Computational Aspects of Manipulation and Control in Judgment Aggregation

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

Abstract

We study computational aspects of various forms of manipulation and control in judgment aggregation, with a focus on the premise-based procedure. For manipulation, we in particular consider incomplete judgment sets and the notions of top-respecting and closeness-respecting preferences introduced by Dietrich and List [DL07]. This complements previous work on the complexity of manipulation in judgment aggregation that focused on Hamming-distance-induced preferences [EGP12, BER11], which we also study here. Regarding control, we introduce the notion of control by bundling judges and show that the premise-based procedure is resistant to it in terms of NP-hardness.

1 Introduction

Judgment Aggregation is the task of aggregating individual judgment sets of possibly interconnected logical propositions (see the surveys by List and Puppe [LP09] and by List [Lis12]). Manipulability and (the game-theoretic concept of) strategy-proofness for the formal framework of judgment aggregation was first introduced by Dietrich and List [DL07]. We focus on their notion of strategy-proofness, since their (non)manipulability condition is not always appropriate in our setting. Manipulation has been studied in a wide variety of settings (voting, mechanism design, game theory, fair division, judgment aggregation, etc.). The incentive of a manipulative attack is always to achieve a "better" result by agents (voters, players, etc.) providing untruthful information. In judgment aggregation, this untruthful information is the manipulator's individual judgment set and the result is the collective outcome of a judgment aggregation procedure. However, it is not at all obvious what a "*better*" result is. To compare two collective judgment sets, a preference over all possible judgment sets would be needed, but such preferences are rarely elicited, and they may be exponentially large in the number of formulas in the agenda (see Section 2 for the notions not defined here). One way to avoid this obstacle, is to derive an order from a given individual judgment set. Based on the notions introduced by Dietrich and List [DL07], we in particular consider incomplete judgment sets and the notions of top-respecting and closeness-respecting preferences. Since most judgment aggregation rules are not strategy-proof, we study the computational complexity of the corresponding decision problems. This complements previous work on the complexity of manipulation in judgment aggregation (initiated by Endriss et al. [EGP12], see also the work of Baumeister et al. [BER11]) that focused on Hamming-distance-induced preferences, which we also study here.

Regarding control in judgment aggregation, we extend previous work by Baumeister et al. [BEER12a, BEER12b] who, inspired by the notion of control in voting (see, e.g., the book chapter [BEH$^+$10] and the references cited therein) studied the complexity of control by adding, deleting, or replacing judges. We introduce a new type of control, *control by bundling judges*, which is well-motivated for judgment aggregation by real-world scenarios and is somewhat reminiscent of control by partitioning voters in voting. We show that one specific judgment aggregation procedure, namely the premise-based procedure, is resistant to this control type in terms of NP-hardness.

This paper is organized as follows. In Section 2, we provide the basic framework of judgment aggregation and define the relevant notions formally. In Section 3, we study the complexity of manipulation in judgment aggregation, and in Section 4 that of the problem modeling control by bundling judges. Finally, Section 5 summarizes our results and presents a number of interesting open problems for future research.

2 Preliminaries

We adopt the framework on judgment aggregation described by Endriss et al. [EGP12] and used also by Baumeister et al. [BER11, BEER12a]. Let $N = \{1, \ldots, n\}$ be a set of judges who have to judge over the formulas in the agenda Φ. We assume that the agenda is a finite, nonempty subset of the set $\mathscr{L}_{PS}$ of all propositional formulas that are built from the boolean constants 1 and 0 and the propositional variables in PS using the boolean connectives $\vee$, $\wedge$, $\rightarrow$, and $\leftrightarrow$. Further, we assume that the agenda does not contain doubly negated formulas. To this end, we denote by $\sim\alpha$ the complement of α: $\sim\alpha = \neg\alpha$ if α is not negated, and $\sim\alpha = \beta$ if $\alpha = \neg\beta$. We also assume that the agenda is closed under complementation (if $\alpha \in \Phi$ then $\sim\alpha \in \Phi$) and under propositional formulas (every literal that occurs in a formula of the agenda is itself contained in the agenda).

An *(individual or collective) judgment set* is a subset of the agenda Φ, where "*individual*" refers to the judgment set of an individual judge and "*collective*" refers to the outcome of a judgment aggregation procedure. A judgment set is said to be *complete* if it contains α or $\sim\alpha$ for all $\alpha \in \Phi$; it is said to be *consistent* if all its formulas can be satisfied by some truth assignment simultaneously; and it is said to be *complement-free* if it does not contain α and $\sim\alpha$ simultaneously for any $\alpha \in \Phi$. Let $\mathscr{J}(\Phi)$ denote the set of all complete and consistent judgment sets.

