a_2, a_3, and a_4. Each voter c_j in C sees the voters in B corresponding to the variables in S_j, and additionally has an arc to a_2. All voters in D see each voter in B and additionally the voter a_3. The voters in E each have arcs to each voter in B, C, and D, and additionally see the voters a_2 and a_3. Finally, each voter $f_{i,j}$ has an edge to a_1, a_2, and a_3, and each voter $f_{i,n}$ has arcs to the voters $f_{i,1}$ to $f_{i,n-1}$.

We base the turn function on the order $\overrightarrow{A} > \overrightarrow{B} > \overrightarrow{C} > \overrightarrow{D} > \overrightarrow{E} > \overrightarrow{F}$ and use the order $z \rhd p \rhd \overrightarrow{X} \rhd \dots$ for tie-breaking.

The following table shows the correct initial poll Δ the polling agency should announce (line 1), and the minimum number of points the polling agency has to give each candidate in a manipulated poll due to the likelihood condition in Definition 2.5 (line 2). All in all, the polling agency has a contingent of (only) k points.

	p	z	$x \in X'$	$x \in X \setminus X'$	$y \in Y$
Δ	$n + k + 1$	n	n	n	3/1
min	$n + 1$	n	n	n	3/1
Δ'	$n + 1$	n	$n + 1$	n	3/1
final	$n + 1$	$n + k + 1$	$n + 1$	n	$2/\leq 1$

We claim that there is a hitting set of size k, i.e., a set $X' \subseteq X$ of size k so that $X' \cap S \neq \emptyset$ for each $S \in \mathcal{S}$, iff the polling agency can publish a plausible Δ' that results in p not winning the election.

($\Rightarrow$) Suppose $(X, \mathcal{S})$ is a yes-instance of Hitting Set and let X' be a hitting set of size k. The polling agency can publish the manipulated initial poll Δ' as described in the table.

The election then proceeds as follows. Voter a_2 changes her ballot to z, whereas the remaining voters in A cannot achieve a better outcome than the current winner p. The voters in B observe the change from a_2 to z and are now convinced that z is winning due to tie-breaking. The k voters corresponding to an $x_i \in X'$ change their ballot to x_i to give the respective x_i the missing point to win, whereas the other $m - k$ voters in B do not think they can change the outcome to their advantage. The voters in C observe the changes in A and B and—since X' is a hitting set—each sees (at least) one x_i gaining a point, so they react by collectively changing their ballot to z to make z the plurality winner by tie-breaking. The voters in D also think an x_i is currently winning by one point after observing the voters in B, and collectively switch to p. After observing all changes made up to this point, the voters in E switch to z—they observe p winning and losing exactly n points for a total of $n+1$ points, z gaining $n+1$ and losing n points for a total of $n+1$ points, and the $x \in X'$ gaining one point for a total of $n + 2$ points. Finally, none of the voters in F change their ballot because they all see z winning and are unable to reach a more favorable result.

All in all, z wins the election. The final scores can be seen in the last line of the table.

($\Leftarrow$) Suppose that each $X' \subseteq X$ of size at most k is disjoint to an $S \in \mathcal{S}$. That means that it is not possible to convince *all* voters in C to change their ballot from p to another candidate. Due to the space constraints, we omit detailed explanations for each possible manipulated poll. However, regardless of how the manipulated poll is set up, p remains the winner of the election. ❑

Note that the above proof also shows that plurality election enforcing is W[2]-hard for both the poll-restricted and unrestricted constructive variant as well as for the destructive variants when parameterized by the distance between the original and the manipulated initial poll. In the constructive cases, z is the target candidate.

The following theorem shows that even a highly restricted acyclic social network is sufficient to show hardness of manipulation for the restricted problem variants. We use a network where the longest path is of length 1 and—in the voter-restricted problem variant—where the maximum outdegree of a node is 6. Furthermore, in the voter-restricted variant, the voters are only inclined to vote for their two most preferred candidates.

Theorem 3.2. *(1) $\mathcal{F}_P$-Voter-Restricted-{Constr., Destr.}-Election-Enforcing is* para-NP*-hard when parameterized by the number of swaps and the length of a longest path in the network.*
(2) $\mathcal{F}_P$-Poll-Restricted-{Constructive, Destructive}-Election-Enforcing is para-NP*-hard when parameterized by the length of a longest path and the maximum outdegree of the social network.*

Due to space constraints, we omit the proof of this theorem.

Next, we investigate whether manipulation becomes easy if we restrict our allowed instances even further.

Proposition 3.3. *If the winner determination is possible in polynomial time, then $\mathcal{F}_P$-{Unrestricted, Poll-Restricted, Voter-Restricted}-{Constr., Destr.}-Election-Enforcing is in* FPT *when parameterized by the number of candidates m.*

Proof. Construct plausible initial polls for each subset $M' \subseteq M$ of the m candidates in the following way. Set M' corresponds to the initial set of potential winners. For each subset $M^* \subseteq M'$, create a plausible poll Δ' if possible so that $\Delta'(x) = \alpha$ for $x \in M^*$, $\Delta'(x) = \alpha - 1$ for $x \in M' \setminus M^*$, and $\Delta'(x) < \alpha - 2$ for $x \in M \setminus M'$, where α is an integer that can differ from poll to poll. For fixed M' and M^*, each poll meeting these requirements will yield the same election result regardless of the value of α and the exact scores of the candidates in $M \setminus M'$. Note that constructing such a poll (resp., ascertaining that a plausible poll satisfying the requirements does not exist) is possible in polynomial time for all problem variants, as the value of α is bounded by the number of voters. Since we construct at most $2^m \cdot 2^m$ initial polls and testing whether they fulfill our requirements and yield the desired election outcome is possible in polynomial time for each poll, our algorithm is an fpt-algorithm when parameterized by m. ❑

A possible further restriction for the network is an empty graph, i.e., a network where voters only rely on the opinion poll. However, this does not seem to simplify the constructive manipulation problem: it can be necessary to include arbitrary many candidates in the initial set of potential winners. While we conjecture that this problem remains NP-hard for all our considered variants for an empty graph, we can only prove this for the poll-restricted variant, leaving the exact complexity open for the unrestricted and voter-restricted variants. The proof uses a reduction from an NP-complete restricted version of the problem X3C [6], where we are given a universe $X = \{x_1, \ldots, x_{3m}\}$ and a collection $\mathcal{S} = \{S_1, \ldots, S_{3m}\}$, $S_j \subseteq X$, $|S_j| = 3$, so that each $x \in X$ is contained in exactly three sets S_j, and we ask whether there exists an exact cover $\mathcal{S}' \subseteq \mathcal{S}$ of size m so that the union of all sets in $\mathcal{S}'$ equals X.

Theorem 3.4. *$\mathcal{F}_P$-Poll-Restricted-Constructive-Election-Enforcing remains* NP*-hard when the social network is empty.*

Sketch of proof. Let $(X, \mathcal{S})$ be an instance of X3C where $X = \{x_1, \ldots, x_{3m}\}$, $\mathcal{S} = \{S_1, \ldots, S_{3m}\}$ and $S_j = \{x_{j,1}, x_{j,2}, x_{j,3}\}$. Construct an instance of $\mathcal{F}_P$-Poll-Restricted-Constr.-Election-Enforcing as follows. Let $X \cup \mathcal{S} \cup \{w, p, z\}$ be the set of candidates, where p is the target candidate. The following table shows the preferences of the voters, partitioned into the parts A to G.

Part	Name	Preference	for
A	a_i :	$w > x_i > z > \overrightarrow{\mathcal{S}} > \overrightarrow{X \setminus \{x_i\}} > p$	$1 \leq i \leq 3m$
B	$b_{j,1}$:	$x_{j,1} > S_j > z > w > \overrightarrow{\mathcal{S} \setminus \{S_j\}} > \overrightarrow{X \setminus \{x_{j,1}\}} > p$	$1 \leq j \leq 3m$
	$b_{j,2}$:	$x_{j,2} > S_j > z > w > \overrightarrow{\mathcal{S} \setminus \{S_j\}} > \overrightarrow{X \setminus \{x_{j,2}\}} > p$	$1 \leq j \leq 3m$
	$b_{j,3}$:	$x_{j,3} > S_j > z > w > \overrightarrow{\mathcal{S} \setminus \{S_j\}} > \overrightarrow{X \setminus \{x_{j,3}\}} > p$	$1 \leq j \leq 3m$
C	c_i	$x_i > z > w > \overrightarrow{\mathcal{S}} > \overrightarrow{X} > p$	$1 \leq i \leq 3m$
D	d_k :	$p > z > \overleftarrow{X} > \overleftarrow{\mathcal{S}} > w$	$1 \leq k \leq 5$
E	e_k :	$z > w > \overrightarrow{\mathcal{S}} > \overrightarrow{X} > p$	$1 \leq k \leq 3$
F	f_k :	$w > z > \overrightarrow{\mathcal{S}} > \overrightarrow{X} > p$	$1 \leq k \leq 4$
G	g_j :	$S_j > z > w > \overrightarrow{\mathcal{S} \setminus S_j} > \overrightarrow{X} > p$	$1 \leq j \leq 3m$

We use the tie-breaking order $z \rhd w \rhd \overrightarrow{\mathcal{S}} \rhd \overrightarrow{X} \rhd p$ and a maximum allowed distance between the correct and the manipulated initial poll of $3m + 1$.

Note that w is currently winning with a score of $3m + 4$, whereas p only has a score of 5 and cannot gain any points regardless of the broadcasted initial poll because each voter but the ones in D rank p last. The only way to make p win the election is therefore convincing the voters approving of w to approve other candidates, but in a way so that p does not lose approvals and so that each other candidate has a final score of at most 4 due to the tie-breaking.

($\Rightarrow$) Suppose $(X, \mathcal{S})$ is a yes-instance of X3C and let $\mathcal{S}'$ be the exact cover of size m. Assign $3m$ points from w in a way so that each $S \in \mathcal{S}'$ gains three points, and assign one point from p to z. This results in each candidate but the $S \notin \mathcal{S}'$ to have a score of 4 so that the voters think z is winning the election. Then the voters in A collectively change their ballot from w to the respective x_i and the voters in B that correspond to the $S_j \in \mathcal{S}'$ change their vote to S_j. Since $\mathcal{S}'$ is an exact cover, each of the $x \in X$ gain and lose exactly one point. None of the remaining voters change their ballot because they cannot improve the election result. Therefore, in the final result, p has kept a score of 5 whereas each other candidate has a score of at most 4, resulting in p winning the election.

($\Leftarrow$) Suppose that there does not exist an exact cover $\mathcal{S}'$ of size at most m and recall that p cannot gain any points. Then for each plausible initial poll, either w does not lose at least $3m$ points, there is an x that is not covered by $\mathcal{S}'$ and therefore receives a final score of at least 5, or the voters in D collectively change their ballot from p to another candidate, all resulting in p losing the election. We omit the details due to space constraints. ❑

In the destructive case, the manipulation problem becomes easy, at least for the unrestricted and poll-restricted variants. Note that without arcs, winner determination is possible in polynomial time.

Proposition 3.5. *$\mathcal{F}_P$-{Unrestricted, Poll-Restricted}-Destr.-Election-Enforcing is in* P *when the social network is empty.*

Our algorithm that solves the election enforcing instance creates plausible initial polls similar to the way of the algorithm in Proposition 3.3, but only uses pairs of candidates (x, y) as potential winners so that we only construct a polynomial number of polls. Nevertheless, this suffices in an empty graph because for each initial poll, at most half of the voters still vote for the target candidate p. If none of the constructed initial polls yield another winner than p, then it is not possible to make p lose the election. However, for the voter-restricted variant, this proof does not work anymore because we do not know how many voters will change their ballot even when at least half of them prefer x to y. In fact, it can be necessary to include up to all candidates in the initial set of potential winners. Therefore, we conjecture that $\mathcal{F}_P$-Voter-Restricted-Destructive-Election-Enforcing remains NP-hard even when the social network is empty.

4 MANIPULATING POLL VETO SCORES

In this section, we investigate the problems of election enforcing for the polling agency under the veto rule and veto-winner dynamics. Due to the nature of the veto-winner deviations, the results differ a bit from those under the plurality rule. In particular, for each agent at each step, the set of potential winners other than the current believed winner is composed of at most one candidate.

Let us denote by V_x the set of voters vetoing candidate x at the initial truthful profile, i.e., $V_x := \{i \in N : \mathbf{b}_i^0 = x\}$. Observe that if the number of vetoes for candidate x announced by polling vector Δ is not sufficiently large, then the voters in V_x will not deviate at step 0, because they would think that they make their worst candidate x win by removing their veto against it, i.e., $PW_i^0 = \{\mathcal{F}_V(\Delta), x\}$

for $i \in V_x$. Therefore, the global idea of the manipulation of the polling agency under veto is to announce enough vetoes against a candidate whose vetoers must deviate. Let us denote by PW the second best candidate announced by Δ with a score difference of one with the announced winner (advantage w.r.t $\rhd$ included), i.e., PW is a potential winner for all voters $i \in V_x$ such that $x \notin PW_i^0$. The problem of enforcing the election of a given candidate p (respectively, ensuring candidate p does not win the election) is intractable even if the social network is relatively sparse. The proof of the following theorem uses a social network where the longest path is only of length 2 (respectively, of length 1 for the voter-restricted variant). Furthermore, in the voter-restricted variant, the voters can even only veto their two least preferred candidates.

Theorem 4.1. *$\mathcal{F}_V$-{Unrestricted, Poll-Restricted, Voter-Restricted}-{Constructive, Destructive}-Election-Enforcing is para-*NP*-hard when parameterized by the length of the longest path and—for the voter-restricted variant—by the number of swaps.*

Sketch of proof. We first prove that $\mathcal{F}_V$-Destructive-Election-Enforcing is NP-hard even when the longest path is of length 2. The hardness of the poll-restricted variant with the same parameters and the hardness of the constructive variants immediately follow. We just need to set the maximum Manhattan distance between the original and the manipulated initial poll to at least $6m + 2$ and—in the constructive variants—the target candidate to z. After the proof, we give a slight modification for the voter-restricted variant that reduces the length of the longest path to 1.

We reduce from X3C. Let $(X, \mathcal{S})$ be an instance of X3C where $X = \{x_1, \ldots, x_{3m}\}$ and $\mathcal{S} = \{S_1, \ldots, S_{3m}\}$ so that $S_j \subseteq X$, $|S_j| = 3$, and each $x \in X$ is contained in exactly three sets $S \in \mathcal{S}$. Construct an instance of $\mathcal{F}_V$-Destructive-Election-Enforcing as follows.

Let $\mathcal{S} \cup \{p, z, d_1, d_2\}$ be the set of candidates, where p is the target candidate. The table below shows the preferences of the voters, partitioned into parts A to G.

Part	Name	Preference	for
A	a_k :	$d_1 \succ p \succ \cdots \succ z \succ d_2$	$1 \leq k \leq m$
	a_{m+1} :	$d_1 \succ z \succ \overrightarrow{\mathcal{S}} \succ p \succ d_2$	
B	b_j :	$\cdots \succ \overrightarrow{\mathcal{S} \setminus \{S_j\}} \succ p \succ z \succ S_j$	$1 \leq j \leq 3m$
C	c_i	$\cdots \succ \overrightarrow{\mathcal{S}} \succ p \succ z$	$1 \leq i \leq 3m$
D	d_k :	$p \succ \cdots \succ d_2 \succ z$	$1 \leq k \leq m$
E	e_j :	$p \succ \cdots \succ z \succ d_2 \succ S_j$	$1 \leq j \leq 3m$
F	f_k :	$p \succ \cdots \succ z \succ d_1 \succ d_2$	$1 \leq k \leq 9m - 1$
G	g_k :	$p \succ \cdots \succ z \succ d_2 \succ d_1$	$1 \leq k \leq 10m$

The social network has the following set of arcs. Each voter c_i in C has an arc to a_1, a_{m+1}, and to each of the three voters in B that correspond to the sets S_j that contain x_i. Furthermore, voter d_m sees each of the other voters in D, the voters in E see a_1 and additionally e_{3m} sees each of the other voters in E, and the voters in F and G have an arc to each voter in A and additionally f_{9m-1} (resp., g_{10m}) has an arc to each of the other voters in F (resp. G).

We base the turn function on the order $\overrightarrow{A} > \overrightarrow{B} > \overrightarrow{C} > \overrightarrow{D} > \overrightarrow{E} > \overrightarrow{F} > \overrightarrow{G}$ and use the order $z \rhd p \rhd \overrightarrow{\mathcal{S}} \rhd \ldots$ for tie-breaking.

The following table shows the correct initial poll Δ the polling agency should announce (line 1), and the minimum number of vetoes the polling agency has to give each candidate in a manipulated poll due to the likelihood condition in Definition 2.5 (line 2).

	p	z	$S \in \mathcal{S}'$	$S \notin \mathcal{S}$	d_1	d_2
Δ	0	$4m$	$3m + 1$	$3m + 1$	$10m$	$10m$
min	0	m	$3m$	$3m$	$10m$	$10m$
Δ'	$3m$	$3m$	$3m + 1$	$3m$	$10m$	$10m$
final	$3m$	$3m$	$3m$	$3m + 1$	$10m$	$9m$

($\Rightarrow$) Suppose $(X, \mathcal{S})$ is a yes-instance of X3C and let $\mathcal{S}' \subseteq \mathcal{S}$ be the exact cover of size m. The polling agency can publish the poll Δ' described in the table. The election then proceeds as follows. All voters think that z is the winner. Therefore, voters a_1 to a_m in A change their ballot to z to make p the winner of the election by tie-breaking. Voters b_j also want to hinder z from winning. However, they only veto z in the case that S_j is part of the exact cover, because otherwise the loss of a veto for S_j would result in S_j being the veto winner with only $3m - 1$ vetoes. Since $\mathcal{S}'$ is an exact cover, each voter in C observes z gaining two vetoes—one from a_1 and one from a voter in C—and reacts by vetoing p. This is possible because z now has enough vetoes to not become the veto winner after losing a veto. None of the voters in D, E, and F change their ballot. All in all, z gains $2m$ vetoes from the voters in A and B and loses $3m$ vetoes from the voters in C, resulting in z winning the election with $3m$ vetoes due to tie-breaking (see last line of the table).

($\Leftarrow$) Suppose that there is no exact cover $\mathcal{S}'$ of size at most m. Then, regardless of the initial poll, there is at least one voter in C who does not change her veto to p so that p does not obtain the necessary number of vetoes to lose the election. We omit the details due to space constraints.

For the voter-restricted variant, the depicted proof obviously works (for a maximum number of $3m + 2$ swaps), but we can even strengthen the result by tightening the parameters: Set the maximum number of swaps to 1, i.e., only allow the voters to veto their two least preferred candidates. Delete the arcs between parts E and A, F and A, as well as G and A. The resulting social network has a longest path of length 1. ❑

However, the manipulation problem can be solved efficiently if the number of candidates is small.

Proposition 4.2. *If the winner determination is possible in polynomial time, then $\mathcal{F}_V$-{Unrestricted, Poll-Restricted, Voter-Restricted}-{Constructive, Destructive}-Election-Enforcing is in* FPT *when parameterized by the number of candidates m.*

The proof works analogously to the proof of Proposition 3.3.

In contrast to plurality, manipulation is easy under veto when the social network is empty, even in the constructive case.

Proposition 4.3. *$\mathcal{F}_V$-Constructive-Election-Enforcing is solvable in polynomial time when the social network is empty.*

Sketch of proof. The idea of our algorithm is to communicate a polling vector Δ that makes the voters removing vetoes against p, or that prevents many deviations from agents vetoing other candidates. Since the set of potential winners other than the current believed winner is composed of at most one candidate, we try all the $O(m^2)$ combinations of pairs of distinct candidates (ω, y) such that ω is the announced winner of Δ and y is the other announced potential winner PW. For any pair of candidates (ω, y), we create a polling vector Δ which fulfills the likelihood condition (Def. 2.5) with a minimum number of vetoes, and then we add the minimum

number of vetoes in order to get ω and y the winner and PW of Δ, respectively. The rest of available vetoes is distributed as follows:
• Case $p \notin \{\omega, y\}$: If p is not at least the second winning candidate in $\mathbf{b}^0$, then vetoes against p must be removed, so we add in Δ the minimum number of vetoes to p in order to "authorize" V_p, i.e., in order to have $p \notin PW_i^0$ (and thus $y \in PW_i^0$) for every $i \in V_p$. Otherwise, we test the two options: authorize V_p or not (still polynomial). With the remaining available vetoes, we "block", as much as possible, the deviation of voters V_x for all other candidates $x \notin \{\omega, y\}$ such that $|V_x \cap N_{\omega > y}| + \mathbb{1}_{\{p \rhd x\}} < \mathbf{s}_{\mathbf{b}^*}(p)$ where $\mathbf{s}_{\mathbf{b}^*}(p) = |V_p|$ if V_p is not authorized or $\mathbf{s}_{\mathbf{b}^*}(p) = |V_p \cap N_{\omega > y}|$ otherwise. The goal is to avoid that the final score of another candidate is lower than the final score of p. For blocking voters V_x, we add the minimum number of vetoes to candidates other than x in order to make a voter $i \in V_x$ believe that $x \in PW_i^0$. If some available vetoes remain, we use them to authorize as much as possible the other voters V_z for all candidates $z \notin \{\omega, y\}$ such that $|V_z \cap N_{\omega > y}| + \mathbb{1}_{\{p \rhd z\}} > |V_p|$, by choosing first the candidates which maximize $|V_z \cap N_{y > \omega}|$.
• Case $y = p$: V_p is already blocked, therefore no veto against p can be removed. To become the iterative winner, p must be the second best in $\mathbf{b}^0$ and the deviations must add enough vetoes against ω, which must be $\mathcal{F}_V(\mathbf{b}^0)$, while the deviations of V_x must be blocked if x could have a smaller score than p. Therefore, we block V_x for all candidates $x \notin \{\omega, y\}$ such that $|V_x \cap N_{\omega > y}| + \mathbb{1}_{\{p \rhd x\}} < |V_p|$. Then, we authorize as much as possible the other voters V_z for all candidates $z \notin \{\omega, y\}$ such that $|V_z \cap N_{\omega > y}| + \mathbb{1}_{\{p \rhd z\}} > |V_p|$ by choosing first the candidates which maximize $|V_z \cap N_{y > \omega}|$.
• Case $\omega = p$: The only possible deviations are vetoes against p and no veto against p can be removed. Therefore, it must hold that p is the actual winner, i.e., $p = \mathcal{F}_V(\mathbf{b}^0)$, and the deviations must be limited as much as possible. We block as many V_x as possible for candidates $x \notin \{\omega, y\}$ by choosing first the candidates which minimize $\min\{\min_{z \neq p}(|V_z| + \mathbb{1}_{\{p \rhd z\}}); |V_x \cap N_{\omega > y}| + \mathbb{1}_{\{p \rhd x\}}\} - |V_x \cap N_{y > \omega}|$. If at some point there are not enough available vetoes to block another set of voters V_x, then we assign the rest of available vetoes to candidate x which maximizes this quantity. ❑

The idea of the algorithm for the destructive variant is similar. The details are omitted due to space restrictions.

Proposition 4.4. *F_V-Destructive-Election-Enforcing is solvable in polynomial time when the social network is empty.*

Note that the algorithms of Propositions 4.3 and 4.4 can be simply adapted for taking into account restrictions in voter manipulations. For the poll-restricted variant, the principle of the algorithms remains the same. The only specific point is how a candidate y is made a (potential) winner. Instead of starting from the minimal vector of scores which satisfies the likelihood condition, we start from the truthful scores. If the imposed distance is d, then we can change a veto to one candidate from another at most $\lfloor d/2 \rfloor$ times. We switch vetoes against y to other candidates x with a smaller score with a priority to those for which V_x must be authorized to deviate.

