Characterizing Non-Manipulability for Multi-agent Argumentation

Weiwei Chen

Institute of Logic and Cognition and Department of Philosophy

Sun Yat-sen University, China

Email: chenww26@mail2.sysu.edu.cn

March 5, 2023

1 Introduction

The field of collective argumentation is concerned with the design and analysis of procedures for combining the opinions of several agents regarding the assessment of either different sets of arguments and/or different attack-relations into a collective outcome [5], for example, when a group of agents are engaged in a debate, each of them indicates that for each argument, whether it is acceptable, and we would like to aggregate individuals' opinions to form a collective decision that is acceptable to the group.

While early work on collective argumentation has focused on applications of the axiomatic method, inspired by closely related work in other areas of social choice theory [19, 8, 3, 11, 22], to derive a host of impossibility and characterization results, more recently attention has shifted towards designing practically useful aggregation rules, a natural question that arises in this context is whether agents have an incentive to misrepresent their own opinion in order to secure an outcome he prefers to the outcome of the function which would have obtained if he had expressed her sincere opinions.

The question of which aggregation rules are manipulable by strategic voting and which are strategy-proof is an important topic in social choice [15, 20] and has gradually received attention in the literature on collective argumentation [19, 9, 4]. Ideally, an aggregation rule that leads individuals to report their opinions truthfully would be considered as a desirable rule. Such rules are crucial for designing systems to support collective argumentation. A manipulable aggregation rule is often viewed negatively, one may find that it is undeserved and undesirable from a social perspective.

Non-manipulability means that there is no opportunity for some individual(s) to manipulate the collective decisions on the acceptance of arguments by expressing untruthful individual opinion(s).

In this paper, we assume that given an argumentation framework, each agent's argumentative stance is represented as a labelling of that argumentation framework. We study that under what condition, an aggregation rule is manipulable. We present two forms of non-manipulability. The first states that expressing untruthful opinions will no change the collective outcome. The second states that an agent expresses untruthful opinions will not lead the collective outcome to her favourable one. By defining these two forms of non-manipulability, we can distinguish the ability of manipulability of different aggregation rules. We also characterize the class of non-manipulable aggregation rules in terms of an independence condition and a monotonicity condition. The aggregation rules we pay attention to in this paper include those that guarantee admissibility, for example, the two aggregation operators introduced in [8], namely the *sceptical (initial) aggregation operator* and the *credulous (initial) aggregation operator*, we show that the sceptical initial aggregation operator is less vulnerable to manipulation than the *credulous initial aggregation operator*. However, both the sceptical aggregation operator and the credulous aggregation operator are manipulable.

Paper overview The paper is organized as follows. Section 2 is a review of relevant concepts from the theory of abstract argumentation and argumentation labelling. Section 3 presents our model, two forms of non-manipulability, and the characterization of non-manipulability of aggregation rules during aggregation of individual labellings. We conclude in Section 5 and point some directions for future work.

2 Argumentation and Labelling

In this section, we will revisit the basic concepts of abstract argumentation models as introduced by Dung [13], as well as the approach to defining argumentation semantics proposed by Caminada [7].

Definition 1 (Argumentation framework)**.** *An argumentation framework is a pair $AF = \langle Arg, \rightharpoonup \rangle$, in which Arg is a set of arguments, and $\rightharpoonup$ is a set of binary relations called attack relations built on Arg.*

Given two arguments $A, B \in Arg$, $A \rightharpoonup B$ represents there is an attack from A to B. When we have an argumentation framework AF that includes a set of arguments and a set of attacks representing conflicts between the arguments, a natural question is how to decide which arguments to accept or reject based on some reasonable criterion. The set of arguments that is chosen according to this criterion is called an extension of AF under the particular semantics being used.

We say that A *attacks* B if $A \rightharpoonup B$ holds for two arguments $A, B \in Arg$. For $\Delta \subseteq Arg$ and $B \in Arg$, we write $\Delta \rightharpoonup B$ (namely, Δ attacks B) in case $A \rightharpoonup B$ for at least one argument $A \in \Delta$. For $\Delta \subseteq Arg$ and $C \in Arg$, we say that Δ *defends* C in the case that Δ attacks all arguments $B \in Arg$ with $B \rightharpoonup C$. We write 2^{Arg} for the powerset of Arg.

Given an argumentation framework $AF = \langle Arg, \rightharpoonup \rangle$, the question arises which subset Δ of the set of arguments Arg to accept. Any such set $\Delta \subseteq Arg$ is called an *extension* of AF. Different criteria have been proposed for choosing an extension. While Dung has defined several semantics, notably complete, grounded, preferred and stable semantics [13], it is worth mentioning that conflict-freeness, being self-defending and admissibility are fundamental properties supposed to be satisfied by extensions of semantics.

Definition 2. *Let $AF = \langle Arg, \rightharpoonup \rangle$ be an argumentation framework and let $\Delta \subseteq Arg$ be a set of arguments. We adopt the following terminology:*

- *Δ is called conflict-free if there are no arguments $A, B \in \Delta$ such that $A \rightharpoonup B$.*
- *Δ is called self-defending if $\Delta \subseteq \{C \mid \Delta$ defends $C\}$.*
- *Δ is called reinstating if $\{C \mid \Delta$ defends $C\} \subseteq \Delta$.*

- Δ *is called admissible if it is both conflict-free and self-defending.*
- Δ *is called complete if it is conflict-free, self-defending, and reinstating.*

Thus, a set of arguments is admissible if it is both conflict-free and self-defending. Furthermore, a set of arguments $\Delta \subseteq Arg$ is a complete extension if it is admissible and includes all arguments it defends, Δ is a preferred extension if it is a maximal complete extension, Δ is the grounded extension if it is the minimal complete extension, and finally, Δ is a stable extension if it is admissible and attacks every argument in $Arg \backslash \Delta$. All of them are admissibility-based in the sense that every extension of such semantics is admissible.