The well-known doctrinal paradox says that under the majority rule the collective outcome may be inconsistent even if all underlying individual judgment sets are consistent. To avoid this, we focus on the *premise-based procedure* (*PBP*) for an odd number of judges, which—under the assumptions made below—always guarantees a complete and consistent outcome. For a given profile $\mathbf{T} = (J_1, \ldots, J_n) \in \mathscr{J}(\Phi)^n$ of individual judgment sets, the agenda Φ is divided into the set of premises Φ_p and the set of conclusions Φ_c. $PBP(\mathbf{T})$ first aggregates the individual judgment sets on the premises Φ_p using the majority rule and then derives the collective outcome for the conclusions Φ_c. Formally, it is a function $PBP : \mathscr{J}(\Phi)^n \rightarrow 2^{\Phi}$ mapping each given profile of individual judgment sets to the collective judgment set $PBP(\mathbf{T}) = \triangle \cup \{\varphi \in \Phi_c \mid \triangle \models \varphi\}$, where $\triangle = \{\varphi \in \Phi_p \mid \|\{i \mid \varphi \in J_i\}\| > n/2\}$. To guarantee complete and consistent outcomes, we follow Endriss et al. [EGP12] and identify the premises with the set of literals from the agenda. Furthermore, we will extend *PBP* to work also for an even number of judges by assuming that in case of a tie the negated literal will be contained in the collective judgment set.

3 Various Forms of Manipulation in Judgment Aggregation

3.1 Definitions

As mentioned in the introduction, we apply the notions introduced by Dietrich and List [DL07] to study various types of preferences. If for two judgment sets $X, Y \in \mathscr{J}(\Phi)$, X is preferred to Y for a given type of preference T and some individual judgment set J, we write $X \succ_T^J Y$.

Definition 1 *Given some individual judgment set J, we define preferences to be (strictly)*

- unrestricted (U) *if there is no restriction on* $\succ_U^J$;
- top-respecting (TR) *if* $J \succ_{TR}^J X$ *for all* $X \in \mathscr{J}(\Phi) \setminus \{J\}$;
- closeness-respecting (CR) *if for all* $X, Y \in \mathscr{J}(\Phi)$, *we have* $X \succ_{CR}^J Y$ *if* $Y \cap J \subset X \cap J$;
- Hamming-distance-induced (HD) *if for all* $X, Y \in \mathscr{J}(\Phi)$, $X \succ_{HD}^J Y$ *if and only if* $HD(X, J) < HD(Y, J)$, *where the Hamming distance* $HD(X, Y)$ *between two (possibly incomplete) judgment sets X and Y is the number of disagreements on propositions that occur in both judgment sets.*

Note that by allowing equalities the Hamming-distance-induced preference is the only complete relation among those defined above. Intuitively, unrestricted preferences capture the setting where we know nothing about the individual preferences. The slightly more restricted case of top-respecting preferences at least requires the given judgment set to be the most preferred one. This also holds for closeness-respecting preferences, but in addition judgment sets that have additional agreement are preferred. In contrast, the Hamming-distance-induced preferences focus only on the total number of disagreements. Hence, for $X, Y \in \mathscr{J}(\Phi)$, if $X \succ^J_{TR} Y$ then it holds that $X \succ^J_{CR} Y$, and if $X \succ^J_{CR} Y$ then it holds that $X \succ^J_{HD} Y$.

Example 2 *Let a, b, c, and d be variables and let the agenda contain the formulas a, b, c, d, $a \lor b$, $b \lor c$, $a \lor c$, $b \lor d$, and their negations. The individual judgment sets of three judges are shown in Table 1, where a* 0 *indicates that the negation of the formula is in the judgment set, and a* 1 *indicates that the formula itself is contained in the judgment set.*

	a	b	c	d		$a \lor b$	$b \lor c$	$a \lor c$	$b \lor d$
Judge 1	1	1	0	0		1	1	1	1
Judge 2	0	0	0	0		0	0	0	0
Judge 3	1	0	1	1		1	1	1	1
PBP	1	0	0	0	$\Rightarrow$	1	0	1	0

Table 1: Applying the premise-based judgment aggregation procedure

The result according to the premise-based procedure is also given in the table. Now assume that the third judge is trying to manipulate and reports the untruthful individual judgment set $\{a, b, c, d\}$ and the corresponding conclusions. Then the collective outcome equals the individual judgment set of the first judge.

- *If the manipulator has unrestricted preferences, we do not know whether she prefers this new outcome or not.*
- *If she has closeness-respecting preferences, we again do not know whether she prefers the new outcome, since the agreement on $\neg b$ is no longer given. However, if she is interested only in the conclusions, then she does prefer the new outcome, since the agreement on $a \lor b$ and $a \lor c$ is preserved and there are the two additional agreements on $b \lor c$ and $b \lor d$.*
- *The same holds for top-respecting preferences: If the manipulator is interested in the whole collective judgment set, we do not know which outcome is better for her, but restricted to the conclusions the new outcome equals her initial individual judgment set and thus is preferred to all other outcomes.*
- *If the manipulator has Hamming-distance-induced preferences, we know that the new outcome is preferred to the old one, since before the manipulation the Hamming distance was* 4*, but now it is only* 3.