Corollary 4.5. *$\mathcal{F}_V$-{Poll-Restr., Voter-Restr.}-{Constr., Destr.}-Election-Enforcing is in* P *when the network is empty.*

5 REAL POLL MANIPULATION: HEURISTICS

Most of our results are complexity results stating that, in the worst case, it may be hard for the polling agency to manipulate. However, it does not prevent manipulation to occur in practice. We thus examine some heuristics for the unrestricted problem of election enforcing and test them by running experiments. All our heuristics follow the same principle: we test all pairs of distinct candidates (ω, y) for announcing them the winner and another potential winner of Δ, respectively, following a given order. The order of test for pairs of candidates varies according to the variant of election enforcing and the voting rule, as described below (we omit further details due to space restrictions):
• Plurality / constructive: We refine a little the heuristic proposed by Wilczynski [20]. The order over pairs is such that $(\omega, y) \geq (\omega', y')$ where $\omega \neq p$ and $\omega' \neq p$ (target candidate p should not lose points) if $y = p$ and $y' \neq p$, or if $y = y' = p$ and $|N_{y > \omega}| \geq |N_{y > \omega'}|$, or if $y \neq p$ and $y' \neq p$ and $|N_{y > \omega}| \leq |N_{y' > \omega'}|$. In such a way, we favor configurations where p can get more points.
• Plurality / destructive: The order is such that $(\omega, y) \geq (\omega', y')$ if $|N_{y > \omega}| \geq |N_{y' > \omega'}|$ where $p \notin \{y, y'\}$ (target candidate p should not get more points). In such a way, we favor configurations where many voters will deviate to favor potential winner y.
• Veto / constructive: The order is such that $(\omega, y) \geq (\omega', y')$ if $|V_p \cap N_{y > \omega}| \geq |V_p \cap N_{y' > \omega'}|$ where $p \notin \{\omega, \omega'\}$ (deviations should not add more vetoes to target candidate p). In such a way, we favor configurations where more vetoes will be removed from p.
• Veto / destructive: For a given pair (ω, y), let x be the candidate which minimizes $\mathbf{s}^*(x) = |V_x \cap N_{\omega > y}|$. The order over pairs is such that $(\omega, y) \geq (\omega', y')$ if $p \notin \{x, x'\}$ and $\mathbf{s}^*(x) \leq \mathbf{s}^*(x')$, or if $\omega = p$ and $\omega' \neq p$, or if $\omega = \omega' = p$ and $|N_{y > \omega} \setminus V_p| \geq |N_{y' > \omega'} \setminus V_p|$. In such a way, we favor configurations where a candidate $x \neq p$ can lose many vetoes, or where many voters will deviate by vetoing p.

We test our heuristics by running 1,000 instances of the poll-confident iterative model with 50 agents and 5 candidates. The preference rankings of the agents are drawn from the impartial culture and the social network is supposed to be acyclic (in order to ensure convergence, to not limit too much manipulation and because our problems are hard for this class of graphs).

We compare the results of heuristics with the results given by the dynamics without manipulation from the polling agency and the results given by the exact algorithm where all possible manipulations of the polling agency that satisfy the likelihood condition are tested. We measure the frequency of election (for the constructive variant) of the target candidate as the iterative winner, or the frequency of non-election (for the destructive variant) of the target candidate as the iterative winner, according to the three different algorithms.

In order to create more challenge for the heuristics, the target candidates for the constructive variant are "bad" candidates: the Condorcet loser, i.e., the candidate which is beaten by all the other candidates in pairwise comparisons (we restrict in this case to a domain where such a candidate exists), or the Borda loser, i.e., the candidate with the lowest Borda score,[2] or the truthful loser, i.e., the candidate with the lowest (resp., highest) score under plurality (resp., veto). In the same vein, the target candidates for the

[2] For computing the Borda score of candidate x, we add $m - j$ points to x for each voter i if x is the j^{th} most preferred candidate of voter i.

destructive variant are “good” candidates: the Condorcet winner, i.e., the candidate which beats all the other candidates in pairwise comparisons (we restrict in this case to a domain where such a candidate exists), or the Borda winner, i.e., the candidate with the highest Borda score, or the truthful winner, i.e., the candidate with the highest (resp., lowest) score under plurality (resp., veto).

The results concerning both variants are presented in Figure 1.

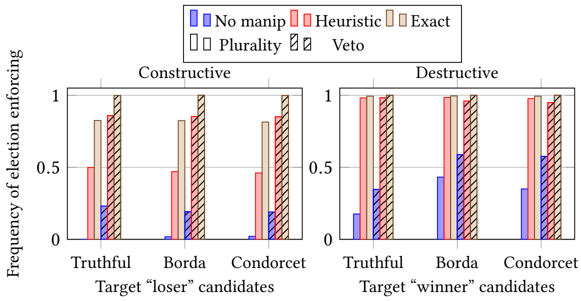


Figure 1: $\mathcal{F}$-{Constr., Destr.}-Election-Enforcing with no poll manipulation, heuristic or exact poll manipulation for different target candidates (truthful/Borda/Condorcet winner/loser) under $\mathcal{F} \in \{plurality, veto\}$ in an acyclic social network for $n = 50$ and $m = 5$.

It turns out that our heuristics for the destructive variant perform very well: the frequency of non-election of the target candidate p is very high and extremely close to the frequency with the exact algorithm. For the constructive variant, the frequency of election of p is very close under veto but the performance of our heuristic is a little bit lower under plurality. Nevertheless, it is always closer to the result of the exact algorithm than to the result where no poll manipulation occurs. This can be explained by the structure of the potential winners set under plurality: in our heuristic we only choose one potential winner to announce as a challenger of the announced winner whereas it could be cleverer to announce as potential winners an appropriate set of candidates.

It seems that even the results with the exact algorithm differ according to the variant of manipulation and the voting rule. In order to have a deeper understanding of this phenomenon, we run further experiments with the exact algorithm where the setting of simulations is the same as previously, except that we vary the number of agents from 10 to 50. The results are presented in Figure 2.

From the results given in Figure 2, two main conclusions can be drawn: (1) the polling agency can successfully manipulate more often for avoiding the election of a candidate than for making a candidate elected, i.e., the frequency of election enforcing is clearly higher for the destructive variant than for the constructive variant, and (2) the polling agency can successfully manipulate more often under veto than under plurality. The highest frequency of successful manipulation occurs for the destructive variant under veto, which seems natural regarding the nature of this voting rule under which a ballot means a disapproval for one candidate.

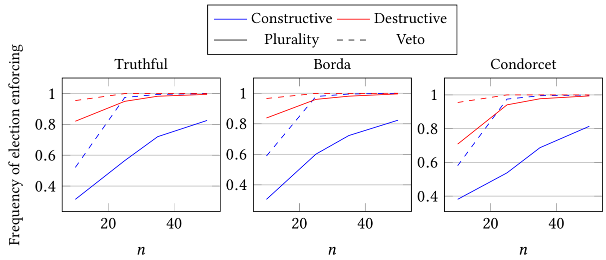


Figure 2: $\mathcal{F}$-{Constr., Destr.}-Election-Enforcing with exact poll manipulation for different target candidates (truthful/Borda/Condorcet winner/loser) under $\mathcal{F} \in \{plurality, veto\}$, in an acyclic social network for $m = 5$.

6 CONCLUSIONS

We have examined the manipulative power of a polling agency announcing preliminary results before an election. The polling agency may manipulate with two different goals in mind: making a given candidate elected (constructive variant) or avoiding the election of a given candidate (destructive variant). However, the polling agency is not totally free regarding how it can manipulate: the announced scores should not be too far from reality to be trusted by voters. Moreover, voters may have a local information by their relatives in a social network, limiting the manipulative power of the polling agency. Our results are summarized in the table below.

Manip.	Variant	Plurality		Veto	
		Acyclic network	Empty network	Acyclic network	Empty network
Constr.	Unrestr.	NP-h ([20]) FPT w.r.t. m (Prop. 3.3)	?	NP-h (Th. 4.1) FPT w.r.t. m (Prop. 4.2)	P (Prop. 4.3)
	Poll-restr.	NP-h (Th. 3.2) FPT w.r.t. m (Prop. 3.3)	NP-h (Th. 3.4)	NP-h (Th. 4.1) FPT w.r.t. m (Prop. 4.2)	P (Cor. 4.5)
	Voter-restr.	NP-h (Th. 3.2) FPT w.r.t. m (Prop. 3.3)	?	NP-h (Th. 4.1) FPT w.r.t. m (Prop. 4.2)	P (Cor. 4.5)
Destr.	Unrestr.	NP-h (Th. 3.1) FPT w.r.t. m (Prop. 3.3)	P (Prop. 3.5)	NP-h (Th. 4.1) FPT w.r.t. m (Prop. 4.2)	P (Prop. 4.4)
	Poll-restr.	NP-h (Th. 3.2) FPT w.r.t. m (Prop. 3.3)	P (Prop. 3.5)	NP-h (Th. 4.1) FPT w.r.t. m (Prop. 4.2)	P (Cor. 4.5)
	Voter-restr.	NP-h (Th. 3.2) FPT w.r.t. m (Prop. 3.3)	?	NP-h (Th. 4.1) FPT w.r.t. m (Prop. 4.2)	P (Cor. 4.5)

When the voters have no local information through the social network, manipulating is easier for the polling agency, especially under veto. Although the manipulative power of the polling agency is mainly computationally limited in theory, we designed efficient heuristics. They perform better for the destructive variant under veto. More generally, it seems that the two variants of manipulation and the two voting rules we consider are not symmetric: the polling agency is more successful in the destructive than in the constructive case, and manipulation is more successful under veto than under plurality. This work can be extended in several directions. Considering more complex voting rules which require the submission of a ranking in ballots could be a challenging perspective. Investigating preference restrictions such as single-peaked preferences could also make sense, as well as supposing that the polling agency only gets partial information about the preferences of the voters.

ACKNOWLEDGMENTS

This work was supported in part by DFG grants BA 6270/1-1 and BR 2312/12-1.

REFERENCES

[1] John J. Bartholdi III, Craig A. Tovey, and Michael A. Trick. 1989. The Computational Difficulty of Manipulating an Election. *Social Choice and Welfare* 6, 3 (1989), 227–241.

[2] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel Procaccia (Eds.). 2016. *Handbook of Computational Social Choice*. Cambridge University Press.

[3] Ulle Endriss, Svetlana Obraztsova, Maria Polukarov, and Jeffrey S. Rosenschein. 2016. Strategic Voting with Incomplete Information. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*. AAAI Press / IJCAI, New York, NY, USA, 236–242.

[4] Piotr Faliszewski, Rica Gonen, Martin Koutecký, and Nimrod Talmon. 2018. Opinion Diffusion and Campaigning on Society Graphs. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-18)*. IJCAI, Stockholm, Sweden, 219–225.

[5] Jörg Flum and Martin Grohe. 2006. *Parameterized Complexity Theory*. Springer-Verlag.

[6] Teofilo F. Gonzalez. 1985. Clustering to Minimize the Maximum Intercluster Distance. *Theoretical Computer Science* 38 (1985), 293–306.

[7] Laurent Gourvès, Julien Lesca, and Anaëlle Wilczynski. 2016. Strategic Voting in a Social Context: Considerate Equilibria. In *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI-16)*. IOS Press, The Hague, The Netherlands, 1423–1431.

[8] Kathrin Konczak and Jérôme Lang. 2005. Voting Procedures with Incomplete Preferences. In *Proceedings of the Multidisciplinary IJCAI-05 Worshop on Advances in Preference Handling*. Edinburgh, Scotland, 124–129.

[9] Omer Lev and Jeffrey S Rosenschein. 2016. Convergence of Iterative Scoring Rules. *Journal of Artificial Intelligence Research* 57 (2016), 573–591.

[10] Reshef Meir. 2017. Iterative Voting. In *Trends in Computational Social Choice*, Ulle Endriss (Ed.). AI Access, Chapter 4, 69–86.

[11] Reshef Meir, Omer Lev, and Jeffrey S. Rosenschein. 2014. A Local-Dominance Theory of Voting Equilibria. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC-14)*. ACM, Palo Alto, California, USA, 313–330.

[12] Reshef Meir, Maria Polukarov, Jeffrey S. Rosenschein, and Nicholas R. Jennings. 2010. Convergence to Equilibria in Plurality Voting. In *Proceedings of the 24th Conference on Artificial Intelligence (AAAI-10)*. AAAI Press, Atlanta, Georgia, USA, 823–828.

[13] Svetlana Obraztsova and Edith Elkind. 2012. Optimal Manipulation of Voting Rules. In *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-12)*. IFAAMAS, Valencia, Spain, 619–626.

[14] Svetlana Obraztsova, Omer Lev, Maria Polukarov, Zinovi Rabinovich, and Jeffrey S. Rosenschein. 2016. Nonmyopic Voting Dynamics: An Optimistic Approach. In *Proceedings of the 10th Multidisciplinary Workshop on Advances in Preference Handling (MPREF-16)*. New York, NY, USA.

[15] Christos Papadimitriou. 1995. *Computational Complexity* (second ed.). Addison-Wesley.

[16] Annemieke Reijngoud and Ulle Endriss. 2012. Voter Response to Iterated Poll Information. In *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-12)*. IFAAMAS, Valencia, Spain, 635–644.

[17] Reyhaneh Reyhani and Mark C. Wilson. 2012. Best Reply Dynamics for Scoring Rules. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI-12)*. IOS Press, Montpellier, France, 672–677.

[18] Sigal Sina, Noam Hazon, Avinatan Hassidim, and Sarit Kraus. 2015. Adapting the Social Network to Affect Elections. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-15)*. IFAAMAS, Istanbul, Turkey, 705–713.

[19] Alan Tsang and Kate Larson. 2016. The Echo Chamber: Strategic Voting and Homophily in Social Networks. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-16)*. IFAAMAS, Singapore, 368–375.

[20] Anaëlle Wilczynski. 2019. Poll-Confident Voters in Iterative Voting. In *Proceedings of the 33rd Conference on Artificial Intelligence (AAAI-19)*. AAAI Press, Honolulu, Hawaii, USA, 2205–2212.

[21] Bryan Wilder and Yevgeniy Vorobeychik. 2018. Controlling Elections through Social Influence. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-18)*. IFAAMAS, Stockholm, Sweden, 265–273.

[22] Bryan Wilder and Yevgeniy Vorobeychik. 2019. Defending Elections against Malicious Spread of Misinformation. In *Proceedings of the 33rd Conference on Artificial Intelligence (AAAI-19)*. AAAI Press, Honolulu, Hawaii, USA, 2213–2220.

Chapter 6

Complexity of Control in Judgment Aggregation for Uniform Premise-Based Quota Rules

This chapter deals with the complexity of influencing the judgment aggregation procedures called uniform (constant) premise-based quota rules by adding, deleting, replacing, and bundling judges. See Section 2.3 for a short introduction to judgment aggregation.

The attached article (Baumeister et al., 2020a) extends several conference versions, including an article I coauthored (Baumeister et al., 2015d).

Baumeister, D., Erdélyi, G., Erdélyi, O. J., Rothe, J., and Selker, A.-K. (2020a). Complexity of control in judgment aggregation for uniform premise-based quota rules. *Journal of Computer and System Sciences*, 112:13–33

Summary

In the attached article, my coauthors and I introduce the notions of adding, deleting, replacing and bundling judges in judgment aggregation. These concepts can, among others, be found in international arbitration procedures.

There are several ways to measure the success of a control action employed by the chair. Here, we ask whether the chair can achieve to include a given (possibly incomplete) judgment set called *desired set* in the new outcome. We call this problem *exact control by control type* $\mathcal{C}$. Further, we use the preference types introduced by Dietrich and List (2007c)—defined in Definition 2.25 in this thesis and adapted to incomplete desired sets by Baumeister et al. (2015b)—to identify better outcomes for the chair. This approach is based on the concept in preference aggregation where an attacker's preference over the candidates is part of the input. However, note that it is not feasible in judgment aggregation to state an explicit preference over different judgment sets since such a preference is exponential in the size of the agenda. Therefore, we assume that the chair's preference over the outcomes belongs to the set of one of the following four preference types: For the set of

- *unrestricted preferences*, we only know that the chair does not differentiate between two different outcomes that both include the desired set;
- *top-respecting preferences*, we additionally know that the chair prefers outcomes that include the desired set over outcomes that do not;
- *closeness-respecting preferences*, it holds that outcomes that contain a subset X of the desired set are preferred to outcomes that only contain a strict subset of X (and no other issues from the desired set);
- *Hamming-distance* preferences we know that the chair prefers outcomes that include more issues from the desired set over outcomes with fewer.

We obtain the following results for the uniform (constant) premise-based quota rules,[1] where we impose certain restrictions on the agenda's premises and conclusions to obtain complete and consistent versions of these judgment aggregation procedures. For the uniform *constant* premise-based quota rules, possible and necessary control by adding, deleting, and replacing judges is NP-complete for each admissible quota q and for exact control as well as nearly all preference types. The only exception is control under unrestricted preferences since the uniform constant premise-based quota rules are immune to control in this case. However, in the presence of a complete desired set, possible control under unrestricted and top-respecting preferences becomes tractable, whereas the complexity of exact control and necessary control under top-respecting preferences is unknown. We do not consider the complexity of bundling judges because this control type does not make sense for uniform constant premise-based quota rules. Further, note that since for control by replacing judges the number of judges remains constant throughout the control action, the results for control by replacing judges also hold for the case of uniform premise-based quota rules.

[1] Given n judges, a quota $q \in \mathbb{N}$, $0 \leq q < n$, and a partition of the agenda into premises and conclusions, the collective outcome under the uniform constant premise-based quota rule for quota q consists of the positive premises that more than q judges accept (and the negations of the remaining premises) as well as the conclusions derived from the included premises. The uniform premise-based quota rules are defined in Definition 2.23 on page 38.

In the case of uniform premise-based quota rules, the results are similar. Possible and necessary control by adding, deleting, and bundling judges is NP-complete for the quota $q = 1/2$ and for exact control as well as nearly all preference types. The only exceptions are again control under unrestricted preferences since the uniform premise-based quota rules are immune to control in this case, and possible control for unrestricted and top-respecting preferences since control is tractable in these cases. The results also hold for a complete desired set, but apart from the immunity and the tractability results, the exact complexity of control for different quotas remains an open problem.

My Contribution

The writing of the attached article was done jointly with my coauthors. I was responsible for the examples (Example 2, Example 6, and Example 7), Definition 8, the results in Section 4 (i.e., Lemmas 9–11 and Propositions 12–14), Theorem 15, Theorem 16, Theorem 18, Theorem 19, Theorem 21, and Theorem 22.

Journal of Computer and System Sciences 112 (2020) 13–33

Contents lists available at ScienceDirect

Journal of Computer and System Sciences

www.elsevier.com/locate/jcss

Complexity of control in judgment aggregation for uniform premise-based quota rules ☆

Dorothea Baumeister [a], Gábor Erdélyi [b], Olivia J. Erdélyi [c], Jörg Rothe [a,*], Ann-Kathrin Selker [a]

[a] Institut für Informatik, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany
[b] School of Mathematics and Statistics, University of Canterbury, Christchurch 8140, New Zealand
[c] School of Law, University of Canterbury, Christchurch 8140, New Zealand

A R T I C L E I N F O

Article history:
Received 23 December 2016
Received in revised form 19 July 2019
Accepted 6 March 2020
Available online 9 March 2020

Keywords:
Judgment aggregation
Control
Computational complexity

A B S T R A C T

The task of aggregating individual judgments over logically interconnected propositions is called judgment aggregation. Manipulation of judgment aggregation procedures has first been studied by List [45] and Dietrich and List [27], and Endriss et al. [30] were the first to study it from a computational perspective. Baumeister et al. [7] extended their results on manipulation and introduced the concept of bribery in judgment aggregation, again focusing on algorithmic and complexity-theoretic properties. Complementing this previous work on strategic scenarios, we introduce the concept of control in judgment aggregation, making use of the preference types introduced by Dietrich and List [27] and studying the class of uniform premise-based quota rules for these control problems in terms of their computational complexity.

1. Introduction

Judgment aggregation is a framework for collective decision making where a number of agents (called judges) aggregate their judgments on logically possibly interconnected propositions in order to determine a collective outcome that appropriately reflects their individual judgments as a whole. The field of judgment aggregation dates back to the work of Kornhauser and Sager [44] who, motivated by legal issues in court proceedings, were the first to discover the *doctrinal paradox*, which we will describe in the next section (see Example 2). Since then, judgment aggregation has been studied from various per-

☆ This paper combines and extends the results regarding control for uniform premise-based quota rules from preliminary conference versions that appear in the proceedings of the *3rd* and *4th International Conference on Algorithmic Decision Theory* (ADT'13 and ADT'15) [6,13], of the *6th European Starting AI Researcher Symposium* (STAIRS'12) [5], and of the *4th International Workshop on Computational Social Choice* (COMSOC'12) [4]. Some early results were presented at the *ESSLLI Workshop on Logical Models of Group Decision Making* (ESSLLI-LMGD'13). This article extends the previous work by adding results about necessary/possible control by bundling judges for closeness-respecting preferences, by adding results regarding Hamming-distance-induced preferences and exact control for complete desired sets, by adding generic reductions for Hamming-distance-induced preferences and exact control, and by adding discussion and examples.

* Corresponding author.
E-mail address: rothe@hhu.de (J. Rothe).
URLs: https://www.cs.hhu.de/lehrstuehle-und-arbeitsgruppen/computational-social-choice/unser-team/baumeister (D. Baumeister), https://www.canterbury.ac.nz/engineering/schools/mathematics-statistics/contact-us/ (G. Erdélyi), https://www.canterbury.ac.nz/law/people/academic-staff/olivia-erdelyi.html (O.J. Erdélyi), https://ccc.cs.uni-duesseldorf.de/~rothe (J. Rothe), https://ccc.cs.uni-duesseldorf.de/~selker (A.-K. Selker).

https://doi.org/10.1016/j.jcss.2020.03.001
0022-0000/

spectives, including legal science, mathematical and philosophical logic, and—more recently—computer science, as surveyed by List and Puppe [47], List [46], Endriss [29], and Baumeister et al. [10,14]. In particular, judgment aggregation has recently evolved to become one of the emerging subfields of computational social choice, which is also concerned with voting theory, coalition formation in cooperative games, fair division, matching under preferences, and other subfields (see the books edited by Brandt et al. [17] and Rothe [57]).

In judgment aggregation, strategic behavior has been studied to a far lesser extent than in voting so far, which is the main motivation for our work. Only starting with the work of List [45], Dietrich and List [27], and Endriss et al. [30] (which subsumes and extends several conference papers) on manipulation in judgment aggregation, the focus of attention has recently shifted toward the study of strategic scenarios where the involved agents seek to influence the outcome of a judgment aggregation procedure to their advantage (for an overview, see, e.g., the book chapter by Baumeister et al. [14]). This line of research follows and runs parallel with the preceding work on manipulative scenarios in voting,[1] carefully taking the differences between voting and judgment aggregation into account. In particular, Dietrich and List [27] established an analogue of the famous Gibbard–Satterthwaite theorem (which, informally, says that every reasonable voting rule is manipulable) in judgment aggregation, Endriss et al. [30] studied manipulation scenarios in judgment aggregation in terms of their computational complexity, and their results have been expanded by Baumeister et al. [7] (also subsuming and extending several previous conference papers), who in addition initiated the study of bribery in judgment aggregation, again focusing on the computational complexity of the associated problems. What has sorely been missing to date is a fundamental study of *control in judgment aggregation*—note that the three basic types of manipulative action studied in voting are manipulation, control, and bribery (as we will explain in more detail below). In a nutshell, we will study the computational complexity of problems formalizing *control by adding, deleting, replacing, and bundling judges*. In Section 3, we will motivate these control actions by real-world examples that can be found in European legislation.

