# Egalitarian Judgment Aggregation

Sirin Botan
University of Amsterdam, The Netherlands
sirin.botan@uva.nl

Ronald de Haan
University of Amsterdam, The Netherlands
me@ronalddehaan.eu

Marija Slavkovik
University of Bergen, Norway
marija.slavkovik@uib.no

Zoi Terzopoulou
University of Amsterdam, The Netherlands
z.terzopoulou@uva.nl

## ABSTRACT

Egalitarian considerations play a central role in many areas of social choice theory. Applications of egalitarian principles range from ensuring everyone gets an equal share of a cake when deciding how to divide it, to guaranteeing balance with respect to gender or ethnicity in committee elections. Yet, the egalitarian approach has received little attention in judgment aggregation—a powerful framework for aggregating logically interconnected issues. We make the first steps towards filling that gap. We introduce axioms capturing two classical interpretations of egalitarianism in judgment aggregation and situate these within the context of existing axioms in the pertinent framework of belief merging. We then explore the relationship between these axioms and several notions of strategyproofness from social choice theory at large. Finally, a novel egalitarian judgment aggregation rule stems from our analysis; we present complexity results concerning both outcome determination and strategic manipulation for that rule.

## KEYWORDS

Social Choice Theory, Judgment Aggregation, Egalitarianism, Strategic Manipulation, Computational Complexity

**ACM Reference Format:**
Sirin Botan, Ronald de Haan, Marija Slavkovik, and Zoi Terzopoulou. 2021. Egalitarian Judgment Aggregation. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 9 pages.

## 1 INTRODUCTION

Judgment aggregation is an area of social choice theory concerned with turning the individual binary judgments of a group of agents over logically related issues into a collective judgment [23]. Being a flexible and widely applicable framework, judgment aggregation provides the foundations for collective decision making settings in various disciplines, like philosophy, economics, legal theory, and artificial intelligence [37]. The purpose of judgment aggregation methods (*rules*) is to find those collective judgments that better represent the group as a whole. Following the utilitarian approach in social choice, an "ideal" such collective judgment has traditionally been considered the will of the majority. In this paper we challenge this perspective, introducing a more egalitarian point of view.

In economic theory, utilitarian approaches are often contrasted with egalitarian ones [51]. In the context of judgment aggregation, an egalitarian rule must take into account whether the collective outcome achieves equally distributed satisfaction among agents and ensure that agents enjoy equal consideration. A rapidly growing application domain of egalitarian judgment aggregation (that also concerns multiagent systems with practical implications like in the construction of self-driving cars) is the aggregation of moral choices [16], where utilitarian approaches do not always offer appropriate solutions [4, 53]. One of the drawbacks of majoritarianism is that a strong enough majority can cancel out the views of a minority, which is questionable in several occasions.

For example, suppose that the president of a student union has secured some budget for the decoration of the union's office and she asks her colleagues for their opinions on which paintings to buy (perhaps imposing some constraints on the combinations of paintings that can be simultaneously selected, due to clashes on style). If the members of the union largely consist of pop-art enthusiasts that the president tries to satisfy, then a few members with diverting taste will find themselves in an office that they detest; an arguably more viable strategy would be to ensure that—as much as possible—no-one is strongly dissatisfied. But then, consider a similar situation in which a kindergarten teacher needs to decide what toys to complement the existing playground with. In that case, the teacher's goal is to select toys that equally (dis)satisfy all kids involved, so that no extra tension is created due to envy, which the teacher will have to resolve—if the kids disagree a lot, then the teacher may end up choosing toys that none of them really likes.

In order to formally capture scenarios like the above, this paper introduces two fundamental properties (also known as *axioms*) of egalitarianism to judgment aggregation, inspired by the theory of justice. The first captures the idea behind the so-called *veil of ignorance* of Rawls [56], while the second speaks about how happy agents are with the collective outcome relative to each other.

Our axioms closely mirror properties in other areas of social choice theory. In *belief merging*, egalitarian axioms and merging operators have been studied by Everaere et al. [28]. The nature of their axioms is in line with the interpretation of egalitarianism in this paper, although the two main properties they study are logically weaker than ours, as we further discuss in Section 3.1. In *resource allocation*, fairness has been interpreted both as maximising the share of the worst off agent [12] as well as eliminating envy between agents [31]. In *multiwinner elections*, egalitarianism is present in diversity [22] and in proportional representation [2, 20] notions.

Unfortunately, egalitarian considerations often come at a cost. A central concern in many areas of social choice theory, of which judgement aggregation does not constitute an exception, is that agents may have incentives to *manipulate*, i.e., to misrepresent their

*Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.*

judgments aiming for a more preferred outcome [18]. Frequently, it is impossible to simultaneously be fair and avoid strategic manipulation. For both variants of fairness in resource allocation, rules satisfying them usually are susceptible to strategic manipulation [1, 10, 14, 50]. The same type of results have recently been obtained for multiwinner elections [45, 54]. It is not easy to be egalitarian while disincentivising agents from taking advantage of it.

Inspired by notions of manipulation stemming from voting theory, we explore how our egalitarian axioms affect the agents' strategic behaviour within judgment aggregation. Our most important result in this vein is showing that the two properties of egalitarianism defined in this paper clearly differ in terms of strategyproofness.

Our axioms give rise to two concrete egalitarian rules—one that has been previously studied, and one that is new to the literature. For the latter, we are interested in exploring how computationally complex its use is in the worst-case scenario. This kind of question, first addressed by Endriss et al. [27], is regularly asked in the literature of judgment aggregation [5, 25, 47]. As Endriss et al. [26] wrote recently, the problem of determining the collective outcome of a given judgment aggregation rule is "the most fundamental algorithmic challenge in this context".

The remainder of this paper is organised as follows. Section 2 reviews the basic model of judgment aggregation, while Section 3 introduces our two original axioms of egalitarianism and the rules they induce. Section 4 analyses the relationship between egalitarianism and strategic manipulation in judgment aggregation, and Section 5 focuses on relevant computational aspects: although the general problems of outcome determination and of strategic manipulation are proven to be very difficult, we propose a way to confront them with the tools of *Answer Set Programming* [35].

## 2 BASIC MODEL

Our framework relies on the standard formula-based model of judgment aggregation [48], but for simplicity we also use notation commonly employed in binary aggregation [36].

Let $\mathbb{N}$ denote the (countably infinite) set of all agents that can potentially participate in a judgment aggregation setting. In every specific such setting, a finite set of agents $N \subset \mathbb{N}$ of size $n \geq 2$ express judgments on a finite and nonempty set of *issues* (formulas in propositional logic) $\Phi = \{\varphi_1, \ldots, \varphi_m\}$, called the *agenda*. $\mathcal{J}(\Phi) \subseteq \{0,1\}^m$ denotes the set of all admissible opinions on $\Phi$. Then, a *judgment* $J$ is a vector in $\mathcal{J}(\Phi)$, with 1 (0) in position $k$ meaning that the issue $\varphi_k$ is accepted (rejected). $\overline{J}$ is the *antipodal* judgment of $J$: for all $\varphi \in \Phi$, $\varphi$ is accepted in $\overline{J}$ if and only if it is rejected in $J$.

A *profile* $\boldsymbol{J} = (J_1, \ldots J_n) \in \mathcal{J}(\Phi)^n$ is a vector of individual judgments, one for each agent in a group $N$. We write $\boldsymbol{J}' =_{-i} \boldsymbol{J}$ when the profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ are the same, besides the judgment of agent $i$. We write $\boldsymbol{J}_{-i}$ to denote the profile $\boldsymbol{J}$ with agent $i$'s judgment removed, and $(\boldsymbol{J}, J) \in \mathcal{J}(\Phi)^{n+1}$ to denote the profile $\boldsymbol{J}$ with judgment $J$ added. A *judgment aggregation rule* $F$ is a function that maps every possible profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$, for every group $N$ and agenda $\Phi$, to a nonempty set $F(\boldsymbol{J})$ of collective judgments in $\mathcal{J}(\Phi)$. Note that a judgment aggregation rule is defined over groups and agendas of variable size, and may return several, tied, collective judgments.

