Control in Judgment Aggregation

Dorothea BAUMEISTER a, 2, Gábor ERDÉLYI b, Olivia J. ERDÉLYI c, and Jörg ROTHE a

a Institut für Informatik, Heinrich-Heine-Universität Düsseldorf, Germany
b Institut für Betriebswirtschaftslehre, Universität Siegen, Germany
c Institut für Versicherungsrecht, Heinrich-Heine-Universität Düsseldorf, Germany

Abstract. In computational social choice, the complexity of changing the outcome of elections via various control actions, such as adding or deleting candidates or voters, has been studied intensely. We introduce the concept of control for judgment aggregation procedures, and study the complexity of changing the outcome of such procedures via control by adding, deleting, or replacing judges.

Keywords. judgment aggregation, computational complexity, control

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. [1,2,3] and the recent surveys and bookchapters by Faliszewski et al. [4,5], Brandt et al. [6], and Baumeister et al. [7]). In particular, (coalitional) manipulation [1,2,8] refers to (a group of) strategic voters casting their votes insincerely to reach their desired outcome; in bribery [9,10] 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 [3,11,12] 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” [13,14] (see also the generalization by Duggan and
many natural voting systems are not “immune” to most or even all of the standard types of control [3, 11, 12], and Dietrich and List [16] 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. This paper is the first to study control in judgment aggregation.

Our work continues the complexity-theoretic study of judgment aggregation initiated recently by Endriss et al. [17, 18], who in particular defined the winner determination problem and the manipulation problem in judgment aggregation [18] and studied their complexity for two important judgment aggregation rules. Baumeister et al. [19] extended their complexity-theoretic investigation for manipulation and also introduced various bribery problems in judgment aggregation. These problems are each closely related to the corresponding winner-determination, manipulation, and bribery problems in voting, yet are specifically tailored to judgment aggregation scenarios. In this paper, we introduce and motivate three types of control in judgment aggregation (namely, control by adding, deleting, or replacing judges), and study their computational complexity.

2. Formal Framework

We adopt the framework stated by Baumeister et al. [19], which follows and extends the judgment aggregation framework described by Endriss et al. [18].

Let \( PS \) be the set of all propositional variables and \( L_{PS} \) the set of propositional formulas built from \( PS \), where the following connections can be used in their usual meaning: disjunction (\( \lor \)), conjunction (\( \land \)), implication (\( \rightarrow \)), equivalence (\( \leftrightarrow \)), and the boolean constants 1 and 0. To avoid double negations, let \( \sim \alpha \) denote the complement of \( \alpha \), i.e., \( \sim \alpha = \neg \alpha \) if \( \alpha \) is not negated, and \( \sim \alpha = \beta \) if \( \alpha = \neg \beta \). The judges have to judge over all formulas in the agenda \( \Phi \), which is a finite, nonempty subset of \( L_{PS} \) 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 \( \Phi \) 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 \( \alpha \in \Phi \), \( \alpha \) 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 \( \mathcal{J}(\Phi) \) we denote the set of all complete and consistent subsets of \( \Phi \).

The famous doctrinal paradox [20] 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. Consider, for example, a controversial penalty situation in a soccer match with three referees having different views of the situation. According to the rules, a team must get a penalty if they have been fouled in the penalty area. The first referee says that there was a foul in the penalty area; the second referee says that what he...
observed in the penalty area in fact was a dive, not a foul, so there is no penalty; and the third one denies a penalty as well, since he has seen a foul outside the penalty area. The three different individual judgments and the evaluation according to the majority rule are shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>penalty area</th>
<th>foul</th>
<th>penalty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referee 1</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Referee 2</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Referee 3</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>majority</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Applying the majority rule here leads to the inconsistent outcome that there was a foul in the penalty area, but there is no penalty. One way of circumventing the doctrinal paradox is to impose restrictions on the agenda. Endriss et al. [17] 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. [18] 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 follow the work of Baumeister et al. [19] who studied the complexity of manipulation and bribery for the more general class of premise-based quota rules as defined by Dietrich and List [21].

