Bribery and Control in Judgment Aggregation[1]

Dorothea Baumeister, Gábor Erdélyi, Olivia J. Erdélyi, and Jörg Rothe

Abstract

In computational social choice, the complexity of changing the outcome of elections via manipulation, bribery, and various control actions, such as adding or deleting candidates or voters, has been studied intensely. Endriss et al. [13, 14] initiated the complexity-theoretic study of problems related to judgment aggregation. We extend their results on manipulation to a whole class of judgment aggregation procedures, and we obtain stronger results by considering not only the classical complexity (NP-hardness) but the parameterized complexity (W[2]-hardness) of these problems with respect to natural parameters. Furthermore, we introduce and study the closely related concepts of bribery and control in judgment aggregation. In particular, we study the complexity of changing the outcome of such procedures via control by adding, deleting, or replacing judges.

1 Introduction

Decision-making processes are often susceptible to various types of interference. In social choice theory and in computational social choice, ways of influencing the outcome of elections—such as manipulation, bribery, and control—have been studied intensely, with a particular focus on the complexity of the related problems (see, e.g., the early work of Bartholdi et al. [2, 1, 3] and the recent surveys and bookchapters by Faliszewski et al. [21, 18], Brandt et al. [5], and Baumeister et al. [4]). In particular, (coalitional) *manipulation* [2, 1, 7] refers to (a group of) strategic voters casting their votes insincerely to reach their desired outcome; in *bribery* [17, 20] an external agent seeks to reach her desired outcome by bribing (without exceeding a given budget) some voters to alter their votes; and in *control* [3, 23, 16] 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.

Decision-making mechanisms or systems that are susceptible to strategic behavior, be it from the agents involved as in manipulation or from external authorities or actors as in bribery and control, are obviously not desirable, as that undermines the trust we have in these systems. We therefore have a strong interest in accurately assessing how vulnerable a system for decision-making processes is to these internal or external influences. Unfortunately, in many concrete settings of social choice, "perfect" systems are impossible to exist. For example, the Gibbard–Satterthwaite theorem says that no reasonable voting system can be "strategyproof" [22, 29] (see also the generalization by Duggan and Schwartz [11]), many natural voting systems are not "immune" to most or even all of the standard types of control [3, 23, 16], and Dietrich and List [9] give an analogue of the Gibbard–Satterthwaite theorem in judgment aggregation. To avoid this obstacle, a common approach in computational social choice is to apply methods from theoretical computer science to show that undesirable strategic behavior is blocked, or at least hindered, by the corresponding task being a computationally intractable problem.

Here we focus on judgment aggregation, which is an important framework for collective decision-making. In a judgment aggregation process, we seek to find a collective judgment set from given individual judgment sets over a set of possibly logically interconnected propositions. For further information on judgment aggregation, we refer the reader to the surveys by List and Puppe [26]

[1]This paper extends the results that appeared in *Proceedings of the 2nd International Conference on Algorithmic Decision Theory*, pages 1–15. Springer-Verlag *LNCS #6992*, 2011. This work was supported in part by DFG grant RO 1202/15-1, NRF (Singapore) grant NRF-RF 2009-08, an SFF grant from HHU, and a DAAD PPP/PROCOPE grant.

and by List [25]. This paper follows up the study of manipulation in judgment aggregation initiated by Endriss et al. [14] and it is the first to study bribery and control in judgment aggregation.

In particular, Endriss et al. [13, 14], defined the winner determination problem and the manipulation problem in judgment aggregation and studied their complexity for two important judgment aggregation rules. We extend their complexity-theoretic investigation for manipulation and also introduce various bribery problems in judgment aggregation. Furthermore, we introduce and motivate three types of control in judgment aggregation (namely, control by adding, deleting, or replacing judges), and study their computational complexity. These problems are each closely related to the corresponding problems in voting, yet are specifically tailored to judgment aggregation scenarios.

2 Formal Framework

We follow and extend the judgment aggregation framework described by Endriss et al. [14].

Let *PS* be the set of all propositional variables and $\mathscr{L}_{PS}$ the set of propositional formulas built from *PS*, where the following connections can be used in their usual meaning: disjunction ($\vee$), conjunction ($\wedge$), implication ($\rightarrow$), equivalence ($\leftrightarrow$), and the boolean constants 1 and 0. To avoid double negations, let $\sim\alpha$ denote the complement of α, i.e., $\sim\alpha = \neg\alpha$ if α is not negated, and $\sim\alpha = \beta$ if $\alpha = \neg\beta$. The judges have to judge over all formulas in the *agenda* Φ, which is a finite, nonempty subset of $\mathscr{L}_{PS}$ without doubly negated formulas. The agenda is required to be closed under complementation, i.e., $\sim\alpha \in \Phi$ if $\alpha \in \Phi$. A *judgment set for an agenda* Φ is a subset $J \subseteq \Phi$. It is said to be an *individual judgment set* if it is the set of propositions in the agenda accepted by an individual judge. A *collective judgment set* is the set of propositions in the agenda accepted by all judges as the result of a judgment aggregation procedure. A judgment set J is *complete* if for all $\alpha \in \Phi$, $\alpha \in J$ or $\sim\alpha \in J$; it is *complement-free* if for no $\alpha \in \Phi$, α and $\sim\alpha$ are in J; and it is *consistent* if there is an assignment that makes all formulas in J true. If a judgment set is complete and consistent, it is obviously complement-free. By $\mathscr{J}(\Phi)$ we denote the set of all complete and consistent subsets of Φ.