Note that, even though judgment aggregation as well as voting procedures are often susceptible to certain kinds of influence, computational complexity can serve as a shield to protect against—or at least hinder—undesirable strategic behavior. In this paper, we will employ NP-hardness as such a complexity barrier to shield judgment aggregation against undesirable strategic behavior, just as NP-hardness has been used to protect elections against such behavior in voting (Footnote 1 collects a number of related book chapters and surveys that cover a vast body of literature on NP-hardness of manipulation, control, and bribery problems in voting). However, one should keep in mind that since NP-hardness is a worst-case complexity measure only, the protection it provides is rather limited. For example, Rothe and Schend [58] have surveyed various challenges to such worst-case complexity shields against manipulation and control in voting. Still, it should be the first step to study whether problems that formalize manipulative behavior can outright be solved in polynomial time or whether they are NP-hard, and this is what we do here for control in judgment aggregation.

Manipulation in voting has been introduced by Bartholdi et al. [2] to model scenarios where some voter acts strategically by casting an insincere vote so as to make some most preferred candidate win. Conitzer et al. [21] considered more general manipulation scenarios, such as weighted coalitional manipulation, and the destructive variants where the goal is to prevent some least preferred candidate's victory (see, e.g., the book chapter by Conitzer and Walsh [22] for an overview of the vast literature on manipulation in voting). *Bribery in voting* is due to Faliszewski et al. [32] (see also the work of Faliszewski et al. [35] and, e.g., the book chapter by Faliszewski and Rothe [37] for an overview). Here, an external agent seeks to bribe the voters so as to make some favorite candidate win (in the constructive case) or prevent some despised candidate's victory (in the destructive case). In a third central type of strategic influence, *control in voting*, an external agent (usually called the "chair") seeks to change the structure of an election (e.g., by adding/deleting/partitioning either candidates or voters) in order to reach her desired outcome. Electoral control has been introduced by Bartholdi et al. [3] in the constructive variant and by Hemaspaandra et al. [41] in the destructive variant (again, see the book chapter by Faliszewski and Rothe [37] for an overview). Note that control scenarios in voting are also related to certain special cases of the possible winner problem [11,19] and to cloning of candidates [28,59].

Subsuming and extending our preceding conference and workshop papers [4–6,13], we here introduce *control in judgment aggregation*,[2] focusing on the important class of uniform premise-based quota rules due to Dietrich and List [26] that will be defined in Section 2.1.[3] A crucial difference to control-in-voting settings is that, while the (constructive) goal in voting is simply to make a most preferred candidate win (according to the chair's preferences over the candidates), we here have to compare several judgment sets in order to find out whether some control action has paid off for the chair. To compare two outcomes of a judgment aggregation procedure, we will use both Hamming-distance-induced preferences and certain notions of preference types induced by the chair's desired set (so-called top-respecting and closeness-respecting preferences), which have been introduced by Dietrich and List [27], have been applied to manipulation and bribery by Baumeister et al. [7], and will be defined in Section 2.2. After giving the needed background on complexity theory in Section 2.3, we then will study the computational complexity of the problems associated with various control types (to be defined in Section 3, with some first results and observations presented in Section 4): control by adding judges in Section 5,

[1] For an overview, see, e.g., the book chapters by Conitzer and Walsh [22], Faliszewski and Rothe [37], Baumeister and Rothe [12], Brandt et al. [16], and Baumeister et al. [8], and the surveys by Faliszewski et al. [33,36].

[2] Weighted voting games are another area where appropriate analogues of electoral control scenarios have been introduced by Rey and Rothe [54,55] who ask how the power indices of players change when other players are added or deleted and what the complexity of the related problems is.

[3] Note that de Haan [24] studied the Kemeny judgment aggregation procedure with respect to manipulation, bribery, and control.

control by deleting judges in Section 6, control by replacing judges in Section 7, and control by bundling judges in Section 8. Finally, in Section 9, we will conclude by summarizing our results and stating some open question.

2. Preliminaries

We first describe the formal framework of judgment aggregation in Section 2.1, introduce the relevant preference types due to Dietrich and List [27] in Section 2.2, illustrating them by suitable examples, and then we will give some background from computational complexity in Section 2.3.

2.1. Formal framework of judgment aggregation

Throughout this article, we will utilize the judgment aggregation framework due to Endriss et al. [30] (see also the book chapters by Endriss [29] and Baumeister et al. [10]). Let $\mathcal{L}_{PS}$ be the set of all propositional formulas that can be built from a set of propositional variables, *PS*, using the common boolean connectives, i.e., *disjunction* ($\vee$), *conjunction* ($\wedge$), *implication* ($\rightarrow$), and *equivalence* ($\leftrightarrow$) as well as the constants 1 (*true*) and 0 (*false*). To avoid double negations, we use $\overline{\alpha}$ to refer to the complement of α, that is, $\overline{\alpha} = \neg\alpha$ if α is not negated, and $\overline{\alpha} = \beta$ if $\alpha = \neg\beta$. A set $\Phi \subseteq \mathcal{L}_{PS}$ is said to be *closed under complementation* if $\overline{\alpha} \in \Phi$ for all $\alpha \in \Phi$, and to be *closed under propositional variables* if $PS \subseteq \Phi$. We call a finite nonempty set $\Phi \subseteq \mathcal{L}_{PS}$ without doubly negated formulas that is closed under complementation an *agenda*, and a subset $J \subseteq \Phi$ a *judgment set for* Φ. If J is the set of propositions accepted by some judge, it is called an *individual judgment set*. Furthermore, J is called *complete* if $\alpha \in J$ or $\overline{\alpha} \in J$ for all $\alpha \in \Phi$, *consistent* if there exists an assignment such that all formulas in J are satisfied, and *rational* if J is both complete and consistent.

Let $\mathcal{J}(\Phi)$ be the set of all rational judgment sets of an agenda Φ and let $N = \{1, \dots, n\}$ be the set of judges. We call $\mathbf{J} = (J_1, \dots, J_n) \in \mathcal{J}(\Phi)^n$ the *profile* of the judges' individual judgment sets. A resolute[4] *(judgment aggregation) procedure* for an agenda Φ and a set of judges N of size n is a function $F : \mathcal{J}(\Phi)^n \to 2^{\Phi}$, where 2^{Φ} denotes the power set of Φ. That means that a procedure maps a profile to a *collective judgment set* or *(collective) outcome*. We will call a procedure *complete* (*consistent*, *rational*) if the collective judgment set is always *complete* (*consistent*, *rational*).

Unfortunately, a judgment aggregation procedure does not always yield a collective outcome with a certain property even if all individual judgment sets satisfy this property. For instance, the famous doctrinal paradox [44] in judgment aggregation says that if the majority rule[5] is used, the collective judgment set can be inconsistent even if each individual judgment set is consistent. We will study the class of uniform premise-based quota rules as defined by Dietrich and List [26] where we preserve consistency (and thus avoid the doctrinal paradox) by first applying a certain quota individually to the premises, and then logically deriving the result for the conclusions from the result of the premises.

Let $|S|$ be the cardinality of the set S and let $\models$ denote the satisfaction relation.

Definition 1 *(Uniform premise-based quota rule).* Let the agenda Φ be closed under propositional variables. Subdivide Φ into the two disjoint subsets Φ_p (the set of *premises*) containing exactly all literals, and Φ_c (the set of *conclusions*), both closed under complementation.

Furthermore, subdivide Φ_p into two disjoint subsets, Φ_1 and Φ_2, satisfying that $\varphi \in \Phi_1$ if and only if $\overline{\varphi} \in \Phi_2$. Assign to each literal $\varphi \in \Phi_1$ a rational quota q, $0 \le q < 1$, and to each literal $\overline{\varphi} \in \Phi_2$ the associated quota $q' = 1 - q$. A *uniform premise-based quota rule with quota q* (denoted by $UPQR_q$) is a procedure mapping each profile $\mathbf{J} = (J_1, \dots, J_n)$ of individual judgment sets for Φ to the collective outcome

$$UPQR_q(\mathbf{J}) = \triangle_q \cup \{\psi \in \Phi_c \mid \triangle_q \models \psi\},$$

where $\triangle_q = \{\varphi \in \Phi_1 \mid |\{i \mid \varphi \in J_i\}| > nq\} \cup \{\varphi \in \Phi_2 \mid |\{i \mid \varphi \in J_i\}| \ge nq'\}$. Throughout the article, we will assume that all literals in Φ_1 are not negated.

Since Φ is closed under propositional variables and Φ_p contains exactly all literals, $UPQR_q$ is rational. The threshold for a literal $\varphi \in \Phi_1$ to be accepted is $\lfloor nq + 1 \rfloor$, i.e., φ is contained in the collective outcome if and only if it is contained in at least $\lfloor nq + 1 \rfloor$ individual judgment sets, whereas literals $\overline{\varphi} \in \Phi_2$ need at least $\lceil nq' \rceil$ affirmations to be accepted. It is possible to determine in polynomial time whether a given formula is an element of the collective outcome of a uniform premise-based quota rule. The special case of $UPQR_{1/2}$ for an odd number of judges is also known as the premise-based procedure (*PBP*).

Example 2. To illustrate the doctrinal paradox, consider a three-person recruiting committee that wants to fill a vacancy in their company. However, the committee members have different opinions about an applicant. According to the company's

[4] There are also irresolute judgment aggregation procedures (i.e., procedures that may output more than one collective judgment set), such as the distance-based procedures introduced by Pigozzi [53] and Miller and Osherson [49], which we won't consider here, though.

[5] The *majority rule* is a resolute judgment aggregation procedure that includes exactly those formulas in the collective outcome that a strict majority of judges have in their individual judgment sets.

Table 1
The (uniform) premise-based quota rule avoids the doctrinal paradox.

	Outstanding degree	Relevant field	Job offer
(a) Doctrinal paradox with the majority rule			
Member 1	1	1	1
Member 2	0	1	0
Member 3	1	0	0
Majority	1	1	0
(b) Different outcome with the uniform premise-based quota rule			
Member 1	1	1	1
Member 2	0	1	0
Member 3	1	0	0
$UPQR_{1/2}$	1	1	1

policy, they will hire an applicant if and only if he or she has an outstanding degree in a field relevant to the company's business sector. However, only the first committee member says that the applicant should be given a job offer, since both requirements are fulfilled. The second committee member thinks that the applicant's degree cannot be called outstanding, while in the third member's opinion the degree was obtained in a nonrelevant field. Table 1(a) shows the three individual judgment sets and the collective outcome according to the majority rule, denoted by a 1 for "yes" and a 0 for "no."

The inconsistent outcome of the majority rule—i.e., the outcome where the applicant is denied the job offer, even though he or she is well suited for the job—can be avoided by using the (uniform) premise-based quota rule, as shown in Table 1(b).

Another way of defining quota rules is to let the quota be a fixed number instead of a portion of the judges. Obviously, if the number of judges is fixed, both definitions yield the same judgment aggregation rule. However, we will study control problems in judgment aggregation where the number of judges can vary (for example, in control by adding judges). In this case the two procedures can output different collective judgment sets. As a real-world example, consider a simplified version of a referendum. Suppose the number of "yes" votes that are needed to put through an issue is 2/3 of all registered voters. Obviously, this quota should not change with the number of people actually going to the polls, so this corresponds to a constant premise-based quota rule. In other situations, however, it is desirable that the quota does depend on the number of judges actually taking part. We now introduce constant premise-based quota rules in addition to the premise-based quota rules defined above.

Definition 3 *(Uniform constant premise-based quota rule).* As in Definition 1, let the agenda Φ be closed under propositional variables, and let Φ_p denote the set of premises, Φ_c the set of conclusions, and Φ_1 and Φ_2 a partition of Φ_p satisfying that $\varphi \in \Phi_1$ if and only if $\overline{\varphi} \in \Phi_2$. Let $q \in \mathbb{N}_0$, $0 \leq q < n$. A *uniform constant premise-based quota rule with quota q* is defined by

$$UCPQR_q(\mathbf{J}) = \triangle'_q \cup \{\psi \in \Phi_c \mid \triangle'_q \models \psi\}, \text{ where } \triangle'_q = \{\varphi \in \Phi_1 \mid |\{i \mid \varphi \in J_i\}| > q\} \cup \{\varphi \in \Phi_2 \mid |\{i \mid \varphi \in J_i\}| \geq (n-q)\}.$$

That is, the number of affirmations that are needed to be in the set $\triangle'_q$ is a fixed constant.

2.2. Types of preferences in judgment aggregation

We will study judgment aggregation problems where some external agent tries to influence a judgment aggregation process in order to obtain a better outcome. In order to compare two outcomes, we will use various notions of preference types induced by an external agent's desired set. These notions have been introduced by Dietrich and List [27]. Formally, this desired set is a subset of a rational judgment set.

Let Φ be an agenda, $X, Y \in \mathcal{J}(\Phi)$, and let $\succsim$ be a weak order over $\mathcal{J}(\Phi)$, i.e., a transitive and total binary relation over rational judgment sets. We say that X *is weakly preferred to* Y whenever $X \succsim Y$, and we say that X *is preferred to* Y, denoted by $X \succ Y$, whenever $X \succsim Y$ and $\neg(Y \succsim X)$. Furthermore, we define $X \sim Y$ by $X \succsim Y$ and $Y \succsim X$.

Based on the notions introduced by Dietrich and List [27], Baumeister et al. [7] define the notions of top-respecting and closeness-respecting preferences in particular for incomplete judgment sets. They and Endriss et al. [30] also consider Hamming-distance-respecting preferences in the context of manipulation and bribery in judgment aggregation. We consider these preferences in the context of control.

Definition 4. Let Φ be an agenda, let U be the set of all weak orders over $\mathcal{J}(\Phi)$, and let D be a possibly incomplete judgment set (typically, D will be the chair's *desired set*). Define

1. the set $U_D \subseteq U$ of *unrestricted D-induced (weak) preferences* by

$$U_D = \{\succsim \ \in U \mid \text{for all } X, Y \in \mathcal{J}(\Phi), X \sim Y \text{ whenever } X \cap D = Y \cap D\};$$

Table 2
Example of the uniform premise-based quota rule.

	a	b	c	$\neg a \vee b$	$a \wedge c$	$b \vee c$
J_1	1	1	0	1	0	1
J_2	0	0	0	1	0	0
J_3	1	0	1	0	1	1
$UPQR_{1/2}$	1	0	0	0	0	0
D	1		1		1	1

2. the set $TR_D \subseteq U_D$ of *top-respecting D-induced (weak) preferences* by

$$TR_D = \{\succsim \in U_D \mid \text{for all } X, Y \in \mathcal{J}(\Phi), X \succ Y \text{ whenever } X \cap D = D \text{ and } Y \cap D \neq D\};$$

3. the set $CR_D \subseteq U_D$ of *closeness-respecting D-induced (weak) preferences* by

$$CR_D = \{\succsim \in U_D \mid \text{for all } X, Y \in \mathcal{J}(\Phi), \text{if } X \cap D \supseteq Y \cap D \text{ then } X \succsim Y\};$$

4. the set $HD_D \subseteq U_D$ of *Hamming-distance D-induced (weak) preferences* by

$$HD_D = \{\succsim \in U_D \mid \text{for all } X, Y \in \mathcal{J}(\Phi), X \succsim Y \text{ if and only if } HD(X, D) \leq HD(Y, D)\},$$

   where the Hamming distance $HD(S, T)$ of two (possibly incomplete) consistent judgment sets S and T denotes the number of $\varphi \in S$ so that $\overline{\varphi} \in T$.

Unrestricted preferences model the scenario where we know nothing about the attacker's preferences, not even whether the desired set D is a subset of her most preferred outcome, which is the only thing known regarding top-respecting preferences. Regarding closeness-respecting preferences, we additionally know that judgment sets with preserved and additional agreements with the desired set are preferred, whereas Hamming-distance-induced preferences are complete orders where the total number of disagreements with the desired set is decisive of a judgment's position in the attacker's preference list. Therefore, it holds that

$$HD_D \subseteq CR_D \subseteq TR_D \subseteq U_D.$$

Definition 5. Let Φ be an agenda, let X and Y be rational judgment sets for Φ, let D be an agent's desired set, and let $T_D \in \{U_D, TR_D, CR_D\}$ be a type of D-induced (weak) preferences. We say that

1. the agent *necessarily/possibly weakly prefers X to Y for type T_D* if $X \succsim Y$ for all/some $\succsim \in T_D$,
2. the agent *necessarily/possibly prefers X to Y for type T_D* if $X \succ Y$ for all/some $\succsim \in T_D$.

Let D be the desired set of an agent. If unrestricted preferences are assumed, a new outcome X is never necessarily preferred to the current outcome Y; by contrast, X is always possibly preferred to Y. If top-respecting preferences are assumed, X is necessarily preferred to Y if and only if D is contained in X and D is not contained in Y, and X is possibly preferred to Y if and only if it is not the case that Y contains all elements from D. In the case of closeness-respecting preferences, the agent necessarily prefers X to Y if and only if she achieves a new agreement with D while preserving the existing agreements. On the other hand, she possibly prefers X to Y if and only if she achieves a new agreement with D regardless of new differences.

Note that the set of Hamming-distance D-induced preferences is a singleton. Therefore, in this case we do not differentiate between possible and necessary preferences and simply say that the agent *(weakly) prefers X to Y under Hamming-distance-induced preferences* exactly if the number of disagreements between X and D is lower (not higher) than the number of disagreements between Y and D.

Example 6. Let $\Phi = \{a, \neg a, b, \neg b, c, \neg c, \neg a \vee b, \neg(\neg a \vee b), a \wedge c, \neg(a \wedge c), b \vee c, \neg(b \vee c)\}$ be an agenda that is closed under the set of propositional variables $PS = \{a, b, c\}$. The profile $\mathbf{J} = (J_1, J_2, J_3)$, the collective outcome $UPQR_{1/2}$, and the external agent's desired set $D = \{a, c, a \wedge c, b \vee c\}$ are shown in Table 2. Here, a 1 denotes that the corresponding formula is contained in the set, whereas a 0 indicates that the negation of the formula is in the set.

Note that D is a consistent judgment set, since it can be reached by setting the propositional variables a and c to true. Let $Y = UPQR_{1/2}(\mathbf{J})$ be the current collective outcome of the uniform premise-based procedure with quota $q = 1/2$ and let $X = \{a, b, \neg c, \neg a \vee b, \neg(a \wedge c), b \vee c\}$ be the new outcome the external agent achieves via exerting control by replacing J_2 with the new individual judgment set $\{\neg a, b, \neg c, \neg a \vee b, \neg(a \wedge c), b \vee c\}$ (i.e., the row for J_2 in Table 2 is replaced by 0 1 0 1 0 1 and the row for $UPQR_{1/2}$, now reflecting X, is replaced by 1 1 0 1 0 1). Note that this control type as well as other specific control types will be formally defined in Section 3.

- Assuming top-respecting preferences, the agent possibly prefers X to Y, since the desired set D is not contained in Y; yet she does not necessarily prefer X to Y, since D is not contained in X either.
- However, if we consider closeness-respecting preferences, the agent does even necessarily (and thus also possibly) prefer X to Y:

$$X \cap D = \{a, b \vee c\} \supset \{a\} = Y \cap D,$$

  i.e., X preserves the existing agreements of Y with D and adds an additional agreement.
- Since the agent necessarily prefers X to Y assuming closeness-respecting preferences and since $HD_D \subseteq CR_D$, she also prefers X to Y assuming Hamming-distance-induced preferences. Alternatively, one can see that X and D disagree on two propositions, whereas Y and D disagree on three. Thus $HD(X, D) < HD(Y, D)$, and it follows that the agent prefers X to Y under Hamming-distance-induced preferences.

2.3. Background on complexity theory

We assume that the reader is familiar with the complexity classes P and NP as well as with the concept of polynomial-time many-one reducibility (denoted by $\leq_m^p$; see, for example, the textbooks by Papadimitriou [52] and Rothe [56]).

We will use the following two NP-complete decision problems in our reductions:

	Dominating-Set
Given:	A graph $G = (V, E)$ and a positive integer k.
Question:	Does G have a dominating set of size at most k, i.e., a subset $V' \subseteq V$, $\|V'\| \leq k$, such that every vertex $v \in V$ belongs to the closed neighborhood of some $v' \in V'$?

	Exact-Cover-by-3-Sets (X3C)
Given:	A set X and a collection C containing 3-element subsets of X.
Question:	Does there exist an exact cover for X, i.e., a subcollection $C' \subseteq C$ such that each element of X is a member of exactly one set in C'?

3. Problem definitions, motivation, and summary of results

Looking at the example from Table 1(b), which illustrates how to avoid the doctrinal paradox by a uniform premise-based quota rule, a typical *control scenario* would be if some external agent (the "chair") had the power to delete, say, the first two recruiting committee members, thus turning the result of $UPQR_{1/2}$ from "yes" to "no" (denying a job offer). Starting with the work of Bartholdi et al. [3], many control scenarios have been studied for various voting systems (see, e.g., [31,34,35,41,42] and the book chapters [8,12,37] for an overview). We introduce control scenarios for judgment aggregation. Specifically, we will consider control by adding, deleting, replacing, and bundling judges, and we will motivate these scenarios by real-world examples from international arbitration procedures as a method of dispute resolution. Note that we will define the decision problems only for uniform premise-based quota rules; the problems for the uniform *constant* premise-based quota rules are defined analogously.

3.1. Control by adding judges

This first control type is analogous to control by adding voters in elections. An example for this control setting can be found in the field of international arbitration, which is becoming increasingly important as an alternative dispute resolution method to litigations conducted by national courts. Parties of arbitration proceedings may choose to entrust a single arbitrator with deciding their dispute. They might, however, also opt for the appointment of several arbitrators and thereby control the arbitral decision-making process by adding judges.[6] Mostly they do so because they feel that due to the complicated nature of the matter or for some other reason, a tribunal with several arbitrators is better suited to arbitrate their case. Their action may also be motivated by the hope of being able to appoint an arbitrator sympathetic to their arguments.

Formally, we define the following two problems for uniform premise-based quota rules and preference type T (as defined in Section 2.2).

[6] See, for instance, Articles 37–40 of the ICSID Convention and Rules 1–4 of the ICSID Rules of Procedure for Arbitration Proceedings [38], Articles 11–12 of the ICC Rules of Arbitration [50], or Articles 7–10 of the UNCITRAL Arbitration Rules [51].

$UPQR_q$-T-Possible-Control-by-Adding-Judges	
Given:	An agenda Φ, two profiles $\mathbf{J} \in \mathcal{J}(\Phi)^n$ and $\mathbf{K} \in \mathcal{J}(\Phi)^m$, a desired set D, and a positive integer k.
Question:	Is there a subprofile $\mathbf{K}' \subseteq \mathbf{K}$ of size at most k such that for the new profile $\mathbf{J}' = \mathbf{J} \cup \mathbf{K}'$, it holds that $UPQR_q(\mathbf{J}') \succ UPQR_q(\mathbf{J})$ for some $\succsim \in T_D$?

Concerning the related problem $UPQR_q$-T-Necessary-Control-by-Adding-Judges, the respective condition must hold for *all* $\succsim$ in T_D for the instance to be accepted. In the case of Hamming-distance-induced preferences, the two problems coincide and will be denoted simply as $UPQR_q$-HD-Control-by-Adding-Judges.

We also consider the exact variants of our problems, in the present control scenario denoted by $UPQR_q$-Exact-Control-by-Adding-Judges, where we ask whether there is a subset $\mathbf{K}' \subseteq \mathbf{K}$, $|\mathbf{K}'| \le k$, such that $D \subseteq UPQR_q(\mathbf{J} \cup \mathbf{K}')$. In this setting, the chair seeks to achieve a *best* outcome for her desired judgment set (rather than just a "better" outcome).