Just as Dietrich and List [DL07], we study settings where the desired judgment set is incomplete, to also capture their "reason-oriented" and "outcome-oriented" preferences. However, we will not generally restrict the desired judgment set to the premises or the conclusions; rather, we allow arbitrary incomplete desired judgment sets (which still must have a consistent extension to the whole agenda). In this case, we restrict the preferences to the formulas that occur in the desired judgment set. Since we want to compare two preferences with each other, but most of the induced preferences will be incomplete, we distinguish the cases where the relation between them is known or unknown. Let $\mathscr{T} \in \{\text{U, TR, CR}\}$ be a type of induced preferences. A judge *necessarily prefers X to Y for* *type* $\mathscr{T}$ if $X >^J_{\mathscr{T}} Y$ for all complete extensions of $\succ^J_{\mathscr{T}}$. A judge *possibly prefers X to Y for type* $\mathscr{T}$ if $X >^J_{\mathscr{T}} Y$ for some complete extension of $\succ^J_{\mathscr{T}}$.

Definition 3 *A judgment aggregation rule F is* necessarily/possibly strategy-proof with respect to induced preferences of type $\mathscr{T} \in \{\text{U, TR, CR}\}$ *if for all profiles* $(J_1,\dots,J_n)$ *and each i,* $1 \le i \le n$, *agent i necessarily/possibly prefers the outcome* $F(J_1,\dots,J_n)$ *to the outcome* $F(J_1,\dots,J_{i-1},J_i^*,J_{i+1},\dots,J_n)$ *(with respect to preferences of type* $\mathscr{T}$ *and the individual judgment set* J_i*) for any* $J_i^* \in \mathscr{J}(\Phi)$.[1]

These notions are remotely inspired by "possible" vs "necessary winner" in voting theory due to Konczak and Lang [KL05] (see also the work of Xia and Conitzer [XC11]), and by "possible" vs "necessary envy-freeness" in fair division due to Bouveret et al. [BEL10] (see also the papers by Brams et al. [BEF04, BK05]). The stronger notion of *necessary strategy-proofness* corresponds to the "strategy-proofness" condition defined by Dietrich and List [DL07], whereas the weaker notion of *possible strategy-proofness* is introduced here. Note that since the Hamming-distance-induced preferences are a complete relation, we simply say that *F* is *strategy-proof with respect to Hamming-distance-induced preferences* if for each individual judge the actual outcome is at least as good as all outcomes obtained by reporting a different individual judgment set.

The result of Dietrich and List [DL07] says that an aggregation rule that satisfies the "universal domain" condition is necessarily strategy-proof with respect to non-strict closeness-respecting preferences if and only if it is independent and monotonic. *Universal domain* is satisfied if the domain of the aggregation function is the set of all possible profiles from $\mathscr{J}(\Phi)^n$, which obviously is true for *PBP*. *Independence* means that the collective decision on each proposition only relies on the individual judgments of this proposition. Since *PBP* derives the outcome for the conclusions from the outcome of the premises, it is not independent and hence not necessarily strategy-proof with respect to non-strict closeness-respecting preferences. An aggregation function is *monotonic* if additional support for some proposition that is currently accepted may never result in a non-acceptance for this formula, provided everything else remains unchanged. In the case where the agenda contains solely premises, *PBP* is independent and monotonic, and hence necessarily strategy-proof also for the case of strict closeness-respecting preferences.

Endriss et al. [EGP12] initiated the study of the complexity of manipulation in judgment aggregation. Their work (and also the follow-up work of Baumeister et al. [BER11]) focuses only on preferences induced by the Hamming distance to the complete desired judgment set of the manipulator. We extend this study to the setting where the manipulator may be interested only in parts of the agenda, so her desired judgment set can be an incomplete subset of her true judgment set. For a given type $\mathscr{T} \in \{\text{U, TR, CR}\}$ of preference induced by the desired judgment set $J \subseteq J_n$ (i.e., judge *n* is the manipulator), we define the manipulation problem $\mathscr{T}$-NECESSARY-MANIPULATION as follows:

$\mathscr{T}$-NECESSARY-MANIPULATION	
Given:	An agenda Φ, a profile $\mathbf{T} = (J_1,\dots,J_n) \in \mathscr{J}(\Phi)^n$, and the manipulator's desired consistent (possibly incomplete) judgment set $J \subseteq J_n$.
Question:	Does there exist a judgment set $J^* \in \mathscr{J}(\Phi)$ such that $PBP(J_1,\dots,J_{n-1},J^*)\vert_J >^J_{\mathscr{T}} PBP(J_1,\dots,J_n)\vert_J$ for all extensions $>^J_{\mathscr{T}}$ that are consistent with $\succ^J_{\mathscr{T}}$?

Here, $PBP(J_1,\dots,J_n)|_J$ denotes the restriction of $PBP(J_1,\dots,J_n)$ to the formulas that occur, negated or not, in the desired judgment set *J*. In $\mathscr{T}$-POSSIBLE-MANIPULATION, we ask whether $PBP(J_1,\dots,J_{n-1},J^*)|_J >^J_{\mathscr{T}} PBP(J_1,\dots,J_n)|_J$ for some extension $>^J_{\mathscr{T}}$ that is consistent with $\succ^J_{\mathscr{T}}$. In the case of Hamming-distance-induced preferences we will simply say HD-MANIPULATION.

[1] This definition applies to complete desired judgment sets J_i only. More generally, the definition can easily be extended to incomplete desired judgment sets $J \subseteq J_i$ as well.