The agents that participate in a judgment aggregation scenario will naturally have preferences over the outcome produced by the aggregation rule. First, given an agent $i$'s truthful judgment $J_i$, we need to determine when agent $i$ would prefer a judgment $J$ over a different judgment $J'$. The most prevalent type of such preferences considered in the judgment aggregation literature is that of *Hamming distance* preferences [6, 7, 9, 59].

The Hamming distance between two judgments $J$ and $J'$ equals the number of issues on which these judgments disagree—concretely, it is defined as $H(J, J') = \sum_{\varphi \in \Phi} |J(\varphi) - J'(\varphi)|$, where $J(\varphi)$ denotes the binary value in the position of $\varphi$ in $J$. For example, $H(100, 111) = 2$. Then, the (weak, and analogously strict) preference of agent $i$ over judgments is defined by the relation $\succeq_i$ (where $J \succeq_i J'$ means that $i$'s utility from $J$ is higher than that from $J'$):

$$J \succeq_i J' \text{ if and only if } H(J_i, J) \leq H(J_i, J').$$

But an aggregation rule often outputs more than one judgment, and thus we also need to determine agents' preferences over sets of judgments.[1] We define two requirements guaranteeing that the preferences of the agents over sets of judgments are consistent with their preferences over single judgments. To that end, let $\mathring{\succeq}_i$ (with strict part $\mathring{\succ}_i$) denote agent $i$'s preferences over sets $X, Y \subseteq \mathcal{J}(\Phi)$. We require that $\mathring{\succeq}_i$ is related to $\succeq_i$ as follows:

- $J \succeq_i J'$ if and only if $\{J\} \mathrel{\mathring{\succeq}_i} \{J'\}$, for any $J, J' \in \mathcal{J}(\Phi)$;
- $X \mathrel{\mathring{\succ}_i} Y$ implies that there exist some $J \in X$ and $J' \in Y$ such that $J \succ_i J'$ and $\{J, J'\} \nsubseteq X \cap Y$.

The above conditions hold for almost all well-known preference extensions. For example, they hold for the *pessimistic* preference ($X \succ^{pess} Y$ if and only if there exists $J' \in Y$ such that $J \succ J'$ for all $J \in X$) and the *optimistic* preference ($X \succ^{opt} Y$ if and only if there exists $J \in X$ such that $J \succ J'$ for all $J' \in Y$) of Duggan and Schwartz [19], as well as the preference extensions of Gärdenfors [32] and Kelly [41]. The results provided in this paper abstract away from specific preference extensions.

## 3 EGALITARIAN AXIOMS AND RULES

This section focuses on two axioms of egalitarianism in judgment aggregation. We examine them in relation to each other and to existing properties from belief merging, as well as to the standard majority property defined below. Most of the well-known judgment aggregation rules return the majority opinion, when that opinion is logically consistent [24].[2]

Let $m(\boldsymbol{J})$ be the judgment that accepts exactly those issues accepted by a strict majority of agents in $\boldsymbol{J}$. A rule $F$ is *majoritarian* when for all profiles $\boldsymbol{J}$, $m(\boldsymbol{J}) \in \mathcal{J}(\Phi)$ implies that $F(\boldsymbol{J}) = \{J\}$.

Our first axiom with an egalitarian flavour is the *maximin property*, suggesting that we should aim at maximising the utility of those agents that will be worst off in the outcome. Assuming that everyone submits their truthful judgment during the aggregation process, this means that we should try to minimise the distance of the agents that are furthest away from the outcome. Formally:

[1] Various approaches have been taken within the area of social choice theory in order to extend preferences over objects to preferences over sets of objects —see Barberà et al. [3] for a review.

[2] A central problem in judgment aggregation concerns the fact that the issue-wise majority is not always logically consistent [48].

- A rule $F$ satisfies the **maximin** property if for all profiles $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ and judgments $J \in F(\boldsymbol{J})$ there do not exist judgment $J' \in \mathcal{J}(\Phi)$ and agent $j \in N$ such that

$$H(J_i, J') < H(J_j, J) \text{ for all } i \in N.$$

Although the maximin property is quite convincing, there are settings like those motivated in the Introduction where it does not offer sufficient egalitarian guarantees. We thus consider a different property next, which we call the *equity property*. This axiom requires that the gaps in the agents' satisfaction be minimised. In other words, no two agents should find themselves in very different distances with respect to the collective outcome. Formally:

- A rule $F$ satisfies the **equity** property if for all profiles $\boldsymbol{J} \in \mathcal{J}^n$ and judgments $J \in F(\boldsymbol{J})$, there do not exist judgment $J' \in \mathcal{J}(\Phi)$ and agents $i', j' \in N$ such that

$$|H(J_i, J') - H(J_j, J')| < |H(J_{i'}, J) - H(J_{j'}, J)| \text{ for all } i, j \in N.$$

No rule that satisfies either the maximin- or equity property can be majoritarian.[3] As an illustration, in a profile of only two agents who disagree on some issues, any egalitarian rule will try to reach a compromise, and this compromise will not be affected if any agents holding one of the two initial judgments are added to the profile—in contrast, a majoritarian rule will simply conform to the crowd.

Proposition 1 shows that it is also impossible for the maximin property and the equity property to simultaneously hold. Therefore, we have established the logical independence of all three axioms discussed so far: maximin, equity, and majoritarianism.

**Proposition 1.** *No judgment aggregation rule can satisfy both the maximin property and the equity property.*

Proof. Take an agenda $\Phi$ where $\mathcal{J}(\Phi)$ consists of the nodes in the graph below and consider the profile $\boldsymbol{J} = (J_1, J_2)$. Each edge is labelled with the Hamming distance between the judgments.

$$\begin{tikzcd}
J_1 : 110000 \arrow[d, dash, "1"'] \arrow[dr, dash, "4"] & \\
J : 010000 \arrow[d, dash, "3"'] & J' : 111111 \\
J_2 : 001100 \arrow[ur, dash, "4"'] &
\end{tikzcd}$$

Every aggregation rule satisfying the maximin property will return $\{J\}$, as this judgment maximises the utility of the worst off agent—in this case, agent 2. However every rule satisfying the equity property will return $\{J'\}$, as this judgment minimises the difference in utility between the best off and worst off agents. Thus, there is no rule that can satisfy the two properties at the same time. □

From Proposition 1, we also know now that the two properties of egalitarianism generate two disjoint classes of aggregation rules. In particular, in this paper we focus on the *maximal* rule that meets each property: a rule $F$ is the maximal one of a given class if, for every profile $\boldsymbol{J}$, the outcomes obtained by any other rule in that class are always outcomes of $F$ too.[4]

[3] This includes popular rules like the median rule [52]—known under a number of other names, notably *distance-based rule* [55], *Kemeny rule* [24], and *prototype rule* [49].

[4] Of course, several natural refinements of these rules can be defrined, with respect to various other axiomatic properties that we may find desirable. Identifying and studying such rules is an interesting direction for future research.

The maximal rule satisfying the maximin property is the rule *MaxHam* (see, e.g., Lang et al., 2011). For all profiles $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$,

$$\text{MaxHam}(\boldsymbol{J}) = \underset{J \in \mathcal{J}(\Phi)}{\text{argmin}} \max_{i \in N} H(J_i, J).$$

Analogously, we define a rule new to the judgment aggregation literature, which is the maximal one satisfying the equity property. For all profiles $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$,

$$\text{MaxEq}(\boldsymbol{J}) = \underset{J \in \mathcal{J}(\Phi)}{\text{argmin}} \max_{i,j \in N} |H(J_i, J) - H(J_j, J)|.$$

To better understand these rules, consider an agenda with six issues: $p, q, r \equiv p \wedge q$, and their negations. Suppose that there are only two agents in a profile $\boldsymbol{J}$, holding judgments $J_1 = (111)$ and $J_2 = (010)$. Then, we have that $\text{MaxHam}(\boldsymbol{J}) = \{(111), (010)\}$, while $\text{MaxEq} = \{(000), (100)\}$. In this example, the difference in spirit between the two rules of our interest is evident. Although the MaxHam rule is able to fully satisfy exactly one of the agents without causing much harm to the other, it still creates greater unbalance than the MaxEq rule, which ensures that the two agents are equally happy with the outcome (under Hamming-distance preferences). In that sense, MaxEq is better suited for a group of agents that do not want any of them to feel particularly put upon, while MaxHam seems more desirable when a minimum level of happiness is asked for.