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

The quota for the literals \( \phi \in \Phi_2 \) is \( q'_\phi = 1 - q_\phi \).

A premise-based quota rule is then defined to be a function \( PQR : \mathcal{J}(\Phi)^n \to 2^\Phi \) such that, for \( \Phi = \Phi_p \cup \Phi_c \), each profile \( \mathbf{J} = (J_1, \ldots, J_n) \) is mapped to the judgment set

\[
PQR(\mathbf{J}) = \triangle_q \cup \{ \phi \in \Phi_c \mid \triangle_q \models \phi \},
\]

where

\[
\triangle_q = \{ \phi \in \Phi_1 \mid \| \{ i \mid \phi \in J_i \} \| > nq_\phi \} \cup \{ \phi \in \Phi_2 \mid \| \{ i \mid \phi \in J_i \} \| > \lceil nq'_\phi - 1 \rceil \}.
\]

To guarantee complete and consistent outcomes for this procedure, it is enough to require that \( \Phi \) is closed under propositional variables and that all literals are in \( \Phi_p \). The number of affirmations needed to be in the collective judgment set is \( [nq_\phi + 1] \) for literals \( \phi \in \Phi_1 \) and \( [nq'_\phi] \) for literals \( \phi \in \Phi_2 \). Note that \( [nq_\phi + 1] + [nq'_\phi] = n + 1 \) ensures that either \( \phi \in PQR(\mathbf{J}) \) or \( \neg \phi \in PQR(\mathbf{J}) \) for every \( \phi \in \Phi \).
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 \( \Phi_1 \), and the quota \( q' = 1 - q \) for every literal in \( \Phi_2 \). 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. [18].

Furthermore, we will consider yet another variant of premise-based procedure, which was introduced by Dietrich and List [21] and is called constant premise-based quota rule and is defined by \( CPQR(J) = \Delta_q' \cup \{ \varphi \in \Phi_1 \mid \Delta_q' \models \varphi \} \). Here, the number of affirmations needed to be in the set \( \Delta_q' \) is a fixed constant. Thus \( q_\varphi \in \mathbb{N} \), \( 0 \leq q_\varphi < n \), and \( \Delta_q' = \{ \varphi \in \Phi_1 \mid \| \{ i \mid \varphi \in J_i \} \| > q_\varphi \} \cup \{ \varphi \in \Phi_2 \mid \| \{ i \mid \varphi \in J_i \} \| > q'_\varphi \} \). Again, to ensure that for every \( \varphi \in \Phi \), either \( \varphi \in CPQR(J) \) or \( \sim \varphi \in CPQR(J) \), 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.

Applying the premise-based procedure with uniform quota \( q = 1/2 \) to the example from Table 1, the outcome for the premises “penalty area” and “foul” is again obtained by applying the majority rule, so the collective decision is that there was a foul in the penalty area. But the outcome for the conclusion “penalty” is now derived from the outcome of the premises, so the collective decision will be “penalty,” which is a consistent outcome.

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., [3,11,7,12]), 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 [22].
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. 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.

**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. 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. [23] on “multimode control attacks.”

---

3See, 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.

4For 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.
4. Problem Definitions

The manipulation problem in voting asks if a manipulator can make her favorite candidate win the election by reporting an untruthful preference. Since there is no single winner in judgment aggregation, Endriss et al. [18] proposed to study the problem of whether the manipulator can report an untruthful judgment set to obtain a collective judgment set with a smaller Hamming distance to her desired judgment set than the original outcome. This approach was followed up and extended by Baumeister et al. [19]. First, they allowed for incomplete desired judgment sets, and they additionally introduced an exact variant of the manipulation problem, where the manipulation is successful only if her desired (possibly incomplete) judgment set is part of the collective judgment set. We will follow this approach and define one variant based on a small Hamming distance and one exact variant for each of our three control problems.

