# On the Complexity of Measuring Disagreement with Interpolants

JANDSON S RIBEIRO, Cardiff University, UK
VIORICA SOFRONIE-STOKKERMANS, University of Koblenz, Germany
MATTHIAS THIMM, University of Hagen, Germany

We consider the problem of quantitatively assessing the conflict between knowledge bases in knowledge merging scenarios. Using the notion of Craig interpolation we define a series of disagreement measures and analyse their compliance with properties proposed in previous work by Potyka. We then focus on a study of computational complexity of various problems that occur when measuring disagreement. In particular, we shed some new lights on the complexity of dealing with interpolants in propositional logic, such as deciding whether a given formula is an interpolant and whether it is a (semantically) strongest or weakest interpolant. We also consider functional problems (determining strongest/weakest interpolants) and counting problems.

**JAIR Associate Editor:** Francesco Calimeri

**JAIR Reference Format:**
Jandson S Ribeiro, Viorica Sofronie-Stokkermans, and Matthias Thimm. 2026. On the Complexity of Measuring Disagreement with Interpolants. *Journal of Artificial Intelligence Research* 87, Article 25 (November 2026), 33 pages. DOI: 10.1613/jair.1.22972

## 1 Introduction

Inconsistencies arise easily when knowledge has to be aggregated from different sources (Baral et al. 1991; Cholvy and Hunter 1997; Konieczny and Pérez 1998). Approaches to *belief merging* address these issues by providing computational approaches for automatically resolving these issues in some sensible way. This field bears a close relationship with the fields of *judgement* and *preference aggregation* (Everaere et al. 2015; Williamson 2009) and also features its own version of Arrow's impossibility result (Arrow 1963), insofar that there cannot be any "rational" belief merging approach (Díaz and Pérez 2017). This calls for semi-automatic methods that take human background knowledge into account when knowledge has to be merged in order not to remove important pieces of information. In order to support the task of semi-automatic merging, we investigate approaches to *analyse* belief merging settings, i. e., approaches that can explain reasons for inconsistency and assess their severity. More specifically, we investigate *disagreement measures* (Potyka 2018), i. e., functions that take a knowledge base profile $P = \langle \Phi_1, \ldots, \Phi_n \rangle$ as input and return a non-negative value that quantifies the severity of the disagreement between the different sources of information modelled by $\Phi_1, \ldots, \Phi_n$. Disagreement measures are closely related to *inconsistency measures* (Grant 1978; Thimm 2019, 2018), which themselves are functions that assess the severity of inconsistency in a single knowledge base. Disagreement and inconsistency measures can be used to help in debugging inconsistencies in semi-automated settings (Corea and Delfmann 2018; Corea and Thimm 2020; Grant and Hunter 2011).

In this paper, we develop novel disagreement measures based on the concept of Craig interpolation (Craig 1957). Given two knowledge bases $\Phi_1$, $\Phi_2$ with $\Phi_1 \cup \Phi_2$ being inconsistent, an *interpolant* is a formula that concisely

Authors' Contact Information: Jandson S Ribeiro, ORCID: 0000-0003-1614-6808, RibeiroJ@cardiff.ac.uk, Cardiff University, UK; Viorica Sofronie-Stokkermans, ORCID: 0000-0002-8486-9955, sofronie@uni-koblenz.de, University of Koblenz, Germany; Matthias Thimm, ORCID: 0000-0002-8157-1053, matthias.thimm@fernuni-hagen.de, University of Hagen, Germany.



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2026 Copyright held by the owner/author(s).
DOI: 10.1613/jair.1.22972