MaxHam generalises minimax approval voting [11], which is the special case without logical constraint on the judgments, meaning agents may approve any subset of issues. Brams et al. [11] show that MaxHam remains manipulable in this special case. As finding the outcome of minimax is computationally hard, Caragiannis et al. [13] provide approximation algorithms that circumvent this problem. They also demonstrate the interplay between manipulability and lower bounds for the approximation algorithm—establishing strategyproofness results for approximations of minimax.

## 3.1 Relations with Egalitarian Belief Merging

A framework closely related to ours is that of belief merging [43], which is concerned with how to aggregate several (possibly inconsistent) sets of beliefs into one consistent belief set.[5] Egalitarian belief merging is studied by Everaere et al. [28], who examine interpretations of the *Sen-Hammond equity condition* [58] and the *Pigou-Dalton transfer principle* [17]—two properties that are logically incomparable.[6] We situate our egalitarian axioms within the context of these egalitarian axioms from belief merging; we reformulate these axioms into our framework.

- Fix an arbitrary profile $\boldsymbol{J}$, agents $i, j$, and any three judgment sets $J, J' \in \mathcal{J}(\Phi)$. An aggregation rule $F$ satisfies the **Sen-Hammond equity property** if whenever

$$H(J_i, J) < H(J_i, J') < H(J_j, J') < H(J_j, J)$$

and $H(J_{i'}, J) = H(J_{i'}, J')$ for all other agents $i' \in N \setminus \{i, j\}$, then $J \in F(\boldsymbol{J})$ implies $J' \in F(\boldsymbol{J})$.

**Proposition 2.** *If a rule satisfies either the maximin property or the equity property, then it will satisfy the Sen-Hammond equity property.*

[5] We refer to Everaere et al. [29] for a detailed comparison of the two frameworks.

[6] Another egalitarian property in belief merging is the *arbitration postulate*. We do not go into detail on this postulate, but refer the reader to Konieczny and Pérez [43].

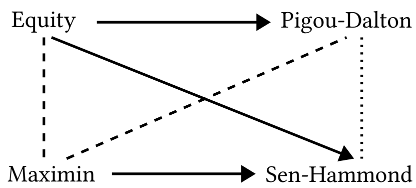


**Figure 1: Dashed lines denote incompatibility, dotted lines incomparability, and arrows implication relations.**

Proof (sketch). Let $\boldsymbol{J} = (J_i, J_j)$ be a profile such that $H(J, J_i) < H(J', J_i) < H(J', J_j) < H(J, J_j)$, and $H(J_{i'}, J) = H(J_{i'}, J')$ for all other agents $i' \in N \setminus \{i, j\}$. Suppose $F$ satisfies the equity property—if there is some agent $i'$ such that $|H(J_i, J) - H(J_{i'}, J)| > |H(J_i, J) - H(J_j, J)|$, then $J \in F(\boldsymbol{J})$ if and only if $J' \in F(\boldsymbol{J})$, as the maximal difference in distance will be the same for the two judgments. If this is not the case, then agents $i$ and $j$ determine the outcome regarding $J$ and $J'$ so clearly $J \in F(\boldsymbol{J})$ implies $J' \in F(\boldsymbol{J})$. The argument for other cases proceeds similarly.

If $F$ satisfies the maximin property, then a similar argument tells us that if membership of $J$ and $J'$ in the outcome is determined by an agent other than $i$ or $j$, we will either have both or neither. If $i$, and $j$ are the determining factor then $J \in F(\boldsymbol{J})$ implies $J' \in F(\boldsymbol{J})$. □

- Given a profile $\boldsymbol{J} = (J_1, \ldots, J_n)$ and agents $i$ and $j$ such that:
  - $H(J_i, J) < H(J_i, J') \leq H(J_j, J') < H(J_j, J)$,
  - $H(J_i, J') - H(J_i, J) = H(J_j, J') - H(J_j, J)$, and
  - $H(J_{i^*}, J) = H(J_{i^*}, J')$ for all other agents $i^* \in N \setminus \{i, j\}$,

  $F$ satisfies the **Pigou-Dalton transfer principle** if $J' \in F(\boldsymbol{J})$ implies $J \notin F(\boldsymbol{J})$.

We refer to these axioms simply as *Sen-Hammond*, and *Pigou-Dalton*. Note that Pigou-Dalton is also a weaker version of our equity property, as it stipulates that the difference between utility in agents should be lessened under certain conditions, while the equity property always aims to minimise this distance.

While we can find a rule that satisfies both the equity property and a weakening of the maximin property, Sen-Hammond, we cannot do the same by weakening the equity property.

**Proposition 3.** *No judgment aggregation rule can satisfy both the maximin property and Pigou-Dalton.*

Proof. Consider the domain $\mathcal{J}(\Phi) = \{J_1, J_2, J_3, J, J'\}$ with the following Hamming distances between judgment sets.[7]

| | $J$ | $J'$ | $J_1$ | $J_2$ | $J_3$ |
|---|---|---|---|---|---|
| $J_1$ | 2 | 4 | 0 | 4 | 8 |
| $J_2$ | 6 | 4 | 4 | 0 | 10 |
| $J_3$ | 6 | 6 | 8 | 10 | 0 |

Let $\boldsymbol{J} = (J_1, J_2, J_3)$. If $F$ satisfies the maximin property, $\{J, J'\} \subseteq F(\boldsymbol{J})$, as we can see from the grey cells. This means Pigou-Dalton is violated in this profile, as $J' \in F(\boldsymbol{J})$ should imply $J \notin F(\boldsymbol{J})$. □

We summarise the observations of this section in Figure 1.

[7] One such domain would be the following, where $J$ = 00000000111, $J'$ = 00000001110, $J_1$ = 00000010011, $J_2$ = 00000111000, and $J_3$ = 11111001111.

# 4 STRATEGIC MANIPULATION

This section provides an account of strategic manipulation with respect to the egalitarian axioms defined in Section 3. We start off with presenting the most general notion of strategic manipulation in judgment aggregation, introduced by Dietrich and List [18].[8] We assume Hamming preferences throughout this section.

**Definition 1.** *A rule $F$ is susceptible to **manipulation** by agent $i$ in profile $\boldsymbol{J}$, if there exists a profile $\boldsymbol{J}' =_{-i} \boldsymbol{J}$ such that $F(\boldsymbol{J}') \mathrel{\mathring{\succ}_i} F(\boldsymbol{J})$.*

We say that $F$ is *strategyproof* in case $F$ is not manipulable by any agent $i \in N$ in any profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$.

Proposition 4 shows an important fact: In judgment aggregation, egalitarianism is incompatible with strategyproofness.[9]

**Proposition 4.** *If an aggregation rule is strategyproof, it cannot satisfy the maximin property or the equity property.*

Proof. We show the contrapositive. Let $\Phi$ be an agenda such that $\mathcal{J}(\Phi) = \{000000, 110000, 111000, 111111\}$. Consider the following two profiles $\boldsymbol{J}$ (left) and $\boldsymbol{J}'$ (right).

| | |
|---|---|
| $J_i$ | 111000 |
| $J_j$ | 000000 |
| $F(\boldsymbol{J})$ | 110000 |

| | |
|---|---|
| $J'_i$ | 111111 |
| $J'_j$ | 000000 |
| $F(\boldsymbol{J}')$ | 111000 |

In profile $\boldsymbol{J}$, both the maximin and the equity properties prescribe that 110000 should be returned as the single outcome, while in profile $\boldsymbol{J}'$ they agree on 111000. Because $\boldsymbol{J}' = (\boldsymbol{J}_{-i}, J'_i)$, and 111000 $\succ_i$ 110000, this is a successful manipulation. Thus, if $F$ satisfies the maximin or the equity property, it fails strategyproofness. □

Strategyproofness according to Definition 1 is a strong requirement, which many known rules fail [9]. We investigate two more nuanced notions of strategyproofness that are novel to judgment aggregation, yet have familiar counterparts in voting theory.