Furthermore, we introduce and study the exact variant, EXACT-MANIPULATION, where the manipulator seeks to achieve not only a better, but a *best* outcome for a given subset of her desired judgment set. Here, the question is whether there is some judgment set $J^* \in \mathscr{J}(\Phi)$ such that $J \subseteq PBP(J_1, \dots, J_{n-1}, J^*)$.

We assume the reader is familiar with complexity classes such as P and NP and the notion of NP-completeness (w.r.t. the polynomial-time many-one reducibility, $\leq_m^p$).

3.2 Results

Theorem 4 EXACT-MANIPULATION *is* NP-*complete, even for only three judges.*

Proof. The proof is by a reduction from the NP-complete satisfiability problem. Let φ be a given formula in conjunctive normal form, where the clauses are built from the set $A = \{\alpha_1, \dots, \alpha_m\}$ of variables. The question is whether there is a satisfying assignment for this formula. Without loss of generality, we may assume that neither setting all variables to true, nor setting all variables to false is a satisfying assignment for φ. Now construct an agenda Φ that consists of the variables in A and their negations, an additional variable β and its negation, and the formula $\varphi \vee \beta$ and its negation. The profile $\mathbf{T}$ consists of three judges. The individual judgment set of the first one contains A and $\neg\beta$ and the individual judgment set of the second one contains $\neg\alpha_i$ for each i, $1 \leq i \leq m$, and $\neg\beta$. The third judge is the manipulative one and his individual judgment set contains A and β. The desired outcome he tries to achieve exactly consists of only the conclusion $\varphi \vee \beta$. It holds that $PBP(\mathbf{T}) = A \cup \{\neg\beta\} \cup \{\neg(\varphi \vee \beta)\}$. Note also that the third judge is decisive for every formula in A, and that independently of the individual judgment set of the manipulator, β is never contained in the collective judgment set. Hence, the only way to obtain the conclusion $\varphi \vee \beta$ in the collective outcome is to evaluate the formula φ to true. This implies that there is a satisfying assignment for φ if and only if the individual judgment set of the third judge can be modified such that $\varphi \vee \beta$ is contained in the collective outcome. ❑

Theorem 5 *1.* EXACT-MANIPULATION $\leq_m^p$ $\mathscr{T}$-NECESSARY-MANIPULATION *for each type* $\mathscr{T} \in \{\text{TR,CR}\}$.

2. EXACT-MANIPULATION $\leq_m^p$ $\mathscr{T}$-POSSIBLE-MANIPULATION *for each type* $\mathscr{T} \in \{\text{U,TR,CR}\}$.

3. EXACT-MANIPULATION $\leq_m^p$ HD-MANIPULATION.

Proof. For the exact problem, we have an agenda Φ, some profile $\mathbf{T} = (J_1, \dots, J_n)$, and some desired judgment set $J = \{\alpha_1, \dots, \alpha_m\} \subseteq J_n$, and we are looking for a modified judgment set J_n^* such that $J \subseteq PBP(J_1, \dots, J_{n-1}, J_n^*)$. In the trivial case that $J \subseteq PBP(\mathbf{T})$, $J_n^* = J_n$ obviously fulfills the requirement, so we can construct an arbitrary yes-instance for the corresponding manipulation problem. We will prove all three assertions via the same reduction, but using different arguments.

Assume that $J \setminus PBP(\mathbf{T}) \neq \emptyset$ and consider the following problem. Fix some $\mathscr{T} \in \{\text{TR,CR,HD}\}$, let the agenda Φ' be the union of Φ, the formula $\varphi = \alpha_1 \wedge \dots \wedge \alpha_m$, and its negation. Let $\mathbf{T}' \in \mathscr{J}(\Phi')^n$ be the consistent extensions of $\mathbf{T}$. In particular, $J_n' = J_n \cup \varphi$. Let the desired judgment set be $J' = \varphi$, and we are looking for a modified judgment set $J_n'^*$ such that for all extensions $>_{\mathscr{T}}^{J'}$ of $\succ_{\mathscr{T}}^{J'}$, we have $PBP(J_1', \dots, J_{n-1}', J_n'^*)|_{J'} >_{\mathscr{T}}^{J'} PBP(J_1', \dots, J_{n-1}', J_n')|_{J'}$. Since J' consists of the single formula φ, there are only two different collective outcomes when restricted to J'. Since $\varphi \subseteq J_n$, it obviously holds that $\varphi \succ_{\mathscr{T}}^{J'} \neg\varphi$ for all $\mathscr{T} \in \{\text{TR,CR,HD}\}$, and since in this case $\succ_{\mathscr{T}}^{J'}$ is complete, there is no difference between the notions of necessary and possible preference. In the case of unrestricted preferences and the possible manipulation problem, we ask whether there is some different outcome, since they all may be possibly preferred. Since there is some J_n^* with $J \subseteq PBP(J_1, \dots, J_{n-1}, J_n^*)$ if and only if there is some $J_n'^*$ with $\varphi \subseteq PBP(J_1', \dots, J_{n-1}', J_n'^*)$, the reduction works in all cases. ❑

This reduction requires a partial desired judgment set for $\mathscr{T}$-NECESSARY-MANIPULATION, $\mathscr{T}$-POSSIBLE-MANIPULATION, and HD-MANIPULATION; together with Theorem 4, this implies NP-completeness of HD-MANIPULATION, $\mathscr{T}$-NECESSARY-MANIPULATION for $\mathscr{T} \in \{\mathrm{TR,CR}\}$, and $\mathscr{T}$-POSSIBLE-MANIPULATION for $\mathscr{T} \in \{\mathrm{U,TR,CR}\}$ whenever the desired judgment set of the manipulator is incomplete. Alternatively, the reduction given by Endriss et al. [EGP12] in fact shows that HD-MANIPULATION remains NP-complete even if the desired judgment set of the manipulator is complete.