3.2. Control by deleting judges

Also very natural is the problem of control by deleting judges as it is a commonly applied method in both jury trials and international arbitration. The empaneling procedure of a jury for a trial is basically a control process via deleting judges and works roughly as follows. First, a certain number of potential jurors[7] is summoned at the place of trial. In the next stage of the selection procedure, all or part of them are subjected to the so-called "voir dire" process, i.e., a questioning by the trial judge and/or the attorneys aiming to obtain information about their person. Admittedly, the purpose of collecting this information is to determine whether they can be impartial, which is a well-justified purpose; but again, attorneys may use it for another reason, namely to indoctrinate prospective jurors laying a foundation for arguments they later intend to make. Driven by good or bad intentions, the lawyers may then challenge jurors for cause, that is, by arguing that and for what reason the juror in question is biased. The trial judge decides over the attorneys' challenges for cause, moreover she may excuse further jurors due to social hardship. Finally, the lawyers may challenge a limited number of potential jurors peremptorily, i.e., without having to justify their reason for doing so. Peremptory challenges are legitimate and useful means of eliminating such jurors that are either presumably biased but the bias cannot be proven to the extent necessary for challenging them for cause, or are for some other reason undesirable. Because their use does not require any explanation, such challenges can also be easily abused; especially until the introduction of the Batson rule, peremptory challenges were often exercised in discriminatory ways, mostly on racial grounds, violating the equal protection rights of jurors. For more details, see the book by Jonakait [43].

Define the following problem for uniform premise-based quota rules and for type T preferences.

$UPQR_q$-T-Possible-Control-by-Deleting-Judges	
Given:	An agenda Φ, a profile $\mathbf{J} \in \mathcal{J}(\Phi)^n$, a desired set D, and a positive integer k.
Question:	Is there a subprofile $\mathbf{J}' \subseteq \mathbf{J}$ of size at most k such that $UPQR_q(\mathbf{J} \smallsetminus \mathbf{J}') \succ UPQR_q(\mathbf{J})$ for some $\succsim \in T_D$?

The other variants are defined analogously to the case of adding judges.

As we can see, deleting judges/jurors is a central part of the empaneling procedure. However, when the total number of jurors is fixed, a new juror needs to be appointed for each deleted juror, which motivates the next scenario to be defined in Section 3.3.

3.3. Control by replacing judges

Control by replacing judges is used in international arbitration when the parties successfully challenge an arbitrator leading to her disqualification and the subsequent appointment of a substitute arbitrator. The institution of challenge is designed to serve as a tool for parties of arbitral proceedings to remove arbitrators posing a possible threat to the integrity of the proceedings. It may be based on several grounds; arbitrators are most commonly challenged because of doubts regarding their impartiality or independence.[8] Challenges are, however, occasionally used as "black art" or "guerrilla tactics" with a view to achieve dishonest purposes, such as eliminating arbitrators that are likely to render an unfavorable award or to delay the proceedings to evade, or at least postpone, an anticipated defeat. Again, see the book by Jonakait [43] for more details.

Control by replacing judges can be seen as a combined action of control by deleting judges and control by adding judges. For a related general model in voting theory, we refer to the work of Faliszewski et al. [34] on multimode control attacks and of Loreggia et al. [48] on electoral control by replacing candidates or voters.

[7] "Juror" here refers to what we call "judge" anywhere else, as we reserve "judge" in this context for "trial judge."

[8] For rules regarding the challenge, disqualification, and replacement of arbitrators, see Articles 56–58 of the ICSID Convention [38], Rules 9–11 of the ICSID Rules of Procedure for Arbitration Proceedings [38], Articles 14–15 of the ICC Rules of Arbitration [50], and Articles 12–14 of the UNCITRAL Arbitration Rules [51].

Define the following problem for uniform premise-based quota rules and for type T preferences.

	$UPQR_q$-T-Possible-Control-by-Replacing-Judges
Given:	An agenda Φ, two profiles $\mathbf{J} \in \mathcal{J}(\Phi)^n$ and $\mathbf{K} \in \mathcal{J}(\Phi)^m$, a desired set D, and a positive integer k.
Question:	Are there subprofiles $\mathbf{J}' \subseteq \mathbf{J}$ and $\mathbf{K}' \subseteq \mathbf{K}$ of size $\lvert\mathbf{J}'\rvert = \lvert\mathbf{K}'\rvert \leq k$ such that for the new profile $\mathbf{S} = (\mathbf{J} \setminus \mathbf{J}') \cup \mathbf{K}'$, it holds that $UPQR_q(\mathbf{S}) \succ UPQR_q(\mathbf{J})$ for some $\succsim\, \in T_D$?

Define the necessary, *HD*, and exact control problems analogously to these variants of the problems modeling control by adding and deleting judges.

3.4. Control by bundling judges

Control by bundling judges is remotely akin to control by partitioning voters in voting. A prominent natural example for control by bundling judges can be found in European legislation. Certain European legislative acts, such as Directives, give considerable freedom to Member States regarding the concrete implementation of these acts. Yet, in some cases uniform implementation is crucial, so the basic act confers implementing powers on the European Commission or the Council of the European Union to adopt the required implementing acts.[9] The exercise of implementing powers through the Commission and Council is controlled by the member states through so-called comitology committees in accordance with previously specified rules.[10] The committees are set up by the basic act in question.[11] Some of these committees are concerned with such a broad range of issues that they are divided into subcommittees, each of which is dealing with different issues. When preparing implementing acts covering several issues, each subcommittee votes on the issues assigned to it, and the implementing act is shaped according to the decisions of the different subcommittees.[12]

The formal definition of control by bundling judges is as follows. In the problem definition below, we will use the notation

$$\Delta_q = \bigcup_{1 \leq i \leq k} UPQR_q(\mathbf{J}|_{\Phi_p^i, N_i}),$$

where $\mathbf{J}|_{\Phi_p^i, N_i}$ is a profile obtained by restricting the set of judges in $\mathbf{J}$ to $N_i \subseteq N$ and their respective judgment sets to the premises Φ_p^i in a given partition (see the problem definition below).

The formal definition is as follows. Recall that the set Φ_c consists of the agenda's conclusions. Define the following problem for uniform premise-based quota rules and for type T preferences.

	$UPQR_q$-T-Possible-Control-by-Bundling-Judges
Given:	An agenda Φ, where the premises are partitioned into k subsets $\Phi_p^1, \dots, \Phi_p^k$ satisfying $\varphi \in \Phi_p^i$ if and only if $\overline{\varphi} \in \Phi_p^i$, a profile $\mathbf{J} \in \mathcal{J}(\Phi)^n$, and a desired set D.
Question:	Does there exist a partition $\{N_1, \dots, N_k\}$ of the n judges such that for some $\succsim \in T_D$, $\Delta_q \cup \{\varphi \in \Phi_c \mid \Delta_q \models \varphi\} \succ UPQR_q(\mathbf{J})$?

As in the case of adding, deleting, and replacing judges, in the related problem $UPQR_q$-T-Necessary-Control-by-Bundling-Judges we ask whether the respective condition holds for *all* $\succsim$ in T_D. Moreover, we define $UPQR_q$-*HD*-Control-by-Bundling-Judges and $UPQR_q$-Exact-Control-by-Bundling-Judges analogously to the corresponding problems of the other control types.

Example 7. Consider the same agenda Φ, the same individual judgment sets, and the same desired set D as in Example 6. Let the quota be $q = 1/2$ for every positive literal in the agenda. Assume that the set of premises is partitioned into $\Phi_1^p = \{a, \neg a, b, \neg b\}$ and $\Phi_2^p = \{c, \neg c\}$. The Hamming distance between the current collective outcome and D is 3. But if we partition the set of judges into two groups, where the first judge forms the first group and the last two judges are in the second group, the outcome is as shown in Table 3, where the individual judgments for premises not belonging to the group deciding over them are marked with $\not1$ or $\not0$ and S denotes $\Delta_{1/2} \cup \{\varphi \in \Phi_c \mid \Delta_{1/2} \models \varphi\}$ as in the problem definition above. Recall that, in case of a tie, a negative literal is contained in the collective judgment set by convention.

After bundling the judges, the Hamming distance between the collective outcome and D has decreased to 2. Hence, this is a positive instance of $UPQR_{1/2}$-*HD*-Control-by-Bundling-Judges. However, since it is not possible to bundle the judges

[9] Article 291 of the Treaty on the Functioning of the European Union.

[10] Regulation (EU) No 182/2011 of the European Parliament and of the Council of 16 February 2011 laying down the rules and general principles concerning mechanisms for control by Member States of the Commission's exercise of implementing powers (Implementing Acts Regulation).

[11] Recital 6 of the Preamble of the Implementing Acts Regulation.

[12] One example is the Customs Code Committee, see Articles 1 (1) and 5 (7) (8) of the Rules of Procedure for the Customs Code Committee.

Table 3
The uniform premise-based quota rule in Example 7 illustrating control by bundling judges.

	a	b	c	$\neg a \vee b$	$a \wedge c$	$b \vee c$
J_1	1	1	$\cancel{0}$	1	0	1
J_2	$\cancel{0}$	$\cancel{0}$	0	1	0	0
J_3	$\cancel{1}$	$\cancel{0}$	1	0	1	1
S	1	1	0	1	0	1
D	1		1		1	1

into two groups to obtain exactly D as a subset of the collective outcome, it is a negative instance of $UPQR_{1/2}$-Exact-Control-by-Bundling-Judges.

Remotely related bundling problems in judgment aggregation have been studied by Alon et al. [1]. However, their setting is different from ours. They consider judgment aggregation over independent variables, and only the variables are bundled in their bundling attacks. It is assumed that then all judges decide over all bundles by deciding uniformly for all variables contained in the same bundle. Furthermore, the goal in their model is to always accept all positive variables, that is, a complete desired set. This setting in fact covers a restriction of judgment aggregation known as optimal lobbying (see the papers by Christian et al. [20], Binkele-Raible et al. [15], and Bredereck et al. [18]).

3.5. Immunity, susceptibility, resistance, and vulnerability

To study the computational complexity of adding, deleting, replacing, and bundling judges, we adopt the terminology introduced by Bartholdi et al. [3] for control problems in voting and adapt it to judgment aggregation and the context of preference types.

Definition 8. Let F be a judgment aggregation procedure, let $\mathcal{C}$ be a given control type, and $T \in \{U, TR, CR, HD\}$ be a given preference type.

1. F is said to be *immune to (exact/possible/necessary) control by $\mathcal{C}$ (under induced preferences of type T)* if it is never possible for the chair to successfully control the judgment aggregation procedure via $\mathcal{C}$.
2. F is said to be *susceptible to (exact/possible/necessary) control by $\mathcal{C}$ (under induced preferences of type T)* if it is not immune to (exact/possible/necessary) control by $\mathcal{C}$ under T.
3. F is said to be *vulnerable to (exact/possible/necessary) control by $\mathcal{C}$ (under induced preferences of type T)* if it is susceptible to (exact/possible/necessary) control by $\mathcal{C}$ under T and the corresponding decision problem is in P.
4. F is said to be *resistant to (exact/possible/necessary) control by $\mathcal{C}$ (under induced preferences of type T)* if it is susceptible to (exact/possible/necessary) control by $\mathcal{C}$ under T and the corresponding decision problem is NP-hard.

Studying the computational complexity of *decision problems* such as those defined in the previous subsections (rather than of *search problems* where one seeks to actually *find* a successful control action) is quite common both in computational social choice (see, e.g., the book chapters [12,37]) and in judgment aggregation (see, e.g., the book chapters [29,10]). Note, however, that the complexity of a search problem can be much harder than the complexity of the corresponding decision problem. For instance, Hemaspaandra et al. [40] have shown that, under the plausible assumption that integer factoring is a hard problem, there exist voting rules for which the decision problems associated with (electoral) manipulation, (electoral) bribery, and some types of (electoral) control can be solved in polynomial time, yet the corresponding search problems cannot be solved in polynomial time. In this paper on control in judgment aggregation, we restrict ourselves to the study of the computational complexity of decision problems, leaving the interesting study of the corresponding search problems to future work.

3.6. Summary of results

Table 4 summarizes our results on the complexity of exact, possible, and necessary control by adding, deleting, and replacing judges for uniform constant premise-based quota rules under various preference types, both for complete and incomplete desired sets (abbreviated by DS).

For control by adding, deleting, and bundling judges, the results for $UPQR_{1/2}$ are shown in Table 5; note that this table refers to uniform premise-based quota rules (and not to uniform *constant* premise-based quota rules, as Table 4 does).[13]

[13] In Section 7 we will explain why uniform constant premise-based quota rules and the corresponding uniform premise-based quota rules coincide for control by replacing judges, and in Section 8 we will explain why it does not make sense to consider uniform constant premise-based quota rules for control by bundling judges.

Table 4
Overview of results for $UCPQR_q$-T-Possible/Necessary-Control-By-$\mathcal{C}$ and $UCPQR_q$-HD/Exact-Control-By-$\mathcal{C}$ for $\mathcal{C} \in \{$Adding-Judges, Deleting-Judges, Replacing-Judges$\}$ and $T \in \{U, TR, CR\}$. NP-c stands for "NP-complete" and a question mark indicates an open problem. Below each result we have noted in parentheses the theorems, lemmas, or propositions stating it.

		U	TR	CR	HD	Exact
Incomplete DS	Possible	NP-c (Lemma 11)	NP-c (Lemma 11)	NP-c (Theorem 15, 18, 21)	NP-c (Lemma 9)	NP-c (Lemma 10)
	Necessary	immune (Proposition 14)	NP-c (Proposition 12)	NP-c (Theorem 15, 18, 21)	NP-c (Lemma 9)	NP-c (Lemma 10)
Complete DS	Possible	P (Proposition 13)	P (Proposition 13)	NP-c (Theorem 15, 18, 21)	NP-c (Lemma 9)	?
	Necessary	immune (Proposition 14)	?	NP-c (Theorem 15, 18, 21)	NP-c (Lemma 9)	?

Table 5
Overview of results for $UPQR_{1/2}$-T-Possible/Necessary-Control-By-$\mathcal{C}$ and $UPQR_{1/2}$-HD/Exact-Control-By-$\mathcal{C}$ for $\mathcal{C} \in \{$Adding-Judges, Deleting-Judges, Bundling-Judges$\}$ and $T \in \{U, TR, CR\}$. Again, NP-c stands for "NP-complete" and below each result we have noted in parentheses the theorems, lemmas, or propositions stating it.

		U	TR	CR	HD	Exact
Incomplete DS	Possible	NP-c (Lemma 11)	NP-c (Lemma 11)	NP-c (Theorem 16, 19, 22)	NP-c (Lemma 9)	NP-c (Lemma 10)
	Necessary	immune (Proposition 14)	NP-c (Proposition 12)	NP-c (Theorem 16, 19, 22)	NP-c (Lemma 9)	NP-c (Lemma 10)
Complete DS	Possible	P (Proposition 13)	P (Proposition 13)	NP-c (Theorem 16, 19, 22)	NP-c (Lemma 9)	NP-c (Theorem 17, 20, 23)
	Necessary	immune (Proposition 14)	NP-c (Proposition 12)	NP-c (Theorem 16, 19, 22)	NP-c (Lemma 9)	NP-c (Theorem 17, 20, 23)

We have restricted our attention in these results to the quota $q = 1/2$, leaving the complexity for a general quota q open for most cases. However, we show that immunity in the case of necessary control and unrestricted preferences as well as the P results in the case of possible control and unrestricted or top-respecting preferences also hold for every admissible quota. Note that since the number of judges varies in these three control scenarios, it does *not* hold that $UPQR_{1/2}$ is a special case of the uniform constant premise-based quota rule. Hence, hardness results cannot be directly transferred from $UPQR_{1/2}$ to $UCPQR_q$.

4. First results and observations

All decision problems defined in the sections above belong to NP. Note that all results in this section also hold for uniform constant premise-based quota rules; we will only state them for *UPQR*, though. Also note that we give the proofs for control by adding, deleting, replacing, and bundling judges at the same time.

Lemma 9. *Let $\mathcal{C}$ be any of the control types defined in this paper and let q, $0 \le q < 1$, be a rational quota. $UPQR_q$-CR-*Necessary-Control-by-$\mathcal{C}$ $\le_m^p$ *$UPQR_q$-HD-*Control-by-$\mathcal{C}$.

Proof. We start with an instance of $UPQR_q$-CR-Necessary-Control-by-$\mathcal{C}$ and construct an instance of $UPQR_q$-HD-Control-by-$\mathcal{C}$ in the following way. Let **J** be the original profile, Φ the original agenda and $D = \{\varphi_1, \dots, \varphi_k, \varphi_{k+1}, \dots, \varphi_t\}$ the original desired set, where $\varphi_i \in UPQR_q(\mathbf{J})$ for $1 \le i \le k$ and $\varphi_i \notin UPQR_q(\mathbf{J})$ for $k+1 \le i \le t$. Without loss of generality, we assume that $t > k \ge 1$, i.e., the collective outcome agrees (disagrees) with at least one formula in the desired set. To obtain the agenda in the new instance, extend Φ by the formula $\psi = \varphi_1 \wedge \dots \wedge \varphi_k$ and by $t-1$ syntactic variations of ψ (and the corresponding negations). This can be seen as giving a weight of t to ψ. Furthermore, extend the profile **J** (and all possibly existing additional profiles, e.g., in the case of control by adding judges) by ψ or $\neg\psi$ (and by the syntactic variations) in a consistent way. Let the new desired set D' be the union of the old desired set D, the formula ψ, and of all its syntactic variations. Note that ψ is contained in the collective outcome of the new instance.

We claim that the chair is able to achieve an additional agreement of his desired set with the new collective outcome in the original instance if and only if the chair is able to reduce the Hamming distance between his desired set and the new collective outcome in the new instance.

From left to right, assume that after exerting control of type $\mathcal{C}$, the chair necessarily prefers the new outcome to the old one assuming closeness-respecting preferences, i.e., $\varphi_1, \dots, \varphi_k$ are still part of the collective outcome and there exists a φ_i, $k+1 \le i \le t$, so that φ_i is part of the new outcome. Then—after employing the same control action in the new instance—the Hamming distance between D' and the corresponding new collective outcome is smaller than before.

From right to left, assume that after exerting control of type $\mathcal{C}$ in the new instance, the Hamming distance between D' and the new collective outcome is reduced, i.e., the distance is smaller than $t-k$. This cannot be achieved without keeping $\varphi_1, \dots, \varphi_k$ in the new collective outcome since an additional Hamming distance of t between the desired set and an outcome containing $\neg\psi$ cannot be compensated. It follows that all agreements between the desired set and the new collective outcome are preserved and a new agreement was added (in the new as well as in the original instance after employing the same control actions), so that the chair necessarily prefers the new collective outcome to the old one in the original instance. □

Note that the desired set in the new instance is complete if and only if the chair's desired set in the original instance is complete as well. Therefore, this result shows that, in the case of a complete desired set, NP-hardness of $UPQR_q$-CR-Necessary-Control-by-$\mathcal{C}$ implies NP-hardness of $UPQR_q$-HD-Control-by-$\mathcal{C}$.

Next, we show that, for uniform premise-based quota rules and any of the control types considered here, exact control is at least as hard as necessary control for closeness-respecting preferences.

Lemma 10. *Let $\mathcal{C}$ be any of the control types defined in this paper and let q, $0 \le q < 1$, be a rational quota. $UPQR_q$-CR-Necessary-Control-by-$\mathcal{C}$ $\le^{\mathrm{p}}_{\mathrm{m}}$ $UPQR_q$-Exact-Control-by-$\mathcal{C}$.*

Proof. We start with an instance of $UPQR_q$-CR-Necessary-Control-by-$\mathcal{C}$ and construct an instance of $UPQR_q$-Exact-Control-by-$\mathcal{C}$ in the following way. Let $\mathbf{J}$ be the original profile, Φ the original agenda and $D = \{\varphi_1, \dots, \varphi_k, \varphi_{k+1}, \dots, \varphi_t\}$ the original desired set, where $\varphi_i \in UPQR_q(\mathbf{J})$ for $1 \le i \le k$ and $\varphi_i \notin UPQR_q(\mathbf{J})$ for $k+1 \le i \le t$. Without loss of generality, we again assume that $t > k \ge 1$, i.e., the collective outcome agrees (disagrees) with at least one formula in the desired set. Extend the agenda by the formulas $\psi_1 = \varphi_1 \wedge \cdots \wedge \varphi_k$ and $\psi_2 = \varphi_{k+1} \vee \cdots \vee \varphi_t$ and the corresponding negations, and extend all involved profiles to the new agenda. Let $D' = \{\psi_1, \psi_2\}$ be the chair's new desired set.

We show that the chair can preserve existing agreements and achieve an additional agreement of her desired set with the new collective outcome in the original instance by exerting control of type $\mathcal{C}$ if and only if she can exert control of type $\mathcal{C}$ to include her entire desired set in the new collective outcome of the new instance.

From left to right, assume that $\varphi_1, \dots, \varphi_k$ (i.e., the existing agreements) and at least one of φ_i, $k+1 \le i \le t$, is contained in the collective outcome after exerting control of type $\mathcal{C}$. If the chair employs the same control action in the new instance, ψ_1 and ψ_2 are both part of the new collective outcome, so the control action is successful.

From right to left, assume that after exerting control of type $\mathcal{C}$, D' is contained in the new collective outcome in the new instance. This can only be the case if $\varphi_1, \dots, \varphi_k$ (i.e., the existing agreements) are preserved to include ψ_1 and at least one of φ_i, $k+1 \le i \le t$, (i.e., an additional agreement) is added to the collective outcome to include ψ_2. As before, after employing the same control action in the original instance the chair necessarily prefers the new to the old outcome under closeness-respecting preferences. □

In contrast to the proof of Lemma 9, the desired set in the new instance is incomplete. Therefore, the complexity of the special case of a complete desired set in the exact problem variant has to be considered separately.

Our next result gives a link between the exact control problem of a given type and the corresponding possible control problem with respect to unrestricted and top-respecting preferences.

Lemma 11. *Let $\mathcal{C}$ be any of the control types defined in this paper and let q, $0 \le q < 1$, be a rational quota. $UPQR_q$-Exact-Control-by-$\mathcal{C}$ $\le^{\mathrm{p}}_{\mathrm{m}}$ $UPQR_q$-T-Possible-Control-by-$\mathcal{C}$ for each preference type $T \in \{U, TR\}$.*

Proof. We give the proof for both claims at the same time. Consider an instance of $UPQR_q$-Exact-Control-by-$\mathcal{C}$, where Φ denotes the agenda, $\mathbf{J}$ the profile of judges before exerting control, and $D = \{\varphi_1, \dots, \varphi_t\}$ the chair's desired set. Without loss of generality, assume that $D \nsubseteq UPQR_q(\mathbf{J})$. To create the new instance, i.e., an instance of $UPQR_q$-U-Possible-Control-by-$\mathcal{C}$ ($UPQR_q$-TR-Possible-Control-by-$\mathcal{C}$), let the agenda Φ' be the union of Φ, the formula $\psi = \varphi_1 \wedge \cdots \wedge \varphi_t$, and its negation. Extend all involved profiles to the new agenda and let $D' = \{\psi\}$ be the new desired set. Note that ψ is not part of the new instance's collective outcome.

We claim that there is a control action of type $\mathcal{C}$ so that D is part of the new collective outcome in the original instance if and only if there is a control action of type $\mathcal{C}$ in the new instance so that the chair possibly prefers the new to the old outcome under unrestricted (top-respecting) preferences induced by D'.

From left to right, assume that it is possible for the chair to include D in the collective outcome, i.e., exert control of type $\mathcal{C}$ so that $\varphi_1, \dots, \varphi_t$ get accepted. Then, the same control action in the new instance results in ψ being part of the collective outcome. Note that under unrestricted as well as under top-respecting preferences, the chair possibly prefers a set containing her desired set to any set without this property, so the control action is successful in the new instance.

From right to left, assume that there is a control action so that the chair possibly prefers the new collective outcome of the new instance to the old one. Under unrestricted preferences, this is the case whenever the corresponding intersections of the outcomes with the desired set, i.e., with D', differ from each other. Since the intersection of the old outcome with D' is empty, this is only the case when the new outcome includes ψ. Under top-respecting preferences, the chair possibly prefers the new outcome to the old one if the intersections of these two sets with the desired set D' both do not equal D' and these intersections differ from each other (which is not possible here, so this case cannot occur) or the new outcome contains D', whereas the old one does not. It follows that the chair was able to include ψ in the collective outcome. But then, the chair can employ the same successful control action in the original instance to include $\varphi_1, \ldots, \varphi_t$ in the new collective outcome. □

Note that the above proof requires the desired set of the preference type variant to be incomplete. Note further that the above proof also works to show that, for any of the control types defined in this paper $\mathcal{C}$, the exact control problem can be reduced to possible control under closeness-respecting preferences, to necessary control under top-respecting and closeness-respecting preferences, and to control under Hamming-distance-induced preferences.