The famous doctrinal paradox [24] in judgment aggregation shows that if the majority rule is used, the collective judgment set can be inconsistent even if all individual judgment sets are consistent. One way of circumventing the doctrinal paradox is to impose restrictions on the agenda. Endriss et al. [13] studied the question of whether one can guarantee for a specific agenda that the outcome is always complete and consistent. They established necessary and sufficient conditions on the agenda to satisfy these criteria, and they studied the complexity of deciding whether a given agenda satisfies these conditions. They also showed that deciding whether an agenda guarantees a complete and consistent outcome for the majority rule is an intractable problem.

Endriss et al. [14] studied the winner and manipulation problem for two specific judgment aggregation procedures that always guarantee consistent outcomes. In the premise-based procedure, this is achieved by applying the majority rule only to the premises of the agenda, and then to derive the outcome for the conclusions from the outcome of the premises. We will study the complexity of manipulation and control also for the more general class of premise-based quota rules as defined by Dietrich and List [8].

Definition 1 (Premise-based Quota Rule) *The agenda* Φ *is divided into two disjoint subsets* $\Phi = \Phi_p \uplus \Phi_c$*, where* Φ_p *is the set of premises and* Φ_c *is the set of conclusions. We assume both* Φ_p *and* Φ_c *to be closed under complementation. The premises* Φ_p *are again divided into two disjoint subsets,* $\Phi_p = \Phi_1 \uplus \Phi_2$*, such that either* $\varphi \in \Phi_1$ *and* $\sim\varphi \in \Phi_2$*, or* $\sim\varphi \in \Phi_1$ *and* $\varphi \in \Phi_2$*. For each literal* $\varphi \in \Phi_1$*, define a quota* $q_\varphi \in \mathbb{Q}$*,* $0 \leq q_\varphi < 1$*. The quota for the literals* $\varphi \in \Phi_2$ *is* $q'_\varphi = 1 - q_\varphi$*.*

A premise-based quota rule is then defined to be a function $PQR : \mathscr{J}(\Phi)^n \rightarrow 2^\Phi$ *such that, for*

$\Phi = \Phi_p \uplus \Phi_c$, *each profile* $\mathbf{J} = (J_1, \ldots, J_n)$ *is mapped to the judgment set*

$$PQR(\mathbf{J}) = \triangle_q \cup \{\varphi \in \Phi_c \mid \triangle_q \models \varphi\}, \text{ where}$$
$$\triangle_q = \{\varphi \in \Phi_1 \mid \|\{i \mid \varphi \in J_i\}\| > nq_\varphi\} \cup \{\varphi \in \Phi_2 \mid \|\{i \mid \varphi \in J_i\}\| > \lceil nq'_\varphi - 1 \rceil\}.$$

To guarantee complete and consistent outcomes for this procedure, it is enough to require that Φ is closed under propositional variables and that Φ_p consists of all literals. The number of affirmations needed to be in the collective judgment set is $\lfloor nq_\varphi + 1 \rfloor$ for literals $\varphi \in \Phi_1$ and $\lceil nq'_\varphi \rceil$ for literals $\varphi \in \Phi_2$. Note that $\lfloor nq_\varphi + 1 \rfloor + \lceil nq'_\varphi \rceil = n+1$ ensures that either $\varphi \in PQR(\mathbf{J})$ or $\sim\varphi \in PQR(\mathbf{J})$ for every $\varphi \in \Phi$. Note that the quota $q_\varphi = 1$ for a literal $\varphi \in \Phi_1$ is not allowed here, as $n+1$ affirmations were then needed for $\varphi \in \Phi_1$ to be in the collective judgment set, which is impossible. However, $q_\varphi = 0$ is allowed, as in that case $\varphi \in \Phi_1$ needs at least one affirmation and $\sim\varphi \in \Phi_2$ needs n affirmations, which is possible. In the special case of *uniform premise-based quota rules*, there is one quota q for every literal in Φ_1, and the quota $q' = 1 - q$ for every literal in Φ_2. We will focus on such rules and denote them by $UPQR_q$. For $q = 1/2$ and the case of an odd number of judges, we obtain the premise-based procedure defined by Endriss et al. [14], and we will denote it by *PBP*.

Furthermore, we will consider yet another variant of premise-based procedure, which was introduced by Dietrich and List [8] and is called *constant premise-based quota rule* and is defined by $CPQR(\mathbf{J}) = \triangle'_q \cup \{\varphi \in \Phi_c \mid \triangle'_q \models \varphi\}$. Here, the number of affirmations needed to be in the set $\triangle'_q$ is a fixed constant. Thus $q_\varphi \in \mathbb{N}$, $0 \le q_\varphi < n$, and $\triangle'_q = \{\varphi \in \Phi_1 \mid \|\{i \mid \varphi \in J_i\}\| > q_\varphi\} \cup \{\varphi \in \Phi_2 \mid \|\{i \mid \varphi \in J_i\}\| > q'_\varphi\}$. Again, to ensure that for every $\varphi \in \Phi$, either $\varphi \in CPQR(\mathbf{J})$ or $\sim\varphi \in CPQR(\mathbf{J})$, we require that $q_\varphi + q'_\varphi = n - 1$ for all $\varphi \in \Phi_1$. The uniform variant, $UCPQR_q$, is defined analogously. If the number of judges who take part in the process is fixed, both classes represent the same judgment aggregation procedures. However, we will study control problems where the number of judges can vary. The constant premise-based quota n can then be seen as an upper bound on the highest number of judges possibly participating in the process. This definition is closely related to (a simplified version of) a referendum. Suppose that there is a fixed number of possible participants who are allowed to go to the polls, and there is a fixed number of affirmations needed for a certain decision, independent of the number of people who are actually participating. Of course, this number may depend on the number of possible participants, for example 20% of them.