The labelling-based approach is an important approach to define argument acceptability criterion (semantics) in the literature [7]. Given an argument, a labelling associated the argument with a label, which can either be `in`, `out`, or `undec`. The label `in` indicates that the argument is explicitly accepted. The label `out` indicates that the argument is explicitly rejected. The label `undec` indicates that the status of the argument is undecided, i.e., someone abstains from an explicit judgment whether the argument is accepted or rejected.

Definition 3 (labelling)**.** *Let $AF = \langle Arg \rangle$ be an argumentation framework, a labelling $\mathcal{L}$ of AF is a total function mapping any argument into a label: $\mathcal{L} : Arg \to \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$.*

Definition 4. *Given an argument $A \in Arg$, $l(A) \in \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$ denotes the label that is assigned to A.*

Definition 5 (Illegal labelling)**.** *Let $AF = \langle Arg, \rightharpoonup \rangle$ be an argumentation framework, given an argument $A \in Arg$ and a labelling $\mathcal{L}$ of AF:*

(1) *A is illegally labeled `in` iff A is labelled `in` and no attacker of A is labeled `out`*

(2) *A is illegally labeled `out` iff A is labeled `out` and it does not have an attacker that is labeled `in`*

(3) *A is illegally labeled `undec` iff A is labeled `undec` and either all of its attacker are labeled `out` or it has an attacker that is labeled `in`.*

Definition 6 (Admissible labelling)**.** *A labelling is call an* admissible labelling *iff (i) no argument is illegal labeled `in`, and (ii) no argument is illegal labeled `out`.*

We can think of a labelling-based semantic as a function that maps a given argumentation framework to a set of labellings, each labelling in the set can be seen as a reasonable position that an agent can take on an argumentation framework.

The extension-based approach and labelling based approach are two main approaches to the definition of argumentation semantics. The extension-based approach identifies which arguments can be accepted together, which gives rise to extensions. Thus, there are two statuses of an argument, whether it is accepted or rejected. The labelling based approach assigns every argument a label that represents such argument is accepted, rejected or left undecided.

3 Aggregation of Labellings

In this section, we consider whether aggregation rules are manipulable during the aggregation of labellings. We first define two forms of manipulability. Then, we introduce the conditions for non-manipulability and study whether several well-known aggregation rules are manipulable.

Fix an argumentation framework $AF = \langle Arg, \rightharpoonup \rangle$. Let $N = \{1, \ldots n\}$ be a set of agents. Every agent provides a labelling that represents her justification on the status of arguments $\mathcal{L}_i$. Then, we obtain a profile of labellings $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_n)$.

Definition 7. *Let $AF = \langle Arg, \rightharpoonup \rangle$ be an argumentation framework. Let $\mathcal{LBL}$ be a set of possible labellings of AF. An aggregation rule F is a function that maps a profile of labellings into a single labelling:*

$$F : 2^{\mathcal{LBL}} \rightarrow \mathcal{LBL}$$

Note that a labelling is function mapping any argument into a label. Thus, given an aggregation rule F, a profile of labelling $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_n)$, an argument $A \in Arg$, we write $F(\boldsymbol{\mathcal{L}})(A)$ to denote the label assignment of the aggregation rule F on profile $\boldsymbol{\mathcal{L}}$ to A.

Definition 8. *Let $\boldsymbol{\mathcal{L}}$ be a profile of labellings. We let $N^{\boldsymbol{\mathcal{L}}}_{l(A)=l}$ denotes the set of agents that assign label $l \in \{\texttt{in}, \texttt{out}, \texttt{undec}\}$ to A in profile $\boldsymbol{\mathcal{L}}$.*

There are some properties of aggregation rules. In social choice theory, such properties are called axioms [21]. We introduce two of them, namely independence and monotonicity.

Definition 9 (Independence)**.** *An aggregation rule F is said to be* independent *if $N^{\boldsymbol{\mathcal{L}}}_{l(A)=l} = N^{\boldsymbol{\mathcal{L}}'}_{l(A)=l}$ implies $F(\boldsymbol{\mathcal{L}})(A) = l \Leftrightarrow F(\boldsymbol{\mathcal{L}}')(A) = l$ for all profiles $\boldsymbol{\mathcal{L}}$, $\boldsymbol{\mathcal{L}}'$, all arguments $A \in Arg$, and all labels $l \in \{\texttt{in}, \texttt{out}, \texttt{undec}\}$.*

Thus, independence states that the collective label to each argument $A \in Arg$ depends on individual labels to A and not on individual labels to other arguments.

Definition 10 (Monotonicity)**.** *An aggregation rule F is said to be* monotonic *if $N^{\boldsymbol{\mathcal{L}}}_{l(A)=l} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=l}$ implies $F(\boldsymbol{\mathcal{L}})(A) = l \Rightarrow F(\boldsymbol{\mathcal{L}}')(A) = l$ for all profiles $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{L}}'$, all arguments $A \in Arg$, and all labels $l \in \{\texttt{in}, \texttt{out}, \texttt{undec}\}$.*

Informally speaking, monotonicity states that an additional individual's support for the label to some argument $A \in Arg$ never changes the collective label to A.

3.1 Two forms of non-manipulability

Before presenting formal definitions of non-manipulability, some definitions are needed. We say that two profiles are i_A-variants of each other if they coincide for all individual labels except possibly $i(A)$. We say that an aggregation rule F is s-manipulable at the profile $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)$ by individual i on $A \in Arg$ if $\mathcal{L}_i(A) \neq F(\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)(A)$, but $\mathcal{L}_i(A) = F(\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)(A)$ for some i_A-variant $F(\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)$, i.e., an agent i s-manipulates successfully only if she has changed the collective assignment of A to her true belief on the assignment of A.

Definition 11 (Non-s-manipulability)**.** *An aggregation rule is said to be* non-s-manipulable *if for any profile $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)$ with $F(\boldsymbol{\mathcal{L}})(A) \neq \mathcal{L}_i(A)$, then $F(\boldsymbol{\mathcal{L}}')(A) \neq \mathcal{L}_i(A)$ holds for every profile $\boldsymbol{\mathcal{L}}' = (\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)$ that is an i_A-variant of $\boldsymbol{\mathcal{L}}$.*

The non-s-manipulability condition states that if an agent i disagrees with the aggregation result on the assignment of argument A, then every variant choice of i still disagrees with the aggregation result, i.e., no agent has the incentive to report her opinion insincerely.