Proposition 6 *For $\mathscr{T} \in \{\mathrm{U,TR}\}$, $\mathscr{T}$-*POSSIBLE-MANIPULATION *can be solved in polynomial time if the desired judgment set of the manipulator is complete.*

Proof. This result holds, since a U-POSSIBLE-MANIPULATION instance is positive exactly if there is some premise from the desired judgment set for which the manipulator is decisive, i.e., the collective outcome depends on the decision of the manipulator. For a TR-POSSIBLE-MANIPULATION instance to be positive, it must additionally be required that the desired judgment set is not the actual outcome. ❑

Proposition 7 *PBP is possibly strategy-proof when closeness-respecting preferences are assumed and the desired judgment set of the manipulator is complete.*

Proof. If closeness-respecting preferences are assumed, a judgment set that is necessarily preferred to the actual collective outcome must preserve all agreements between the desired judgment set and the actual outcome. Now consider a premise α that is contained in the collective judgment set, but $\sim\alpha$ is contained in the desired judgment set. It can obviously never be the case that a switch from the manipulator to α causes $\sim\alpha$ to be in the collective judgment set. Hence there can be no additional agreement among the premises. Since the desired judgment set is complete and the outcome for the conclusions depends solely on the outcome of the premises, *PBP* is possibly strategy-proof in this case. ❑

Note that this does not contradict the results of Dietrich and List [DL07], since they impose different conditions on nonmanipulability and non-strict preferences.

4 Control by Bundling Judges

Previous work on control in judgment aggregation (see [BEER12a, BEER12b]) considered the problems of control by adding, deleting, or replacing judges. Although adding and deleting judges is inspired by the corresponding control problems in voting, explicit examples for such control actions in judgment aggregation are given, and the third type, control by replacing judges, was motivated by real-world examples from international arbitration. We here introduce another type of control motivated by real-world scenarios, *control by bundling judges*, which is remotely akin to control by partitioning voters in voting. A prominent natural example for control by bundling judges can be found in European legislation. Certain European legislative acts, such as Directives, give considerable freedom to Member States regarding the concrete implementation of these acts. Yet, in some cases uniform implementation is crucial, so the basic act confers implementing powers on the European Commission or the Council of the European Union to adopt the required implementing acts.[2] The exercise of implementing powers through the Commission and Council is controlled by the member states through so-called comitology committees in accordance with previously specified rules.[3] The committees are set up by the basic act in question.[4] Some of these committees are

[2]Article 291 of the Treaty on the Functioning of the European Union.

[3]Regulation (EU) No 182/2011 of the European Parliament and of the Council of 16 February 2011 laying down the rules and general principles concerning mechanisms for control by Member States of the Commission's exercise of implementing powers (Implementing Acts Regulation).

[4]Recital 6 of the Preamble of Implementing Acts Regulation.

concerned with such a broad range of issues that they are divided into subcommittees, each of which is dealing with different issues. When preparing implementing acts covering several issues, each subcommittee votes on the issues assigned to it, and the implementing act is shaped according to the decisions of the different subcommittees.[5]

4.1 Definitions

The problem EXACT CONTROL BY ADDING JUDGES asks, given an agenda Φ, two complete profiles $\mathbf{T} \in \mathscr{J}(\Phi)^n$ and $\mathbf{S} \in \mathscr{J}(\Phi)^{\|S\|}$, a positive integer k, and a desired judgment set J (which may be incomplete, i.e., $J \subseteq J'$ for some $J' \in \mathscr{J}(\Phi)$), whether there is a subset $\mathbf{S}' \subseteq \mathbf{S}$ of the potential new judges of size at most k, which can be added such that $J \subseteq PBP(\mathbf{T} \cup \mathbf{S}')$. The variant of this problem asking for a preferred outcome when Hamming-distance-induced preferences are assumed will be denoted by CONTROL BY ADDING JUDGES. The problem EXACT CONTROL BY DELETING JUDGES asks, given an agenda Φ, a complete profile $\mathbf{T} \in \mathscr{J}(\Phi)^n$, a positive integer k, and a desired (possibly incomplete) judgment set J, whether it is possible to delete at most k judges from $\mathbf{T}$ such that J is a collective outcome, and the corresponding problem CONTROL BY DELETING JUDGES asks, for the same input, whether there is a preferred outcome when Hamming-distance-induced preferences are assumed.