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. As in [19], we define the Hamming distance between a complete and complement-free judgment set $J$ and a complement-free judgment set $J'$ to be the number of formulas in $J'$ on which $J$ does not agree: $H(J,J') = \Vert \{ \varphi \mid \varphi \in J' \land \varphi \not\in J \} \Vert$. If $J'$ is also complete, this yields the natural definition of the Hamming distance between two judgment sets as used by Endriss et al. [18].

Now we are ready to define our control problems. For a given judgment aggregation procedure $F$, the problem of control by adding judges is defined as follows:

\[
\text{\textit{F-CONTROL BY ADDING JUDGES}}
\]

\textbf{Given:} An agenda $\Phi$, complete profiles $T \in \mathcal{F}(\Phi)^n$ and $S \in \mathcal{F}(\Phi)^{|S|}$, a positive integer $k$, and a consistent and complement-free judgment set $J$ (not necessarily complete).

\textbf{Question:} Is there a subset $S' \subset S$, $\|S'\| \leq k$, such that $H(J,F(T \cup S')) \leq H(J,F(T))$?

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

For a given judgment aggregation procedure $F$, \textit{F-CONTROL BY DELETING JUDGES} is defined as follows: Given an agenda $\Phi$, a complete profile $T \in \mathcal{F}(\Phi)^n$, a positive integer $k$, and a consistent and complement-free judgment set $J$ (not necessarily complete), is there a subset $T' \subset T$ with $\|T'\| \leq k$ such that $H(J,F(T \setminus T')) < H(J,F(T))$? 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). For a given judgment aggregation procedure $F$, \textit{F-CONTROL BY REPLACING JUDGES} is defined as follows: Given an agenda $\Phi$, complete profiles $T \in \mathcal{F}(\Phi)^n$ and $S \in \mathcal{F}(\Phi)^{|S|}$, a positive integer $k$, and a consistent and complement-free judgment set $J$ (not necessarily complete), are there subsets $T' \subset T$ and $S' \subset S$, with $\|T'\| = \|S'\| \leq k$, such that $H(J,F((T \setminus T') \cup S')) < H(J,F(T))$? Define \textit{F-EXACT CONTROL BY REPLACING JUDGES} analogously to the exact variants of the adding and deleting judges problems.

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 $\mathcal{C}$ be a given control type. $F$ is said to be \textit{immune} to control by $\mathcal{C}$ if it is never possible for an external
person to successfully control the judgment aggregation procedure via $C$-control. $F$ is said to be susceptible to control by $C$ if it is not immune. $F$ is said to be resistant to control by $C$ if it is susceptible and the corresponding decision problem is NP-hard. $F$ is said to be vulnerable to control by $C$ if it is susceptible and the corresponding decision problem is in P. For the complexity-theoretic study of the problems defined above, we assume that the reader is familiar with the basic complexity classes, such as P and NP, and we refer to the textbook [24] for more background. The NP-hardness proofs will be from the NP-complete problems EXACT COVER BY 3-SETS and DOMINATING SET (see, e.g., [25]). An instance for EXACT COVER BY 3-SETS (X3C for short) consists of a given set $X = \{x_1, \ldots, x_m\}$ 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 $C'$. In the DOMINATING SET problem, given a graph $G = (V,E)$ and a positive integer $k$, we ask 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 $\{v, w\} \in E$.

5. Results

In the manipulation and bribery problems studied in [18,19], 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 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 in the constant premise-based quota rule the number of affirmations remains the same regardless of the number of judges participating. Since the number of judges participating varies for both adding and 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 hardness for UCPQR$_q$ for control by adding and deleting judges in the Hamming distance based and in the exact variant.

Theorem 2 For each admissible value of $q$, UCPQR$_q$ is resistant to CONTROL BY ADDING JUDGES and to EXACT CONTROL BY ADDING JUDGES.