We now introduce another form of the notion of manipulability. We say that an aggregation rule F is t-manipulable at the profile $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)$ by individual i on $A \in Arg$ if $\mathcal{L}_i(A) \neq F(\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)(A)$, but $F(\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)(A) \neq F(\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)(A)$ for some i_A-variant $F(\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)$, i.e., an agent i t-manipulates successfully only if she has changed the collective assignment of A to the label that is different from the original one.

Definition 12 (Non-t-manipulability)**.** *An aggregation rule is said to be* non-t-manipulable *if for any profile* $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)$ *with* $F(\boldsymbol{\mathcal{L}})(A) \neq \mathcal{L}_i(A)$*, then* $F(\boldsymbol{\mathcal{L}}')(A) = F(\boldsymbol{\mathcal{L}})(A)$ *holds for every profile* $\boldsymbol{\mathcal{L}}' = (\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)$ *that is an* i_A*-variant of* $\boldsymbol{\mathcal{L}}$*.*

The non-t-manipulability condition states that if an agent i disagrees with the aggregation result on the assignment of argument A, then every variant choice of i will not change the aggregation result.

The connection between non-t-manipulability and non-s-manipulability is straightforward, as described in the following proposition.

Proposition 1. *Given an aggregation rule* F*, if* F *is non-t-manipulable, then* F *is non-s-manipulable.*

3.2 Examples of aggregation rules

We now introduce examples of aggregation rules considered in this section, including the sceptical and credulous initial aggregation operator, the sceptical and credulous aggregation operator and the argument-wise plurity rule.

Definition 13 (Caminada and Pigozzi, 2011)**.** *An aggregation rule* F *is the* sceptical initial aggregation operator *if for all profiles* $\boldsymbol{\mathcal{L}}$*:*

$$\begin{aligned} F(\boldsymbol{\mathcal{L}}) = \{(A, \texttt{in}) \mid \forall i \in N : \mathcal{L}_i(A) = \texttt{in}\} \cup \\ \{(A, \texttt{out}) \mid \forall i \in N : \mathcal{L}_i(A) = \texttt{out}\} \cup \\ \{(A, \texttt{undec}) \mid \exists i \in N : \mathcal{L}_i(A) \neq \texttt{in} \wedge \exists j \in N : \mathcal{L}_j(A) \neq \texttt{out}\}. \end{aligned}$$

The skeptical initial aggregation operator assigns `in` to an argument only if each agent assigns `in` to it and assigns `out` to an argument only if each agent assigns `out` to it. Otherwise, it will be assigned `undec`. In other words, this aggregation rule will only assign `in` or `out` for an argument if there is a unanimous agreement among all agents, otherwise it will label the argument as `undec`.

Definition 14 (Caminada and Pigozzi, 2011)**.** *An aggregation rule* F *is the* credulous initial aggregation operator *if for all profiles* $\boldsymbol{\mathcal{L}}$*:*

$$\begin{aligned} F(\boldsymbol{\mathcal{L}}) = \{(A, \texttt{in}) \mid \exists i \in N : \mathcal{L}_i(A) = \texttt{in} \wedge \nexists j \in N : \mathcal{L}_j(A) = \texttt{out}\} \cup \\ \{(A, \texttt{out}) \mid \exists i \in N : \mathcal{L}_i(A) = \texttt{out} \wedge \nexists j \in N : \mathcal{L}_j(A) = \texttt{in}\} \cup \\ \{(A, \texttt{undec}) \mid \forall i \in N : \mathcal{L}_i(A) = \texttt{undec} \\ \vee (\exists i \in N : \mathcal{L}_i(A) = \texttt{in} \wedge \exists j \in N : \mathcal{L}_j(A) = \texttt{out})\}. \end{aligned}$$

The credulous initial aggregation operator assigns `in` to an argument A if there is at least one agent who assigns `in` to A, and no one assigns `out` to A, assigns `out` to A if there is at least one agent who assigns `out` to A, and no one assigns `in` to A. It assigns `undec` to A if all agents assign `undec` to A or if there is at least one agent assigns `in` and at least one agent assigns `out` to A.

So, the main difference between the two definitions lies in their approach towards assigning an argument as `in` or `out`. The skeptical initial aggregation operator requires unanimity, while the credulous initial aggregation operator only requires a single vote in favor of `in` or `out`. Both the sceptical initial aggregation operator and the credulous initial aggregation operator assign exactly one label to each argument, thus, they are well-defined. On the downside, both of them fail to satisfy admissibility.

Definition 15. *Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two labellings of argumentation framework AF. We say that $\mathcal{L}_1$ is less of equally committed as $\mathcal{L}_2$ iff $in(\mathcal{L}_1) \subseteq in(\mathcal{L}2)$ and $out(\mathcal{L}_1) \subseteq out(\mathcal{L}2)$.*

The relation between labelling defined in Defintion 15 is a partial order [8]. Thus, we say that labelling $\mathcal{L}_2$ is bigger or equal to labelling $\mathcal{L}_1$ if $\mathcal{L}_1$ is less or equally committed as $\mathcal{L}_2$.

Theorem 2 (Caminada and Pigozzi, 2011)**.** *Let $\mathcal{L}$ be a labelling of argumentation framework AF, the set of admissible labelling that are smaller or equal ot $\mathcal{L}$ has a (unique) biggest element.*

Definition 16. *Let $\mathcal{L}$ be a labelling of argumentation framework AF. The down-admissible labelling of $\mathcal{L}$ is the biggest element of the set of all admissible labellings that are smaller or equal to $\mathcal{L}$.*

Definition 17. *An aggregation rule F is the* sceptical aggregation operator *if it is the down-admissible labelling of the sceptical initial aggregation operator.*

Definition 18. *An aggregation rule F is the* credulous aggregation operator *if it is the down-admissible labelling of the credulous initial aggregation operator.*

Thus, the sceptical aggregation operator and the credulous aggregation operator trivially preserve admissibility. Further, it is guaranteed that if an argument is accepted by the sceptical aggregation operator, it is also accepted by every individual agent.