First, *no-show manipulation* happens when an agent can achieve a preferable outcome simply by not submitting any judgment, instead of reporting a truthful or an untruthful one.

**Definition 2.** *A rule $F$ is susceptible to **no-show manipulation** by agent $i$ in profile $\boldsymbol{J}$ if $F(\boldsymbol{J}_{-i}) \mathrel{\mathring{\succ}_i} F(\boldsymbol{J})$.*

We say that $F$ satisfies *participation* if it is not susceptible to no-show manipulation by any agent $i \in N$ in any profile.[10]

Second, *antipodal strategyproofness* poses another barrier against manipulation, by stipulating that an agent cannot change the outcome towards a better one for herself by reporting a totally untruthful judgment. This is a strictly weaker requirement than full strategyproofness, serving as a protection against excessive lying.

**Definition 3.** *A rule $F$ is susceptible to **antipodal manipulation** by agent $i$ in profile $\boldsymbol{J}$ if $F(\boldsymbol{J}_{-i}, \overline{J_i}) \mathrel{\mathring{\succ}_i} F(\boldsymbol{J})$.*

We say that $F$ satisfies *antipodal strategyproofness* if it not susceptible to antipodal manipulation by any agent $i \in N$ in any profile. As is the case for participation, antipodal strategyproofness is a weaker notion of strategyproofness as far as the MaxHam and the MaxEq rules are concerned.

[8] The original definition of Dietrich and List [18] concerned single-judgment collective outcomes, and a type of preferences that covers Hamming-distance ones.

[9] This in in line with Brams et al.'s work on the minimax rule in approval voting.

[10] cf. the no-show paradox in voting [30].

In voting theory, Sanver and Zwicker [57] show that participation implies antipodal strategyproofness (or *half-way monotonicity*, as called in that framework) for rules that output a single winning alternative. Notably, this is not always the case in our model (see Example 1). This is not surprising, as obtaining such a result independently of the preference extension would be significantly stronger than the result by Sanver and Zwicker [57]. We are, however, able to reproduce this relationship between participation and strategyproofness in Theorem 1, for a specific type of preferences.

**Example 1.** *We present a rule that satisfies participation but violates antipodal strategyproofness. The other direction admits a similar example, and is thus omitted. Note that the rule demonstrated is quite unnatural for simplicity of the presentation.*

*Consider an agenda $\Phi$ with $\mathcal{J}(\Phi) = \{00, 01, 11\}$.*[11] *We construct an anonymous rule $F$ that is only sensitive to* which *judgments are submitted and not to their quantity:*

$F(00) = F(11) = F(01, 00) = F(00, 11) = \{01, 11\}$;
$F(01) = \{00, 11\}$; $F(01, 11) = F(01, 00, 11) = \{01\}$.

*For the pessimistic preference, no agent can be strictly better off by abstaining. However, compare the profiles $(01, 00)$ and $(01, 11)$: agent 2 with truthful judgment $00$ can move from outcome $\{01, 11\}$ to outcome $\{01\}$, which is strictly better for her.*

While the two axioms are independent in the general case, participation implies antipodal strategyproofness (Theorem 1) if we stipulate that

- $X \mathring{\succ}_i Y$ **if and only if** there exist some $J \in X$ and $J' \in Y$ such that $J \succ_i J'$ and $\{J, J'\} \not\subseteq X \cap Y$.

If a preference satisfies the above condition, we say that it is *decisive*. This condition gives rise to a preference extension equivalent to the *large preference extension* of Kruger and Terzopoulou [44]. Note that a decisive preference is not necessarily acyclic—in fact, it may even be symmetric. The interpretation of such a preference extension is slightly different than the usual one; when we say that a rule is strategyproof for a decisive preference where both $J \mathring{\succ} J'$ and $J' \mathring{\succ} J$ hold, we mean that no agent $i$ with $J \mathring{\succ}_i J'$ and no agent $j \neq i$ with $J' \mathring{\succ}_j J$ will ever have an incentive to manipulate.

Using Lemma 1, we can now prove a result analogous to the one in voting theory, to give a complete picture of how these axioms relate to each other in judgment aggregation.

**Lemma 1.** *For judgment sets $J, J'$ and $J''$: $H(J, J') > H(J, J'')$, if and only if $H(\bar{J}, J') < H(\bar{J}, J'')$.*

Proof. For judgment sets $J, J' \in \mathcal{J}(\Phi)$, $H(\bar{J}, J') = m - H(J, J')$. Suppose $H(J, J') > H(J, J'')$. Then $H(\bar{J}, J') = m - H(J, J') < m - H(J, J'') = H(\bar{J}, J'')$. The other direction is analogous. □

**Theorem 1.** *For decisive preferences over sets of judgments, participation implies antipodal strategyproofness.*

Proof. Working on the contrapositive, suppose that $F$ is susceptible to antipodal manipulation. We will prove that $F$ is susceptible to no-show manipulation too. We know that there exists $i \in N$ such that $F(\boldsymbol{J}_{-i}, \overline{J_i}) \mathring{\succ}_i F(\boldsymbol{J}_{-i}, J_i)$, for some profile $\boldsymbol{J}$. This means that there exist $J' \in F(\boldsymbol{J}_{-i}, \overline{J_i})$ and $J \in F(\boldsymbol{J}_{-i}, J_i)$ with $J' \succ_i J$. Equivalently,

$$H(J_i, J') < H(J_i, J) \tag{1}$$

Next, consider a judgment $J'' \in F(\boldsymbol{J}_{-i})$.

If $H(\overline{J_i}, J'') < H(\overline{J_i}, J')$, then $F$ is susceptible to no-show manipulation by agent $i$ in the profile $(\boldsymbol{J}_{-i}, \overline{J_i})$.

Otherwise, $H(\overline{J_i}, J') \leq H(\overline{J_i}, J'')$. Then Lemma 1 implies that $H(J_i, J'') \leq H(J_i, J')$. So, together with Inequality (1), we have that $H(J_i, J'') < H(J_i, J)$. This means that $F$ is susceptible to no-show manipulation by agent $i$ in the profile $(\boldsymbol{J}_{-i}, J_i)$. □

We next prove that any rule satisfying the maximin property is immune to both no-show manipulation and antipodal manipulation (Theorem 2), while this is not true for the equity property (Proposition 5).[12] We emphasise that the theorem holds for *all* preference extensions. These results—holding for two independent notions of strategyproofness—are significant for two reasons. First, they bring to light the conditions under which we can have our cake and eat it too, simultaneously satisfying an egalitarian property and a degree of strategyproofness. In addition, they provide a further way to distinguish between the properties of maximin and equity: the former is better suited in contexts where we may worry about the agents' strategic behaviour.

**Theorem 2.** *The maximin property implies participation and antipodal strategyproofness.*

Proof. We prove the participation case; the proof for antipodal strategyproofness is analogous, and utilises Lemma 1.

Suppose for contradiction that $F$ is a rule that satisfies the maximin property but violates participation. Then there must exist agent $i \in N$ and profile $\boldsymbol{J}$ where $J_i$ is agent $i$'s truthful judgment, such that $F(\boldsymbol{J}_{-i}) \mathring{\succ}_i F(\boldsymbol{J})$. This means there must exist judgments $J \in F(\boldsymbol{J})$ and $J' \in F(\boldsymbol{J}_{-i})$ such that $J' \succ_i J$ and $\{J, J'\} \not\subseteq F(\boldsymbol{J}) \cap F(\boldsymbol{J}_{-i})$. Because agent $i$ strictly prefers $J'$ to $J$, this means that $H(J_i, J) > H(J_i, J')$. We consider two cases.