When analyzing the complexity of these problems, Baumeister et al. [BEER12a, BEER12b] follow the terminology introduced in [BTT92] for control problems in voting. For a given judgment aggregation procedure F (such as PBP) and a given control type $\mathscr{C}$ (such as those defined above), if it is never possible to successfully exert this type of control, F is said to be *immune to control by* $\mathscr{C}$; otherwise, F is said to be *susceptible to control by* $\mathscr{C}$. If F is susceptible to control by $\mathscr{C}$, then F is said to be *vulnerable to control by* $\mathscr{C}$ whenever the corresponding decision problem is in P, and F is said to be *resistant to control by* $\mathscr{C}$ whenever the corresponding decision problem is NP-hard. Baumeister et al. [BEER12a, BEER12b] have shown that the premise-based procedure is resistant to control by adding judges, to control by deleting judges, and to control by replacing judges (which in some sense combines control by deleting with control by adding judges) when preferences are assumed to be Hamming-distance-induced and in the exact variant. We will study the new problem of CONTROL BY BUNDLING JUDGES also in these two variants for the premise-based procedure. The formal definition for the Hamming-distance-induced version is as follows. In the problem definition below, we will use the notation $\Delta = \bigcup_{1 \le i \le k} PBP(\mathbf{T}|_{\Phi_p^i, N_i})$, where $PBP(\mathbf{T}|_{\Phi_p^i, N_i})$ is the collective judgment set obtained by restricting the agenda to Φ_p^i and the set of judges to $N_i \subseteq N$.

CONTROL BY BUNDLING JUDGES	
Given:	An agenda Φ, where the premises are partitioned into k subsets $\Phi_p^1, \dots, \Phi_p^k$, a complete profile $\mathbf{T} \in \mathscr{J}(\Phi)^n$, and a consistent and complement-free judgment set J (not necessarily complete).
Question:	Is there a partition $N_1, \dots, N_k$ of the n judges such that $H(J, \Delta \cup \{\varphi \in \Phi_c \mid \Delta \models \varphi\}) < H(J, PBP(\mathbf{T}))$?

The problem EXACT CONTROL BY BUNDLING JUDGES asks, for the same input, whether there is a partition $N_1, \dots, N_k$ of the n judges such that $J \subseteq \Delta \cup \{\varphi \in \Phi_c \mid \Delta \models \varphi\}$.

4.2 Results

We show that the exact variant and the Hamming-distance-induced variant defined above are closely related. In fact, the proof of Lemma 8 below applies to all the control problems in judgment agree-

[5]One example is the Customs Code Committee, see Articles 1 (1) and 5 (7) (8) of the Rules of procedure for the Customs Code Committee.

ment studied in the literature (control by adding, deleting, replacing, or bundling judges; for the formal definition of control by replacing judges, see [BEER12a, BEER12b]).

Lemma 8 *Let $\mathscr{C}$ be a control type.* EXACT CONTROL BY $\mathscr{C}$ $\leq_m^p$*-reduces to* CONTROL BY $\mathscr{C}$.

Proof. In the exact problem variant, we have an agenda Φ, some profile $\mathbf{T}$, and some desired judgment set $J = \{\alpha_1, \dots, \alpha_m\}$, and we are looking for a modified profile $\mathbf{U}$ such that $PBP(\mathbf{U}) = J$. Now consider the following problem. Let the agenda Φ' be the union of Φ, the formula $\varphi = \alpha_1 \wedge \cdots \wedge \alpha_m$, and its negation. Let $\mathbf{T}'$ and $\mathbf{U}'$ (both in $\mathscr{J}(\Phi')$) be the consistent extensions of, respectively, $\mathbf{T}$ and $\mathbf{U}$, and let $J' = \varphi$. In the trivial case that $PBP(\mathbf{T}) = J$, we have $H(J, PBP(\mathbf{T}')) = 0$. In the nontrivial case that $PBP(\mathbf{T}) \neq J$, we have $H(J, PBP(\mathbf{T}')) = 1$. This implies $H(J, PBP(\mathbf{U}')) < H(J, PBP(\mathbf{T}'))$ if and only if $PBP(\mathbf{U}') = J'$, and hence $H(J, PBP(\mathbf{U}')) < H(J, PBP(\mathbf{T}'))$ is equivalent to $PBP(\mathbf{U}) = J$. ❑

Note that the above proof requires the desired judgment set of the Hamming-distance-induced variant to be incomplete. Note further that Lemma 8 implies that NP-hardness of a Hamming-distance-induced variant is inherited from NP-hardness of the corresponding exact problem variant.

The problem CONTROL BY BUNDLING JUDGES is somewhat similar to the problem of CONTROL BY DELETING JUDGES. We will exploit this in the proof of the following theorem.

Theorem 9 *PBP is resistant to* EXACT CONTROL BY BUNDLING JUDGES *and to* CONTROL BY BUNDLING JUDGES.