3 Motivation for Control in Judgment Aggregation

We study three types of control for judgment aggregation. So far control has been studied extensively for voting systems (see, e.g., [3, 23, 4, 16]), where control is normally perceived as dishonest and thus as an undesired behavior. Therefore, this research focuses on finding ways to avoid it. Looking at real-world examples, this point of view is not always justified; in fact, some "control" attempts may be justified by fairly decent considerations (e.g., excluding children from elections is some reasonable kind of exerting control). Nevertheless, one is well advised to be aware of control attempts, since their objective is indeed frequently enough abusive (e.g., excluding voters from elections based on racial or gender grounds, as is still common in certain countries, is abusive and unacceptable). If control is generally possible, one way of circumventing it is to study the computational complexity of the underlying decision problems. If it turns out to be NP-hard, the desired control action can, in general, not be performed in polynomial time, unless P = NP. For practical purposes, showing hardness in appropriate typical-case models is even more useful, but also more challenging [28]. As motivation for studying control in judgment aggregation, we will now illustrate the three different control types for judgment aggregation considered in this paper with some examples from the American jury trial system and international arbitration.

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.[2] 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.

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 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 impartial. 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 proved 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. As we can see, deleting judges/jurors is a central part of the empaneling procedure. However, since the total number of jurors is fixed, a new juror needs to be appointed for each deleted juror, which motivates the next scenario.

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.[3] 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.

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. [19] on "multimode control attacks."

4 Problem Definitions

Bribery problems in voting theory, as introduced by Faliszewski et al. [17] (see also, e.g., [12, 20]), model scenarios in which an external actor seeks to bribe some of the voters to change their votes such that a distinguished candidate becomes the winner of the election. In judgment aggregation it is not the case that one single candidate wins, but there is a decision for every formula in the agenda.

[2] See, for instance, Articles 37–40 of the ICSID Convention and Rules 1–4 of the ICSID Arbitration Rules, Articles 11–12 of the ICC Arbitration Rules, or Articles 7–10 of the UNCITRAL Arbitration Rules.

[3] For rules regarding the challenge, disqualification, and replacement of arbitrators, see Articles 56–58 of the ICSID Convention, Rules 9–11 of the ICSID Arbitration Rules, Articles 14–15 of the ICC Arbitration Rules, and Articles 12–14 of the UNCITRAL Arbitration Rules.

So the external actor might seek to obtain exactly his or her desired collective outcome by bribing the judges, or he or she might be interested only in the desired outcome of some formulas in Φ. The exact bribery problem is then defined as follows for a given aggregation procedure F.

F-EXACT BRIBERY	
Given:	An agenda Φ, a profile $\mathbf{T} \in \mathscr{J}(\Phi)^n$, a consistent and complement-free judgment set J (not necessarily complete) desired by the briber, and a positive integer k.
Question:	Is it possible to change up to k individual judgment sets in $\mathbf{T}$ such that for the resulting new profile $\mathbf{T}'$ it holds that $J \subseteq F(\mathbf{T}')$?

Note that if J is a complete judgment set then the question is whether $J = F(\mathbf{T}')$.

Since in the case of judgment aggregation there is no winner, we also adopt the approach Endriss et al. [14] used to define the manipulation problem in judgment aggregation. In their definition, an outcome (i.e., a collective judgment set) is more desirable for the manipulator if its Hamming distance to the manipulator's desired judgment set is smaller, where for an agenda Φ the Hamming distance $H(J,J')$ between two complete and consistent judgment sets $J,J' \in \mathscr{J}(\Phi)$ is defined as the number of positive formulas in Φ on which J and J' differ. The formal definition of the manipulation problem in judgment aggregation is as follows, for a given aggregation procedure F.

F-MANIPULATION	
Given:	An agenda Φ, a profile $\mathbf{T} \in \mathscr{J}(\Phi)^{n-1}$, and a consistent and complete judgment set J desired by the manipulator.
Question:	Does there exist a judgment set $J' \in \mathscr{J}(\Phi)$ such that $H(J,F(\mathbf{T},J')) < H(J,F(\mathbf{T},J))$?

A specific judgment aggregation procedure is called *strategyproof* if a manipulator can never benefit from reporting an insincere preference. Now, we can give the formal definition of bribery in judgment aggregation, where the briber seeks to obtain a collective judgment set having a smaller Hamming distance to the desired judgment set, then the original outcome has. In bribery scenarios, we extend the above approach of Endriss et al. [14] by allowing that the desired outcome for the briber may be an incomplete (albeit consistent and complement-free) judgment set. This reflects a scenario where the briber may be interested only in some part of the agenda. The definition of Hamming distance is extended accordingly as follows. Let Φ be an agenda, $J \in \mathscr{J}(\Phi)$ be a complete and consistent judgment set, and $J' \subseteq \Phi$ be a consistent and complement-free judgment set. The *Hamming distance $H(J,J')$ between J and J'* is defined as the number of formulas from J' on which J does not agree: $H(J,J') = \|\{\varphi \mid \varphi \in J' \wedge \varphi \notin J\}\|$. Observe that if J' is also complete, this extended notion of Hamming distance coincides with the notion Endriss et al. [14] use.