Proof. Membership in NP is obvious for both problems. We prove only resistance to EXACT CONTROL BY ADDING JUDGES and note that the (omitted) proof for CONTROL BY ADDING JUDGES requires only slight modifications in the construction. The reduction to EXACT CONTROL BY ADDING JUDGES is from the NP-complete problem DOMINATING SET. For a given graph $G = (V,E)$ with vertex set $V = \{v_1, \ldots, v_n\}$, let $N(v_i)$ denote the closed neighborhood of vertex $v_i$, i.e., the set of all vertices that are adjacent to $v_i$ and the vertex $v_i$ itself. Then, a subset $V'$ of the vertices is a dominating set for $G$ if and only if for each $i$, $1 \leq i \leq n$, $N(v_i) \cap V' \neq \emptyset$. Let $(G,k)$ be a given DOMINATING SET instance. For the judgment aggregation instance, let the agenda $\Phi$ contain the variables $v_1, \ldots, v_n, y$ and their negations, for each $i$, $1 \leq i \leq n$, the formula $\varphi_i = v_i \lor \cdots \lor v_n \lor y$, where $\{v_1, \ldots, v_n\} = N(v_i)$, and its negation, the formula $\psi = v_1 \lor \cdots \lor v_n$, its negation, and $n-2$ syntactic variations of this formula and its negation. This can be seen as giving the weight $n-1$ to the formula $\psi$. The quota
for every positive literal is \( q \), hence \( q + 1 \) affirmations are needed to be in the collective judgment set. The set \( T \) of judges who initially take part contains \( q \) judgment sets that contain the literals \( v_1, \ldots, v_n \), and hence all formulas \( \varphi_i, 1 \leq i \leq n \), and the formula \( \psi \), and its syntactic variations, and the negation of all formulas in \( \Phi \) not mentioned here. Furthermore, there is one judgment set \( T \) that contains all negated formulas from \( \Phi \). The desired judgment set \( J \) is incomplete and contains the formulas \( \varphi_i, 1 \leq i \leq n \). Observe that \( J \) is consistent, since setting \( y \) to true and all \( v_i, 1 \leq i \leq n \), to false results in the desired evaluation. The profile \( S \) of judges who may be added contains \( n \) judges, with the individual judgment sets \( J_i, 1 \leq i \leq n \), where \( J_i \) contains the variable \( v_i \), the negation of all \( v_j, 1 \leq j \leq n, j \neq i \), the negation of \( y \), and the corresponding conclusions.

We claim that there is a dominating set of size at most \( k \) for \( G \) if and only if we can ensure that the outcome contains all formulas from \( J \) by adding at most \( k \) judges from \( S \). From left to right, if there is a dominating set \( V' \), we can ensure that the formulas from \( J \) are part of the collective judgment set by adding those judges \( J_i \) with \( v_i \in V' \). Thus, all formulas \( \varphi_i, 1 \leq i \leq n \), evaluate to true. Conversely, assume that all formulas \( \varphi_i, 1 \leq i \leq n \), evaluate to true. It is not possible to achieve this by having \( y \) in the collective outcome, since there are no individual judgment sets containing \( y \). Hence, the collective outcome for \( v_i, 1 \leq i \leq n \), makes all \( \varphi_i \) 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 \( S \) that is added. Hence, the vertices \( v_i \) corresponding to the judges \( J_i \) from \( S \) that have been added must form a dominating set for graph \( G \).

**Theorem 3** For each admissible value of \( q \), \( UCPQR_q \) is resistant to **Control by Deleting Judges** and to **Exact Control by Deleting Judges**.