characterises an aspect of this inconsistency (we will provide formal definitions later). Thus, interpolants play a similar role in analysing the disagreement between two knowledge bases as minimal inconsistent subsets do in analysing the inconsistency within a single knowledge base. As a matter of fact, minimal inconsistent subsets play a dominant role in many approaches to measuring inconsistency (Hunter and Konieczny 2008; Mu et al. 2011; Xiao and Ma 2012) and, therefore, it seems plausible to explore the use of interpolants in measuring disagreement. In fact, a concrete scenario for our setting is the aforementioned semi-automatic merging of business rule bases, where rules maintained by different departments are integrated into a single process model and conflicts have to be resolved before merging (Corea and Delfmann 2018; Corea and Thimm 2020). Here an interpolant identifies exactly the shared vocabulary on which two rule bases disagree, while abstracting from the internal structure of each source, so that a modeller can rank pairs of departments by how severely and in how many distinct ways they conflict, rather than inspecting the union of all rules. Example 5 at the end of Section 4 illustrates this on a small instance. In order to formalise this, we consider the set of all interpolants (up to semantical equivalence) and define measures based on the size of that set and on the information content of the weakest/strongest interpolants (which are well-defined concepts due to the fact that the set of interpolants forms a complete lattice). We show that our approaches provide meaningful results and comply with many of the rationality postulates introduced by Potyka (2018). However, the focus of this paper is the study of the computational complexity of various problems within this context. In particular, we show that the problem of deciding whether a given formula is a weakest or strongest interpolant is $\Pi_2^p$-complete, respectively. We also consider the functional problems of determining some interpolant, a (representative of) a weakest interpolant and a strongest interpolant, and discuss that these problems are inherently hard to characterise. Furthermore, we also consider *counting problems* and show, in particular, that both counting the number of all (equivalence classes of) interpolants and counting the number of weakest interpolants is #·coNP-complete and counting the number of strongest interpolants is #·NP-complete. Finally, we apply these results to our new measures and obtain results for the complexities of computing upper/lower bounds and the exact values.

In summary, the contributions of this paper are as follows:

(1) We make some general observations on interpolants in order to establish a framework suitable for measuring disagreement (Section 3).
(2) We present novel disagreement measures based on interpolants (Section 4).
(3) We investigate the compliance of these disagreement measures with rationality postulates from the literature (Section 5).
(4) We investigate the computational complexity of several tasks pertaining to our disagreement measures (Section 6).

We introduce necessary preliminaries in Section 2 and conclude in Section 7.

This paper is an extended version of the conference paper (Ribeiro et al. 2020). This paper greatly extends the analysis of computational complexity and provides novel results, in particular pertaining to lower upper bounds and several hardness results (the conference paper only established weak upper bounds in ExpSpace and FExpSpace).

## 2 Preliminaries

We recall some background on propositional logic and quantified Boolean formulae.

### 2.1 Propositional Logic

Let At be some fixed propositional signature, i. e., a (possibly infinite) set of propositions, and let $\mathcal{L}(\text{At})$ be the corresponding propositional language constructed using the usual connectives $\wedge$ (*conjunction*), $\vee$ (*disjunction*), $\rightarrow$ (*implication*), and $\neg$ (*negation*). A literal is a proposition $p$ or a negated proposition $\neg p$. We denote with $\top$

an arbitrary tautology and with $\bot$ an arbitrary contradiction. A clause is a disjunction of literals. A formula is in conjunctive normal form (CNF) if the formula is a conjunction of clauses. If $X$ is a formula or a set of formulae we write $\mathsf{At}(X)$ to denote the set of propositions appearing in $X$. For a finite set $\Phi = \{\phi_1, \ldots, \phi_n\}$ let $\bigwedge \Phi = \phi_1 \wedge \ldots \wedge \phi_n$ and $\neg\Phi = \{\neg\phi \mid \phi \in \Phi\}$. Such a finite set $\Phi$ is also called a *knowledge base*. Let $\mathbb{K}$ be the set of all knowledge bases.

Semantics for a propositional language is given by *interpretations* where an *interpretation* $\omega$ on At is a function $\omega : \mathsf{At} \to \{\mathsf{true}, \mathsf{false}\}$. Let $\Omega(\mathsf{At})$ denote the set of all interpretations for At. An interpretation $\omega$ *satisfies* (or is a *model* of) an atom $a \in \mathsf{At}$, denoted by $\omega \models a$, if and only if $\omega(a) = \mathsf{true}$. The satisfaction relation $\models$ is extended to formulae in the usual way. When the exact signature is of no concern, we sometimes restrict attention to the propositions appearing in the formula/formulae under consideration, i. e., we then consider interpretations of the form $\omega : \mathsf{At}(\phi) \to \{\mathsf{true}, \mathsf{false}\}$. For $\Phi \subseteq \mathcal{L}(\mathsf{At})$ we also define $\omega \models \Phi$ if and only if $\omega \models \phi$ for every $\phi \in \Phi$.