F-BRIBERY	
Given:	An agenda Φ, a profile $\mathbf{T} \in \mathscr{J}(\Phi)^n$, a consistent and complement-free judgment set J (not necessarily complete) desired by the briber, and a positive integer k.
Question:	Is it possible to change up to k individual judgment sets in $\mathbf{T}$ such that for the resulting new profile $\mathbf{T}'$ it holds that $H(F(\mathbf{T}'),J) < H(F(\mathbf{T}),J)$?

Faliszewski et al. [20] introduced microbribery for voting systems. We adopt their notion so as to apply to judgment aggregation. In microbribery for judgment aggregation, if the briber's budget is k, he or she is not allowed to change up to k entire judgment sets but instead can change up to k premise entries in the given profile (the conclusions change automatically if necessary). We will denote this problem by F-MICROBRIBERY, and the exact variant by F-EXACT MICROBRIBERY.

We will now formally define the underlying decision problems for the complexity-theoretic study of control in judgment aggregation, closely related to the corresponding problems in elections. For a given judgment aggregation procedure F, the problem of control by adding judges is defined as follows:

F-CONTROL BY ADDING JUDGES	
Given:	An agenda Φ, complete profiles $\mathbf{T} \in \mathscr{J}(\Phi)^n$ and $\mathbf{S} \in \mathscr{J}(\Phi)^{\|\mathbf{S}\|}$, a positive integer k, and a consistent and complement-free judgment set J (not necessarily complete).
Question:	Is there a subset $\mathbf{S}' \subset \mathbf{S}$, $\|\mathbf{S}'\| \leq k$, such that $H(J, F(\mathbf{T} \cup \mathbf{S}')) < H(J, F(\mathbf{T}))$?

If we consider the variant F-EXACT CONTROL BY ADDING JUDGES, we ask if there is a subset $\mathbf{S}' \subset \mathbf{S}$, $\|\mathbf{S}'\| \leq k$, such that $J \subseteq F(\mathbf{T} \cup \mathbf{S}')$.

Control by deleting judges is defined as follows for a given judgment aggregation procedure F:

F-CONTROL BY DELETING JUDGES	
Given:	An agenda Φ, a complete profile $\mathbf{T} \in \mathscr{J}(\Phi)^n$, a positive integer k, and a consistent and complement-free judgment set J (not necessarily complete).
Question:	Is there a subset $\mathbf{T}' \subset \mathbf{T}$ with $\|\mathbf{T}'\| \leq k$ such that $H(J, F(\mathbf{T} \setminus \mathbf{T}')) < H(J, F(\mathbf{T}))$?

The exact variant is defined analogously to the case of adding judges.

The new control problem we introduce here is specific to judgment aggregation. It considers the case where some judges may be replaced (see our motivating examples in Section 3):

F-CONTROL BY REPLACING JUDGES	
Given:	An agenda Φ, complete profiles $\mathbf{T} \in \mathscr{J}(\Phi)^n$ and $\mathbf{S} \in \mathscr{J}(\Phi)^{\|\mathbf{S}\|}$, a positive integer k, and a consistent and complement-free judgment set J (not necessarily complete).
Question:	Are there subsets $\mathbf{T}' \subset \mathbf{T}$ and $\mathbf{S}' \subset \mathbf{S}$, with $\|\mathbf{T}'\| = \|\mathbf{S}'\| \leq k$ such that $H(J, F((\mathbf{T} \setminus \mathbf{T}') \cup \mathbf{S}')) < H(J, F(\mathbf{T}))$?

Define F-EXACT CONTROL BY REPLACING JUDGES analogously to the exact variants of the adding and deleting judges problems. To study the computational complexity of adding, deleting, and replacing judges, we adopt the terminology introduced in [3] for control problems in voting and adapt it to judgment aggregation. Let F be an aggregation procedure and let $\mathscr{C}$ be a given control type. F is said to be *immune* to control by $\mathscr{C}$ if it is never possible for an external person to successfully control the judgment aggregation procedure via $\mathscr{C}$-control. F is said to be *susceptible* to control by $\mathscr{C}$ if it is not immune. F is said to be *resistant* to control by $\mathscr{C}$ if it is susceptible and the corresponding decision problem is NP-hard. F is said to be *vulnerable* to control by $\mathscr{C}$ if it is susceptible and the corresponding decision problem is in P.

We assume that the reader is familiar with the basic concepts of complexity theory and with complexity classes such as P and NP; see, e.g., [27]. Downey and Fellows [10] introduced *parameterized* complexity theory; in their framework it is possible to do a more fine-grained multi-dimensional complexity analysis. In particular, NP-complete problems may be easy (i.e., fixed-parameter tractable) with respect to certain parameters confining the seemingly unavoidable combinatorial explosion. If this parameter is reasonably small, a fixed-parameter tractable problem can be solved efficiently in practice, despite its NP-hardness. Formally, a *parameterized decision problem* is a set $L \subseteq \Sigma^* \times N$, and we say it is *fixed-parameter tractable* (FPT) if there is a constant c such that for each input (x,k) of size $n = |(x,k)|$ we can determine in time $O(f(k) \cdot n^c)$ whether (x,k) is in L, where f is a function depending only on the parameter k. The main hierarchy of parameterized complexity classes is: $\mathrm{FPT} = \mathrm{W}[0] \subseteq \mathrm{W}[1] \subseteq \mathrm{W}[2] \subseteq \cdots \subseteq \mathrm{W}[\ell] \subseteq \mathrm{XP}$.