Proof. The proof will be by a reduction from the related problem EXACT CONTROL BY DELETING JUDGES. Given an agenda $\Phi = \Phi_p \cup \Phi_c$, a complete profile $\mathbf{T} \in \mathscr{J}(\Phi)^n$, and a positive integer k that is the bound on the number of judges that may be deleted. We assume that the individual judgment set of the manipulator is J_n, and $J \subseteq J_n$ is the desired judgment set. Now, we construct an instance of EXACT CONTROL BY BUNDLING JUDGES, resistance for CONTROL BY BUNDLING JUDGES then follows from Lemma 8. Without loss of generality, we assume that $n \geq k+2$. The agenda is $\Phi' = \Phi \cup \{\alpha, \neg\alpha\}$, and is divided into two subsets. The first one consists of Φ_p, and the second one is $\{\alpha, \neg\alpha\}$. The profile $\mathbf{S} \in \mathscr{J}(\Phi')^{n+k+1}$ contains the individual judgment sets from $\mathbf{T}$, each extended by $\neg\alpha$. Furthermore, there are $k+1$ new individual judgment sets that each contain $\varphi \in \Phi_p$ if and only if $\sim\varphi \in J$, they each contain α, and the conclusions are evaluated accordingly. These $k+1$ new judges will be denoted by N'. The desired judgment set is $J' = J \cup \{\alpha\}$. Now we show that it is possible to obtain the desired judgment set J by deleting at most k judges from $\mathbf{T}$ if and only if the judges from $\mathbf{S}$ can be bundled into two groups such that the desired outcome is J'.

For the direction from left to right, assume that there is a subset $\mathbf{T}' \subseteq \mathbf{T}$, $\|\mathbf{T}'\| \leq k$, such that $PBP(\mathbf{T} \setminus \mathbf{T}') = J$. Then the judges can be bundled as follows. The $k+1$ new judges and the judges corresponding to $\mathbf{T}'$ decide over α. Then obviously α is contained in the collective outcome, hence the constructed instance is a positive one for EXACT CONTROL BY BUNDLING JUDGES.

For the direction from right to left, assume that the judges can be bundled into N_1 and N_2 such that the collective outcome is J'. Hence, it holds that $PBP(\mathbf{S}|_{\Phi, N_1}) = J$. We will show that $\|N_2 \setminus N'\| \leq k$ and $PBP(\mathbf{S}|_{\Phi, N_1 \setminus N'}) = J$. Since α is contained in the collective judgment set and since there are only $k+1$ judges having α in their individual judgment set, at most k of the initial judges can be in N_2. Due to the premise-based procedure, it is enough to show that $PBP(\mathbf{S}|_{\Phi_p, N_1}) = PBP(\mathbf{S}|_{\Phi_p, N_1 \setminus N'})$. This holds trivially, since for all judges from N_1 it holds that $\varphi \in \Phi_p$ is contained in the individual judgment set if and only if $\sim\varphi \in J$. ❑

5 Conclusions and Future Work

To conclude, we investigated various forms of manipulation in judgment aggregation that originate from different assumptions on the incentives and the type of preferences of the manipulator. Our

$\mathscr{T}$	U	TR	CR	HD	Exact
$\mathscr{T}$-Possible-Manipulation for incomplete DJS	NP-c	NP-c	NP-c	NP-c	NP-c
$\mathscr{T}$-Necessary-Manipulation for incomplete DJS	?	NP-c	NP-c		
$\mathscr{T}$-Possible-Manipulation for complete DJS	in P	in P	?	NP-c [EGP12]	strategy-proof
$\mathscr{T}$-Necessary-Manipulation for complete DJS	?	?	possibly strategy-proof		

Table 2: Overview of results for various manipulation problems

results show that whether one considers a judgment aggregation rule to be (necessarily or possibly) strategy-proof or manipulable crucially depends on the given setting. Table 2 summarizes our results for the various manipulation problems. The last two columns consider the Hamming-distance-induced preferences and the exact variant; note that there is no distinction between the possible and necessary manipulation problem for these preference types. The first two rows concern the general problem with an incomplete desired judgment set (abbreviated by DJS in the table), whereas the last two rows show the results for the restricted problem where the desired judgment set is required to be complete. We abbreviate "NP-complete" by "NP-c." All results stated in the table are new to this paper, except for the one for Hamming-distance-induced preferences with a complete desired judgment set, which is due to Endriss et al. [EGP12].

We propose to launch a systematic study of the computational aspects of manipulation in judgment aggregation for complete and incomplete desired judgment sets, in particular by solving the open problems indicated by question marks in Table 2. Furthermore, the concepts studied here for manipulation can be transferred to other forms of interference as well, such as bribery and control [BEER12a, BEER12b, BER11]. Regarding the latter, we have proposed a new control type, *control by bundling judges*, which in some sense corresponds to control by partitioning voters in voting. We have argued why this control type models a natural real-world scenario and showed that the premise-based procedure is resistant to it. It would be interesting to complement such worst-case complexity results by typical-case studies, or with respect to parameterized complexity, as has been done successfully in voting (see, e.g, the papers by Betzler and Uhlmann [BU09], Erdélyi et al. [EFRS12], Liu et al. [LFZL09, LZ10], and Rothe and Schend [RS12a, RS12b, RS12c] for control and the papers by Isaksson et al. [IKM12], Friedgut et al. [FKKN11], Walsh [Wal09, Wal10, Wal11], and Xia and Conitzer [XC08a, XC08b] for manipulation).

6 Acknowledgments

This work was supported in part by DFG grants RO 1202/15-1 and ER 738/1-1 and NRF (Singapore) grant NRF-RF 2009-08. Work done in part while the second and third author were visiting the University of Rochester.