**Proof.** Both problems are easily seen to be in NP, and we will show NP-hardness by a similar construction as in the proof of Theorem 2. Again, we consider only the problem **Exact Control by Deleting Judges**. For a given **Dominating Set** instance \((G,k)\), we construct the following judgment aggregation scenario. Let the agenda \( \Phi \) be the same as in the proof of Theorem 2 plus an additional variable \( z \), its negation, and \( n - 1 \) syntactic variations of this variable and its negation. This can again be seen as if \( z \) has a weight of \( n \). The quota is \( q \) for \( \neg v_i, 1 \leq i \leq n \), and all remaining positive literals. The profile \( T \) contains \( q \) individual judgment sets that each contain \( z \) and the negation of all remaining formulas in \( \Phi \), one judgment set containing \( v_1, \ldots, v_n, z \) and the negation of all remaining formulas, and for each \( i \) one judgment set \( J_i \) that contains all \( v_i, i \neq j \), and the negation of all remaining formulas. The judgment set desired by the chair is incomplete and contains \( z \) (and its syntactic variations) and the formulas \( \varphi_i, 1 \leq i \leq n \).

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. If there is a dominating set \( V' \) for \( G \), then the desired judgment set \( J \) is obtained by deleting those \( J_i \) with \( v_i \in V' \). Conversely, assume that it is possible that the desired formulas are in the collective judgment set by deleting at most \( k \) judges. Since \( z \) is in the collective outcome, only judges of the form \( J_i, 1 \leq i \leq n \), may be deleted. The deletion of a judge \( J_i \) has the effect that \( v_i \) is in the collective outcome, hence at most \( k \) different \( v_i \) are contained in the collective outcome, and since they evaluate all formulas \( \varphi_i, 1 \leq i \leq n \) to true, those \( v_i \) must form a dominating set of size at most \( k \) for \( G \).

We omit the similar proof of NP-hardness for **Control by Deleting Judges**. \(\square\)
Now we turn to the results for the uniform premise-based quota rule in the case of adding and deleting judges. Here we only consider UPQR\(_{1/2}\), which equals the premise-based procedure defined in [18]. We show NP-hardness for adding and deleting judges in both problem variants.

**Theorem 4** 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 \(\Phi\) contains \(\{α_0, α_1, \ldots, α_{3m}\}\) and their negations. The quota is \(\frac{1}{2}\) for every positive literal. The profile of the individual judgment sets initially taking part in the process is \(T = \{T_1, \ldots, T_{m-1}\}\) with \(T_i = \{α_0, α_1, \ldots, α_{3m}\}\), \(T_i = \{¬α_0, α_1, \ldots, α_{3m}\}\), \(2 ≤ i ≤ m\), and \(T_{m-1} = \{¬α_0, ¬α_1, \ldots, ¬α_{3m}\}\). The profile of the judges who can be added is \(S = (S_1, \ldots, S_n)\) with \(S_i = \{α_0, α_j, ¬α_0 \mid x_j ∈ C_i, x_j /∈ C_i, 1 ≤ j, ℓ ≤ 3m\}\). The maximum number of judges from \(S\) who can be added is \(m\). The desired outcome of the external person is \(J = \{α_0, α_1, \ldots, α_{3m}\}\). Then it holds, that there is a profile \(S′ ⊆ S\), \(∥S′∥ ≤ m\), such that \(H(J, F(T ∪ S′)) < H(J, F(T))\) if and only if there is an exact cover for the given X3C instance. The collective judgment set for UPQR\(_{1/2}\)(\(T\)) is \(\{¬α_0, α_1, \ldots, α_{3m}\}\). Observe that \(H(J, F(T)) = 1\), since the only difference lies in \(α_0\). Hence, \(F(T ∪ S′)\) must be exactly \(J\), and the reduction will hold for both problems at hand.

\((⇒)\) Assume that there is an exact cover \(C′ ⊆ C\) for the given X3C instance \((X,C)\). Then the profile \(S′\) contains those judges \(S_i\) with \(C_i ∈ C′\). 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 \(T\) and \(m\) affirmations from the judges in \(S′\). Every \(α_i\), \(1 ≤ i ≤ 3m\), gets \(m\) affirmations from the judges in \(T\) and one affirmation from a judge in \(S′\). Hence, the collective judgment set is \(J\), as desired.