In our results, we will focus on only the class W[2], which refers to problems that are considered to be fixed-parameter intractable. In order to show that a parameterized problem is W[2]-hard, we will give a parameterized reduction from the W[2]-complete problem k-DOMINATING SET (see [10]). We say that a parameterized problem A *parameterized reduces* to a parameterized problem B if each instance (x,k) of A can be transformed in time $\mathscr{O}(g(k) \cdot |x|^c)$ (for some function g and some constant c) into an instance (x',k') of B such that $(x,k) \in A$ if and only if $(x',k') \in B$, where $k' = g(k)$. Note that $g(k) \equiv c$ may also be a constant function not depending on k.

In our proofs we will make use of three different problems. First, we will use the NP-complete problem EXACT COVER BY 3-SETS (X3C for short), where an instance consists of a given set $X = \{x_1, \ldots, x_{3m}\}$ and a collection $C = \{C_1, \ldots, C_n\}$ of 3-element subsets of X, and the question is whether there is an *exact cover for X*, i.e., a subcollection $C' \subseteq C$ such that every element of X occurs in exactly one member of C'. We will also use the DOMINATING SET problem, where we are given a graph $G = (V,E)$ and a positive integer k, and the question is whether there is a *dominating set for G of size at most k*, i.e., whether there is a subset $V' \subseteq V$, $\|V\| \leq k$, such that for each $v \in V$, either $v \in V'$ or there is a $w \in V'$ with $\{v,w\} \in E$. DOMINATING SET is NP-complete and, when parameterized by the upper bound k on the size of the dominating set, its parameterized variant (denoted by k-DOMINATING SET, to be explicit) is W[2]-complete [10]. Finally, we will also use the following problem for our parameterized complexity results:

OPTIMAL LOBBYING	
Given:	An $m \times n$ 0-1 matrix L (whose rows represent the voters, whose columns represent the referenda, and whose 0-1 entries represent No/Yes votes), a positive integer $k \leq m$, and a target vector $x \in \{0,1\}^n$.
Question:	Is there a choice of k rows in L such that by changing the entries of these rows the resulting matrix has the property that, for each j, $1 \leq j \leq n$, the jth column has a strict majority of ones (respectively, zeros) if and only if the jth entry of the target vector x of The Lobby is one (respectively, zero)?

OPTIMAL LOBBYING has been introduced and, parameterized by the number k of rows The Lobby can change, shown to be W[2]-complete by Christian et al. [6] (see also [15] for a more general framework and more W[2]-hardness results).

Note that a multiple referendum as in OPTIMAL LOBBYING can be seen as the special case of a judgment aggregation scenario where the agenda is closed under complementation and propositional variables and contains only premises and where the majority rule is used for aggregation. For illustration, consider the following simple example of a multiple referendum. Suppose the citizens of a town are asked to decide by a referendum whether two projects, A and B (e.g., a new hospital and a new bridge), are to be realized. Suppose the building contractor (who, of course, is interested in being awarded a contract for both projects) sets some money aside to attempt to influence the outcome of the referenda, by bribing some of the citizens without exceeding this budget. Observe that an *PBP*-EXACT BRIBERY instance with only premises in the agenda and with a complete desired judgment set J is nothing other than an OPTIMAL LOBBYING instance, where J corresponds to The Lobby's target vector.[4] Requiring the citizens to give their opinion only for the premises A and B of the referendum and not for the conclusion (whether both projects are to be realized) again avoids the doctrinal paradox. Again, the citizens might also vote strategically in these referenda. Both projects will cost money, and if both projects are realized, the amount available for each must be reduced. Some citizens may wish to support some project, say A, and may be unhappy with reducing the amount for A due to both projects being realized. They might even prefer none of the projects being realized over only B being realized. For them it is natural to consider the possibility of reporting insincere votes (provided they know how the others will vote); this may turn out to be more advantageous for them, as then they can possibly prevent that both projects are realized.

[4] Although exact bribery in judgment aggregation generalizes optimal lobbying in the sense of Christian et al. [6] (which is different from bribery in voting, as defined by Faliszewski et al. [17]), we will use the term "bribery" rather than "lobbying" in the context of judgment aggregation.

5 Results

We start by extending the result of Endriss et al. [14] that *PBP*-MANIPULATION is NP-complete. We study a parameterized version of the manipulation problem and establish a W[2]-hardness result with respect to the uniform premise-based quota rule. Due to space restrictions all proofs except one will be omitted.

Theorem 2 *For each rational quota q, $0 \leq q < 1$ and for any fixed number $n \geq 3$ of judges, $UPQR_q$-*MANIPULATION *is* W[2]*-hard when parameterized by the maximum number of changes in the premises needed in the manipulator's judgment set.*

Since the reduction is from the NP-complete problem DOMINATING SET, NP-completeness of $UPQR_q$-MANIPULATION, $0 \leq q < 1$, for any fixed number $n \geq 3$ of judges follows immediately from the proof of Theorem 2. Note that NP-hardness of $UPQR_q$-MANIPULATION could have also been shown by a modification of the proof of Theorem 2 in [14], but this reduction would not be appropriate to establish W[2]-hardness, since the corresponding parameterized version of SAT is not known to be W[2]-hard.

As mentioned above, studying the case of a fixed total number of judges is very natural. The second parameter we have considered for the manipulation problem in Theorem 2 is the "maximum number of changes in the premises needed in the manipulator's judgment set." Hence this theorem shows that the problem remains hard even if the number of premises the manipulator can change is bounded by a fixed constant. This is also very natural, since the manipulator may wish to report a judgment set that is as close as possible to his or her sincere judgment set, because for a completely different judgment set it might be discovered too easily that he was judging strategically.