We now consider another aggregation rule called argument-wise plurality rule, introduced by Rahwan and Tohmé [19]. The argument-wise plurality rule accepts a label if it receives the most support from the agents in the group.

Definition 19 (Rahwan and Tohmé, 2011)**.** *An aggregation rule F is the* argument-wise plurality rule *if for any argument $A \in Arg$, for all profiles $\boldsymbol{\mathcal{L}}$:*

$$F(\boldsymbol{\mathcal{L}})(A) = l \in \{\texttt{in}, \texttt{out}, \texttt{undec}\} \text{ if and only if } |i : \mathcal{L}_i(A) = l| > max_{l' \neq l}|i : \mathcal{L}_i(A) = l'|.$$

Otherwise, $F(\boldsymbol{\mathcal{L}})(A) = \emptyset$.

It should be noted that even if there are ties, F can be considered well-defined because the empty set is an element of every set. However, if $F(\mathcal{L})(A) = \emptyset$ for a certain $A \in Arg$, then the output of F is clearly not a legal labelling.

3.3 Characterization results

Proposition 3. *The sceptical initial aggregation operator is monotonic.*

By comparison, note that the credulous initial aggregation operator is not monotonic. This can be easily verified by the following example. Suppose that F is the credulous initial aggregation operator. Consider two profiles $\boldsymbol{\mathcal{L}}$, $\boldsymbol{\mathcal{L}}'$. Suppose that $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\texttt{in}} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\texttt{in}}$ and $F(\boldsymbol{\mathcal{L}})(A) = \texttt{in}$. In addition, we suppose that $\mathcal{L}'_i(A) = \texttt{out}$ for some agent $i \in N$. Thus, $F(\boldsymbol{\mathcal{L}}')(A) = \texttt{undec}$. Therefore, in this case, $F(\boldsymbol{\mathcal{L}})(A)$ is different from $F(\boldsymbol{\mathcal{L}}')(A)$.

Proposition 4. *The sceptical initial aggregation operator is independent.*

However, the credulous initial aggregation operator is not independent. To see this, consider a pair of profiles $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{L}}'$, and F be the credulous initial aggregation operator. If $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\texttt{in}} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\texttt{in}}$ and $F(\boldsymbol{\mathcal{L}})(A) = \texttt{in}$, then there is some agent $i \in N$ with $\mathcal{L}_i(A) = \texttt{in}$ and not agent $j \in N$ with $\mathcal{L}_j(A) = \texttt{out}$. Suppose that there is some agent j who labels A to $\texttt{out}$ in $\boldsymbol{\mathcal{L}}'$, i.e., $\mathcal{L}_j(A)' = \texttt{out}$. Thus, $\mathcal{L}_i(A)' = \texttt{in}$, $\mathcal{L}_j(A)' = \texttt{out}$ implies that $F(\boldsymbol{\mathcal{L}}')(A) = \texttt{undec}$.

We now introduce a simple condition of non-t-manipulability and characterize the class of non-manipulable aggregation rules, i.e., any rule that satisfies independence and monotonicity is non-t-manipulable.

Theorem 5. *An aggregation rule F is non-t-manipulable iff F is independent and monotonic.*

Theorem 5 is adapted from [12]. The reader may have noticed that the concept of manipulation is non-t-manipulation. When we turn to non-s-manipulation, only one direction is the case.

Theorem 6. *Given an aggregation rule F, if F is independent and monotonic, then F is non-s-manipulable.*

Theorem 6 is an immediate consequence of Proposition 1 and Theorem 5 (the right to the left direction). The other direction of Theorem is not the case, to show this, we show that there is an aggregation rule F that is non-s-manipulable, but not monotonic. Take a pair of profiles $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)$ and $\boldsymbol{\mathcal{L}}' = (\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)$. Assume $F(\boldsymbol{\mathcal{L}})(A) = \texttt{out}$ and $\mathcal{L}_i(A) \neq \texttt{out}$. Then, according to the condition of non-s-manipulation, $F(\boldsymbol{\mathcal{L}}')(A) \neq \texttt{out}$ for every i_A-variant $F(\boldsymbol{\mathcal{L}}')$. In the meantime, $F(\boldsymbol{\mathcal{L}})(A) \neq F(\boldsymbol{\mathcal{L}}')(A)$ is possible due to the fact that the condition of non-s-manipulation does not require the collective outcome of the i_A-variant profile be the same as the original collective outcome. But we have assumed $N^{\boldsymbol{\mathcal{L}}}_{l(A)=l} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=l}$. Thus, we are done.

According to Proposition 3 and Proposition 4, we get that the sceptical initial aggregation operator is monotonic and independent. Thus, a corollary of Theorem 5 is obtained.

Corollary 7. *The sceptical initial aggregation operator is non-t-manipulable and non-s-manipulable.*

Note that using a non-manipulable sceptical initial aggregation operator does not guarantee an admissible outcome. In the paper by Caminada and Pigozzi [8], they proposed a sceptical aggregation operator as a subsequent step to ensure admissibility. However, it's worth noting that this new operator can be manipulated, as demonstrated in the following example.