\((⇒)\) Assume that there is a profile \(S′\) with \(∥S′∥ ≤ m\) such that UPQR\(_{1/2}\)(\(T ∪ S′\)) = \(J\). Since \(α_0\) is contained in the collective judgment set it must receive enough affirmations of the judges in \(S′\). Adding less than \(m\) new affirmations for \(α_0\) is not enough, since \(m − 1 ≤ (2m)(1/2)\), 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 \(α_i\), \(1 ≤ i ≤ 3m\), receive only \(m\) affirmations from \(T\), they must all get one additional affirmation from \(S′\). Since \(∥S′∥ ≤ m\) and every judge affirms of exactly four formulas, including \(α_0\), the sets \(C_i\) corresponding to the judges in \(S′\) must form an exact cover for the given X3C instance.

One important point regarding the proof of Theorem 4 is that the agenda contains only premises. Baumeister et al. [19] showed that if the agenda contains only premises, UPQR\(_q\)-MANIPULATION is in P for each rational quota \(q\), \(0 ≤ q < 1\), and they showed that EXACT-UPQR\(_{1/2}\)-MICROBRIBERY is also in P if the desired judgment set contains only premises. We now show that this is not the case for UPQR\(_{1/2}\)-EXACT CONTROL BY DELETING JUDGES.
Theorem 5 *UPQR*$_{1/2}$ is resistant to Exact Control by Deleting Judges and Control by Deleting Judges.

**Proof.** Membership in NP is obvious for both problems. We prove NP-hardness for *UPQR*$_{1/2}$-Exact Control by Deleting Judges only. A slight modification of the construction (omitted here due to space) works for *UPQR*$_{1/2}$-Control by Deleting Judges. The proof of NP-hardness for the exact problem is by a reduction from the NP-complete X3C problem. Given an X3C instance $(X, C)$ with $X = \{x_1, \ldots, x_{3m}\}$ and $C = \{C_1, \ldots, C_n\}$, we assume that every element from $X$ occurs in at least one set from $C$. If this is not the case, it is a no-instance for X3C (and we then map to an easily constructed no-instance of our exact control problem in judgment aggregation). Now, we construct the following judgment aggregation scenario. The agenda $\Phi$ contains $\beta, \alpha_0, \ldots, \alpha_{3m}$ and their negations, and the quota is $1/2$ for every positive literal. The complete profile is $T = T_1 \cup T_2$, where $T_1 = \{J_1, \ldots, J_{n+m}\}$ and $T_2 = \{L_1, \ldots, L_n\}$, so $|T| = 2n + m$. The individual judgment sets are $J_i$ contains the set $\{\alpha_j, \neg \alpha_i \mid m + d_i \geq i , m + d_i < i , 1 \leq j , \ell \leq 3m\}$, for $1 \leq i \leq n + m$, where $d_i$ is the number of sets $C_j$ in which $x_i$ occurs. Note that $\alpha_0$ is contained in $J_i$ if $i \leq n + 1$, and $\neg \alpha_0$ otherwise; and that $\beta$ is contained in $J_i$ if $i \leq m$, and $\neg \beta$ otherwise. Also, $L_j = \{\beta, \neg \alpha_0, \alpha_j, \neg \alpha_i \mid x_j \in C_j, x_i \in C_i, 1 \leq j , \ell \leq 3m\}$, $1 \leq i \leq n$. The desired outcome is $J = \{\beta, \alpha_0, \ldots, \alpha_{3m}\}$, and at most $m$ judges can be deleted. The outcome of *UPQR*$_{1/2}(T)$ is $\{\beta, \neg \alpha_0, \ldots, \alpha_{3m}\}$, since $\alpha_0$ receives $n + 1$ affirmations and $\beta$ and all $\alpha_j, 1 \leq i \leq 3m$, receive $n + m$ affirmations each. The only difference between the actual outcome and the desired judgment set is that $\alpha_0$ and $\neg \beta$ are not contained in the collective judgment set.