<u>Case 1:</u> Suppose that $J' \notin F(\boldsymbol{J})$. Let $k$ be the distance between the worst off agent's judgment in $\boldsymbol{J}$ and any judgment in $F(\boldsymbol{J})$. Then,

$$H(J_{j'}, J) \leq k \text{ for all } j' \in N. \tag{2}$$

We know that $H(J_i, J') < k$ because $H(J_i, J) \leq k$, and agent $i$ strictly prefers $J'$ to $J$. From Inequality (2), this means that if $J'$ is not among the outcomes in $F(\boldsymbol{J})$, there has to be some $j \in N \setminus \{i\}$ such that $H(J_j, J') > k$. But all judgments submitted to profile $(\boldsymbol{J}_{-i})$ by agents in $N \setminus \{i\}$ are at most at distance $k$ from $J$ by Inequality (2), so $J$ would be selected by any rule satisfying the maximin property will select $J$ as an outcome of $F(\boldsymbol{J}_{-i})$—instead of $J'$, a contradiction.

<u>Case 2:</u> Suppose that $J' \in F(\boldsymbol{J})$, meaning that $J \notin F(\boldsymbol{J}_{-i})$. Analogously to the first case, let $k'$ be the distance between the worst off agent's judgment in $\boldsymbol{J}_{-i}$ and any judgment in $F(\boldsymbol{J}_{-i})$. Then,

$$H(J_{j'}, J') \leq k' \text{ for all } j' \in N \setminus \{i\}. \tag{3}$$

Moreover, since $J \notin F(\boldsymbol{J}_{-i})$, it is the case that

$$H(J_j, J) > k' \text{ for some } j \neq i. \tag{4}$$

[11] For other agendas we can simply take the rule to be constant.

[12] Note that antipodal strategyproofness is not so weak a requirement that is immediately satisfied by all "utilitarian" aggregation rules. For example, the Copeland voting rule fails the analogous axiom of half-way monotonicity [60].

In profile $\boldsymbol{J}$, Inequalities (3) and (4) still hold. In addition, we have that $H(J_i, J) > H(J_i, J')$ because agent $i$ strictly prefers $J'$ to $J$. So, for any rule satisfying the maximin property, judgment $J'$ will be better as an outcome of $F(\boldsymbol{J})$ than $J$, a contradiction. □

**Corollary 1.** *The rule MaxHam satisfies antipodal strategyproofness and participation.*

**Proposition 5.** *No rule that satisfies the equity property can satisfy participation or antipodal strategyproofness .*

Proof. The following is a counterexample for antipodal strategyproofness. A similar one exists for participation.

Consider the following profiles $\boldsymbol{J} = \{Ji, J_j\}$ and $\boldsymbol{J}' = (\boldsymbol{J}_{-i}, \overline{J_i})$. We give a visual representation of the profiles as well as the outcomes under an arbitrary rule $F$ that satisfies the equity principle. We specify that $\mathcal{J}(\Phi) = \{00110, 00000, 01110, 10000, 11111\}$.

$$
\begin{array}{ccccccc}
 & & \boldsymbol{J} & & & & \boldsymbol{J}' \\
F(\boldsymbol{J}) & 2 & J_i : 00000 & 1 & F(\boldsymbol{J}') & 4 & J_i' : 11111 \\
00110 & & & & 10000 & & \\
 & 1 & J_j : 01110 & 4 & & 4 & J_j' : 01110
\end{array}
$$

Each edge from an individual judgment to a collective one is labelled with the Hamming distance between the two. It is clear that agent $i$ will benefit from her antipodal manipulation, as her true judgment is much closer to the singleton outcome in $\boldsymbol{J}'$ than the singleton outcome in $\boldsymbol{J}$. □

**Corollary 2.** *The rule MaxEq does not satisfy participation or antipodal strategyproofness.*

## 5 COMPUTATIONAL ASPECTS

We have discussed two aggregation rules that reflect desirable egalitarian principles—i.e., the MaxHam and MaxEq rules—and examined whether they give agents incentives to misrepresent their truthful judgments. In this section we consider how complex it is, computationally, to employ these rules, and the complexity of determining whether an agent can manipulate the collective outcome.

The MaxHam rule has been considered from a computational perspective before [38–40]. Here, we extend this analysis to the MaxEq rule, and we compare the two rules with each other on their computational properties. Concretely, we primarily establish some computational complexity results; motivated by these results, we then illustrate how some computational problems related to these rules can be solved using the paradigm of Answer Set Programming.

### 5.1 Computational Complexity

We investigate some computational complexity aspects of the judgment aggregation rules that we have considered. Due to space constraints, we will only describe the main lines of these results—for full details, we refer to the accompanying Appendix.[13]

Consider the problem of outcome determination (for a rule $F$). This is most naturally modelled as a search problem, where the input consists of an agenda $\Phi$ and a profile $\boldsymbol{J} = (J_1, \ldots, J_n) \in \mathcal{J}(\Phi)^n$. The problem is to produce some judgment set $J^* \in F(\boldsymbol{J})$. We will show that for the MaxEq rule, this problem can be solved in polynomial time with a logarithmic number of calls to an oracle for NP search problems (where the oracle also produces a witness for yes answers—also called an FNP witness oracle). Said differently, the outcome determination problem for the the MaxEq rule lies in the complexity class $\mathsf{FP}^{\mathsf{NP}}[\text{log,wit}]$. We also show that the problem is complete for this class (using the standard type of reductions used for search problems: polynomial-time Levin reductions).

**Theorem 3.** *The outcome determination problem for the MaxEq rule is* $\mathsf{FP}^{\mathsf{NP}}[\text{log,wit}]$*-complete under polynomial-time Levin reductions.*

Proof (sketch). Membership in $\mathsf{FP}^{\mathsf{NP}}[\text{log,wit}]$ can be shown by giving a polynomial-time algorithm that solves the problem by querying an FNP witness oracle a logarithmic number of times. The algorithm first finds the minimum value $k$ of $\max_{J',J''\in\boldsymbol{J}} |H(J, J') - H(J, J'')|$ by means of binary search—requiring a logarithmic number of oracle queries. Then, with one additional oracle query, the algorithm can produce some $J^* \in \mathcal{J}(\Phi)$ with $\max_{J',J''\in\boldsymbol{J}} |H(J^*, J') - H(J^*, J'')| = k$.

To show $\mathsf{FP}^{\mathsf{NP}}[\text{log,wit}]$-hardness, we reduce from the problem of finding a satisfying assignment of a (satisfiable) propositional formula $\psi$ that sets a maximum number of variables to true [15, 42]. This reduction works roughly as follows. Firstly, we produce 3CNF formulas $\psi_1, \ldots, \psi_v$ where each $\psi_i$ is 1-in-3-satisfiable if and only if there exists a satisfying assignment of $\psi$ that sets at least $i$ variables to true. Then, for each $i$, we transform $\psi_i$ to an agenda $\Phi_i$ and a profile $\boldsymbol{J}_i$ such that there is a judgment set with equal Hamming distance to each $J \in \boldsymbol{J}_i$ if and only if $\psi_i$ is 1-in-3-satisfiable. Finally, we put the agendas $\Phi_i$ and profiles $\boldsymbol{J}_i$ together into a single agenda $\Phi$ and a single profile $\boldsymbol{J}$ such that we can—from the outcomes selected by the MaxEq rule—read off the largest $i$ for which $\psi_i$ is 1-in-3-satisfiable, and thus, the maximum number of variables set to true in any truth assignment satisfying $\psi$. This last step involves duplicating issues in $\Phi_1, \ldots, \Phi_v$ different numbers of times, and creating logical dependencies between them. Moreover, we do this in such a way that from any outcome selected by the MaxEq rule, we can reconstruct a truth assignment satisfying $\psi$ that sets a maximum number of variables to true. □

The result of Theorem 3 means that the computational complexity of computing outcomes for the MaxEq rule lies at the $\Theta_2^{\mathsf{p}}$-level of the Polynomial Hierarchy. This is in line with previous results on the computational complexity of the outcome determination problem for the MaxHam rule—De Haan and Slavkovik [39] showed that a decision variant of the outcome determination problem for the MaxHam rule is $\Theta_2^{\mathsf{p}}$-complete. Notably, our proof (presented in detail in the Appendix) brings out an intriguing fact about a problem that is at first glance simpler than outcome determination for MaxEq: Given an agenda $\Phi$ and a profile $\boldsymbol{J}$, deciding whether the minimum value of $\max_{i,j\in N} |H(J_i, J) - H(J_j, J)|$ for $J \in \mathcal{J}(\Phi)$—the value that the MaxEq rule minimizes—is divisible by 4, is $\Theta_2^{\mathsf{p}}$-complete (Proposition 6). Intuitively, merely computing the minimum value that is relevant for MaxEq is $\Theta_2^{\mathsf{p}}$-hard.