Proposition 12. *Let $\mathcal{C}$ be any of the control types defined in this paper and let q, $0 \leq q < 1$, be a rational quota. $UPQR_q$-Exact-Control-by-$\mathcal{C}$ and $UPQR_q$-TR-Necessary-Control-by-$\mathcal{C}$ are equal in terms of complexity.*

Proof. We only show one proof direction (i.e., $UPQR_q$-Exact-Control-by-$\mathcal{C}$ $\leq_m^p$ $UPQR_q$-TR-Necessary-Control-by-$\mathcal{C}$); the reduction in the other direction works analogously. In the case that the desired set is already a part of the collective outcome in the instance of $UPQR_q$-Exact-Control-by-$\mathcal{C}$, construct an arbitrary no-instance for $UPQR_q$-TR-Necessary-Control-by-$\mathcal{C}$. Otherwise, note that in the case of necessary control under top-respecting preferences, the external agent necessarily prefers a judgment set X to Y only if her desired set is a subset of X, but not of Y. Therefore, the chair is able to include her desired set in the collective outcome after exerting control of type $\mathcal{C}$ if and only if there is a control action of type $\mathcal{C}$ so that she necessarily prefers the new collective outcome to the old one in the same instance under top-respecting preferences. □

Proposition 13. *Let $\mathcal{C}$ be any of the control types defined in this paper, let $T \in \{U, TR\}$ be a preference type, and let the desired set be complete. For each rational quota q, $0 \leq q < 1$, $UPQR_q$ is vulnerable to possible control by $\mathcal{C}$ under T.*

Proof. In the case of unrestricted preferences, the chair possibly prefers every new outcome to the current outcome. Since her desired set is complete, she only has to check if she can change a premise so as to change the collective judgment set. This is possible in polynomial time for every $\mathcal{C}$.

In the case of top-induced preferences, the chair possibly prefers every new outcome to the current outcome as long as the latter is not identical to her desired set. Therefore, it also suffices to change some premise if possible. □

Proposition 14. *Let $\mathcal{C}$ be any of the control types defined in this paper. For each rational quota q, $0 \leq q < 1$, $UPQR_q$ is immune to necessary control by $\mathcal{C}$ under unrestricted preferences.*

Proof. In the case of unrestricted preferences, there is never a new outcome that is necessarily preferred to the old outcome. It follows that the chair cannot exert control by $\mathcal{C}$ in a way that he necessarily prefers the new outcome to the old one, so $UPQR_q$ is immune to control by $\mathcal{C}$ under unrestricted preferences. □

Proposition 14 explains why we do not further consider the combination of necessary control with unrestricted preferences. In the following sections, we show that possible and necessary control of type $\mathcal{C}$ under closeness-respecting preferences is NP-hard for uniform (constant) premise-based quota rules with a certain quota, even for a complete desired set of the external agent.[14] Lemma 10 then proves NP-hardness of the corresponding exact control problem, and Lemmas 9 and 11 show NP-hardness of most other problem variants.

5. Control by adding judges

In the manipulation and bribery problems studied by Endriss et al. [30] and Baumeister et al. [9], the number of judges participating is constant and hence uniform premise-based quota rules and uniform constant premise-based quota rules describe the same judgment aggregation procedures. However, this is not the case if the number of judges participating is not fixed as in control by adding or deleting judges. For uniform premise-based quota rules the number of affirmations needed to be in the collective judgment set varies with the number of judges, whereas for uniform *constant* premise-based quota rules the number of affirmations remains the same regardless of the number of judges participating. Therefore, we will study these problems with respect to both judgment aggregation procedures separately.

5.1. Uniform constant premise-based quota rules

We start with uniform constant premise-based quota rules.

[14] Note that NP-hardness of a problem's special case where the desired set is complete immediately implies NP-hardness of the more general problem with an incomplete desired set.

Table 6
Construction for the first part of the proof of Theorem 15.

Judgment set	v_1	$\cdots$	v_n	β	$\psi \vee \beta$
$J_1, \dots, J_q$	1	$\cdots$	1	0	1
J_{q+1}	0	$\cdots$	0	0	1
$UCPQR_q$	0	$\cdots$	0	0	1
D	1	$\cdots$	1	1	1

Table 7
Construction for the second part of the proof of Theorem 15.

Judgment set	v_1	$\cdots$	v_n	β	$\psi' \vee \beta$
$J'_1, \dots, J'_q$	1	$\cdots$	1	0	1
J'_{q+1}	0	$\cdots$	0	0	0
$UCPQR_q$	0	$\cdots$	0	0	0
D'	0	$\cdots$	0	1	1

Theorem 15. *For each admissible value of q, $UCPQR_q$ is resistant to necessary and possible control by adding judges under closeness-respecting preferences, even for a complete desired set.*

Proof. The proof works by a reduction from the problem Dominating-Set (recall the definition from Section 2.3). Let (G, k) with $G = (V, E)$ and $V = \{v_1, \dots, v_n\}$ be a Dominating-Set instance. The neighbors of vertex v_i (including v_i itself) will be denoted by $v_i^1, v_i^2, \dots, v_i^{j_i}$ where j_i indicates the size of the closed neighborhood of v_i.

For the first part of the theorem (i.e., for showing NP-hardness of $UCPQR_q$-CR-Necessary-Control-by-Adding-Judges), we construct an instance of the control problem as follows. The agenda Φ contains the literals $v_1, \dots, v_n$, β, the formula $\psi \vee \beta$, where $\psi = (\varphi_1 \wedge \cdots \wedge \varphi_n) \vee (\neg\varphi_1 \wedge \cdots \wedge \neg\varphi_n)$ and $\varphi_i = v_i^1 \vee \cdots \vee v_i^{j_i}$, and all corresponding negations. The quota for every positive literal is q, hence $q+1$ affirmations are needed to be in the collective judgment set. The profile $\mathbf{J} = (J_1, \dots, J_{q+1})$, the outcome, and the chair's desired set D can be seen in Table 6. Note that each v_i, $1 \le i \le n$, needs one additional affirmation to be contained in the collective outcome.

The chair can choose at most k judgment sets from the profile $\mathbf{K} = (K_1, \dots, K_n)$ to add to $\mathbf{J}$, where each K_i, $1 \le i \le n$, contains $\neg\beta$, v_i, $\neg v_j$ for each $j \ne i$, $1 \le j \le n$, and the corresponding conclusion. Note that all existing agreements with the chair's desired set D have to be preserved and at least one new agreement with D has to arise for the chair to necessarily prefer the new outcome to the old one under closeness-respecting preferences.

We claim that there is a dominating set of size k for G if and only if there is a successful control action by the chair.

From left to right, let V' be a dominating set of size k for G. We can ensure that all formulas φ_i, $1 \le i \le n$—and, therefore, the conclusion $\psi \vee \beta$—evaluate to true by adding those judges with judgment sets K_i where $v_i \in V'$. There are additional agreements of the new collective outcome with the desired set D (namely, the $v_i \in V'$ that are now contained in the collective outcome), while the already existing agreement (namely, the conclusion $\psi \vee \beta$) has been preserved, so the control action was successful.

Conversely, assume that the formula $\psi \vee \beta$ and at least one v_i or β evaluate to true. It is not possible to achieve this by having β in the collective outcome, since there are no individual judgment sets containing β. Hence, the collective outcome for v_i, $1 \le i \le n$, makes all φ_i and, therefore, the conclusion $\psi \vee \beta$ true. The maximum number of judges that can be added is k, and exactly one literal v_i is contained in the collective judgment set for each judge from $\mathbf{K}$ that is added. Hence, the vertices v_i corresponding to the judgment sets K_i from $\mathbf{K}$ that have been added must form a dominating set of size k for graph G.

We prove the second part of the theorem (i.e., NP-hardness of $UCPQR_q$-CR-Possible-Control-by-Adding-Judges) in a similar way. Unlike in the first part of the proof, the chair only has to achieve an additional agreement regardless of new differences, as pointed out right after Definition 5. In the agenda, replace the formula $\psi \vee \beta$ with the formula $\psi' \vee \beta$ where $\psi' = \varphi_1 \wedge \cdots \wedge \varphi_n$. All required changes in the profile $\mathbf{J}'$, the outcome, and the desired set D' can be seen in Table 7.

To obtain the profiles $\mathbf{K}'$ of judgment sets to choose from, the premises of the judgment sets in $\mathbf{K}$ restricted to the corresponding new agenda remain unchanged and the new conclusion is evaluated accordingly. As above, the chair is allowed to add k judgment sets. No judge accepts β, so the chair has to achieve an additional agreement with the conclusion. This is only possible if the chair adds the judgment sets from $\mathbf{K}'$ that correspond to the vertices in a dominating set of G. It follows that the control action is successful if and only if G has a dominating set of size k. □

5.2. Uniform premise-based quota rules

We now turn to the results for the uniform premise-based quota rules, where we only consider $UPQR_{1/2}$, which equals the premise-based procedure *PBP* whenever there are an odd number of judges but which unlike *PBP* is also defined for an even number of judges. We show resistance to control by adding judges.

Table 8
Construction for the proof of Theorem 16.

Judgment set	α_0	α_1	$\cdots$	α_{3m}	β	$\varphi \vee \beta$
J_1	1	1	$\cdots$	1	0	1
$J_2, \dots, J_m$	0	1	$\cdots$	1	0	0
J_{m+1}	0	0	$\cdots$	0	0	0
$UPQR_{1/2}$	0	1	$\cdots$	1	0	0
D	0	1	$\cdots$	1	1	1

Theorem 16. *$UPQR_{1/2}$ is resistant to possible and necessary control by adding judges under closeness-respecting preferences, even for a complete desired set.*

Proof. The proof works by a reduction from X3C (recall the definition from Section 2.3). Let (X, C) be an X3C instance, where $X = \{x_1, \dots, x_{3m}\}$ and $C = \{C_1, \dots, C_n\}$. For the first part of the theorem (i.e., for showing NP-hardness of $UPQR_{1/2}$-CR-Possible-Control-by-Adding-Judges), let the agenda Φ contain the literals $\alpha_0, \alpha_1, \dots, \alpha_{3m}, \beta$, the formula $\varphi \vee \beta$ with $\varphi = \alpha_0 \wedge \dots \wedge \alpha_{3m}$, and the corresponding negations. The profile $\mathbf{J} = (J_1, \dots, J_{m+1})$, the collective judgment set $UPQR_{1/2}(\mathbf{J})$, and the desired set D can be seen in Table 8.

Let $\mathbf{K} = (K_1, \dots, K_n)$ be the profile containing the individual judgment sets to be added, where $K_i = \{\alpha_0, \alpha_j, \neg\alpha_l, \neg\beta \mid x_j \in C_i, x_l \notin C_i, 1 \le j, l \le 3m\}$. The chair is allowed to add m judgment sets from $\mathbf{K}$.

We claim that there is a profile $\mathbf{K}' \subseteq \mathbf{K}$, $|\mathbf{K}'| \le m$, such that the chair possibly prefers the outcome in $\mathbf{J} \cup \mathbf{K}'$ to the one in $\mathbf{J}$ if and only if there is an exact cover for the given X3C instance.

From right to left, assume that there is an exact cover $\mathcal{C}' \subseteq \mathcal{C}$ for the given X3C instance $(X, \mathcal{C})$. Then there is a profile $\mathbf{K}'$ containing those judges K_i with $C_i \in \mathcal{C}'$. The total number of judges is then $2m + 1$. The number of affirmations needed to be in the collective judgment set is strictly greater than $m + (1/2)$, so $m + 1$ affirmations are needed. Note that α_0 gets one affirmation from the judges in $\mathbf{J}$ and m affirmations from the judges in $\mathbf{K}'$. Every α_i, $1 \le i \le 3m$, gets m affirmations from the judges in $\mathbf{J}$ and one affirmation from a judge in $\mathbf{K}'$. Hence, $\varphi \vee \beta$ evaluates to true and the new collective outcome has an additional agreement with D, as desired.

From left to right, assume that there is a successful control action, i.e., the chair can achieve an additional agreement of the new outcome with D. Since no judge accepts β, the additional agreement can only occur for the formula $\varphi \vee \beta$. To add α_0, the chair has to add at least m judges for a total of $2m + 1$ judges. But then every α_i, $1 \le i \le 3m$, needs at least one additional affirmation. Therefore, the sets C_i corresponding to the judges in $\mathbf{K}'$ must form an exact cover for the given X3C instance, since this is the only possibility to achieve the additional affirmations for each α_i while not exceeding the allowed number of judges to add. This shows that $UPQR_{1/2}$-CR-Possible-Control-by-Adding-Judges is NP-hard.

Concerning the proof of the second part (i.e., NP-hardness of $UPQR_{1/2}$-CR-Necessary-Control-by-Adding-Judges), let the agenda Φ' contain only α_0, α_1, $\dots$, α_{3m} and the corresponding negations. Let $\mathbf{J}^*$ and $\mathbf{K}^*$ be the corresponding profiles restricted to Φ' and let $D' = \{\alpha_0, \alpha_1, \dots, \alpha_{3m}\}$ be the chair's desired set. The only difference between the desired set and the collective outcome is α_0. Since the chair has to preserve the initial agreements with D', by a similar argumentation as above, there is a successful control action if and only if there is an exact cover for the given X3C instance. Thus $UPQR_{1/2}$-CR-Necessary-Control-by-Adding-Judges is NP-hard. □

Note that the second proof part of Theorem 16 uses an agenda that only consists of premises and can also be used immediately to prove the next theorem since the only successful control action is to fully include the desired set in the new collective outcome. Furthermore, it is an alternative proof to show that $UPQR_{1/2}$ is resistant to control by adding judges under Hamming-distance-induced preferences.

Theorem 17. *$UPQR_{1/2}$ is resistant to exact control by adding judges, even for a complete desired set.*

6. Control by deleting judges

Next, we turn to control by deleting judges.

6.1. Uniform constant premise-based quota rules

Again, we start with the results for the uniform constant premise-based quota rules.

Theorem 18. *For each admissible value of q, $UCPQR_q$ is resistant to necessary and possible control by deleting judges under closeness-respecting preferences, even for a complete desired set.*

Proof. We will show NP-hardness by a reduction from Dominating-Set. For a given Dominating-Set instance (G, k), we construct the following judgment aggregation scenarios.

Table 9
Construction for the first part of the proof of Theorem 18.

Judgment set	v_1	v_2	v_3	$\cdots$	v_{n-2}	v_{n-1}	v_n	β	γ	$\psi \vee \beta$
$J_1, \ldots, J_q$	1	1	1	$\cdots$	1	1	1	0	1	1
J_{q+1}	0	0	0	$\cdots$	0	0	0	0	1	1
L_1	1	0	0	$\cdots$	0	0	0	0	0	?
L_2	0	1	0	$\cdots$	0	0	0	0	0	?
L_3	0	0	1	$\cdots$	0	0	0	0	0	?
$\cdots$	$\cdots$									
L_{n-2}	0	0	0	$\cdots$	1	0	0	0	0	?
L_{n-1}	0	0	0	$\cdots$	0	1	0	0	0	?
L_n	0	0	0	$\cdots$	0	0	1	0	0	?
$UCPQR_q$	1	1	1	$\cdots$	1	1	1	0	1	1
D	0	0	0	$\cdots$	0	0	0	1	1	1

Table 10
Construction for the second part of the proof of Theorem 18.

Judgment set	v_1	v_2	v_3	$\cdots$	v_{n-2}	v_{n-1}	v_n	β	γ	$\psi' \vee \beta$
$J_1, \ldots, J_q$	1	1	1	$\cdots$	1	1	1	0	1	0
J_{q+1}	0	0	0	$\cdots$	0	0	0	0	1	1
L_1	1	0	0	$\cdots$	0	0	0	0	0	?
L_2	0	1	0	$\cdots$	0	0	0	0	0	?
L_3	0	0	1	$\cdots$	0	0	0	0	0	?
$\cdots$	$\cdots$									
L_{n-2}	0	0	0	$\cdots$	1	0	0	0	0	?
L_{n-1}	0	0	0	$\cdots$	0	1	0	0	0	?
L_n	0	0	0	$\cdots$	0	0	1	0	0	?
$UCPQR_q$	1	1	1	$\cdots$	1	1	1	0	1	0
D'	1	1	1	$\cdots$	1	1	1	1	1	1

For the first part of the theorem (i.e., for showing NP-hardness of $UCPQR_q$-CR-Necessary-Control-by-Deleting-Judges), let the agenda Φ contain the premises v_i, $1 \le i \le n$, β, γ, and the conclusion $\psi \vee \beta$, where $\psi = (\varphi_1 \wedge \cdots \wedge \varphi_n) \vee (\neg\varphi_1 \wedge \cdots \wedge \neg\varphi_n)$ and $\varphi_i = \neg v_i^1 \vee \cdots \vee \neg v_i^{j_i}$ (where j_i denotes the size of the closed neighborhood of v_i), and all corresponding negations. The quota is q. Let $\mathbf{J} = \mathbf{T} \cup \mathbf{L}$ be the profile to delete judges from, where $\mathbf{T} = (J_1, \ldots, J_{q+1})$ and $\mathbf{L} = (L_1, \ldots, L_n)$, as stated in Table 9 with the chair's desired set D and the collective outcome. Note that each L_i, $1 \le i \le n$, accepts only the premise v_i and evaluates the conclusion accordingly (which therefore is indicated by a question mark in Table 9).

We claim that there is a dominating set of size at most k for G if and only if there is a successful control action, i.e., deleting at most k judges so that the new collective outcome preserves all existing agreements with D and adds an additional agreement.

Let V' be a dominating set for G of size at most k. Then the outcome obtained by deleting those L_i with $v_i \in V'$ preserves all agreements with D (because the conclusion still evaluates to true) and adds a new agreement, namely at least one $\neg v_i$, $1 \le i \le n$.

Conversely, assume that a successful control action is possible by deleting at most k judges. No judge accepts β, so this formula will never be in the collective outcome. Since the agreement with γ has to be preserved, only judges of the form L_i, $1 \le i \le n$, may be deleted. The deletion of a judge L_i has the effect that v_i is not in the collective outcome (hence, at most k different v_i are not contained in the collective outcome), and since they evaluate all formulas φ_i, $1 \le i \le n$ to true, those v_i must form a dominating set of size at most k for G.

For the second part of the theorem (i.e., for showing NP-hardness of $UCPQR_q$-CR-Possible-Control-by-Deleting-Judges), replace the formula $\psi \vee \beta$ with the formula $\psi' \vee \beta$, where $\psi' = \varphi_1 \wedge \cdots \wedge \varphi_n \wedge \gamma$, and adjust the profile accordingly. The profile, the chair's desired set D', and the collective outcome can be seen in Table 10.

The only possible additional agreement of a new collective outcome with the desired set D' is the conclusion, since no judge accepts β. Therefore, only judges from $\mathbf{L}$ can be deleted for γ to be contained in the new collective outcome. Again, there is a dominating set of size at most k for G if and only if there is a successful control action. $\square$

6.2. Uniform premise-based quota rules

We now turn to the results for uniform premise-based quota rules, where we again consider $UPQR_{1/2}$ only. Again, we will show resistance to control by deleting judges for the following problem variants.

Theorem 19. *$UPQR_{1/2}$ is resistant to possible and necessary control by deleting judges under closeness-respecting preferences, even for a complete desired set.*

Proof. Let (X, C) be an X3C instance, where $X = \{x_1, \ldots, x_{3m}\}$ and $C = \{C_1, \ldots, C_n\}$. We assume that each x_i is contained in exactly three sets C_j, since X3C is known to be NP-complete even under this restriction (as shown by Gonzalez [39]).

For the first part (i.e., for showing NP-hardness of *UPQR$_q$-CR*-Possible-Control-by-Deleting-Judges), let Φ be the agenda containing the literals $\alpha_0, \alpha_1, \ldots, \alpha_{3m}, \beta, \gamma$, the formula $\varphi \vee \beta$ with $\varphi = \alpha_0 \wedge \cdots \wedge \alpha_{3m} \wedge \neg\gamma$, and all corresponding negations. Let $\mathbf{J} = \mathbf{T} \cup \mathbf{L}$ be a profile, where $\mathbf{T} = (J_1, \ldots, J_{n+m})$ and $\mathbf{L} = (L_1, \ldots, L_n)$ for a total of $2n + m$ judges. For each i, $1 \le i \le n + m$, let J_i contain $\neg\beta$, α_j if $i \le 3m$ (and $\neg\alpha_j$ otherwise) for $1 \le j \le 3m$, α_0 if $i \le n + 1$ (and $\neg\alpha_0$ otherwise), γ if $i \le m$ (and $\neg\gamma$ otherwise), and the corresponding conclusion $\varphi \vee \beta$ (respectively, $\neg(\varphi \vee \beta)$). Furthermore, for $1 \le i \le n$, define

$$L_i = \{\neg\beta, \gamma, \neg\alpha_0, \alpha_j, \neg\alpha_l, \neg(\varphi \vee \beta) \mid x_j \notin C_i, x_l \in C_i, 1 \le j, l \le 3m\}.$$

The threshold for a premise to be included in the collective outcome is $n + m/2$. Since β has no affirmation, γ and every a_k, $1 \le k \le 3m$, each have $n + m$ affirmations, and since α_0 has $n + 1$ affirmations, it follows that

$$UPQR_{1/2}(\mathbf{J}) = \{\neg\alpha_0, \alpha_1, \ldots, \alpha_{3m}, \neg\beta, \gamma, \neg(\varphi \vee \beta)\}.$$

Let the chair's desired set be $D = \{\neg\alpha_0, \alpha_1, \ldots, \alpha_{3m}, \beta, \gamma, \varphi \vee \beta\}$. She is able to delete m individual judgment sets from the profile $\mathbf{J}$. We claim that there is an exact cover for the given X3C instance if and only if there is a successful control action by the chair.

From left to right, assume that there is an exact cover for the given X3C instance. Delete the judges from $\mathbf{L}$ whose individual judgment sets correspond to this exact cover. The new threshold is $n + 1$, so γ is not contained in the collective outcome anymore—since it lost m affirmations for a new total of n affirmations—but each α_i, $0 \le i \le 3m$ and therefore the conclusion $\varphi \vee \beta$ is part of the outcome. The chair achieves a new agreement of the desired set with the new outcome, namely the conclusion, so that the control action was successful.

From right to left, assume that there exists a successful control action, i.e., the chair is able to delete up to m judges so that the new collective outcome has a new agreement with her desired set. Since no judge accepts β, it will never be in the collective outcome. Therefore, the new agreement of the desired set with the new outcome has to occur in the conclusion. Since φ is a conjunction including α_0, this is only possible if the chair manages to include α_0 in the new collective outcome, even though α_0 is not part of the chair's desired set. To include α_0, the chair has to delete m judges to lower the acceptance threshold to $n + 1$. These judges' individual judgment sets have to contain γ, so that γ loses m affirmations and is not contained in the collective outcome anymore, and cannot contain α_0 in order for α_0 to have the needed number of $n + 1$ affirmations to be part of the outcome. Therefore, the judges' sets have to be deleted from $\mathbf{L}$. The α_i, $1 \le i \le 3m$, are only allowed to lose $m - 1$ affirmations so that they are still contained in the new collective outcome. If some x_i is not contained in one of the sets C_j that match the individual judgment sets of the deleted judges, the corresponding α_i loses too many affirmations and is thus rejected in the new collective outcome. The control action is successful (i.e., $\varphi \vee \beta$ is contained in the new collective outcome) if and only if the sets C_j corresponding to the deleted individual judgment sets form an exact cover of X. This shows that *UPQR$_{1/2}$-CR*-Possible-Control-by-Deleting-Judges is NP-hard.