Assume that there is an exact cover $C' \subseteq C$ with $\|C'\| = m$. By deleting the judges $L_i$ corresponding to this exact cover, we have that $\alpha_0$ still receives $n + 1$ affirmations, and each $\alpha_i, 1 \leq i \leq 3m$ loses exactly $m - 1$ affirmations and has $n + 1$ affirmations. Also, $\beta$ loses $m$ affirmations and has only $n$ affirmations left. The number of participating judges is now $2n$, so $n + 1$ affirmations are needed to be in the collective judgment set. As desired, the collective outcome is $\{\neg \beta, \alpha_0, \alpha_1, \ldots, \alpha_m\}$.

Conversely, assume that by deleting at most $m$ judges, the collective outcome is $\{\neg \beta, \alpha_0, \alpha_1, \ldots, \alpha_m\}$. Observe that exactly $m$ judges must be deleted, since $\alpha_0$ has only $n + 1$ affirmations and hence can be in the collective judgment set only if at least $m$ voters that do not affirm $\alpha_0$ are deleted. Furthermore, $\beta$ should not be contained in the collective outcome. The number of initial affirmations for $\beta$ is $n + m$. If the number of judges participating is reduced by $m$, $n + 1$ affirmations suffice to be in the collective judgment set. This implies that only judges having $\beta$ in their individual judgment sets can be deleted. In total, this means that all $m$ judges must be deleted from the set $T_2$. If there is one $x_i$ which is not contained in any of sets $C_j$ corresponding to the deleted judges, $\alpha_i$ loses $m$ affirmations, and would no longer be contained in the collective judgment set. But if every $\alpha_i$ is deleted exactly $m - 1$ times, they all have $n + 1$ affirmations and are all contained in the collective judgment set. Hence the sets $C_i$ corresponding to the deleted judges must form an exact cover for the given X3C instance.

In the proof of NP-hardness for *UPQR*$_{1/2}$-Exact Control by Deleting Judges it again is the case that the agenda contains only premises, but for *UPQR*$_{1/2}$-Control by Deleting Judges it remains open whether the problem is still NP-complete if the agenda contains only premises.
Furthermore, there are no natural restrictions that turns one of our control problem to be polynomial time solvable, as this was the case for manipulation and bribery in [19].

Finally, we consider Control by Replacing Judges. In contrast to the problems of adding and deleting judges, the number of judges here is constant, just like in the corresponding manipulation and bribery problems for judgment aggregation (see the work of Endriss et al. [18,19]). 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 6** For each rational quota $q$, $0 < q < 1$, $UPQR_q$ is resistant to Exact Control by Replacing Judges and Control by Replacing Judges.

**Proof.** It can easily be seen that both problems belong to NP. Hardness for NP will be shown by a slight modification of the construction from the proof of Theorem 2. To make the reduction from this proof work for the case of replacing judges, we add $k$ judges to the profile $T$ that potentially will be replaced by those in $S$. These $k$ new judgment sets from $T$ contain all negated formulas from the agenda. To ensure that only those judges can be replaced, we introduce one new formula, $d$, its negation, and $n$ syntactic variations of it into the agenda. This formula is contained in all $q + 1$ individual judgment sets from the initial set $T$. The judges from $S$ all contain $\neg z$. Furthermore, the desired judgment set $J$ also contains $z$. For the exact variant, the same arguments as in the proof of Theorem 2 hold. For Control by Replacing Judges, note that the Hamming distance of the original outcome to $J$ is still $n$, and that replacing a judge that has $z$ in its individual judgment set always results in a Hamming distance that is greater than $n + 1$. Again, only those judges that contain all negated formulas from $\Phi$ can be replaced, and the same arguments as in the proof of Theorem 2 apply.

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 [22]). 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 [11] 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 like the desired outcome. Note that defining destructive variants of control by adding, deleting, or replacing judges would lead to the same definitions as for their constructive counterparts: We have an undesired (possibly partial) judgment set $J \in \mathcal{J}(\Phi)$ 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 definition, as the complement of a partial judgment set $J$ is simply the negation of the formulas in $J$.

**Acknowledgments:** We thank the reviewers for their helpful comments.
References