**Proposition 6.** *Given an agenda $\Phi$ and a profile $\boldsymbol{J}$, deciding whether the minimal value of* $\max_{J',J''\in\boldsymbol{J}} |H(J^*, J') - H(J^*, J'')|$ *for* $J^* \in \mathcal{J}(\Phi)$*, is divisible by 4, is a* $\Theta_2^{\mathsf{p}}$*-complete problem.*

[13] The appendix is available here [8].

Interestingly, we found that the problem of deciding if there exists a judgment set $J^* \in \mathcal{J}(\Phi)$ that has the exact same Hamming distance to each judgment set in the profile is NP-hard, even when the agenda consists of logically independent issues.

**Proposition 7.** *Given an agenda $\Phi$ and a profile $\boldsymbol{J}$, the problem of deciding whether there is some $J^* \in \mathcal{J}(\Phi)$ with $\max_{J',J'' \in \boldsymbol{J}} |H(J^*, J') - H(J^*, J'')| = 0$ is* NP*-complete. Moreover,* NP*-hardness holds even for the case where $\Phi$ consists of logically independent issues—i.e., the case where $\mathcal{J}(\Phi) = \{0,1\}^m$ for some $m$.*

This is also in line with previous results for the MaxHam rule—De Haan [38] showed that computing outcomes for the MaxHam rule is computationally intractable even when the agenda consists of logically independent issues.

Next, we turn our attention to the problem of strategic manipulation. Specifically, we show that—for the case of decisive preferences over sets of judgment sets—the problem of deciding if an agent $i$ can strategically manipulate is in the complexity class $\Sigma_2^{\mathrm{p}}$.

**Proposition 8.** *Let $\succeq$ be a preference relation over judgment sets that is polynomial-time computable, and let $\mathring{\succeq}$ be a decisive extension over sets of judgment sets. Then the problem of deciding if a given agent $i$ can strategically manipulate under the MaxEq rule—i.e., given $\Phi$ and $\boldsymbol{J}$, deciding if there exists some $\boldsymbol{J}' =_{-i} \boldsymbol{J}$ with MaxEq$(\boldsymbol{J}')$ $\mathring{\succ}_i$ MaxEq$(\boldsymbol{J})$—is in the complexity class $\Sigma_2^{\mathrm{p}}$.*

Proof (sketch). To show membership in $\Sigma_2^{\mathrm{p}} = \mathrm{NP}^{\mathrm{NP}}$, we describe a nondeterministic polynomial-time algorithm with access to an NP oracle that solves the problem. The algorithm firstly guesses a new judgment set $J_i'$ for agent $i$ in the new profile $\boldsymbol{J}'$, and guesses a truth assignment witnessing that $J_i'$ is consistent. Then, using the NP oracle, it computes the values $k = \max_{J',J'' \in \boldsymbol{J}} |H(J, J') - H(J, J'')|$ and $k' = \max_{J',J'' \in \boldsymbol{J}'} |H(J, J') - H(J, J'')|$, for $J \in \mathcal{J}(\Phi)$. Finally, it guesses some $J, J' \in \mathcal{J}(\Phi)$, together with truth assignments witnessing consistency, and it verifies that $J' \succ_i J$, that $J' \in$ MaxEq$(\boldsymbol{J}')$, that $J \in$ MaxEq$(\boldsymbol{J})$, and that $\{J, J'\} \not\subseteq$ MaxEq$(\boldsymbol{J}) \cap$ MaxEq$(\boldsymbol{J}')$. Since these final checks can all be done in polynomial time—using the previously guessed and computed information—one can verify that this can be implemented by an $\mathrm{NP}^{\mathrm{NP}}$ algorithm. □

This $\Sigma_2^{\mathrm{p}}$-membership result can straightforwardly be extended to other variants of the manipulation problem (e.g., no-show manipulation and antipodal manipulation) and to other preferences, as well as to the MaxHam rule. Due to space constraints, we omit further details on this. Still, we shall mention that results demonstrating that strategic manipulation is very complex are generally more welcome than analogous ones regarding outcome determination. If manipulation is considered a negative side-effect of the agents' strategic behaviour, knowing that it is hard for the agents to materialise it is good news.[14] In Section 5.2 we will revisit these concerns from a different angle.

## 5.2 ASP Encoding for the MaxEq Rule

The complexity results in Section 5.1 leave no doubt that applying our egalitarian rules is computationally difficult. Nevertheless, they also indicate that a useful approach for computing outcomes of the MaxEq rule in practice would be to encode this problem into the paradigm of Answer Set Programming (ASP) [35], and to use ASP solving algorithms. ASP offers an expressive automated reasoning framework that typically works well for problems at the $\Theta_2^{\mathrm{p}}$ level of the Polynomial Hierarchy. In this section, we will show how this encoding can be done—similarly to an ASP encoding for the Max-Ham rule [40]. Due to space restrictions, we refer to the literature for details on the syntax and semantics of ASP—e.g., [33, 35].

We use the same basic setup that De Haan and Slavkovik [40] use to represent judgment aggregation scenarios—with some simplifications and modifications for the sake of readability. In particular, we use the predicate `voter/1` to represent individuals, we use `issue/1` to represent issues in the agenda, and we use `js/2` to represent judgment sets—both for the individual voters and for a dedicated agent `col` that represents the outcome of the rule.

With this encoding of judgment aggregation scenarios, one can add further constraints on the predicate `js/2` that express which judgment sets are consistent, based on the logical relations between the issues in the agenda $\Phi$—as done by De Haan and Slavkovik [40]. We refer to their work for further details on how this can be done.

Now, we show how to encode the MaxEq rule into ASP, similarly to the encoding of the MaxHam rule by De Haan and Slavkovik [40]. We begin by defining a predicate `dist/2` to capture the Hamming distance `D` between the outcome and the judgment set of an agent `A`.

```
1 dist(A,D) :- voter(A),
      D = #count { X : issue(X), js(col,X), js(A,-X) }.
```

Then, we define predicates `maxdist/1`, `mindist/1` and `inequity/1` that capture the maximum Hamming distance from the outcome to any judgment set in the profile, the minimum such Hamming distance, and the difference between the maximum and minimum (or *inequity*), respectively.

```
2 maxdist(Max) :- Max = #max { D : dist(A,D) }.
3 mindist(Min) :- Min = #min { D : dist(A,D) }.
4 inequity(Max-Min) :- maxdist(Max), mindist(Min).
```

Finally, we add an optimization constraint that states that only outcomes should be selected that minimize the inequity.[15]

```
5 #minimize { I@30 : inequity(I) }.
```

For any answer set program that encodes a judgment aggregation setting, combined with Lines 1–5, it then holds that the optimal answer sets are in one-to-one correspondence with the outcomes selected by the MaxEq rule.

Interestingly, we can readily modify this encoding to capture refinements of the MaxEq rule. An example of this is the refinement that selects (among the outcomes of the MaxEq rule) the outcomes that minimize the maximum Hamming distance to any judgment set in the profile. We can encode this example refinement by adding the following optimization statement that works at a lower priority level than the optimization in Line 5.

```
6 #minimize { Max@20 : maxdist(Max) }.
```

## 5.3 Encoding Strategic Manipulation

We now show how to encode the problem of strategic manipulation into ASP. The value of this section's contribution should be viewed from the perspective of the modeller rather than from that of the

[14] Note though that hardness results regarding manipulation of our egalitarian rules remain an open question.

[15] The expression "`@30`" in Line 5 indicates the priority level of this optimization statement (we used the arbitrary value of 30, and priority levels lexicographically).

agents. That is, even if we do not wish for the agents to be able to easily check whether they can be better off by lying, it may be reasonable, given a profile of judgments, to externally determine whether a certain agent can benefit from being untruthful.