To prove the second part (i.e., NP-hardness of *UPQR$_q$-CR*-Necessary-Control-by-Replacing-Judges), we create a new agenda Φ' from Φ by removing β, $\varphi \vee \beta$, and the corresponding negations, and by adding the formula $\psi = (\neg\alpha_0 \wedge \gamma) \vee (\alpha_0 \wedge \neg\gamma)$ and its negation. Let $\mathbf{J}^* = \mathbf{T}^* \cup \mathbf{L}^*$ be the resulting profile that is obtained by restricting $\mathbf{T}$ and $\mathbf{L}$ to Φ' and by adding the corresponding conclusions to all J_i and L_j. Then it holds that

$$UPQR_{1/2}(\mathbf{J}^*) = \{\neg\alpha_0, \alpha_1, \ldots, \alpha_{3m}, \gamma, \psi\}.$$

Let $D' = \{\alpha_0, \alpha_1, \ldots, \alpha_{3m}, \neg\gamma, \psi\}$ be the chair's desired set and let the chair be able to delete m judgment sets.

We claim that there is an exact cover for the given X3C instance if and only if there is a successful control action by the chair, i.e., it is possible for the chair to delete at most m judges so that the new collective outcome has an additional agreement with his desired set while preserving the existing agreements.

From left to right, delete the judges from $\mathbf{L}^*$ that correspond to the exact cover. As argued above, it follows that the new collective outcome contains $\neg\gamma$, each α_i, $0 \le i \le 3m$, and therefore also the conclusion ψ, i.e., the new collective outcome is identical to the desired set and the control action was successful.

From right to left, assume that there is a successful control action. To preserve the agreements on the conclusion, the chair has to change the collective outcome in regard to α_0 as well as γ. Again, the chair has to delete exactly m judgment sets (for α_0 to meet the acceptance threshold), can only delete judgment sets from $\mathbf{L}^*$ (to ensure that α_0 is accepted in the new collective outcome, but γ is not), and, therefore, can preserve the agreements concerning the α_i if and only if the sets C_j corresponding to the deleted individual judgment sets form an exact cover of X. Thus *UPQR$_{1/2}$-CR*-Necessary-Control-by-Deleting-Judges is NP-hard. □

We already proved the following theorem in the second part of the proof of Theorem 19, since in this proof the chair has to make sure that the new collective outcome is identical to his desired set.

Theorem 20. *UPQR$_{1/2}$ is resistant to exact control by deleting judges, even for a complete desired set.*

Table 11
Construction for the first part of the proof of Theorem 21.

Judgment set	v_1	$\cdots$	v_n	β	γ	$\psi \vee \beta$
(a) Rational quota q with $0 \leq q < 1/2$						
$J_1, \ldots, J_{\lfloor m \cdot q \rfloor}$	1	$\cdots$	1	0	1	1
$J_{\lfloor m \cdot q \rfloor + 1}$	0	$\cdots$	0	0	1	0
$J_{\lfloor m \cdot q \rfloor + 2}, \ldots, J_m$	0	$\cdots$	0	0	0	0
$UPQR_q$	0	$\cdots$	0	0	1	0
D	0	$\cdots$	0	1	1	1
Judgment set	v_1	$\cdots$	v_n	β	γ	$\psi' \vee \neg\beta$
(b) Rational quota q with $1/2 \leq q < 1$						
$J'_1, \ldots, J'_{\lceil m \cdot (1-q) \rceil - 1}$	0	$\cdots$	0	1	0	1
$J'_{\lceil m \cdot (1-q) \rceil}$	1	$\cdots$	1	1	0	0
$J'_{\lceil m \cdot (1-q) \rceil + 1}, \ldots, J'_m$	1	$\cdots$	1	1	1	0
$UPQR_q$	1	$\cdots$	1	1	0	0
D'	1	$\cdots$	1	0	0	1

7. Control by replacing judges for uniform premise-based quota rules

Turning now to control by replacing judges, note that—in contrast to the problems of control by adding and deleting judges—the number of judges here is constant. Thus there is no difference between uniform constant premise-based quota rules and the corresponding uniform premise-based quota rules. The following result thus establishes resistance for both classes of rules at one fell swoop; we will state it only for $UPQR_q$.

Theorem 21. *For each rational quota q, $0 \leq q < 1$, $UPQR_q$ is resistant to possible and necessary control by replacing judges under closeness-respecting preferences, even for a complete desired set.*

Proof. The proof works by a reduction from the problem Dominating-Set. Let (G, k) with $G = (V, E)$ and $V = \{v_1, \ldots, v_n\}$ be a Dominating-Set instance. The neighbors of vertex v_i (including v_i itself) will be denoted by $v_i^1, v_i^2, \ldots, v_i^{j_i}$, where j_i indicates the size of the closed neighborhood of v_i.

For the first part of the theorem (i.e., for showing NP-hardness of $UPQR_q$-CR-Possible-Control-by-Replacing-Judges), first assume that the quota q is lower than $1/2$. We construct an instance of the control problem as follows. The agenda Φ contains the literals $v_1, \ldots, v_n$, β, γ, the formula $\psi \vee \beta$, where $\psi = \varphi_1 \wedge \cdots \wedge \varphi_n \wedge \gamma$ and $\varphi_i = v_i^1 \vee \cdots \vee v_i^{j_i}$, and all corresponding negations. The profile $\mathbf{J} = \mathbf{T_1} \cup \mathbf{T_2}$ ($\mathbf{T_1} = (J_1, \ldots, J_{\lfloor m \cdot q \rfloor + 1}), \mathbf{T_2} = (J_{\lfloor m \cdot q \rfloor + 2}, \ldots, J_m)$) with $m = 2k + 1$ judges, the outcome, and the chair's desired set D can be seen in Table 11(a).

The chair can choose at most k judgment sets from the profile $\mathbf{K} = (K_1, \ldots, K_n)$ with $K_i = \{\neg\beta, \neg\gamma, v_i, \neg v_j, \neg(\psi \vee \beta) \mid 1 \leq j \leq n, i \neq j\}$ to replace judgment sets in $\mathbf{J}$.

We claim that there is a dominating set of size at most k for G if and only if there is a successful control action.

From left to right, replace arbitrary judges in $\mathbf{T_2}$ with the judges from $\mathbf{K}$ whose individual judgment sets correspond to the dominating set. This control action results in a new agreement of the desired set with the new collective outcome, since the v_i corresponding to the dominating set now meet the acceptance threshold and therefore the conclusion evaluates to true.

From right to left, assume that there is a successful control action that achieves a new agreement of the desired set with the new collective outcome. The formula β will never be contained in the outcome because no judge accepts it. In order to achieve the desired additional agreement between the new outcome and D, the chair has to get the conclusion—and, therefore, ψ—be accepted. Each v_i needs at least one additional affirmation to be contained in the new outcome. Note that only judgment sets in $\mathbf{T_2}$ can be replaced (or else γ would lose an affirmation, would not be contained in the collective outcome anymore, and thus ψ cannot be evaluated to true). Since $\psi \vee \beta$ is contained in the new outcome if and only if the accepted v_i form a dominating set, and since only k judgment sets can be replaced, the control action is successful under closeness-respecting preferences if and only if G has a dominating set of size k. This completes the proof of NP-hardness of $UPQR_q$-CR-Possible-Control-by-Replacing-Judges for the case of $q < 1/2$.

In the case of a quota q greater than or equal to $1/2$, the agenda changes slightly. Instead of the formula $\psi \vee \beta$ and its negation the new agenda Φ' contains the formula $\psi' \vee \neg\beta$ with $\psi' = \varphi'_1 \wedge \cdots \wedge \varphi'_n \wedge \neg\gamma$ and $\varphi'_i = \neg v_i^1 \vee \cdots \vee \neg v_i^{j_i}$, and its negation, $\neg(\psi' \vee \neg\beta)$. The profile $\mathbf{J'} = \mathbf{T'_1} \cup \mathbf{T'_2}$ ($\mathbf{T'_1} = (J'_1, \ldots, J'_{\lceil m \cdot (1-q) \rceil}), \mathbf{T'_2} = (J'_{\lceil m \cdot (1-q) \rceil + 1}, \ldots, J'_m)$) with $m = 2k + 1$ judges, the outcome, and the chair's desired set D' can be seen in Table 11(b).

Let $\mathbf{K'} = (K'_1, \ldots, K'_n)$ be a profile, where

$$K'_i = \{\beta, \gamma, \neg v_i, v_j, \neg(\psi' \vee \neg\beta) \mid 1 \leq j \leq n, i \neq j\}$$

Table 12
Construction for the second part of the proof of Theorem 21.

Judgment set	v_1	$\cdots$	v_n	γ	Ψ
(a) Rational quota q with $0 \le q < 1/2$					
$J_1^*, \dots, J_{\lfloor m \cdot q \rfloor}^*$	1	$\cdots$	1	1	1
$J_{\lfloor m \cdot q \rfloor + 1}^*$	0	$\cdots$	0	1	1
$J_{\lfloor m \cdot q \rfloor + 2}^*, \dots, J_m^*$	0	$\cdots$	0	0	1
$UPQR_q$	0	$\cdots$	0	1	1
D^*	1	$\cdots$	1	1	1

Judgment set	v_1	$\cdots$	v_n	γ	Ψ'
(b) Rational quota q with $1/2 \le q < 1$					
$J_1'^*, \dots, J_{\lceil m \cdot (1-q) \rceil - 1}'^*$	0	$\cdots$	0	0	1
$J_{\lceil m \cdot (1-q) \rceil}'^*$	1	$\cdots$	1	0	1
$J_{\lceil m \cdot (1-q) \rceil + 1}'^*, \cdots J_m'^*$	1	$\cdots$	1	1	1
$UPQR_q$	1	$\cdots$	1	0	1
D'^*	0	$\cdots$	0	0	1

for $1 \le i \le n$. Again, the chair is able to replace k judgment sets from $\mathbf{J}'$ with k judgment sets from $\mathbf{K}'$. A formula needs at least $\lceil m(1-q) \rceil$ rejections in order to not be accepted. Since every judge accepts β, its negation will never be contained in the collective outcome. Thus the chair has to get ψ' accepted so as to achieve the desired additional agreement of the new outcome with D'. The argumentation then follows from the first case: Since ψ' is true if and only if the rejected v_i form a dominating set and since the k replaceable judgment sets must be from $\mathbf{T}_2'$, the control action is successful under closeness-respecting preferences if and only if G has a dominating set of size k.

We prove the second part of the theorem (i.e., NP-hardness of $UPQR_q$-CR-Necessary-Control-by-Replacing-Judges) in a similar way. Unlike in the first part of the proof, the chair now has to necessarily prefer the new outcome to the current one. That means that all existing agreements have to be preserved. Remove β from the former agenda Φ (respectively, Φ') and replace all appearances of ψ (respectively, ψ') with the formula $\Psi = \psi \vee (\neg v_1 \wedge \cdots \wedge \neg v_n)$ (respectively, $\Psi' = \psi' \vee (v_1 \wedge \cdots \wedge v_n)$). All required changes in the resulting profiles $\mathbf{J}^* = \mathbf{T}_1^* \cup \mathbf{T}_2^*$ (respectively, $\mathbf{J}'^* = \mathbf{T}_1'^* \cup \mathbf{T}_2'^*$), the outcome, and the desired set D^* (respectively, D'^*) can be seen in Table 12(a) (respectively, in Table 12(b)).

To obtain the profiles $\mathbf{K}^*$ (respectively, $\mathbf{K}'^*$) of judgment sets to choose from, the premises of the judgment sets in $\mathbf{K}$ (respectively, $\mathbf{K}'$) restricted to the corresponding new agenda remain unchanged and the new conclusion is evaluated accordingly. As above, the chair is allowed to replace k judgment sets.

We claim that there is a dominating set of size k for G if and only if the chair achieves an additional agreement of her desired set with the new collective outcome while preserving all existing agreements.

From left to right, replace arbitrary judges in $\mathbf{T}_2^*$ (respectively, $\mathbf{T}_2'^*$) with the judges from $\mathbf{K}^*$ (respectively, $\mathbf{K}'^*$) that correspond to a dominating set in G. This results in additional agreements of the desired set with the new collective outcome—namely the v_i (respectively, $\neg v_i$) from the dominating set that gain the needed additional affirmation—while preserving the agreements with γ and the conclusion, i.e., this action is a successful control action by the chair.

From right to left, assume that there is a successful control action by the chair. The chair has to change some premise different from γ in order to achieve a new agreement. But after this action the second part of Ψ (respectively, Ψ') is not satisfied anymore. In order to preserve the agreement of the outcome with her desired set regarding the conclusion, the chair has to replace some judgment sets from $\mathbf{T}_2^*$ (respectively, $\mathbf{T}_2'^*$) with the judgment sets from $\mathbf{K}^*$ (respectively, $\mathbf{K}'^*$) that correspond to the vertices in a dominating set of G. It follows that the control action is successful if and only if G has a dominating set of size k. $\square$

8. Control by bundling judges for uniform premise-based quota rules

Finally, we consider control by bundling judges. Note that $UPQR_q$-Control-by-Bundling-Judges is somewhat similar to $UPQR_q$-Control-by-Deleting-Judges. We will exploit this in the following proof. Note that it does not make sense to consider uniform *constant* premise-based quota rules for control by bundling judges: If we have a constant number of judges and then partition the group of judges, bundling them to smaller groups, it would not be reasonable to have the original constant number of judges carry over to the smaller groups.

Theorem 22. *$UPQR_{1/2}$ is resistant to possible and necessary control by bundling judges under closeness-respecting preferences, even for a complete desired set.*

Proof. The proofs will be by a reduction from the related problem $UPQR_{1/2}$-Exact-Control-by-Deleting-Judges. We are given an agenda $\Phi = \Phi_p \cup \Phi_c$,[15] a profile $\mathbf{J} \in \mathcal{J}(\Phi)^n$, and a positive integer k as a bound on the number of judges that may be deleted. The quota $1/2$ holds for every positive literal in the agenda. Let D be the desired set. Note that we can assume that D is complete since we showed in Theorem 20 that this restriction does not change the NP-hardness of the problem. Without loss of generality, we assume that $n \ge k + 2$.

We begin by showing that $UPQR_{1/2}$-CR-Necessary-Control-by-Bundling-Judges is NP-hard. Let $\bigwedge X$ denote the conjunction of each formula in the set X. First, we construct an agenda $\Phi' = \Phi \cup \{\alpha, \neg\alpha\} \cup \{\varphi, \neg\varphi, \psi, \neg\psi\}$, where α is a newly

15 Recall that Φ_p denotes the set of premises and Φ_c the set of conclusions of Φ.

introduced premise, $\varphi = (\bigwedge UPQR_{1/2}(\mathbf{J})) \vee \alpha)$, and $\psi = (\bigwedge D) \vee \neg\alpha$. The profile $\mathbf{S} \in \mathcal{J}(\Phi')^{n+k+1}$ of the bundling instance contains the original n judgment sets of $\mathbf{J}$ extended by $\neg\alpha$ and the corresponding new conclusions, and $k+1$ new individual judgment sets that each contain $\gamma \in \Phi_p$ if and only if $\overline{\gamma} \in D$, they each contain α, and the conclusions are evaluated accordingly. The premises are divided into the following two subsets: The first one consists of Φ_p, and the second one is $\{\alpha, \neg\alpha\}$. The set of the $k+1$ new judges will be denoted by N'. The desired set is $D' = D \cup \{\alpha, \varphi, \psi\}$. We show that it is possible to obtain the desired set D by deleting at most k judges from $\mathbf{J}$ if and only if the judges from $\mathbf{S}$ can be bundled into two groups such that an additional agreement between D' and the new outcome can be achieved without losing an existing agreement.

From left to right, assume that there is a subset $\mathbf{T}' \subseteq \mathbf{J}$, $|\mathbf{T}'| \leq k$, such that $UPQR_{1/2}(\mathbf{J} \smallsetminus \mathbf{T}') = D$. Then the judges can be bundled as follows. The $k+1$ new judges and the judges corresponding to $\mathbf{T}'$ decide over α. Obviously the resulting outcome is identical to D', so the constructed instance is a positive one for $UPQR_{1/2}$-CR-Necessary-Control-by-Bundling-Judges.

From right to left, assume that the judges in the bundling instance can be bundled into N_1 and N_2 such that the collective outcome changes, but does preserve all existing agreements. This is only possible by adding α to the outcome because of the formula φ. However, to preserve the agreement with ψ, it holds that $UPQR_{1/2}(\mathbf{S}|_{\Phi_p, N_1}) = D$. Note that this implies that the chair obtains a new outcome identical to her desired set. Since α is contained in the collective judgment set and since there are only $k+1$ judges having α in their individual judgment set, at most k of the initial judges can be in N_2. Due to the uniform premise-based procedure, it is enough to show that in the case of $UPQR_{1/2}(\mathbf{S}|_{\Phi_p, N_1}) = D$, we have $UPQR_{1/2}(\mathbf{S}|_{\Phi_p, N_1 \smallsetminus N'}) = D$. But this holds trivially, since for all judges from N' we have that $\gamma \in \Phi_p$ is contained in the individual judgment set if and only if $\overline{\gamma} \in D$.

The proof that $UPQR_{1/2}$-CR-Possible-Control-by-Bundling-Judges is NP-hard uses the following agenda. Add the premises α and β, the conclusion $\varphi' = (\bigwedge D \wedge \alpha) \vee \beta$, as well as the corresponding negations to Φ and adapt the profile of the proof's first part to the new agenda by adding α or $\neg\alpha$ as before and adding $\neg\beta$ and the corresponding evaluation of φ' to each individual judgment. The premises are divided into the following two subsets: The first one consists of Φ_p, and the second one is $\{\alpha, \neg\alpha, \beta, \neg\beta\}$. Let $D'' = UPQR_{1/2}(\mathbf{J}) \cup \{\neg\alpha, \beta, \varphi'\}$ be the chair's desired set in the bundling instance. D'' only differs from the collective outcome in β and φ', i.e., the desired set contains β and φ', whereas the collective outcome does not. Since the chair cannot achieve to add β to the collective outcome, she has to obtain an agreement with φ' to possibly prefer the new outcome to the old one. This shows that it is possible to obtain the desired set D by deleting at most k judges from $\mathbf{J}$ if and only if the judges from the new profile can be bundled into two groups such that both $\bigwedge D$ and α are contained in the outcome. $\square$

The first part of the proof of Theorem 22 already proves the next theorem.

Theorem 23. *$UPQR_{1/2}$ is resistant to exact control by bundling judges, even for a complete desired set.*

9. Conclusions and open questions

We have introduced four fundamental control scenarios in judgment aggregation, inspired by the corresponding notions of electoral control that have been intensively studied in voting (see, e.g., [3,8,31,34,35,37,41,42]). We studied the complexity of control by adding, deleting, and replacing judges for the uniform constant premise-based quota rule. In the case of control by replacing judges this rule coincides with the uniform premise-based quota rule. The only open question remaining is the complexity of exact control when the chair's desired set is complete.[16]

Since it is not clear how to reasonably define the problem of control by bundling judges for uniform *constant* premise-based quota rules (see the remark right before Theorem 23), we only considered the uniform premise-based quota rules in this control scenario.

In this article, we have focused on classical computational complexity. However, as noted in the Introduction, a worst-case complexity measure like NP-hardness has its limitations, and it would therefore be interesting to study, for instance, the typical-case complexity of the problems considered here (see, e.g., the survey by Rothe and Schend [58] for such approaches to manipulation and control problems in voting).

Another interesting future research question is whether our hardness results also hold in restricted cases, e.g., when the number of judges or issues are small, and in particular to see whether they are fixed-parameter tractable or hard in terms of *parameterized* complexity—an approach applied, for example, by de Haan [23] to judgment aggregation. We have started looking at special cases by considering problems restricted to complete desired sets, and we even obtained some results when the agenda only consists of premises, as mentioned after these results. Relatedly, all our results transfer to the case where the desired set only consists of conclusions.[17]

---

[16] Note that the problem of exact control is equivalent to necessary control under top-respecting preferences in terms of complexity (Proposition 12).

[17] Proof sketch: We can slightly modify the proofs by adding the conclusion $\alpha \vee \delta$ for each premise α that occurs in the desired set, where δ is a newly introduced premise that each judge rejects, and replace each occurrence of α in the desired set with $\alpha \vee \delta$. The resulting desired set consists only of conclusions and each control action in this modified instance is successful if and only if the same control action is successful in the instance presented in our proofs.

The control scenarios we introduced here all take influence on the set of judges. Finally, another interesting direction for future research is to extend the study of control in judgment aggregation to actions on the agenda (see, e.g., the manuscript of Dietrich [25]), or to actions on the aggregation rule itself.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are grateful to the anonymous JCSS, STAIRS'12, COMSOC'12, ADT'13, ADT'15, and ESSLLI-LMGD'13 reviewers for helpful comments and literature pointers. This work was supported in part by DFG grants RO 1202/14-2, RO 1202/15-1, BA 6270/1-1, ER 738/1-1, and ER 738/2-1 and NRF (Singapore) grant NRF-RF 2009-08. This work was done in part while the second and third author were visiting University of Rochester and while the fourth author was visiting Stanford University, and we thank the hosts, Lane A. Hemaspaandra and Yoav Shoham, for their warm hospitality.