In the following, let $\Phi, \Phi_1, \Phi_2$ be formulae or sets of formulae. Define the set of models $\mathsf{Mod}(\Phi) = \{\omega \in \Omega(\mathsf{At}) \mid \omega \models \Phi\}$. We also write $\mathsf{Mod}_A(\phi) = \{\omega : A \to \{\mathsf{true}, \mathsf{false}\} \mid \omega \models \phi\}$ as the set of models over the signature $A$. We write $\Phi_1 \models \Phi_2$ if $\mathsf{Mod}(\Phi_1) \subseteq \mathsf{Mod}(\Phi_2)$. $\Phi_1, \Phi_2$ are *equivalent*, denoted by $\Phi_1 \equiv \Phi_2$, if and only if $\mathsf{Mod}(\Phi_1) = \mathsf{Mod}(\Phi_2)$. Define the closure $\mathsf{Cn}(\Phi)$ of a formula or set of formulae $\Phi$ via $\mathsf{Cn}(\Phi) = \{\phi \mid \Phi \models \phi\}$. If $\mathsf{Mod}(\Phi) = \emptyset$ we also write $\Phi \models \bot$ and say that $\Phi$ is *inconsistent* (or *unsatisfiable*). The power set of a set $A$ is denoted as $2^A$.

## 2.2 Quantified Boolean Formulae

In this paper we also consider formulae in which (some of) the propositions are quantified. Let $x_1, \ldots, x_n \in \mathsf{At}$ be distinct propositions. A *quantified Boolean formula (QBF)* is an expression of the form $Q_1 x_1 \ldots Q_n x_n \psi$, where $\psi \in \mathcal{L}(\mathsf{At})$, $Q_i \in \{\exists, \forall\}$ and $\{x_1, \ldots, x_n\} \subseteq \mathsf{At}$. A *totally quantified Boolean formula* is a quantified Boolean formula in which all propositional variables are quantified.

Let $\psi \in \mathcal{L}(\{x, y_1, \ldots, y_m\})$. An interpretation $\omega : \{y_1, \ldots, y_m\} \to \{\mathsf{true}, \mathsf{false}\}$ is a model of the QBF $\exists x\, \psi$ if it can be extended to a model $\omega' : \{x, y_1, \ldots, y_m\} \to \{\mathsf{true}, \mathsf{false}\}$ of $\psi$; $\omega$ is a model of the QBF $\forall x\, \psi$ if all its extensions $\omega' : \{x, y_1, \ldots, y_m\} \to \{\mathsf{true}, \mathsf{false}\}$ are models of $\psi$. Therefore, for every formula $\psi \in \mathcal{L}(\{x, y_1, \ldots, y_m\})$, $\exists x\, \psi$ has the same models as the propositional formula $\psi[x \mapsto \bot] \vee \psi[x \mapsto \top] \in \mathcal{L}(\{y_1, \ldots, y_m\})$, and $\forall x\, \psi$ has the same models as the propositional formula $\psi[x \mapsto \bot] \wedge \psi[x \mapsto \top] \in \mathcal{L}(\{y_1, \ldots, y_m\})$, where $\psi[x \mapsto \top]$ is the formula obtained by replacing all occurrences of $x$ in $\psi$ by $\top$ and $\psi[x \mapsto \bot]$ is the formula obtained by replacing all occurrences of $x$ in $\psi$ by $\bot$.[1]

From this it follows that for every QBF $Q_1 x_1 \ldots Q_n x_n \psi$, where $Q_i \in \{\exists, \forall\}$ and $\psi \in \mathcal{L}(\{x_1, \ldots, x_n, y_1, \ldots, y_m\})$ there exists a propositional formula $\psi' \in \mathcal{L}(\{y_1, \ldots, y_m\})$ such that $Q_1 x_1 \ldots Q_n x_n \psi$ and $\psi'$ have the same models. In this case we say that they are equivalent and denote this by:

$$Q_1 x_1 \ldots Q_n x_n \psi \equiv \psi'.$$

Clearly, a totally quantified Boolean formula is equivalent to $\top$ or to $\bot$. In this paper we will sometimes not distinguish between quantified Boolean formulae and the propositional formulae they are equivalent to, and use the following abbreviations, where $\phi$ is a formula and $x$ a propositional symbol:

- $\exists x\, \phi := \phi[x \mapsto \bot] \vee \phi[x \mapsto \top]$;
- $\forall x\, \phi := \phi[x \mapsto \bot] \wedge \phi[x \mapsto \top]$.

Let $\Phi$ be a finite set of propositional formulae, and let $\phi$ be the conjunction of all the formulae in $\Phi$. For every $x \in \mathsf{At}(\phi)$, we use the following notation $\exists x \Phi := \exists x \phi$, $\forall x \Phi := \forall x \phi$.

Since a propositional formula equivalent to a quantified Boolean formula $Q_1 x_1 \ldots Q_n x_n \psi$ might have size exponential in the size of $\psi$, in Section 6—where the computational complexity of the problems we consider is

[1] For instance, for the formula $\psi = p \vee \neg q$, we get $\psi[p \mapsto \top] \equiv \top \vee \neg q$.

discussed—we make a distinction between quantified Boolean formulae and the propositional formulae they are equivalent to.