We achieve this with the meta-programming techniques developed by Gebser et al. [34]. Their meta-programming approach allows one to additionally express optimization statements that are based on subset-minimality, and to transform programs with this extended expressivity to standard (disjunctive) answer set programs. We use this to encode the problem of strategic manipulation.

Due to space reasons, we will not spell out the full ASP encoding needed to do so. Instead, we will highlight the main steps, and describe how these fit together. We will use the example of MaxEq, but the exact same approach would work for any other judgment aggregation rule that can be expressed in ASP efficiently using regular (cardinality) optimization constraints—in other words, for all rules for which the outcome determination problem lies at the $\Theta_2^{\mathrm{p}}$ level of the Polynomial Hierarchy. Moreover, we will use the example of a decisive preference $\mathring{\succ}$ over sets of judgment sets that is based on a polynomial-time computable preference $\succ$ over judgment sets. The approach can be modified to work with other preferences as well.

We begin by guessing a new judgment set $J_i'$ for the individual $i$ that is trying to manipulate—and we assume, w.l.o.g., that $i = 1$.

```
voter(prime(1)).
1 { js(prime(1),X), js(prime(1),-X) } 1 :- issue(X).
```

Then, we express the outcomes of the MaxEq rule, both for the non-manipulated profile $\boldsymbol{J}$ and for the manipulated profile $\boldsymbol{J}'$, using the dedicated agents `col` (for $\boldsymbol{J}$) and `prime(col)` (for $\boldsymbol{J}'$). This is done exactly as in the encoding of the problem of outcome determination (so for the case of MaxEq, as described in Section 5.2)—with the difference that optimization is expressed in the right format for the meta-programming method of Gebser et al. [34].

We express the following subset-minimality minimization statement (at a higher priority level than all other optimization constraints used so far). This will ensure that every possible judgment set $J_i'$ will be considered as a subset-minimal solution.

```
_criteria(40,1,js(prime(1),X)) :- js(prime(1),S).
_optimize(40,1,incl).
```

To encode whether or not the guessed manipulation was successful, we have to define a predicate `successful/0` that is true if and only if (i) $J' \succ_i J$ and (ii) $J$ and $J'$ are not both selected as outcome by the MaxEq rule for both $\boldsymbol{J}$ and $\boldsymbol{J}'$, where $J'$ is the outcome encoded by the statements `js(prime(col),X)` and $J$ is the outcome encoded by the statements `js(col,X)`. Since we assume that $\succ_i$ is computable in polynomial time, and since we can efficiently check using statements in the answer set whether $J$ and $J'$ are selected by the MaxEq rule for $\boldsymbol{J}$ and $\boldsymbol{J}'$, we know that we can define the predicate `successful/0` correctly and succinctly in our encoding. For space reasons, we omit further details on how to do this.

Then, we express another minimization statement (at a lower priority level than all other optimization statements used so far), that states that we should make `successful` true whenever possible. Intuitively, we will use this to filter our guessed manipulations that are unsuccessful.

```
unsuccessful :- not successful.
successful :- not unsuccessful.
_criteria(10,1,unsuccessful) :- unsuccessful.
_optimize(10,1,card).
```

Finally, we feed the answer set program $P$ that we constructed so far into the meta-programming method, resulting in a new (disjunctive) answer set program $P'$ that uses no optimization statements at all, and whose answer sets correspond exactly to the (lexicographically) optimized answer sets of our program $P$. Since the new program $P'$ does not use optimization, we can add additional constraint to $P'$ to remove some of the answer sets. In particular, we will filter out those answer sets that correspond to an unsuccessful manipulation—i.e., those containing the statement `unsuccessful`. Effectively, we add the following constraint to $P'$:

```
:- unsuccessful.
```

As a result the only answer sets of $P'$ that remain correspond exactly to successful manipulations $J_i'$ for agent $i$.

The meta-programming technique that we use uses the full disjunctive answer set programming language. For this full language, finding answer sets is a $\Sigma_2^{\mathrm{p}}$-complete problem [21]. This is in line with our result of Proposition 8 where we show that the problem of strategic manipulation is in $\Sigma_2^{\mathrm{p}}$.

The encoding that we described can straightforwardly be modified for various variants of strategic manipulation (e.g., antipodal manipulation). To make this work, one needs to express additional constraints on the choice of the judgment set $J_i'$. To adapt the encoding for other preference relations $\mathring{\succ}$, one needs to adapt the definition of `successful/0`, expressing under what conditions an act of manipulation is successful.

Our encoding using meta-programming is relatively easily understandable, since we do not need to tinker with the encoding of complex optimization constraints in full disjunctive answer set programming ourselves—this we outsource to the meta-programming method. If one were to do this manually, there is more space for tailor-made optimizations, which might lead to a better performance of ASP solving algorithms for the problem of strategic manipulation. It is an interesting topic for future research to investigate this, and possibly to experimentally test the performance of different encodings, when combined with ASP solving algorithms.

## 6 CONCLUSION

We have introduced the concept of egalitarianism into the framework of judgment aggregation and have presented how egalitarian and strategyproofness axioms interact in this setting. Importantly, we have shown that the two main interpretations of egalitarianism give rise to rules with differing levels of protection against manipulation. In addition, we have looked into various computational aspects of the egalitarian rules that arise from our axioms, in a twofold manner: First, we have provided worst-case complexity results; second, we have shown how to solve the relevant hard problems using Answer Set Programming.

While we have axiomatised two prominent egalitarian principles, it remains to be seen whether other egalitarian axioms can provide stronger barriers against manipulation. For example, in parallel to majoritarian rules, one could define rules that minimise the distance to some egalitarian ideal. Moreover, as is the case in judgment aggregation, there is an obvious lack of voting rules designed with egalitarian principles in mind. We hope this paper opens the door for similar explorations in voting theory.

# REFERENCES


[1] Georgios Amanatidis, Georgios Birmpas, and Evangelos Markakis. 2016. On Truthful Mechanisms for Maximin Share Allocations. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*.

[2] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. 2017. Justified Representation in Approval-based Committee Voting. *Social Choice and Welfare* 48, 2 (2017), 461–485.

[3] Salvador Barberà, Walter Bossert, and Prasanta K Pattanaik. 2004. Ranking Sets of Objects. In *Handbook of Utility Theory*. Springer, 893–977.

[4] Seth D. Baum. 2017. Social Choice Ethics in Artificial Intelligence. *AI & Society* (2017), 1–12.

[5] Dorothea Baumeister, Gábor Erdélyi, Olivia Johanna Erdélyi, and Jörg Rothe. 2013. Computational aspects of manipulation and control in judgment aggregation. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*.

[6] Dorothea Baumeister, Gábor Erdélyi, Olivia J Erdélyi, and Jörg Rothe. 2015. Complexity of Manipulation and Bribery in Judgment Aggregation for Uniform Premise-Based Quota Rules. *Mathematical Social Sciences* 76 (2015), 19–30.

[7] Dorothea Baumeister, Jörg Rothe, and Ann-Kathrin Selker. 2017. Strategic Behavior in Judgment Aggregation. In *Trends in Computational Social Choice*. Lulu. com, 145–168.

[8] Sirin Botan, Ronald de Haan, Marija Slavkovik, and Zoi Terzopoulou. 2021. Egalitarian Judgment Aggregation. *CoRR* abs/2102.02785 (2021). arXiv:2102.02785 http://arxiv.org/abs/2102.02785

[9] Sirin Botan and Ulle Endriss. 2020. Majority-Strategyproofness in Judgment Aggregation. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

[10] Steven J Brams, Michael A Jones, and Christian Klamler. 2008. Proportional Pie-Cutting. *International Journal of Game Theory* 36, 3-4 (2008), 353–367.

[11] Steven J Brams, D Marc Kilgour, and M Remzi Sanver. 2007. A Minimax Procedure for Electing Committees. *Public Choice* 132, 3 (2007), 401–420.

[12] Eric Budish. 2011. The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes. *Journal of Political Economy* 119, 6 (2011), 1061–1103.