In contrast to the hardness results stated in Theorem 2, the following proposition shows that, depending on the agenda, there are cases in which manipulation for $UPQR_q$, $0 \leq q < 1$, is outright impossible, and thus $UPQR_q$-MANIPULATION is trivially in P.

Proposition 3 *If the agenda contains only premises then $UPQR_q$, $0 \leq q < 1$, is strategyproof.*

NP-completeness for $UPQR_q$-MANIPULATION with a fixed number of judges, which is stated in Theorem 2, implies that there is little hope to find a polynomial-time algorithm for the general problem even when the number of participating judges is fixed. However, Proposition 3 tells us that if the agenda is simple and contains no conclusions, $UPQR_q$ is even strategyproof.

Now we will study the complexity of various bribery problems for the premise-based procedure *PBP*, i.e., $UPQR_{1/2}$ for an odd number of judges. We will establish NP-completeness for bribery, microbribery, and exact microbribery, and a W[2]-hardness result for exact bribery with respect to a natural parameter. We start with bribery.

Theorem 4 *PBP*-BRIBERY *is* NP*-complete, even when the total number of judges ($n \geq 3$ odd) or the number of judges that can be bribed is a fixed constant.*

Next, we turn to microbribery. Here the briber can change only up to a fixed number of entries in the individual judgment sets. We again prove NP-completeness when the number of judges or the number of microbribes allowed is a fixed constant.

Theorem 5 *PBP*-MICROBRIBERY *is* NP*-complete, even when the total number of judges ($n \geq 3$ odd) or the number of microbribes allowed is a fixed constant.*

Theorem 6 *PBP*-EXACT BRIBERY *is* W[2]*-hard when parameterized by the number of judges that can be bribed.*

This result follows from the fact that OPTIMAL LOBBYING is a special case of *PBP*-EXACT BRIBERY. Note that W[2]-hardness with respect to any parameter directly implies NP-hardness for the corresponding unparameterized problem, so *PBP*-EXACT BRIBERY is also NP-complete; all (unparameterized) problems considered here are easily seen to be in NP.

Theorem 7 *PBP*-EXACT MICROBRIBERY *is* NP-*complete, even when the total number of judges ($n \geq 3$ odd) or the number of microbribes allowed is a fixed constant.*

As for the manipulation problem, Theorems 4, 5, and 7 are concerned with a fixed number of judges. It turns out that even in this case BRIBERY, MICROBRIBERY, and EXACT MICROBRIBERY are NP-complete for *PBP*. Furthermore, we consider the case of a fixed number of judges allowed to bribe for *PBP*-BRIBERY, the corresponding parameter for its exact variant, and the case where the number of microbribes allowed is a fixed constant for *PBP*-MICROBRIBERY and its exact variant. Both parameters concern the budget of the briber. Since the briber aims at spending as little money as possible, it is also natural to consider these cases. But again, NP-completeness was shown even when the budget is a fixed constant and in one case W[2]-hardness for this parameter, so bounding the budget does not help to solve the problem easily. Although the exact microbribery problem is computationally hard in general for the aggregation procedure *PBP*, there are some interesting naturally restricted instances where it is computationally easy.

Theorem 8 *If the desired judgment set J is complete or if the desired judgment set is incomplete but contains all of the premises or only premises, then PBP*-EXACT MICROBRIBERY *is in* P.

In the last part of this section we study control in judgment aggregation. In the manipulation and bribery problems studied in this paper the number of participating judges 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 participating judges is *not* fixed, as in control by adding or deleting judges. For the uniform premise-based quota rule the number of affirmations needed to be in the collective judgment set varies with the number of judges, whereas for the constant premise-based quota rule the number of affirmations remains the same regardless of the number of judges participating. Since the number of participating judges varies for both control by adding and by deleting judges, we study these problems with respect to both judgment aggregation procedures.

We will first consider the uniform constant premise-based quota rule and show NP-hardness of $UCPQR_q$ for control by adding and by deleting judges in the Hamming distance based and in the exact variant.

Theorem 9 *For each admissible value of q, $UCPQR_q$ is resistant to* CONTROL BY ADDING JUDGES *and to* EXACT CONTROL BY ADDING JUDGES.

Theorem 10 *For each admissible value of q, $UCPQR_q$ is resistant to* CONTROL BY DELETING JUDGES *and to* EXACT CONTROL BY DELETING JUDGES.

Now we turn to the results for the uniform premise-based quota rule in the case of control by adding and by deleting judges. Here we only consider $UPQR_{1/2}$, which equals the premise-based procedure *PBP* defined by Endriss et al. [14] for an odd number of judges. We show NP-hardness for control by adding and by deleting judges in both problem variants.

Theorem 11 *$UPQR_{1/2}$ is resistant to* EXACT CONTROL BY ADDING JUDGES *and to* CONTROL BY ADDING JUDGES.