**Remark 1.** Let $\forall x_1 \dots x_n \exists y_1 \dots y_m \psi$ be a totally quantified Boolean formula which is equivalent to $\top$. Then the QBF $\exists y_1 \dots y_m \psi \in \mathcal{L}(\{x_1, \dots, x_n\})$ is equivalent to $\top$ for all interpretations of the variables $x_1, \dots, x_n$, so if $\psi' \in \mathcal{L}(\{x_1, \dots, x_n\})$ is a propositional formula equivalent to $\exists y_1 \dots y_m \psi$, then $\psi'$ is valid. If we do not distinguish between quantified Boolean formulae and the propositional formulae they are equivalent to, and use the notation $\exists y_1 \dots y_m \psi$ to refer to the equivalent propositional formula, we, in this case, say that $\exists y_1 \dots y_m \psi$ is valid.

**Remark 2.** If $\phi, \psi \in \mathcal{L}(\mathsf{At})$ and $x_1, \dots, x_n \in \mathsf{At}$ then:

- $\phi \models \exists x_1 \dots \exists x_n \psi$ if for every model $\omega$ of $\phi$ there exists a model $\omega'$ of $\psi$ such that $\omega$ and $\omega'$ agree on all propositions in $\mathsf{At} \backslash \{x_1, \dots, x_n\}$;
- $\phi \models \forall x_1 \dots x_n \psi$ if for every model $\omega$ of $\phi$, every model $\omega'$ which agrees on all propositions in $\mathsf{At} \backslash \{x_1, \dots, x_n\}$ with $\omega$ is a model of $\psi$;
- $\phi \models \forall x_1 \dots x_n \exists y_1 \dots y_m \psi$ if for every model $\omega$ of $\phi$, and for every model $\omega'$ which agrees with $\omega$ on all propositions in $\mathsf{At} \backslash \{x_1, \dots, x_n\}$ there exists a model $\omega''$ of $\psi$ which agrees with $\omega'$ on all propositions except possibly $y_1, \dots, y_m$.

Since according to this definition $\phi \models \exists x_1 \dots \exists x_n \psi$ (resp. $\phi \models \forall x_1 \dots x_n \psi$, resp. $\phi \models \forall x_1 \dots x_n \exists y_1 \dots y_m \psi$) iff $\phi \models \psi'$, where $\psi'$ is the propositional formula corresponding to the respective QBF, identifying quantified Boolean formulae with the propositional formulae they are equivalent to is unproblematic.

## 3 Craig Interpolants

We consider the following notion of an *interpolant* in propositional logic.[2]

**Definition 1.** Let $\Phi$ and $\Phi'$ be finite sets of propositional logic formulae. A formula $\phi$ is called an *interpolant* of $\Phi$ w.r.t. $\Phi'$ if

(1) $\Phi \models \phi$,
(2) $\Phi' \cup \{\phi\} \models \bot$, and
(3) $\mathsf{At}(\phi) \subseteq \mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')$

**Example 1.** Consider, for instance, two sets $\Phi_1 = \{r \vee \neg p, \neg r \vee \neg q\}$ and $\Phi_2 = \{p, q\}$. The formula $p \rightarrow \neg q$ is an interpolant of $\Phi_1$ w.r.t. $\Phi_2$, as $\Phi_1 \models p \rightarrow \neg q$, $\Phi_2 \cup \{p \rightarrow \neg q\} \models \bot$ and $\mathsf{At}(\{p \rightarrow \neg q\}) = \{p, q\} \subseteq \mathsf{At}(\Phi_1) \cap \mathsf{At}(\Phi_2)$.

Clearly, two finite sets $\Phi$ and $\Phi'$ of formulae in propositional logic have an interpolant if and only if $\Phi \cup \Phi'$ is unsatisfiable. Let $\mathbb{I}(\Phi, \Phi')$ denote the set of interpolants of $\Phi$ and $\Phi'$.