Example 1. To see that the sceptical aggregation operator is manipulable, let us consider the argumentation framework $AF = \langle \{A, B, C, D\}, \{A \rightharpoonup B, B \rightharpoonup A, A \rightharpoonup C, B \rightharpoonup C, C \rightharpoonup D\} \rangle$ as illustrated

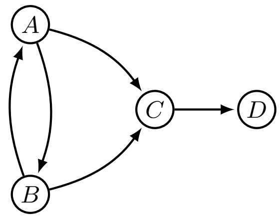

Figure 1: Scenarios used in Example 1

in Figure 1. Let F be the sceptical initial aggregation operator and let F' be the sceptical aggregation operator. Take a pair of profiles $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \mathcal{L}_2)$ and $\boldsymbol{\mathcal{L}'} = (\mathcal{L}_1, \mathcal{L}'_2)$ in which $\mathcal{L}_2$ is the truthful labelling of agent 2 and $\mathcal{L}'_2$ is a variant of $\mathcal{L}_2$. Assume that:

$$\mathcal{L}_1(A) = \texttt{in}, \mathcal{L}_1(B) = \texttt{out}, \mathcal{L}_1(C) = \texttt{out}, \mathcal{L}_1(D) = \texttt{in};$$

$$\mathcal{L}_2(A) = \texttt{out}, \mathcal{L}_2(B) = \texttt{in}, \mathcal{L}_2(C) = \texttt{out}, \mathcal{L}_2(D) = \texttt{in}.$$

We get that $F(\boldsymbol{\mathcal{L}})(A) = \texttt{undec}, F(\boldsymbol{\mathcal{L}})(B) = \texttt{undec}, F(\boldsymbol{\mathcal{L}})(C) = \texttt{out}, F(\boldsymbol{\mathcal{L}})(D) = \texttt{in}$. The down-admissible labelling of the sceptical initial aggregation operator F is:

$$F'(\boldsymbol{\mathcal{L}})(A) = \texttt{undec}, F'(\boldsymbol{\mathcal{L}})(B) = \texttt{undec}, F'(\boldsymbol{\mathcal{L}})(C) = \texttt{undec}, F'(\boldsymbol{\mathcal{L}})(D) = \texttt{undec}.$$

Now assume that agent 2 falsely submits the labelling $\mathcal{L}'_2$ in which:

$$\mathcal{L}'_2(A) = \texttt{in}, \mathcal{L}'_2(B) = \texttt{out}, \mathcal{L}'_2(C) = \texttt{out}, \mathcal{L}'_2(D) = \texttt{in}.$$

The result of the sceptical initial aggregation operator F is $F(\boldsymbol{\mathcal{L}'})(A) = \texttt{in}, F(\boldsymbol{\mathcal{L}'})(B) = \texttt{out}, F(\boldsymbol{\mathcal{L}'})(C) = \texttt{out}, F(\boldsymbol{\mathcal{L}'})(D) = \texttt{in}$. The down-admissible labelling of the sceptical initial aggregation operator F is:

$$F'(\boldsymbol{\mathcal{L}'})(A) = \texttt{in}, F'(\boldsymbol{\mathcal{L}'})(B) = \texttt{out}, F'(\boldsymbol{\mathcal{L}'})(C) = \texttt{out}, F'(\boldsymbol{\mathcal{L}'})(D) = \texttt{in}.$$

Note that $F'(\boldsymbol{\mathcal{L}})(C) = \texttt{undec}, \mathcal{L}_2(C) = \texttt{out}, F'(\boldsymbol{\mathcal{L}})(C) \neq \mathcal{L}_2(C)$. However,

$$F'(\boldsymbol{\mathcal{L}'})(C) = \texttt{out}, F'(\boldsymbol{\mathcal{L}})(C) = \mathcal{L}_2(C).$$

Thus, agent 2 s-manipulates successfully. Furthermore, $F'(\boldsymbol{\mathcal{L}})(C) \neq F'(\boldsymbol{\mathcal{L}'})(C)$. Thus, agent 2 t-manipulates successfully as well. $\triangle$

In Example 1, if we replace the sceptical initial aggregation operator with the *credulous initial aggregation operator* and replace the sceptical aggregation operator with the *credulous aggregation operator*, we can still get that the credulous aggregation operator is manipulable. This example shows that the sceptical aggregation operator and the credulous aggregation operator sometimes incentivise agents to submit their untruthful labellings, because doing so may result in a more preferred outcome.

We now consider another aggregation rule called argument-wise plurality rule, introduced by Rahwan and Tohmé [19]. The argument-wise plurality rule accepts a label if it receives the most support from the agents in the group.

Proposition 8. *The argument-wise plurality rule is not non-t-manipulable.*

Proof. Let F be the argument-wise plurality rule. Given an argument $A \in \mathit{Arg}$. Assume that there are four agents, two of them assign `in` to A, one assign `out` to A, and one assign `undec` to A, then $F(\boldsymbol{\mathcal{L}}) = \texttt{in}$. However, if the agent who assigns `out` to A changes her opinion and reassigns `undec` to A, then the label of A will be $\emptyset$. Thus, the agent t-manipulable successfully. □

Proposition 9. *The argument-wise plurality rule is non-s-manipulable.*

Proof. Assume the argument-wise plurality rule is denoted by F. Let i be an arbitrary agent with labelling $\mathcal{L}_i$. Let $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)$ be a profile of labellings in which $\mathcal{L}_i$ is the truthful labelling of agent i. We also define $\boldsymbol{\mathcal{L}'} = (\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)$ when i reports $\mathcal{L}'_i$ (given the same labellings reported by others). By the definition of non-s-manipulability (Definition 11), the goal is to prove that $F(\boldsymbol{\mathcal{L}})(A) \neq \mathcal{L}_i(A)$ implies $F(\boldsymbol{\mathcal{L}'})(A) \neq \mathcal{L}_i(A)$.

If $F(\boldsymbol{\mathcal{L}}) \in \{\texttt{in}, \texttt{out}, \texttt{undec}\}$, assume $F(\boldsymbol{\mathcal{L}}) \neq \mathcal{L}_i(A)$ for some A. From the definition of the plurality rule, it follows that the plurality agreed on $F(\boldsymbol{\mathcal{L}})(A)$ as opposed to any other labelling of A. Formally:

$$|j : \mathcal{L}_j(A) = l| > \max_{l' \neq l} |j : \mathcal{L}_j(A) = l'|$$

Since $F(\boldsymbol{\mathcal{L}})(A) \neq \mathcal{L}_i(A)$, it follows that the plurality disagrees with i's labelling of A:

$$|j : \mathcal{L}_j(A) = l| > |j : \mathcal{L}_j(A) = \mathcal{L}_i(A)|$$

We now show that the above inequality does not change with any $\mathcal{L}'_i \neq \mathcal{L}_i$. We consider three possible cases:

(i) $\mathcal{L}'_i(A) = \mathcal{L}_i(A)$ but $\mathcal{L}'_i(B) \neq \mathcal{L}_i(B)$ for one or more $B \neq A$. However, this does not affect the inequality in question, since the value of $F(\boldsymbol{\mathcal{L}})(A)$ is solely determined by the votes cast for A.