Proof. Membership in NP is obvious for both problems. Again, we show NP-hardness for $UPQR_{1/2}$-EXACT CONTROL BY ADDING JUDGES only and $UPQR_{1/2}$-CONTROL BY ADDING

JUDGES at the same time, by a reduction from the NP-complete problem X3C. Given an X3C instance (X,C) with $X=\{x_1,\ldots,x_{3m}\}$ and $C=\{C_1,\ldots,C_n\}$, define the following judgment aggregation scenario. The agenda Φ contains $\{\alpha_0,\alpha_1,\ldots,\alpha_{3m}\}$ and their negations. The quota is $1/2$ for every positive literal. The profile of the individual judgment sets initially taking part in the process is $\mathbf{T}=(T_1,\ldots,T_{m+1})$ with $T_1=\{\alpha_0,\alpha_1,\ldots,\alpha_{3m}\}$, $T_i=\{\neg\alpha_0,\alpha_1,\ldots,\alpha_{3m}\}$, $2\leq i\leq m$, and $T_{m+1}=\{\neg\alpha_0,\neg\alpha_1,\ldots,\neg\alpha_{3m}\}$. The profile of the judges who can be added is $\mathbf{S}=(S_1,\ldots,S_n)$ with $S_i=\{\alpha_0,\alpha_j,\neg\alpha_\ell \mid x_j\in C_i, x_\ell\notin C_i, 1\leq j,\ell\leq 3m\}$. The maximum number of judges from $\mathbf{S}$ who can be added is m. The desired outcome of the external person is $J=\{\alpha_0,\alpha_1,\ldots,\alpha_{3m}\}$. Then it holds, that there is a profile $\mathbf{S}'\subseteq\mathbf{S}$, $\|\mathbf{S}'\|\leq m$, such that $H(J,F(\mathbf{T}\cup\mathbf{S}'))<H(J,F(\mathbf{T}))$ if and only if there is an exact cover for the given X3C instance. The collective judgment set for $UPQR_{1/2}(\mathbf{T})$ is $\{\neg\alpha_0,\alpha_1,\ldots,\alpha_{3m}\}$. Observe that $H(J,F(\mathbf{T}))=1$, since the only difference lies in α_0. Hence, $F(\mathbf{T}\cup\mathbf{S}')$ must be exactly J, and the reduction will hold for both problems at hand.

($\Leftarrow$) Assume that there is an exact cover $C'\subseteq C$ for the given X3C instance (X,C). Then the profile $\mathbf{S}'$ contains those judges S_i with $C_i\in C'$. 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 $\mathbf{T}$ and m affirmations from the judges in $\mathbf{S}'$. Every α_i, $1\leq i\leq 3m$, gets m affirmations from the judges in $\mathbf{T}$ and one affirmation from a judge in $\mathbf{S}'$. Hence, the collective judgment set is J, as desired.

($\Rightarrow$) Assume that there is a profile $\mathbf{S}'$ with $\|\mathbf{S}'\|\leq m$ such that $UPQR_{1/2}(\mathbf{T}\cup\mathbf{S}')=J$. Since α_0 is contained in the collective judgment set it must receive enough affirmations of the judges in $\mathbf{S}'$. Adding less than m new affirmations for α_0 is not enough, since $m-1\leq(2m)(1/2)$, but since $(2m+1)(1/2)<m+1$, m new affirmations are enough. As above, if there is a total number of $2m+1$ judges then the number of affirmations needed for a positive formula to be in the collective judgment set is $m+1$. Since the α_i, $1\leq i\leq 3m$, receive only m affirmations from $\mathbf{T}$, they must all get one additional affirmation from $\mathbf{S}'$. Since $\|\mathbf{S}'\|\leq m$ and every judge affirms of exactly four formulas, including α_0, the sets C_i corresponding to the judges in $\mathbf{S}'$ must form an exact cover for the given X3C instance. ❑

One important point regarding the proof of Theorem 11 is that the agenda contains only premises. For $UPQR_{1/2}$-EXACT CONTROL BY DELETING JUDGES, the proof of Theorem 12 below also establishes NP-hardness even if the agenda contains only premises. By contrast, in Proposition 3 we showed that if the agenda contains only premises then $UPQR_q$ is strategyproof (thus, $UPQR_q$-MANIPULATION is in P) for each rational quota q, $0\leq q<1$, and in Theorem 5 we showed that $UPQR_{1/2}$-EXACT MICROBRIBERY is also in P if the desired judgment set contains only premises.

Theorem 12 *$UPQR_{1/2}$ is resistant to* EXACT CONTROL BY DELETING JUDGES *and* CONTROL BY DELETING JUDGES.

In contrast to $UPQR_{1/2}$-CONTROL BY ADDING JUDGES it remains open whether $UPQR_{1/2}$-CONTROL BY DELETING JUDGES is still NP-complete if the agenda contains only premises.

Unlike for manipulation and bribery, we have not been able to identify natural restrictions for which one of our NP-hard control problems can be solved in polynomial time.

Finally, we consider CONTROL BY REPLACING JUDGES. In contrast to the problems of control by adding and by deleting judges, the number of judges here is constant, just as in the corresponding manipulation and bribery problems for judgment aggregation. Thus, there is no difference between the uniform constant premise-based quota rule and the uniform premise-based quota rule. The following theorem implies NP-completeness for both classes of rules.

Theorem 13 *For each rational quota q, $0\leq q<1$, $UPQR_q$ is resistant to* EXACT CONTROL BY REPLACING JUDGES *and* CONTROL BY REPLACING JUDGES.

To conclude, we mention some possible future research questions. First, we have introduced some very natural control problems for judgment aggregation. Are there any others? Second, it would be very interesting to complement our NP-hardness results by typical-case analyses, as has been done for voting problems (see the survey [28]). Third, from all W[2]-hardness results we immediately obtain the corresponding NP-hardness results, and since all problems considered are easily seen to be in NP, we have NP-completeness results. It remains open, however, whether one can also obtain matching upper bounds in terms of parameterized complexity. We suspect that all W[2]-hardness results in this paper in fact can be strengthened to W[2]-completeness results. Finally, note that we have considered only "constructive" control scenarios. For voting problems, constructive control means that the Chair's goal is to make some candidate win, whereas "destructive" control [23] refers to making any other than the most hated candidate win the election. Constructive control in judgment aggregation, however, means that we seek an outcome *closer to the desired outcome*, or *exactly the desired outcome*. Note that defining destructive variants of control by adding, deleting, or replacing judges would thus lead to the same definitions as for their constructive counterparts: We have an undesired (possibly partial) judgment set $J \in \mathscr{J}(\Phi)$ and seek an outcome with a smaller Hamming distance to the complement of J than from the original outcome to the complement of J, but replacing the (partial) judgment set J with its complement leads to essentially the same question, as the complement of a partial judgment set J is simply the negation of the formulas in J. Therefore, it does not make sense to distinguish between constructive and desctructive control.