References

[1] N. Alon, D. Falik, R. Meir, M. Tennenholtz, Bundling attacks in judgment aggregation, in: Proceedings of the 27th AAAI Conference on Artificial Intelligence, AAAI Press, 2013, pp. 39–45.
[2] J. Bartholdi III, C. Tovey, M. Trick, The computational difficulty of manipulating an election, Soc. Choice Welf. 6 (3) (1989) 227–241.
[3] J. Bartholdi III, C. Tovey, M. Trick, How hard is it to control an election?, Math. Comput. Model. 16 (8/9) (1992) 27–40.
[4] D. Baumeister, G. Erdélyi, O. Erdélyi, J. Rothe, Bribery and control in judgment aggregation, in: F. Brandt, P. Faliszewski (Eds.), Proceedings of the 4th International Workshop on Computational Social Choice, AGH University of Science and Technology, Kraków, Poland, September 2012, pp. 37–48.
[5] D. Baumeister, G. Erdélyi, O. Erdélyi, J. Rothe, Control in judgment aggregation, in: Proceedings of the 6th European Starting AI Researcher Symposium, IOS Press, August 2012, pp. 23–34.
[6] D. Baumeister, G. Erdélyi, O. Erdélyi, J. Rothe, Computational aspects of manipulation and control in judgment aggregation, in: Proceedings of the 3rd International Conference on Algorithmic Decision Theory, in: Lecture Notes in Artificial Intelligence, vol. 8176, Springer-Verlag, November 2013, pp. 71–85.
[7] D. Baumeister, G. Erdélyi, O. Erdélyi, J. Rothe, Complexity of manipulation and bribery in judgment aggregation for uniform premise-based quota rules, Math. Soc. Sci. 76 (2015) 19–30.
[8] D. Baumeister, G. Erdélyi, E. Hemaspaandra, L. Hemaspaandra, J. Rothe, Computational aspects of approval voting, in: J. Laslier, R. Sanver (Eds.), Handbook on Approval Voting, Springer, 2010, pp. 199–251, chapter 10.
[9] D. Baumeister, G. Erdélyi, J. Rothe, How hard is it to bribe the judges? A study of the complexity of bribery in judgment aggregation, in: Proceedings of the 2nd International Conference on Algorithmic Decision Theory, in: Lecture Notes in Artificial Intelligence, vol. 6992, Springer-Verlag, October 2011, pp. 1–15.
[10] D. Baumeister, G. Erdélyi, J. Rothe, Judgment aggregation, in: J. Rothe (Ed.), Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division, Springer-Verlag, 2015, pp. 361–391, chapter 6.
[11] D. Baumeister, M. Roos, J. Rothe, L. Schend, L. Xia, The possible winner problem with uncertain weights, in: Proceedings of the 20th European Conference on Artificial Intelligence, IOS Press, August 2012, pp. 133–138.
[12] D. Baumeister, J. Rothe, Preference aggregation by voting, in: J. Rothe (Ed.), Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division, Springer-Verlag, 2015, pp. 197–325, chapter 4.
[13] D. Baumeister, J. Rothe, A. Selker, Complexity of bribery and control for uniform premise-based quota rules under various preference types, in: Proceedings of the 4th International Conference on Algorithmic Decision Theory, in: Lecture Notes in Artificial Intelligence, vol. 9346, Springer-Verlag, September 2015, pp. 432–448.
[14] D. Baumeister, J. Rothe, A. Selker, Strategic behavior in judgment aggregation, in: U. Endriss (Ed.), Trends in Computational Social Choice, AI Access Foundation, 2017, pp. 145–168, chapter 8.
[15] D. Binkele-Raible, G. Erdélyi, H. Fernau, J. Goldsmith, N. Mattei, J. Rothe, The complexity of probabilistic lobbying, Discrete Optim. 11 (1) (2014) 1–21.
[16] F. Brandt, V. Conitzer, U. Endriss, Computational social choice, in: G. Weiß (Ed.), Multiagent Systems, second edition, MIT Press, 2012, pp. 213–283.
[17] F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016.
[18] R. Bredereck, J. Chen, S. Hartung, S. Kratsch, R. Niedermeier, O. Suchý, G. Woeginger, A multivariate complexity analysis of lobbying in multiple referenda, J. Artif. Intell. Res. 50 (2014) 409–446.
[19] Y. Chevaleyre, J. Lang, N. Maudet, J. Monnot, L. Xia, New candidates welcome! Possible winners with respect to the addition of new candidates, Math. Soc. Sci. 64 (1) (2012) 74–88.
[20] R. Christian, M. Fellows, F. Rosamond, A. Slinko, On complexity of lobbying in multiple referenda, Rev. Econ. Des. 11 (3) (2007) 217–224.
[21] V. Conitzer, T. Sandholm, J. Lang, When are elections with few candidates hard to manipulate?, J. ACM 54 (3) (2007) 14.
[22] V. Conitzer, T. Walsh, Barriers to manipulation in voting, in: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016, pp. 127–145, chapter 6.
[23] R. de Haan, Parameterized complexity results for the Kemeny rule in judgment aggregation, in: Proceedings of the 22nd European Conference on Artificial Intelligence, 2016, pp. 1502–1510.
[24] R. de Haan, Complexity results for manipulation, bribery and control of the Kemeny judgment aggregation procedure, in: Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2017, pp. 1151–1159.
[25] F. Dietrich, Judgment aggregation and agenda manipulation, Manuscript, 2014.
[26] F. Dietrich, C. List, Judgment aggregation by quota rules: majority voting generalized, J. Theor. Polit. 19 (4) (2007) 391–424.
[27] F. Dietrich, C. List, Strategy-proof judgment aggregation, Econ. Philos. 23 (3) (2007) 269–300.
[28] E. Elkind, P. Faliszewski, A. Slinko, Cloning in elections: finding the possible winners, J. Artif. Intell. Res. 42 (2011) 529–573.
[29] U. Endriss, Judgment aggregation, in: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016, pp. 399–426, chapter 17.
[30] U. Endriss, U. Grandi, D. Porello, Complexity of judgment aggregation, J. Artif. Intell. Res. 45 (2012) 481–514.

[31] G. Erdélyi, M. Fellows, J. Rothe, L. Schend, Control complexity in Bucklin and fallback voting: a theoretical analysis, J. Comput. Syst. Sci. 81 (4) (2015) 632–660.
[32] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, How hard is bribery in elections?, J. Artif. Intell. Res. 35 (2009) 485–532.
[33] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, Using complexity to protect elections, Commun. ACM 53 (11) (2010) 74–82.
[34] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, Multimode control attacks on elections, J. Artif. Intell. Res. 40 (2011) 305–351.
[35] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, J. Rothe, Llull and Copeland voting computationally resist bribery and constructive control, J. Artif. Intell. Res. 35 (2009) 275–341.
[36] P. Faliszewski, A. Procaccia, AI's war on manipulation: are we winning?, AI Mag. 31 (4) (2010) 53–64.
[37] P. Faliszewski, J. Rothe, Control and bribery in voting, in: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016, pp. 146–168, chapter 7.
[38] International Centre for Settlement of Investment Disputes, ICSID rules of procedure for arbitration proceedings, http://icsidfiles.worldbank.org/icsid/icsid/StaticFiles/basicdoc/partF.htm, 2006.
[39] T. Gonzalez, Clustering to minimize the maximum intercluster distance, Theor. Comput. Sci. 38 (1985) 293–306.
[40] E. Hemaspaandra, L. Hemaspaandra, C. Menton, Search versus decision for election manipulation problems, in: Proceedings of the 30th Annual Symposium on Theoretical Aspects of Computer Science, in: LIPIcs, vol. 20, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, February 2013, pp. 377–388.
[41] E. Hemaspaandra, L. Hemaspaandra, J. Rothe, Anyone but him: the complexity of precluding an alternative, Artif. Intell. 171 (5–6) (2007) 255–285.
[42] E. Hemaspaandra, L. Hemaspaandra, J. Rothe, Hybrid elections broaden complexity-theoretic resistance to control, Math. Log. Q. 55 (4) (2009) 397–424.
[43] R. Jonakait, The American Jury System, Yale University Press, 2006.
[44] L. Kornhauser, L. Sager, Unpacking the court, Yale Law J. 96 (1) (1986) 82–117.
[45] C. List, The discursive dilemma and public reason, Ethics 116 (2) (2006) 362–402.
[46] C. List, The theory of judgment aggregation: an introductory review, Synthese 187 (1) (2012) 179–207.
[47] C. List, C. Puppe, Judgment aggregation, in: P. Anand, P. Pattanaik, C. Puppe (Eds.), The Handbook of Rational and Social Choice, Oxford University Press, 2009, pp. 457–482, chapter 19.
[48] A. Loreggia, N. Narodytska, F. Rossi, B. Venable, T. Walsh, Controlling elections by replacing candidates or votes (extended abstract), in: Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2015, pp. 1737–1738.
[49] M. Miller, D. Osherson, Methods for distance-based judgment aggregation, Soc. Choice Welf. 32 (4) (2009) 575–601.
[50] International Chamber of Commerce, ICC rules of arbitration, https://cdn.iccwbo.org/content/uploads/sites/3/2017/01/ICC-2017-Arbitration-and-2014-Mediation-Rules-english-version.pdf, 2017.
[51] United Nations Commission on International Trade Law, UNCITRAL arbitration rules, https://www.uncitral.org/pdf/english/texts/arbitration/arb-rules/arb-rules.pdf, 2013.
[52] C. Papadimitriou, Computational Complexity, second edition, Addison-Wesley, 1995.
[53] G. Pigozzi, Belief merging and the discursive dilemma: an argument-based account of paradoxes of judgment, Synthese 152 (2) (2006) 285–298.
[54] A. Rey, J. Rothe, Structural control in weighted voting games, in: Proceedings of the 41st International Symposium on Mathematical Foundations of Computer Science, in: LIPIcs, vol. 58, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, August 2016, pp. 80:1–80:15, An extended version is to appear in B.E. J. Theor. Econ.
[55] A. Rey, J. Rothe, Structural control in weighted voting games (extended abstract), in: Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2016, pp. 1371–1372.
[56] J. Rothe, Complexity Theory and Cryptology. An Introduction to Cryptocomplexity, EATCS Texts in Theoretical Computer Science, Springer-Verlag, 2005.
[57] J. Rothe (Ed.), Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division, Springer Texts in Business and Economics, Springer-Verlag, 2015.
[58] J. Rothe, L. Schend, Challenges to complexity shields that are supposed to protect elections against manipulation and control: a survey, Ann. Math. Artif. Intell. 68 (1–3) (2013) 161–193.
[59] N. Tideman, Independence of clones as a criterion for voting rules, Soc. Choice Welf. 4 (3) (1987) 185–206.

Chapter 7

Conclusions and Future Work

This thesis covers several topics in the field of computational social choice. The goal is to further the understanding of axiomatic and complexity theoretic properties of decision making procedures.

First, in Chapter 3, my coauthors and I studied a new type of ballot called ℓ-ballot that combines the concept of ordinal and cardinal preferences, and defined two types of committee election rules tailored to these type of ballots. Our proposed minisum and minimax rules were designed to minimize the dissatisfaction that voters have with the winning committee. We then modified several existing properties of single- and multiwinner voting rules to fit our type of ballots and rules and studied the axiomatic properties of the minisum and minimax rules. Further, we were able to show that although the winner determination is NP-hard for the minimax rules, an auxiliary problem asking whether there exists a committee with a voter's maximum dissatisfaction of at most d is fixed-parameter tractable when parameterized by d. As an outlook, we proposed a type of ballot called (a,b)-ballot that is based on cardinal preferences and allows more flexibility for voters to express their underlying preferences. In contrast to existing cardinal-based ballots, here, voters can express their dissatisfaction for a candidate being in a winning committee (a) as well as not being in a winning committee (b), without the restriction that those two values need to be related (e.g., always add up to a fixed constant). Future work includes a characterization of our rules in the context of ℓ-group rules. Furthermore, it would be interesting to define fairness criteria to evaluate the outcome's quality for the voters. In regard to (a,b)-ballots, experiments are needed to determine optimal bounds for the values and study whether the added expressiveness of the model leads to a higher satisfaction of the voters with the election outcome. It would also be interesting to define a cardinal-based variant of the Chamberlin-Courant rules as introduced in Definition 2.14, or to focus on representation in this context, for example by modifying the axiom *justified representation* (Definition 2.16) to allow for cardinal ballots.

My respective coauthors and I were also able to identify several barriers to strategic behavior in voting and judgment aggregation. In Chapter 4, we closed a gap by showing that shift bribery is also hard for several iterative scoring rules, i.e., scoring rules that proceed in rounds where in each round, candidates are eliminated. By allowing the campaign manager to exploit the nonmonotonicity of most of our considered rules, we further showed by using Hare and plurality with runoff as an example that this hardness does not result from restricting the briber to shift the designated candidate forwards in the constructive case (respectively, backwards in the destructive case). Based on our results, Zhou and Guo (2020) started the study of parameterized complexity for the iterative scoring rules considered in this thesis for the parameters number of voters, number of candidates, and budget. For future work, we propose to extend the study of parameterized complexity and to investigate the effect of exploiting nonmonotonicity in-depth. We conjecture that the complexity of shift bribery for all nonmonotonic rules considered by us remains unchanged, but it would be interesting to identify a rule for which shift bribery becomes tractable in these circumstances. Furthermore, domain restrictions as defined on page 13 might also lead to a complexity shift for iterative scoring rules.

For iterative voting, i.e., voting where voters are allowed to update their ballots repeatedly, we studied the manipulative power of the polling agency that announces a dishonest opinion poll to reach a desired outcome of the election. Chapter 5 extended the research on manipulation by the polling agency by introducing a best-response model for the voting rule veto, by studying destructive manipulation, by conducting experiments on efficient heuristics, and most importantly by introducing distance-based problem variants and providing parameterized tractability and intractability results. In particular, we showed that manipulation is para-NP-hard for all considered problems even for very restricted underlying social networks. However, we were able to show that all considered problem variants for veto are tractable when the social network contains no edges, which can be seen as a case where voters are not influenced by their neighbors. Here, future work includes completing the complexity results for plurality in the case of a social network without edges and further the study of parameterized complexity for more natural parameters, for example parameters that describe the underlying social network. We also propose to define best-response dynamics for voting rules that require more complex ballots than plurality and veto and experimentally study how these dynamics affect the quality of the outcome and the possibilities to manipulate by the polling agency. Currently, the polling agency has complete information over the voters' preferences, the voters have no memory and are myopic, they only deviate in the rare cases that they are pivotal, voters trust the polling agency, and they communicate truthfully to their neighbors which candidate they currently vote for. It would therefore be interesting to incorporate changes regarding these aspects into the model.

Finally, in Chapter 6, my coauthors and I introduced control in judgment aggregation. We defined the concepts of control by adding, deleting, replacing, and bundling judges, and proved that these types of control are intractable for the uniform (constant) premise-based quota rules and for several types of the chair's preferences. Our results hold for each rational quota in the case of the uniform constant premise-based quota rules, and for the quota $q = 1/2$ in the case of the uniform premise-based quota rules. Future work includes completing the classical complexity results for all quotas, studying parameterized complexity in our context, and considering new rules. The question whether restricted domains such as the domain of unidimensionally aligned profiles, a variant of the single-crossing domain in preference aggregation, have an impact of the complexity of control in judgment aggregation is also still an open problem. Further, it would be interesting to define new types of control in judgment aggregation, especially types that influence the agenda or the aggregation rule itself instead of the set of judges.

Bibliography

Ågotnes, T., van der Hoek, W., and Wooldridge, M. (2011). On the logic of preference and judgment aggregation. *Journal of Autonomous Agents and Multi-Agent Systems*, 22(1):4–30.

Alcantud, J. C. R. and Laruelle, A. (2014). Dis&approval voting: A characterization. *Social Choice and Welfare*, 43(1):1–10.

Alon, N., Falik, D., Meir, R., and Tennenholtz, M. (2013). Bundling attacks in judgment aggregation. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, pages 39–45. AAAI Press.

Arora, S. and Barak, B. (2009). *Computational Complexity: A Modern Approach*. Cambridge University Press.

Arrow, K. (1951 (revised edition 1963)). *Social Choice and Individual Values*. John Wiley and Sons.

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346.

Aziz, H., Brill, M., Conitzer, V., Elkind, E., Freeman, R., and Walsh, T. (2017a). Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485.

Aziz, H., Elkind, E., Faliszewski, P., Lackner, M., and Skowron, P. (2017b). The Condorcet principle for multiwinner elections: From shortlisting to proportionality. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 84–90. IJCAI.

Aziz, H., Elkind, E., Huang, S., Lackner, M., Sánchez-Fernández, L., and Skowron, P. (2018). On the complexity of extended and proportional justified representation. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 902–909. AAAI Press.

Aziz, H. and Lee, B. E. (2020). The expanding approvals rule: Improving proportional representation and monotonicity. *Social Choice and Welfare*, 54(1):1–45.

Baldwin, J. (1926). The technique of the Nanson preferential majority system of election. *Transactions and Proceedings of the Royal Society of Victoria*, 39:42–52.

Balinski, M. and Laraki, R. (2011). *Majority judgment*. Cambridge/Mass.

Ballester, M. and Rey-Biel, P. (2006). Sincere voting with cardinal preferences: Approval voting. Manuscript.

Barberà, S., Bossert, W., and Pattanaik, P. K. (2004). Ranking sets of objects. In *Handbook of Utility Theory*, pages 893–977. Springer.

Barberà, S. and Coelho, D. (2008). How to choose a non-controversial list with k names. *Social Choice and Welfare*, 31(1):79–96.

Bartholdi, III, J. and Trick, M. A. (1986). Stable matching with preferences derived from a psychological model. *Operations Research Letters*, 5(4):165–169.

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

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

Bartholdi III, J., Tovey, C., and Trick, M. (1992). How hard is it to control an election? *Mathematical and Computer Modelling*, 16(8/9):27–40.

Bartholdi III, J., Tovey, C. A., and Trick, M. A. (1989b). Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165.

Baumeister, D., Boes, L., and Weishaupt, R. (2021a). Complexity of sequential rules in judgment aggregation. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, pages 187–195. IFAAMAS.

Baumeister, D., Böhnlein, T., Rey, L., Schaudt, O., and Selker, A.-K. (2016). Minisum and minimax committee election rules for general preference types. In *Proceedings of the 22nd European Conference on Artificial Intelligence*, pages 1656–1657. IOS Press. Extended Abstract.

Baumeister, D. and Dennisen, S. (2015). Voter dissatisfaction in committee elections. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS. Extended abstract.

Baumeister, D., Dennisen, S., and Rey, L. (2015a). Winner determination and manipulation in minisum and minimax committee elections. In *Proceedings of the 4th International Conference on Algorithm Decision Theory*, pages 469–485. Springer-Verlag *Lecture Notes in Computer Science #9346*.

Baumeister, D., Erdélyi, G., Erdélyi, O., and Rothe, J. (2012). Control in judgment aggregation. In *Proceedings of the 6th European Starting AI Researcher Symposium*, pages 23–34. IOS Press.

Baumeister, D., Erdélyi, G., Erdélyi, O., and Rothe, J. (2013). Computational aspects of manipulation and control in judgment aggregation. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory*, pages 71–85. Springer-Verlag *Lecture Notes in Artificial Intelligence #8176*.

Baumeister, D., Erdélyi, G., Erdélyi, O. J., and Rothe, J. (2015b). Complexity of manipulation and bribery in judgment aggregation for uniform premise-based quota rules. *Mathematical Social Sciences*, 76:19–30.

Baumeister, D., Erdélyi, G., Erdélyi, O. J., Rothe, J., and Selker, A.-K. (2020a). Complexity of control in judgment aggregation for uniform premise-based quota rules. *Journal of Computer and System Sciences*, 112:13–33.

Baumeister, D., Erdélyi, G., and Rothe, J. (2015c). Judgment aggregation. In Rothe, J., editor, *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, chapter 6, pages 361–391. Springer-Verlag.

Baumeister, D., Hogrebe, T., and Rey, L. (2019). Generalized distance bribery. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pages 1764–1771. AAAI Press.

Baumeister, D. and Rothe, J. (2015). Preference aggregation by voting. In Rothe, J., editor, *Economics and Computation. An Introduction to Algorithmic Game Theory,* *Computational Social Choice, and Fair Division*, chapter 4, pages 197–325. Springer-Verlag.

Baumeister, D., Rothe, J., and Selker, A.-K. (2015d). Complexity of bribery and control for uniform premise-based quota rules under various preference types. In *Proceedings of the 4th International Conference on Algorithmic Decision Theory*, pages 432–448. Springer-Verlag *Lecture Notes in Artificial Intelligence #9346*.

Baumeister, D., Rothe, J., and Selker, A.-K. (2017). Strategic behavior in judgment aggregation. In Endriss, U., editor, *Trends in Computational Social Choice*, chapter 8, pages 145–168. AI Access.

Baumeister, D., Selker, A.-K., and Wilczynski, A. (2020b). Manipulation of opinion polls to influence iterative elections. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, pages 132–140. IFAAMAS.

Baumeister, D., Selker, A.-K., and Wilczynski, A. (2021b). Manipulation of opinion polls to influence iterative elections. In *Proceedings of the 8th International Workshop on Computational Social Choice*.

Betzler, N., Slinko, A., and Uhlmann, J. (2013). On the computation of fully proportional representation. *Journal of Artificial Intelligence Research*, 47:475–519.

Black, D. (1958). *The Theory of Committees and Elections*. Cambridge University Press.

Botan, S., de Haan, R., Slavkovik, M., and Terzopoulou, Z. (2021). Egalitarian judgment aggregation. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, pages 214–222. IFAAMAS.

Botan, S., Novaro, A., and Endriss, U. (2016). Group manipulation in judgment aggregation. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pages 411–419. IFAAMAS.

Brams, S., Kilgour, D., and Sanver, R. (2007). A minimax procedure for electing committees. *Public Choice*, 132:401–420.

Brandt, F., Brill, M., Hemaspaandra, E., and Hemaspaandra, L. A. (2015). Bypassing combinatorial protections: Polynomial-time algorithms for single-peaked electorates. *Journal of Autonomous Agents and Multi-Agent Systems*, 53:439–496.

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

Brânzei, S., Caragiannis, I., Morgenstern, J., and Procaccia, A. D. (2013). How bad is selfish voting? In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, pages 138–144. AAAI Press.

Bredereck, R., Chen, J., Faliszewski, P., Guo, J., Niedermeier, R., and Woeginger, G. J. (2014a). Parameterized algorithmics for computational social choice: Nine research challenges. *Tsinghua Science and Technology*, 19(4):358–373.

Bredereck, R., Chen, J., Faliszewski, P., Nichterlein, A., and Niedermeier, R. (2016a). Prices matter for the parameterized complexity of shift bribery. *Information and Computation*, 251:140–164.

Bredereck, R., Chen, J., Hartung, S., Kratsch, S., Niedermeier, R., Suchy, O., and Woeginger, G. J. (2014b). A multivariate complexity analysis of lobbying in multiple referenda. *Journal of Artificial Intelligence Research*, 50:409–446.

Bredereck, R., Chen, J., Niedermeier, R., and Walsh, T. (2017). Parliamentary voting procedures: Agenda control, manipulation, and uncertainty. *Journal of Artificial Intelligence Research*, 59:133–173.

Bredereck, R., Chen, J., and Woeginger, G. J. (2013). A characterization of the single-crossing domain. *Social Choice and Welfare*, 41(4):989–998.

Bredereck, R., Faliszewski, P., Igarashi, A., Lackner, M., and Skowron, P. (2018). Multi-winner elections with diversity constraints. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 933–940. AAAI Press.

Bredereck, R., Faliszewski, P., Kaczmarczyk, A., Niedermeier, R., Skowron, P., and Talmon, N. (2021). Robustness among multiwinner voting rules. *Artificial Intelligence*, 290:103403.

Bredereck, R., Faliszewski, P., Niedermeier, R., Skowron, P., and Talmon, N. (2015). Elections with few candidates: Prices, weights, and covering problems. In *Proceedings of the 4th International Conference on Algorithmic Decision Theory*, pages 414–431. Springer-Verlag *Lecture Notes in Artificial Intelligence #9346*.

Bredereck, R., Faliszewski, P., Niedermeier, R., and Talmon, N. (2016b). Complexity of shift bribery in committee elections. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 2452–2458. AAAI Press.

Bredereck, R., Faliszewski, P., Niedermeier, R., and Talmon, N. (2016c). Large-scale election campaigns: Combinatorial shift bribery. *Journal of Artificial Intelligence Research*, 55:603–652.

Brill, M., Faliszewski, P., Sommer, F., and Talmon, N. (2019). Approximation algorithms for balancedCC multiwinner rules. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pages 494–502. IFAAMAS.

Chamberlin, J. R. and Courant, P. N. (1983). Representative deliberations and representative decisions: Proportional representation and the Borda rule. *American Political Science Review*, 77(3):718–733.

Chen, J., Faliszewski, P., Niedermeier, R., and Talmon, N. (2017). Elections with few voters: Candidate control can be easy. *Journal of Artificial Intelligence Research*, 60:937–1002.

Christian, R., Fellows, M., Rosamond, F., and Slinko, A. (2007). On complexity of lobbying in multiple referenda. *Review of Economic Design*, 11(3):217–224.

Condorcet, M. d. (1785). *Essay sur l'application de l'analyse de la probabilité des decisions: Redues et pluralité des voix*. l'Imprimerie Royale. Translated by McLean and Urken (1995).

Conitzer, V. and Sandholm, T. (2003). Universal voting protocol tweaks to make manipulation hard. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, pages 781–788.

Conitzer, V., Sandholm, T., and Lang, J. (2007). When are elections with few candidates hard to manipulate? *Journal of the ACM*, 54(3):Article 14.

Conitzer, V. and Walsh, T. (2016). Barriers to manipulation in voting. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., 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 AAAI Conference on Artificial Intelligence*, pages 638–643. AAAI Press.

Cook, S. A. (1971). The complexity of theorem-proving procedures. In *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing*, pages 151–158. ACM Press.

Copeland, A. H. (1951). A reasonable social welfare function. Technical report, Mimeo, University of Michigan.

Cygan, M., Fomin, F. V., Kowalik, Ł., Lokshtanov, D., Marx, D., Pilipczuk, M., Pilipczuk, M., and Saurabh, S. (2015). *Parameterized algorithms*. Springer.