(ii) $\mathcal{L}'_i(A) = F(\boldsymbol{\mathcal{L}})(A)$. Clearly, in this case, one additional agent is counted as voting for the label $F(\boldsymbol{\mathcal{L}})(A)$, while one fewer agent is counted as voting for the label $\mathcal{L}_i(A)$.

(iii) $\mathcal{L}'_i(A) = \{\texttt{in}, \texttt{out}, \texttt{undec}\} \backslash F(\boldsymbol{\mathcal{L}})(A) \cup \mathcal{L}_i(A)$. In this situation, the number of agents who vote for label $F(\boldsymbol{\mathcal{L}})(A)$ remains unchanged, while the number of agents who vote for label $\mathcal{L}_i(A)$ decreases by one.

In any case, the above inequality remains the same, so it will always be the case that $F(\boldsymbol{\mathcal{L}'})(A) \neq \mathcal{L}_i(A)$.

If $F(\boldsymbol{\mathcal{L}})(A) = \emptyset$ for some A, then there exist three distinct labels l, l', and l'' such that the number of agents who have label l or l' in argument A is greater than the number of agents who have label l'' in argument A, i.e., $|j : \mathcal{L}_j(A) = l| = |j : \mathcal{L}_j(A) = l'| > |j : \mathcal{L}_j(A) = l''|$. Suppose that $\mathcal{L}_i(A) = l$, then, if $\mathcal{L}'_i(A) = l'$ or $\mathcal{L}'_i(A) = l''$, then $F(\boldsymbol{\mathcal{L}'})(A) \neq \mathcal{L}_i(A)$. □

The relationship between non-s-manipulability and non-t-manipulability is such that the second condition is a sufficient condition for the first as shown by Proposition 1. However, Proposition 8 and Proposition 9 show that a particular aggregation rule (namely, the argument-wise plurality rule) satisfies the first but not the second. Therefore, it is possible to differentiate between these two forms of manipulability not only by their implication relation but also by the use of aggregation rules. In other words, there exist aggregation rules that can effectively discriminate between these two forms of manipulability.

4 Related work

The condition of non-manipulability and for characterizing the class of non-manipulable aggregation rules is similar to the one introduced by Gibbard [15] and Satterthwaite [20] for preference aggregation. They independently prove that no voting rule that satisfies some fairly weak conditions is immune to strategic manipulation. The conditions include that the preferences reported by individual agents are complete, as well as the outcome is resolute in the sense that there is only a single winner. Dietrich and List [12] study strategic manipulation in judgment aggregation [17, 16, 14], a branch of social choice that studies the aggregatio of logically connected propositions. They show that only rules that cannot guarantee the consistency of outcomes are immune to manipulation [12]. Our characterization of non-manipulability in terms of monotonicity and independence during aggregation of individual extensions may be considered a conceptual—albeit not technical—generalisation of the theorem by Dietrich and List [12].

Caminada et al., [9] also consider the non-manipulability of aggregation operators in the context of collective argumentation. The aggregation rules they focus on include sceptical aggregation operator and sceptical aggregation operator introduced by Caminada and Pigozzi [8]. These two aggregation rules were built from sceptical initial aggregation operator and sceptical initial aggregation operator respectively. Caminada et al., [9] investigate the relationship between Pareto optimality and strategyproofness. Notably, they derive the preferences for agents over permissible labellings by using a notion of distance. This work has later been continued by Awad et al., [4].

Rahwan and Tohmé [19] also considered the problem of manipulability in the context of labelling aggregation. They defined the condition of strategy-proofness based on the preferences of agents and showed that the argument-wise plurality rule, which is based on the idea that an assessment of an argument is chosen if it is submitted by the majority regardless of what the minority thinks, is strategy-proof.

Botan and Endriss [6] propose a weaken notion of strategyproofness, where immunity to manipulation only applies when either the truthful profile of individual opinions or the outcome of a potential manipulator is majority-consistent. Botan and Endriss argue that this definition represents a balance for aggregation methods in practical use and demonstrate that several significant methods are strategyproof under this definition. Specifically, all rules within the additive majority family [18], including Kemeny and Slater rules, are strategyproof.

The aggregation of individual labellings into a collective labelling by a group of individuals is a topic of increasing discussion in the literature [5]. It expands on earlier issues in social choice, particularly preference aggregation in the Condorcet-Arrow approach [1, 2]. Recently, Chen [10] examines aggregation of individual extensions, focusing on ensuring admissibility during the process. He assumes that the semantic properties of individual extensions may differ from the collective extension obtained by the aggregation rule. Preserving admissibility is crucial, as the paper investigates, if aggregation rules can be manipulated while still maintaining admissible extensions or labellings. Rahwan and Tohmé's work [19] investigate the "argument-wise plurality rule" which is similar in spirit to the majority rule. They note that the rule does not preserve the completeness of labellings, which is an important property that we focus on in our paper. Awad et al. [3] continued this line of work and identified special classes of argumentation frameworks for better results.

5 Conclusion

In this paper, we have explored how non-manipulability of aggregation rules can be achieved in multi-agent argumentation problems. Non-manipulability means that agents have no incentive to report insincere opinions on the acceptance of arguments. This is important as collective argumentation problems arise in many real-world decision-making situations.

We have used a three-valued labelling-based approach to represent the stances of individual agents and characterized non-manipulability conditions by monotonicity and independence. We have also introduced two forms of non-manipubility to measure the degree of manipulability.