References

[BEER12a] D. Baumeister, G. Erdélyi, O. Erdélyi, and J. Rothe. Bribery and control in judgment aggregation. In F. Brandt and P. Faliszewski, editors, *Proceedings of the 4th Interna-*

tional Workshop on Computational Social Choice, pages 37–48. AGH University of Science and Technology, Kraków, Poland, September 2012.

[BEER12b] D. Baumeister, G. Erdélyi, O. Erdélyi, and J. Rothe. Control in judgment aggregation. In *Proceedings of the 6th European Starting AI Researcher Symposium*, pages 23–34. IOS Press, August 2012.

[BEF04] S. Brams, P. Edelman, and P. Fishburn. Fair division of indivisible items. *Theory and Decision*, 5(2):147–180, 2004.

[BEH$^+$10] 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.

[BEL10] S. Bouveret, U. Endriss, and J. Lang. Fair division under ordinal preferences: Computing envy-free allocations of indivisible goods. In *Proceedings of the 19th European Conference on Artificial Intelligence*, pages 387–392. IOS Press, August 2010.

[BER11] D. Baumeister, G. Erdélyi, and J. Rothe. How hard is it to bribe the judges? A study of the complexity of bribery in judgment aggregation. In *Proceedings of the 2nd International Conference on Algorithmic Decision Theory*, pages 1–15. Springer-Verlag *Lecture Notes in Artificial Intelligence #6992*, October 2011.

[BK05] S. Brams and D. King. Efiicient fair division—help the worst off or avoid envy? *Rationality and Society*, 17(4):387–421, 2005.

[BTT92] 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.

[BU09] N. Betzler and J. Uhlmann. Parameterized complexity of candidate control in elections and related digraph problems. *Theoretical Computer Science*, 410(52):5425–5442, 2009.

[DL07] F. Dietrich and C. List. Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(3):269–300, 2007.

[EFRS12] G. Erdélyi, M. Fellows, J. Rothe, and L. Schend. Control complexity in Bucklin and fallback voting. Technical Report arXiv:1103.2230 [cs.CC], Computing Research Repository, arXiv.org/corr/, March 2012. March, 2011. Revised August, 2012.

[EGP12] U. Endriss, U. Grandi, and D. Porello. Complexity of judgment aggregation. *Journal of Artificial Intelligence Research*, 45:481–514, 2012.

[FKKN11] E. Friedgut, N. Keller, G. Kalai, and N. Nisan. A quantitative version of the Gibbard–Satterthwaite theorem for three alternatives. *SIAM Journal on Computing*, 40(3):934–952, 2011.

[IKM12] Marcus Isaksson, Guy Kindler, and Elchan Mossel. The geometry of manipulation – a quantitative proof of the Gibbard-Satterthwaite theorem. *Combinatorica*, 32(2):221–250, 2012.

[KL05] K. Konczak and J. Lang. Voting procedures with incomplete preferences. In *Proceedings of the Multidisciplinary IJCAI-05 Workshop on Advances in Preference Handling*, pages 124–129, July/August 2005.

[LFZL09] H. Liu, H. Feng, D. Zhu, and J. Luan. Parameterized computational complexity of control problems in voting systems. *Theoretical Computer Science*, 410(27–29):2746–2753, 2009.

[Lis12] C. List. The theory of judgment aggregation: An introductory review. *Synthese*, 187(1):179–207, 2012.

[LP09] 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.

[LZ10] H. Liu and D. Zhu. Parameterized complexity of control problems in maximin election. *Information Processing Letters*, 110(10):383–388, 2010.

[RS12a] J. Rothe and L. Schend. Control complexity in Bucklin, fallback, and plurality voting: An experimental approach. In *Proceedings of the 11th International Symposium on Experimental Algorithms*, pages 356–368. Springer-Verlag *Lecture Notes in Computer Science #7276*, June 2012.

[RS12b] J. Rothe and L. Schend. Control complexity in Bucklin, fallback, and plurality voting: An experimental approach. Technical Report arXiv:1203.3967 [cs.GT], Computing Research Repository, arXiv.org/corr/, March 2012. March, 2012. Revised August, 2012.

[RS12c] 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.

[Wal09] T. Walsh. Where are the really hard manipulation problems? The phase transition in manipulating the veto rule. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, pages 324–329. IJCAI, July 2009.

[Wal10] T. Walsh. An empirical study of the manipulability of single transferable voting. In *Proceedings of the 19th European Conference on Artificial Intelligence*, pages 257–262. IOS Press, August 2010.

[Wal11] T. Walsh. Is computational complexity a barrier to manipulation? *Annals of Mathematics and Artificial Intelligence*, 62(1–2):7–26, 2011.

[XC08a] L. Xia and V. Conitzer. Generalized scoring rules and the frequency of coalitional manipulability. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 109–118. ACM Press, June 2008.

[XC08b] L. Xia and V. Conitzer. A sufficient condition for voting rules to be frequently manipulable. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 99–108. ACM Press, June 2008.

[XC11] L. Xia and V. Conitzer. Determining possible and necessary winners given partial orders. *Journal of Artificial Intelligence Research*, 41:25–67, 2011.

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

Gábor Erdélyi
School of Economic Disciplines
University of 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