Cygan, M., Kowalik, Ł., Socała, A., and Sornat, K. (2018). Approximation and parameterized complexity of minimax approval voting. *Journal of Artificial Intelligence Research*, 63:495–513.

de Haan, R. (2016). Parameterized complexity results for the Kemeny rule in judgment aggregation. In *Proceedings of the 22nd European Conference on Artificial Intelligence*, pages 1502–1510. IOS Press.

de Haan, R. (2017). Complexity results for manipulation, bribery and control of the Kemeny judgment aggregation procedure. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems*, pages 1151–1159. IFAAMAS.

de Haan, R. (2018). Hunting for tractable languages for judgment aggregation. In *Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning*, pages 194–203. AAAI Press.

de Haan, R. and Slavkovik, M. (2017). Complexity results for aggregating judgments using scoring or distance-based procedures. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems*, pages 952–961. IFAAMAS.

de Haan, R. and Slavkovik, M. (2019). Answer set programming for judgment aggregation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 1668–1674. IJCAI.

Debord, B. (1992). An axiomatic characterization of Borda's k-choice function. *Social Choice and Welfare*, 9(4):337–343.

Dey, P. (2021). Local distance constrained bribery in voting. *Theoretical Computer Science*, 849:1–21.

Dey, P., Misra, N., and Narahari, Y. (2017). Frugal bribery in voting. *Theoretical Computer Science*, 676:15–32.

Dey, P., Misra, N., and Narahari, Y. (2018). Complexity of manipulation with partial information in voting. *Theoretical Computer Science*, 726:78–99.

Dey, P. and Narahari, Y. (2015). Estimating the margin of victory of an election using sampling. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 1120–1126. AAAI Press/IJCAI.

Dietrich, F. (2007). A generalised model of judgment aggregation. *Social Choice and Welfare*, 28(4):529–565.

Dietrich, F. (2014). Scoring rules for judgment aggregation. *Social Choice and Welfare*, 42(4):873–911.

Dietrich, F. (2016). Judgment aggregation and agenda manipulation. *Games and Economic Behavior*, 95:113–136.

Dietrich, F. and List, C. (2007a). Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, 29(1):19–33.

Dietrich, F. and List, C. (2007b). Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics*, 19(4):391–424.

Dietrich, F. and List, C. (2007c). Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(3):269–300.

Dietrich, F. and List, C. (2010). The impossibility of unbiased judgment aggregation. *Theory and Decision*, 68(3):281–299.

Dietrich, F. and Mongin, P. (2010). The premiss-based approach to judgment aggregation. *Journal of Economic Theory*, 145(2):562–582.

Dodgson, C. (1876). A method of taking votes on more than two issues. *The Theory of Committees and Elections*. Reprinted by McLean and Urken (1995).

Dokow, E. and Holzman, R. (2010). Aggregation of binary evaluations. *Journal of Economic Theory*, 145(2):495–511.

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

Downey, R. G. and Fellows, M. R. (1995a). Fixed-parameter tractability and completeness I: Basic results. *SIAM Journal on Computing*, 24(4):873–921.

Downey, R. G. and Fellows, M. R. (1995b). Fixed-parameter tractability and completeness II: On completeness for W[1]. *Theoretical Computer Science*, 141(1-2):109–131.

Dummett, M. (1984). *Voting procedures*. Oxford University Press.

Elkind, E. and Faliszewski, P. (2010). Approximation algorithms for campaign management. In *International Workshop on Internet and Network Economics*, pages 473–482. Springer.

Elkind, E., Faliszewski, P., Gupta, S., and Roy, S. (2020a). Algorithms for swap and shift bribery in structured elections. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, pages 366–374. IFAAMAS.

Elkind, E., Faliszewski, P., and Skowron, P. (2020b). A characterization of the single-peaked single-crossing domain. *Social Choice and Welfare*, 54(1):167–181.

Elkind, E., Faliszewski, P., Skowron, P., and Slinko, A. (2017a). Properties of multiwinner voting rules. *Social Choice and Welfare*, 48(3):599–632.

Elkind, E., Faliszewski, P., and Slinko, A. (2009). Swap bribery. In *Proceedings of the 2nd International Symposium on Algorithmic Game Theory*, pages 299–310. Springer-Verlag *Lecture Notes in Computer Science #5814*.

Elkind, E., Faliszewski, P., and Slinko, A. (2011). Cloning in elections: Finding the possible winners. *Journal of Artificial Intelligence Research*, 42:529–573.

Elkind, E., Faliszewski, P., and Slinko, A. (2012). Clone structures in voters' preferences. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 496–513.

Elkind, E. and Lackner, M. (2015). Structure in dichotomous preferences. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 2019–2025. AAAI Press/IJCAI.

Elkind, E., Lackner, M., and Peters, D. (2017b). Structured preferences. In Endriss, U., editor, *Trends in Computational Social Choice*, chapter 10, pages 187–207. AI Access.

Elkind, E., Lang, J., and Saffidine, A. (2015). Condorcet winning sets. *Social Choice and Welfare*, 44(3):493–517.

Endriss, U. (2016). Judgment aggregation. In *Handbook of Computational Social Choice*, chapter 17, pages 399–426. Cambridge University Press.

Endriss, U., editor (2017). *Trends in Computational Social Choice*. AI Access.

Endriss, U. and de Haan, R. (2015). Complexity of the winner determination problem in judgment aggregation: Kemeny, Slater, Tideman, Young. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pages 117–125. IFAAMAS.

Endriss, U., de Haan, R., Lang, J., and Slavkovik, M. (2020). The complexity landscape of outcome determination in judgment aggregation. *Journal of Artificial Intelligence Research*, 69:687–731.

Endriss, U., de Haan, R., and Szeider, S. (2015). Parameterized complexity results for agenda safety in judgment aggregation. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pages 127–136. IFAAMAS.

Endriss, U., Grandi, U., de Haan, R., and Lang, J. (2016a). Succinctness of languages for judgment aggregation. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning*, pages 176–185. AAAI Press.

Endriss, U., Grandi, U., and Porello, D. (2012). Complexity of judgment aggregation. *Journal of Artificial Intelligence Research*, 45:481–514.

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

Fairstein, R., Lauz, A., Meir, R., and Gal, K. (2019). Modeling people's voting behavior with poll information. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pages 1422–1430. IFAAMAS.

Faliszewski, P., Hemaspaandra, E., and Hemaspaandra, L. (2009a). How hard is bribery in elections? *Journal of Artificial Intelligence Research*, 35:485–532.

Faliszewski, P., Hemaspaandra, E., Hemaspaandra, L. A., and Rothe, J. (2009b). Llull and Copeland voting computationally resist bribery and constructive control. *Journal of Artificial Intelligence Research*, 35:275–341.

Faliszewski, P., Hemaspaandra, E., Hemaspaandra, L. A., and Rothe, J. (2011). The shield that never was: Societies with single-peaked preferences are more open to manipulation and control. *Information and Computation*, 209(2):89–107.

Faliszewski, P., Manurangsi, P., and Sornat, K. (2021). Approximation and hardness of shift-bribery. *Artificial Intelligence*, 298:103520.

Faliszewski, P. and Procaccia, A. D. (2010). AI's war on manipulation: Are we winning? *AI Magazine*, 31(4):53–64.

Faliszewski, P., Rothe, I., and Rothe, J. (2016). Noncooperative game theory. In *Economics and Computation*, pages 41–134. Springer.

Faliszewski, P. and Rothe, J. (2016). Control and bribery in voting. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A., editors, *Handbook of Computational Social Choice*, chapter 7, pages 146–168. Cambridge University Press.

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2017a). Multiwinner rules on paths from k–Borda to Chamberlin–Courant. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 192–198. IJCAI.

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2017b). Multiwinner voting: A new challenge for social choice theory. In Endriss, U., editor, *Trends in Computational Social Choice*, chapter 2, pages 27–47. AI Access.

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2019). Committee scoring rules: Axiomatic characterization and hierarchy. *ACM Transactions on Economics and Computation*, 7(1):1–39.

Faliszewski, P., Skowron, P., and Talmon, N. (2017c). Bribery as a measure of candidate success: Complexity results for approval-based multiwinner rules. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems*, pages 6–14. IFAAMAS.

Faliszewski, P., Slinko, A., and Talmon, N. (2020). Multiwinner rules with variable number of winners. In *Proceedings of the 24th European Conference on Artificial Intelligence*, pages 67–74. IOS Press.

Felsenthal, D. and Maoz, Z. (1992). Normative properties of four single-stage multi-winner electoral procedures. *Behavioral Science*, 37(2):109–127.

Fishburn, P. C. (1974). Paradoxes of voting. *The American Political Science Review*, 68(2):537–546.

Flum, J. and Grohe, M. (2003). Describing parameterized complexity classes. *Information and Computation*, 187(2):291–319.

Flum, J. and Grohe, M. (2006). *Parameterized Complexity Theory*. EATCS Texts in Theoretical Computer Science. Springer.

Garey, M. R. and Johnson, D. S. (1979). *Computers and intractability*, volume 174. freeman San Francisco.

Gehrlein, W. (1985). The Condorcet criterion and committee selection. *Mathematical Social Sciences*, 10(3):199–209.

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

Gourvès, L., Lesca, J., and Wilczynski, A. (2016). Strategic voting in a social context: Considerate equilibria. In *Proceedings of the 22nd European Conference on Artificial Intelligence*, pages 1423–1431. IOS Press.

Grandi, U., Loreggia, A., Rossi, F., Venable, K. B., and Walsh, T. (2013). Restricted manipulation in iterative voting: Condorcet efficiency and Borda score. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory*, pages 181–192. Springer-Verlag *Lecture Notes in Artificial Intelligence #8176*.

Gupta, S., Jain, P., Roy, S., Saurabh, S., and Zehavi, M. (2019). Gehrlein stability in committee selection: Parameterized hardness and algorithms. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pages 511–519. IFAAMAS.

Hartmanis, J. and Stearns, R. E. (1965). On the computational complexity of algorithms. *Transactions of the American Mathematical Society*, 117:285–306.

Hemaspaandra, E., Hemaspaandra, L. A., and Rothe, J. (1997). Exact analysis of Dodgson elections: Lewis Carroll's 1876 voting system is complete for parallel access to NP. *Journal of the ACM*, 44(6):806–825.

Hemaspaandra, E., Hemaspaandra, L. A., and Rothe, J. (2007). Anyone but him: The complexity of precluding an alternative. *Artificial Intelligence*, 171(5-6):255–285.

Hemaspaandra, E., Spakowski, H., and Vogel, J. (2005). The complexity of Kemeny elections. *Theoretical Computer Science*, 349(3):382–391.

Izsak, R., Talmon, N., and Woeginger, G. (2018). Committee selection with intraclass and interclass synergies. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 1071–1078. AAAI Press.

Johnson, D. S. (1974). Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9(3):256–278.

Kaczmarczyk, A. and Faliszewski, P. (2019). Algorithms for destructive shift bribery. *Journal of Autonomous Agents and Multi-Agent Systems*, 33(3):275–297.

Kagita, V. R., Pujari, A. K., Padmanabhan, V., Aziz, H., and Kumar, V. (2021). Committee selection using attribute approvals. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, pages 683–691. IFAAMAS.

Karp, R. M. (1972). Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer.

Kemeny, J. (1959). Mathematics without numbers. *Daedalus*, 88(4):577–591.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2):81–93.

Kilgour, D., Brams, S., and Sanver, R. (2006). How to elect a representative committee using approval balloting. In Pukelsheim, F. and Simeone, B., editors, *Mathematics and Democracy: Recent Advances in Voting Systems and Collective Choice*, pages 83–95. Springer.

Kilgour, D. M. (2016). Approval elections with a variable number of winners. *Theory and Decision*, 81(2):199–211.

Knop, D., Koutecký, M., and Mnich, M. (2020). Voting and bribing in single-exponential time. *ACM Transactions on Economics and Computation*, 8(3):1–28.

Kocot, M., Kolonko, A., Elkind, E., Faliszewski, P., and Talmon, N. (2019). Multigoal committee selection. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 385–391. IJCAI.

Kornhauser, L. and Sager, L. (1986). Unpacking the court. *Yale Law Journal*, 96(1):82–117.

Lackner, M. and Maly, J. (2021). Approval-based shortlisting. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, pages 737–745. IFAAMAS.

Lackner, M. and Skowron, P. (2018). Approval-based multi-winner rules and strategic voting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 340–346. IJCAI.

Lackner, M. and Skowron, P. (2019). A quantitative analysis of multi-winner rules. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 407–413. IJCAI.

Lackner, M. and Skowron, P. (2021). Consistent approval-based multi-winner rules. *Journal of Economic Theory*, 192:105173.

Lang, J., Pigozzi, G., Slavkovik, M., and Torre, J. (2011). Judgment aggregation rules based on minimization. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 238–246. ACM Press.

Lang, J., Pigozzi, G., Slavkovik, M., van der Torre, L., and Vesic, S. (2017). A partial taxonomy of judgment aggregation rules and their properties. *Social Choice and Welfare*, 48(2):327–356.

Lang, J. and Slavkovik, M. (2013). Judgment aggregation rules and voting rules. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory*, pages 230–243. Springer-Verlag *Lecture Notes in Artificial Intelligence #8176*.

Lang, J. and Slavkovik, M. (2014). How hard is it to compute majority-preserving judgment aggregation rules? In *Proceedings of the 21st European Conference on Artificial Intelligence*, pages 501–506. IOS Press.

Laslier, J.-F. and Sanver, M. R. (2010). *Handbook on Approval Voting*. Springer Science & Business Media.

LeGrand, R. (2004). Analysis of the minimax procedure. Technical Report WUCSE-2004-67, Department of Computer Science and Engineering, Washington University, St. Louis, Missouri.

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

Levin, J. and Nalebuff, B. (1995). An introduction to vote-counting schemes. *The Journal of Economic Perspectives*, 9(1):3–26.

List, C. (2003). A possibility theorem on aggregation over multiple interconnected propositions. *Mathematical Social Sciences*, 45(1):1–13.

List, C. (2004). A model of path-dependence in decisions over multiple propositions. *American Political Science Review*, 93(3):495–513.

List, C. and Pettit, P. (2002). Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(1):89–110.

List, C. and Pettit, P. (2004). Aggregating sets of judgments: Two impossibility results compared. *Synthese*, 140(1–2):207–235.

List, C. and Puppe, C. (2009). Judgment aggregation. In Anand, P., Pattanaik, P., and Puppe, C., editors, *The Handbook of Rational and Social Choice*, chapter 19, pages 457–482. Oxford University Press.

Liu, H. and Guo, J. (2016). Parameterized complexity of winner determination in minimax committee elections. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pages 341–349. IFAAMAS.

Magiera, K. and Faliszewski, P. (2017). How hard is control in single-crossing elections? *Journal of Autonomous Agents and Multi-Agent Systems*, 31(3):606–627.

Maushagen, C., Neveling, M., Rothe, J., and Selker, A.-K. (2018a). Complexity of shift bribery in hare, coombs, baldwin, and nanson elections. In *Proceedings of the 15th International Symposium on Artificial Intelligence and Mathematics*.

Maushagen, C., Neveling, M., Rothe, J., and Selker, A.-K. (2018b). Complexity of shift bribery in iterative elections. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, pages 1567–1575. IFAAMAS.

Maushagen, C., Neveling, M., Rothe, J., and Selker, A.-K. (2021). Complexity of shift bribery for iterative voting rules. Submitted to *Journal of Artificial Intelligence Research*.

May, K. (1952). A set of independent, necessary and sufficient conditions for simple majority décisions. *Econometrica*, 20:680–684.

McLean, I. and Urken, A. B. (1995). *Classics of Social Choice*. University of Michigan Press.

Meir, R. (2017). Iterative voting. In Endriss, U., editor, *Trends in Computational Social Choice*, chapter 4, pages 69–86. AI Access.

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

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

Meir, R., Procaccia, A., Rosenschein, J., and Zohar, A. (2008). Complexity of strategic behaviour in multi-winner elections. *Journal of Artificial Intelligence Research*, 33:149–178.

Miasko, T. and Faliszewski, P. (2016). The complexity of priced control in elections. *Annals of Mathematics and Artificial Intelligence*, 77(3-4):225–250.

Miller, M. and Osherson, D. (2009). Methods for distance-based judgment aggregation. *Social Choice and Welfare*, 32(4):575–601.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208.

Misra, N., Nabeel, A., and Singh, H. (2015). On the parameterized complexity of minimax approval voting. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pages 97–105. IFAAMAS.

Mongin, P. (2012). The doctrinal paradox, the discursive dilemma, and logical aggregation theory. *Theory and Decision*, 73(3):315–355.

Monroe, B. L. (1995). Fully proportional representation. *American Political Science Review*, 89(4):925–940.

Nanson, E. (1882). Methods of election. *Transactions and Proceedings of the Royal Society of Victoria*, 19:197–240.

Obraztsova, S. and Elkind, E. (2012). Optimal manipulation of voting rules. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, pages 2141–2147. AAAI Press.

Obraztsova, S., Elkind, E., Polukarov, M., and Rabinovich, Z. (2015a). Doodle poll games. In *Proceedings of the First IJCAI-Workshop on Algorithmic Game Theory*.

Obraztsova, S., Lev, O., Polukarov, M., Rabinovich, Z., and Rosenschein, J. S. (2016). Nonmyopic voting dynamics: An optimistic approach. In *Proceedings of the 10th Multidisciplinary Workshop on Advances in Preference Handling*.

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

Obraztsova, S., Polukarov, M., Rabinovich, Z., and Elkind, E. (2017). Doodle poll games. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems*, pages 876–884. IFAAMAS.

Obraztsova, S., Zick, Y., and Elkind, E. (2013). On manipulation in multi-winner elections based on scoring rules. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, pages 359–366. IFAAMAS.

Peters, D. (2018a). Proportionality and strategyproofness in multiwinner elections. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, pages 1549–1557. IFAAMAS.

Peters, D. (2018b). Single-peakedness and total unimodularity: New polynomial-time algorithms for multi-winner elections. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 1169–1176. AAAI Press.

Pettit, P. (2001). Deliberative democracy and the discursive dilemma. *Philosophical Issues*, 11(1):268–299.

Pigozzi, G. (2006). Belief merging and the discursive dilemma: An argument-based account of paradoxes of judgment. *Synthese*, 152(2):285–298.

Procaccia, A. and Rosenschein, J. (2006). The distortion of cardinal preferences in voting. In *Proceedings of the 10th International Workshop on Cooperative Information Agents*, pages 317–331. Springer-Verlag *Lecture Notes in Computer Science #4149*.

Procaccia, A., Rosenschein, J., and Zohar, A. (2007). Multi-winner elections: Complexity of manipulation, control, and winner-determination. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, pages 1476–1481.

Procaccia, A. D. and Rosenschein, J. S. (2007). Average-case tractability of manipulation in voting via the fraction of manipulators. In *Proceedings of the 6th International Conference on Autonomous Agents and Multiagent Systems*, pages 718–720. IFAAMAS.

Ratliff, T. C. (2003). Some startling inconsistencies when electing committees. *Social Choice and Welfare*, 21(3):433–454.

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*, pages 635–644, Valencia, Spain. IFAAMAS.

Reisch, Y., Rothe, J., and Schend, L. (2014). The margin of victory in Schulze, cup, and Copeland elections: Complexity of the regular and exact variants. In *Proceedings of the 7th European Starting AI Researcher Symposium*, pages 250–259. IOS Press.

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

Rothe, J. and Schend, L. (2013). Challenges to complexity shields that are supposed to protect elections against manipulation and control: A survey. *Annals of Mathematics and Artificial Intelligence*, 68(1–3):161–193.

Sánchez-Fernández, L., Elkind, E., Lackner, M., Fernández, N., Fisteus, J. A., Val, P. B., and Skowron, P. (2017). Proportional justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 670–676. AAAI Press.

Sánchez-Fernández, L. and Fisteus, J. A. (2019). Monotonicity axioms in approval-based multi-winner voting rules. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pages 485–493. IFAAMAS.

Saporiti, A. and Tohmé, F. (2006). Single-crossing, strategic voting and the median choice rule. *Social Choice and Welfare*, 26(2):363–383.

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(2):187–217.

Schlotter, I., Faliszewski, P., and Elkind, E. (2017). Campaign management under approval-driven voting rules. *Algorithmica*, 77(1):84–115.

Shiryaev, D., Yu, L., and Elkind, E. (2013). On elections with robust winners. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, pages 415–422. IFAAMAS.

Sina, S., Hazon, N., Hassidim, A., and Kraus, S. (2015). Adapting the social network to affect elections. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pages 705–713. IFAAMAS.

Skowron, P. (2018). Proportionality degree of multiwinner rules. Technical Report arXiv: 1810.08799v2 [cs.GT], ACM Computing Research Repository (CoRR).

Skowron, P. and Faliszewski, P. (2017). Chamberlin–Courant rule with approval ballots: Approximating the MaxCover problem with bounded frequencies in FPT time. *Journal of Artificial Intelligence Research*, 60:687–716.

Skowron, P., Faliszewski, P., and Lang, J. (2016). Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence*, 241:191–216.

Skowron, P., Faliszewski, P., and Slinko, A. (2015a). Achieving fully proportional representation: Approximability results. *Artificial Intelligence*, 222:67–103.

Skowron, P., Faliszewski, P., and Slinko, A. (2019). Axiomatic characterization of committee scoring rules. *Journal of Economic Theory*, 180:244–273.

Skowron, P., Yu, L., Faliszewski, P., and Elkind, E. (2015b). The complexity of fully proportional representation for single-crossing electorates. *Theoretical Computer Science*, 569:43–57.

Slavkovik, M. (2014). Not all judgment aggregation should be neutral. In *Proceedings of the European Conference on Social Intelligence*, pages 198–211.

Smith, W. D. (2000). Range voting. `http://scorevoting.net/WarrenSmithPages/homepage/rangevote.pdf`.

Taylor, A. D. (2005). *Social Choice and the Mathematics of Manipulation*. Cambridge University Press.

Terzopoulou, Z. and Endriss, U. (2018). Modelling iterative judgment aggregation. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 1234–1241. AAAI Press.

Terzopoulou, Z. and Endriss, U. (2019a). Rethinking the neutrality axiom in judgment aggregation. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pages 2212–2214. IFAAMAS.

Terzopoulou, Z. and Endriss, U. (2019b). Strategyproof judgment aggregation under partial information. *Social Choice and Welfare*, 53(3):415–442.

Terzopoulou, Z. and Endriss, U. (2020). Neutrality and relative acceptability in judgment aggregation. *Social Choice and Welfare*, 55(1):25–49.

Terzopoulou, Z., Endriss, U., and de Haan, R. (2018). Aggregating incomplete judgments: Axiomatisations for scoring rules. In *Proceedings of the 7th International Workshop on Computational Social Choice*.

Tideman, T. (1987). Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206.

Tsang, A. and Larson, K. (2016). The echo chamber: Strategic voting and homophily in social networks. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pages 368–375. IFAAMAS.

Vazirani, V. V. (2013). *Approximation Algorithms*. Springer Science & Business Media.

Wilczynski, A. (2019). Poll-confident voters in iterative voting. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pages 2205–2212. AAAI Press.

Xia, L. (2012). Computing the margin of victory for various voting rules. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 982–999. ACM Press.

Yang, Y. (2019). Complexity of manipulating and controlling approval-based multiwinner voting. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 637–643. IJCAI.

Yang, Y. (2020). On the complexity of destructive bribery in approval-based multi-winner voting. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, pages 1584–1592. IFAAMAS.

Zhou, A. and Guo, J. (2020). Parameterized complexity of shift bribery in iterative elections. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, pages 1665–1673. IFAAMAS.

Zwicker, W. S. (2016). Introduction to the theory of voting. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A., editors, *Handbook of Computational Social Choice*, chapter 1, pages 23–56. Cambridge University Press.

Eidesstattliche Erklärung

entsprechend §5 der Promotionsordnung vom 15.06.2018

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Desweiteren erkläre ich, dass ich eine Dissertation in der vorliegenden oder in ähnlicher Form noch bei keiner anderen Institution eingereicht habe.

Teile dieser Dissertation wurden bereits in Form folgender Zeitschriftenartikel und Konferenzberichte veröffentlicht oder zur Begutachtung eingereicht und sind entsprechend gekenn- zeichnet: Baumeister et al. (2016), Maushagen et al. (2021), Baumeister et al. (2020b), Baumeister et al. (2020a)

Ort, Datum	Ann-Kathrin Selker