Appendix

Proposition 10. *The sceptical initial aggregation operator is monotonic.*

Proof. Let F the sceptical initial aggregation operator, and let $\boldsymbol{\mathcal{L}}$, $\boldsymbol{\mathcal{L}}'$ be two profiles. Take an argument $A \in Arg$. Suppose that $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{in}} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{in}}$. If $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{in}$. Then, it must be the case that all agents label A $\mathtt{in}$ in profile $\boldsymbol{\mathcal{L}}$, i.e., $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{in}} = \{1, \cdots, n\}$. Thus, all agents assign $\mathtt{in}$ to A in $\boldsymbol{\mathcal{L}}'$ as well, i.e., $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{in}$. If $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{out}} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{out}}$ and $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{out}$. This case is similar to the case that $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{in}$, which will guarantee $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{out}$

It remains to show that for the case that $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{undec}} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{undec}}$ and $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{undec}$, $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{undec}$ is the case. By $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{undec}$, we get that (i) there is some agent $i \in N$ with $\mathcal{L}_i(A) = \mathtt{in}$ and some agent $j \in N$ with $\mathcal{L}_j(A) = \mathtt{out}$, or (ii) there is some agent $k \in N$ with $\mathcal{L}_k(A) = \mathtt{undec}$. For (i), we know that $\mathcal{L}_i(A) = \mathtt{in}$ and $\mathcal{L}_j(A) = \mathtt{out}$ by the assumption that $N^{\boldsymbol{\mathcal{L}}}_{l(A)} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)}$, which lead to $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{undec}$ according to the definition of sceptical initial aggregation operator. For (ii), since we have $N^{\boldsymbol{\mathcal{L}}}_{l(A)} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)}$, we know that $\mathcal{L}_k(A) = \mathtt{undec}$ as well, which will lead to $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{undec}$ as expected. $\square$

Proposition 11. *The sceptical initial aggregation operator is independent.*

Proof. Let F be the sceptical initial aggregation operator. Let us consider a pair of profile $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{L}}'$. If $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{in}} = N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{in}}$ and $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{in}$. Then, it must be the case that all agents label A $\mathtt{in}$ in profile $\boldsymbol{\mathcal{L}}$, i.e., $N^{\boldsymbol{\mathcal{L}}}_{l(A)} = \{1, \cdots, n\}$. Thus, all agents assign $\mathtt{in}$ to A in $\boldsymbol{\mathcal{L}}'$ as well, i.e., $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{in}$. If $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{out}} = N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{out}}$ and $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{out}$. This case is similar to the case that $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{in}$, which will guarantee $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{out}$

It remains to show that for the case that $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{undec}} = N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{undec}}$ and $F(\boldsymbol{\mathcal{L}})(A) = \mathtt{undec}$, $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{undec}$ is the case. For this scenario, there are two possible cases, namely (i) there is some agent $i \in N$ with $\mathcal{L}_i(A) = \mathtt{in}$ and some agent j with $\mathcal{L}_j = \mathtt{out}$, or (ii) there is some agent k with $\mathcal{L}_k(A) = \mathtt{undec}$. Recall that $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{undec}} = N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{undec}}$. For (i), we know that $\mathcal{L}'_i(A) = \mathtt{in}$ and $\mathcal{L}'_j(A) = \mathtt{out}$ in $\boldsymbol{\mathcal{L}}'$ as well. For (ii), we get that $\mathcal{L}_k(A) = \mathtt{undec}$ in $\boldsymbol{\mathcal{L}}'$. Thus, both cases will lead to $F(\boldsymbol{\mathcal{L}}')(A) = \mathtt{undec}$ as expected. $\square$

Theorem 12. *An aggregation rule F is non-t-manipulable iff F is independent and monotonic.*

Proof. We first demonstrate that the direction from the right to the left is the case. Let F be an independent and monotonic aggregation rule. We need to show that no agent has the incentive to vote insincerely. Consider profile $\boldsymbol{\mathcal{L}} = \{\mathcal{L}_1, \ldots, \mathcal{L}_n\}$, an argument $A \in Arg$, and an agent $i \in N$ whose sincere ballot is $\mathcal{L}_i$, and $\mathcal{L}_i(A) \neq F(\boldsymbol{\mathcal{L}})(A)$. Consider another profile $\boldsymbol{\mathcal{L}'} = \{\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n\}$ for which the ballot for i has been replaced by $\mathcal{L}'_i$. We are going to show that $F(\boldsymbol{\mathcal{L}'})(A) = F(\boldsymbol{\mathcal{L}})(A)$ holds, i.e., the aggregation result still disagrees with agent i on the label assignment of A. If $\mathcal{L}_i(A) = \mathcal{L}_i(A)'$, then $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{1}} = N^{\boldsymbol{\mathcal{L}}'}_{l(A)=l}$ for all $l \in \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$. It follows that $F(\boldsymbol{\mathcal{L}})(A) = F(\boldsymbol{\mathcal{L}'})(A)$ by the independence (so we can consider every argument independently) of F. If $\mathcal{L}_i(A) \neq \mathcal{L}_i(A)'$, then there can be two possibilities, namely (i) $\mathcal{L}_i(A)' = F(\boldsymbol{\mathcal{L}})(A)$ and (ii) $\mathcal{L}_i(A)' \neq F(\boldsymbol{\mathcal{L}})(A)$. For (i), $N^{\boldsymbol{\mathcal{L}}}_{l(A)=l} \subset N^{\boldsymbol{\mathcal{L}}'}_{l(A)=l}$ is the case; for (ii), $N^{\boldsymbol{\mathcal{L}}}_{l(A)=l} = N^{\boldsymbol{\mathcal{L}}'}_{l(A)=l}$ is the case. Thus, for both scenarios, $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{1}} \subseteq N^{\boldsymbol{\mathcal{L}}'}_{l(A)=\mathtt{1}}$ is the case for all $l \in \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$. Then, by monotonicity on $\boldsymbol{\mathcal{L}}$, we get that $F(\boldsymbol{\mathcal{L}})(A) = F(\boldsymbol{\mathcal{L}'})(A)$. We are done.