[13] Ioannis Caragiannis, Dimitris Kalaitzis, and Evangelos Markakis. 2010. Approximation algorithms and mechanism design for minimax approval voting. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*, Vol. 24.

[14] Yiling Chen, John K Lai, David C Parkes, and Ariel D Procaccia. 2013. Truth, Justice, and Cake Cutting. *Games and Economic Behavior* 77, 1 (2013), 284–297.

[15] Zhi-Zhong Chen and Seinosuke Toda. 1995. The Complexity of Selecting Maximal Solutions. *Information and Computation* 119 (1995), 231–239. Issue 2.

[16] Vincent Conitzer, Walter Sinnott-Armstrong, Jana Schaich Borg, Yuan Deng, and Max Kramer. 2017. Moral Decision Making Frameworks for Artificial Intelligence. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*.

[17] H Dalton. 1920. The Measurement of the Inequality of Incomes. *Economic Journal* 30, 119 (1920), 348–461.

[18] Franz Dietrich and Christian List. 2007. Strategy-proof Judgment Aggregation. *Economics & Philosophy* 23, 3 (2007), 269–300.

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

[20] Michael Dummet. 1984. *Voting Procedures*. Oxford University Press.

[21] Thomas Eiter and Georg Gottlob. 1995. On the Computational Cost of Disjunctive Logic Programming: Propositional Case. *Annals of Mathematics and Artifficial Intelligence* 15, 3-4 (1995), 289–323.

[22] Edith Elkind, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko. 2017. Properties of Multiwinner Voting Rules. *Social Choice and Welfare* 48, 3 (2017), 599–632.

[23] Ulle Endriss. 2016. Judgment Aggregation. In *Handbook of Computational Social Choice*, F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia (Eds.). Cambridge University Press.

[24] Ulle Endriss. 2016. Judgment Aggregation. In *Handbook of Computational Social Choice*, F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia (Eds.). Cambridge University Press.

[25] Ulle Endriss and Ronald de Haan. 2015. Complexity of the Winner Determination Problem in Judgment Aggregation: Kemeny, Slater, Tideman, Young. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

[26] Ulle Endriss, Ronald de Haan, Jérôme Lang, and Marija Slavkovik. 2020. The Complexity Landscape of Outcome Determination in Judgment Aggregation. *Journal of Artificial Intelligence Research* (2020).

[27] Ulle Endriss, Umberto Grandi, and Daniele Porello. 2012. Complexity of Judgment Aggregation. *Journal of Artificial Intelligence Research* 45 (2012), 481–514.

[28] Patricia Everaere, Sébastien Konieczny, and Pierre Marquis. 2014. On Egalitarian Belief Merging. In *Proceedings of the 14th International Conference on the Principles of Knowledge Representation and Reasoning (KR)*.

[29] Patricia Everaere, Sébastien Konieczny, and Pierre Marquis. 2017. Belief Merging and its Links with Judgment Aggregation. In *Trends in Computational Social Choice*, Ulle Endriss (Ed.). AI Access Foundation, 123–143.

[30] Peter C Fishburn and Steven J Brams. 1983. Paradoxes of Preferential Voting. *Mathematics Magazine* 56, 4 (1983), 207–214.

[31] Duncan K Foley. 1967. Resource Allocation and the Public Sector. *Yale Economic Essays* 7, 1 (1967), 45–98.

[32] Peter Gärdenfors. 1976. Manipulation of Social Choice Functions. *Journal of Economic Theory* 13, 2 (1976), 217–228.

[33] Martin Gebser, Roland Kaminski, Benjamin Kaufmann, and Torsten Schaub. 2012. *Answer Set Solving in Practice*. Morgan & Claypool Publishers.

[34] Martin Gebser, Roland Kaminski, and Torsten Schaub. 2011. Complex Optimization in Answer Set Programming. *Theory and Practics of Logic Programming* 11, 4-5 (2011), 821–839.

[35] Michael Gelfond. 2006. Answer Sets. In *Handbook of Knowledge Representation*, Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter (Eds.). Elsevier.

[36] Umberto Grandi and Ulle Endriss. 2010. Lifting Rationality Assumptions in Binary Aggregation. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*.

[37] Davide Grossi and Gabriella Pigozzi. 2014. *Judgment Aggregation: A Primer*. Synthesis Lectures on Artificial Intelligence and Machine Learning, Vol. 8. Morgan & Claypool Publishers. 1–151 pages.

[38] Ronald de Haan. 2018. Hunting for Tractable Languages for Judgment Aggregation. In *Proceedings of the 16th International Conference on the Principles of Knowledge Representation and Reasoning (KR)*.

[39] Ronald de Haan and Marija Slavkovik. 2017. Complexity Results for Aggregating Judgments using Scoring or Distance-Based Procedures. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS*.

[40] Ronald de Haan and Marija Slavkovik. 2019. Answer Set Programming for Judgment Aggregation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*.

[41] Jerry S Kelly. 1977. Strategy-proofness and Social Choice Functions without Singlevaluedness. *Econometrica* 45, 2 (1977), 439–446.

[42] Johannes Köbler and Thomas Thierauf. 1990. Complexity Classes with Advice. In *Proceedings of the 5th Annual Structure in Complexity Theory Conference*.

[43] Sébastien Konieczny and Ramón Pino Pérez. 2011. Logic Based Merging. *Journal of Philosophical Logic* 40, 2 (2011), 239–270.

[44] Justin Kruger and Zoi Terzopoulou. 2020. Strategic Manipulation with Incomplete Preferences: Possibilities and Impossibilities for Positional Scoring Rules. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

[45] Martin Lackner and Piotr Skowron. 2018. Approval-Based Multi-Winner Rules and Strategic Voting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*.

[46] Jérôme Lang, Gabriella Pigozzi, Marija Slavkovik, and Leendert van der Torre. 2011. Judgment Aggregation Rules Based on Minimization. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*.

[47] Jérôme Lang and Marija Slavkovik. 2014. How Hard is it to Compute Majority-Preserving Judgment Aggregation Rules?. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI)*.

[48] Christian List and Philip Pettit. 2002. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy* 18, 1 (2002), 89–110.

[49] Michael K. Miller and Daniel Osherson. 2009. Methods for Distance-based Judgment Aggregation. *Social Choice and Welfare* 32, 4 (2009), 575–601.

[50] Elchanan Mossel and Omer Tamuz. 2010. Truthful Fair Division. In *Proceedings of the 3rd International Symposium on Algorithmic Game Theory (SAGT)*.

[51] Hervé Moulin. 1988. *Axioms of Cooperative Decision Making*. Econometric Society Monographs, Vol. 15. Cambridge University Press.

[52] Klaus Nehring, Marcus Pivato, and Clemens Puppe. 2014. The Condorcet Set: Majority Voting over Interconnected Propositions. *Journal of Economic Theory* 151 (2014), 268–303.

[53] Ritesh Noothigattu, Snehalkumar S Gaikwad, Edmond Awad, Sohan Dsouza, Iyad Rahwan, Pradeep Ravikumar, and Ariel D Procaccia. 2018. A Voting-Based System for Ethical Decision Making. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*.

[54] Dominik Peters. 2018. Proportionality and Strategyproofness in Multiwinner Elections. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

[55] Gabriella Pigozzi. 2006. Belief Merging and the Discursive Dilemma: An Argument-based Account to Paradoxes of Judgment Aggregation. *Synthese* 152, 2 (2006), 285–298.

[56] John Rawls. 1971. *A Theory of Justice*. Belknap Press.

[57] M Remzi Sanver and William S Zwicker. 2009. One-way Monotonicity as a Form of Strategy-proofness. *International Journal of Game Theory* 38, 4 (2009), 553–574.

[58] Amartya Sen. 1997. *Choice, Welfare and Measurement*. Harvard University Press.

[59] Zoi Terzopoulou and Ulle Endriss. 2018. Modelling Iterative Judgment Aggregation. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*.

[60] William S. Zwicker. 2016. Introduction to the Theory of Voting. In *Handbook of Computational Social Choice*, F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia (Eds.). Cambridge University Press.