Acknowledgments: We thank the reviewers for helpful comments.

References

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

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

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

[4] D. Baumeister, G. Erdélyi, E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Computational aspects of approval voting. In J. Laslier and R. Sanver, editors, *Handbook on Approval Voting*, chapter 10, pages 199–251. Springer, 2010.

[5] F. Brandt, V. Conitzer, and U. Endriss. Computational social choice. In G. Weiß, editor, *Multiagent Systems*. MIT Press, second edition, 2012. To appear.

[6] R. Christian, M. Fellows, F. Rosamond, and A. Slinko. On complexity of lobbying in multiple referenda. *Review of Economic Design*, 11(3):217–224, 2007.

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

[8] F. Dietrich and C. List. Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics*, 19(4):391–424, 2007.

[9] F. Dietrich and C. List. Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(3):269–300, 2007.

[10] R. Downey and M. Fellows. *Parameterized Complexity*. Springer-Verlag, 1999.

[11] J. Duggan and T. Schwartz. Strategic manipulability without resoluteness or shared beliefs: Gibbard–Satterthwaite generalized. *Social Choice and Welfare*, 17(1):85–93, 2000.

[12] E. Elkind, P. Faliszewski, and A. Slinko. Swap bribery. In *Proceedings of the 2nd International Symposium on Algorithmic Game Theory*, pages 299–310. Springer-Verlag *Lecture Notes in Computer Science #5814*, October 2009.

[13] U. Endriss, U. Grandi, and D. Porello. Complexity of judgment aggregation: Safety of the agenda. In *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 359–366. IFAAMAS, May 2010.

[14] U. Endriss, U. Grandi, and D. Porello. Complexity of winner determination and strategic manipulation in judgment aggregation. In V. Conitzer and J. Rothe, editors, *Proceedings of the 3rd International Workshop on Computational Social Choice*, pages 139–150. Universität Düsseldorf, September 2010.

[15] G. Erdélyi, H. Fernau, J. Goldsmith, N. Mattei, D. Raible, and J. Rothe. The complexity of probabilistic lobbying. In *Proceedings of the 1st International Conference on Algorithmic Decision Theory*, pages 86–97. Springer-Verlag *Lecture Notes in Artificial Intelligence #5783*, October 2009.

[16] G. Erdélyi, L. Piras, and J. Rothe. The complexity of voter partition in Bucklin and fallback voting: Solving three open problems. In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 837–844. IFAAMAS, May 2011.

[17] P. Faliszewski, E. Hemaspaandra, and L. Hemaspaandra. How hard is bribery in elections? *Journal of Artificial Intelligence Research*, 35:485–532, 2009.

[18] P. Faliszewski, E. Hemaspaandra, and L. Hemaspaandra. Using complexity to protect elections. *Communications of the ACM*, 53(11):74–82, 2010.

[19] P. Faliszewski, E. Hemaspaandra, and L. Hemaspaandra. Multimode control attacks on elections. *Journal of Artificial Intelligence Research*, 40:305–351, 2011.

[20] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Llull and Copeland voting computationally resist bribery and constructive control. *Journal of Artificial Intelligence Research*, 35:275–341, 2009.

[21] P. Faliszewski and A. Procaccia. AI's war on manipulation: Are we winning? *AI Magazine*, 31(4):53–64, 2010.

[22] A. Gibbard. Manipulation of voting schemes. *Econometrica*, 41(4):587–601, 1973.

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

[24] L. Kornhauser and L. Sager. Unpacking the court. *Yale Law Journal*, 96(1):82–117, 1986.

[25] C. List. The theory of judgment aggregation: An introductory review. *Synthese*. To appear.

[26] C. List and C. Puppe. Judgment aggregation: A survey. In P. Anand, P. Pattanaik, and C. Puppe, editors, *Oxford Handbook of Rational and Social Choice*, chapter 19. Oxford University Press, 2009.

[27] C. Papadimitriou. *Computational Complexity*. Addison-Wesley, second edition, 1995.

[28] J. Rothe and L. Schend. Typical-case challenges to complexity shields that are supposed to protect elections against manipulation and control: A survey. In *Website Proceedings of the Special Session on Computational Social Choice at the 12th International Symposium on Artificial Intelligence and Mathematics*, January 2012.

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

Dorothea Baumeister and Jörg Rothe
Institut für Informatik
Heinrich-Heine-Universität Düsseldorf
40225 Düsseldorf, Germany
Email: `{baumeister,rothe}@cs.uni-duesseldorf.de`

Gábor Erdélyi
Institut für Betriebswirtschaftslehre
Universität Siegen
57076 Siegen, Germany
Email: `erdelyi@wiwi.uni-siegen.de`

Olivia J. Erdélyi
Institut für Versicherungsrecht
Heinrich-Heine-Universität Düsseldorf
40225 Düsseldorf, Germany
Email: `olivia.erdelyi@uni-duesseldorf.de`