We then prove the direction from the left to the right. We first show that if F is non-t-manipulable, then F is monotonic. Given an argument $A \in \mathit{Arg}$ and two profiles $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_i, \ldots, \mathcal{L}_n)$ and $\boldsymbol{\mathcal{L}'} = (\mathcal{L}_1, \ldots, \mathcal{L}'_i, \ldots, \mathcal{L}_n)$. Suppose that $F(\boldsymbol{\mathcal{L}})(A) = l$, $\mathcal{L}_i(A) \neq l$, and $\mathcal{L}'_i(A) = l$, we are going to show that $F(\boldsymbol{\mathcal{L}'})(A) = l$ holds. If $F(\boldsymbol{\mathcal{L}})(A) = l$, then $\mathcal{L}_i(A) \neq F(\boldsymbol{\mathcal{L}})(A)$, which implies $\mathcal{L}_i(A) \neq F(\boldsymbol{\mathcal{L}'})(A)$ (if not, the agent will manipulate successfully by submitting her insincere ballot $\mathcal{L}_j(A)$). Thus, $F(\boldsymbol{\mathcal{L}'})(A) = l$ by non-t-manipulation.

We then demonstrate that if F is non-t-manipulable, then F is independent. Given a pair of profiles $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{L}'}$, as well as an argument $A \in \mathit{Arg}$. Assume $N^{\boldsymbol{\mathcal{L}}}_{l(A)=\mathtt{1}} = N^{\boldsymbol{\mathcal{L}'}}_{l(A)=l}$, to demonstrate that F is independent, we need to show that $F(\boldsymbol{\mathcal{L}})(A) = F(\boldsymbol{\mathcal{L}'})(A)$. Initially, we replace $\mathcal{L}_1$ of $\boldsymbol{\mathcal{L}}$ with $\mathcal{L}'_1$ to obtain a new profile $\boldsymbol{\mathcal{L}_1}$, which is the same as $\boldsymbol{\mathcal{L}}$ except for $\mathcal{L}_1$. We now prove $F(\boldsymbol{\mathcal{L}'})(A) = F(\boldsymbol{\mathcal{L}_1})(A)$ is the case. If not, then it allows agent i to manipulate: (i) agent i has the incentive to submit $\mathcal{L}_1$ to replace her sincere choice $\mathcal{L}$ to manipulate, or (ii) agent i submit $\mathcal{L}$ to replace her sincere choice $\mathcal{L}_1$ to manipulate. We continue the above process, and in this time we replace $\mathcal{L}_2$ of $\mathcal{L}_1$ of $\boldsymbol{\mathcal{L}_1}$ with $\mathcal{L}'_2$ to obtain a new profile $\boldsymbol{\mathcal{L}_2}$. Using the same construction, we get that $F(\boldsymbol{\mathcal{L}_1})(A) = F(\boldsymbol{\mathcal{L}_2})(A)$. We repeat the above process and learn that $F(\boldsymbol{\mathcal{L}_i})(A) = F(\boldsymbol{\mathcal{L}_{i+1}})(A)$ is the case for all $i \in \{1, \ldots, n-1\}$. Combining the above facts we get that $F(\boldsymbol{\mathcal{L}})(A) = F(\boldsymbol{\mathcal{L}'})(A)$, which proves independence of F. $\square$

References

[1] K. J. Arrow. *Social Choice and Individual Values.* John Wiley and Sons, 2nd edition, 1963. First edition published in 1951.

[2] K. J. Arrow, A. K. Sen, and K. Suzumura, editors. *Handbook of Social Choice and Welfare.* North-Holland, 2002.

[3] E. Awad, R. Booth, F. Tohmé, and I. Rahwan. Judgement aggregation in multi-agent argumentation. *Journal of Logic and Computation*, 27(1):227–259, 2017.

[4] E. Awad, M. W. Caminada, G. Pigozzi, M. Podlaszewski, and I. Rahwan. Pareto optimality and strategy-proofness in group argument evaluation. *Journal of Logic and Computation*, 27(8): 2581–2609, 2017.

[5] G. A. Bodanza, F. A. Tohmé, and M. R. Auday. Collective argumentation: A survey of aggregation issues around argumentation frameworks. *Argument & Computation*, 8(1):1–34, 2017.

[6] S. Botan and U. Endriss. Majority-strategyproofness in judgment aggregation. 2020.

[7] M. Caminada. On the issue of reinstatement in argumentation. In *Proc. 10th European Conference on Logics in Artificial Intelligence (JELIA)*. Springer, 2006.

[8] M. Caminada and G. Pigozzi. On judgment aggregation in abstract argumentation. *Journal of Autonomous Agents and Multiagent Systems*, 22(1):64–102, 2011.

[9] M. Caminada, G. Pigozzi, and M. Podlaszewski. Manipulation in group argument evaluation. In *Twenty-Second International Joint Conference on Artificial Intelligence*, pages 121–126, 2011.

[10] W. Chen. Guaranteeing admissibility of abstract argumentation frameworks with rationality and feasibility constraints. *Journal of Logic and Computation*, 31(8):2133–2158, 2021.

[11] W. Chen and U. Endriss. Aggregating alternative extensions of abstract argumentation frameworks: Preservation results for quota rules. In *Proceedings of the 7th International Conference on Computational Models of Argument (COMMA)*. IOS Press, 2018.

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

[13] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995.

[14] U. Endriss. Judgment aggregation. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 17, pages 399–426. Cambridge University Press, 2016.

[15] A. Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41:587–601, 1973.

[16] D. Grossi and G. Pigozzi. *Judgment Aggregation: A Primer*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2014.

[17] C. List and P. Pettit. Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(1):89–110, 2002.

[18] K. Nehring and M. Pivato. Majority rule in the absence of a majority. *Journal of Economic Theory*, 183:213–257, 2019.

[19] I. Rahwan and F. A. Tohmé. Collective argument evaluation as judgement aggregation. In *Proc. 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS, 2010.

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

[21] A. K. Sen. Social choice theory. In K. J. Arrow and M. D. Intriligator, editors, *Handbook of Mathematical Economics*, volume 3, pages 1073–1181. North-Holland, 1986.

[22] F. A. Tohmé, G. A. Bodanza, and G. R. Simari. Aggregation of attack relations: A social-choice theoretical analysis of defeasibility criteria. In *Proc. 5th International Symposium on Foundations of Information and Knowledge Systems (FoIKS)*. Springer, 2008.