Propositional logic allows uniform interpolation: For every formula $\phi$ and every set $\{y_1, \dots, y_m\} \subseteq \mathsf{At}(\phi)$ there exists a propositional formula $I_\phi$ (the *uniform interpolant* of $\phi$ w.r.t. $\{y_1, \dots, y_m\}$) such that $\mathsf{At}(I_\phi) \subseteq \{y_1, \dots, y_m\}$ with the property that for every formula $\phi'$ such that $\phi \wedge \phi' \models \bot$ and $\{y_1, \dots, y_m\} = \mathsf{At}(\phi) \cap \mathsf{At}(\phi')$, $I_\phi$ is an interpolant of $\phi$ and $\phi'$. This follows from the following results (here formulated for finite sets of propositional formulae), which give a description for the uniform interpolant of a formula w.r.t. a set of variables.

Lemma 1. *Let $\phi$ be a propositional formula and $x$ one of the propositional variables occurring in $\phi$. Then the following hold:*

*(1) $\phi \models \exists x \phi$;*

[2]Note that the term *interpolant* is sometimes interpreted in another fashion. Craig (1957) introduces an interpolant $I$ for formulae $\phi$ and $\psi$ as a formula containing only propositional variables occurring in both $\phi$ and $\psi$ such that $\phi \models I$ and $I \models \psi$. Kovács and Voronkov (2009) used the name "reversed interpolant" for formulae satisfying the conditions in Definition 1. In this paper we will refer to them as interpolants, however.

*(2) Let $\psi$ be a formula with $\mathrm{At}(\psi) \subseteq \mathrm{At}\backslash\{x\}$ such that $\phi \models \psi$. Then $\exists x\phi \models \psi$.*
*(3) $\forall x\phi \models \phi$;*
*(4) Let $\psi$ be a formula with $\mathrm{At}(\psi) \subseteq \mathrm{At}\backslash\{x\}$ such that $\psi \models \phi$. Then $\psi \models \forall x\phi$.*

Proof. (1) Let $\omega : \mathrm{At} \to \{\text{true}, \text{false}\}$ be such that $\omega \models \phi$. If $\omega(x) = \text{false}$ then $\omega \models \phi[x \mapsto \bot]$; if $\omega(x) = \text{true}$ then $\omega \models \phi[x \mapsto \top]$. Thus, $\omega \models \exists x\phi$.

(2) Let $\psi$ be a formula with $\mathrm{At}(\psi) \subseteq \mathrm{At}\backslash\{x\}$ such that $\phi \models \psi$. Let $\omega : \mathrm{At} \to \{\text{true}, \text{false}\}$ be such that $\omega \models \exists x\phi$. Then $\omega \models \phi[x \mapsto \bot] \vee \phi[x \mapsto \top]$, so $\omega \models \phi[x \mapsto \bot]$ or $\omega \models \phi[x \mapsto \top]$. Assume that $\omega \models \phi[x \mapsto \bot]$ (the other case is analogous). Let $\omega' : \mathrm{At} \to \{\text{true}, \text{false}\}$ defined by $\omega'(x) = \text{false}$ and $\omega'(y) = \omega(y)$ for all $y \neq x$. Then $\omega' \models \phi$, so as $\phi \models \psi$ we have $\omega' \models \psi$. But since $x$ does not occur in $\psi$ and $\omega$ and $\omega'$ agree for all atoms but $x$, we have $\omega \models \psi$.

(3) Let $\omega : \mathrm{At} \to \{\text{true}, \text{false}\}$ be such that $\omega \models \forall x\phi$. Then $\omega \models \phi[x \mapsto \bot] \wedge \phi[x \mapsto \top]$. Assume $\omega \not\models \phi$. If $\omega(x) = \text{false}$ it follows that $\omega \not\models \phi[x \mapsto \bot]$; if $\omega(x) = \text{true}$ it follows that $\omega \not\models \phi[x \mapsto \top]$. Contradiction.

(4) Let $\psi$ be a formula with $\mathrm{At}(\psi) \subseteq \mathrm{At}\backslash\{x\}$ such that $\psi \models \phi$. Let $\omega : \mathrm{At} \to \{\text{true}, \text{false}\}$ be such that $\omega \models \psi$. Since $\psi \models \phi$ it follow that $\omega \models \phi$. Clearly, $\omega(x) = \text{false}$ or $\omega(x) = \text{true}$. Assume $\omega(x) = \text{false}$ (the other case is analogous). Then $\omega \models \phi[x \mapsto \bot]$. Let $\omega' : \mathrm{At} \to \{\text{false}, \text{true}\}$ defined by $\omega'(x) = \text{true}$ and $\omega'(y) = \omega(y)$ for all $y \neq x$. Since $x$ does not occur in $\psi$ and $\omega$ and $\omega'$ agree for all atoms but $x$, we have $\omega' \models \psi$, so $\omega' \models \phi$, i. e. $\omega' \models \phi[x \mapsto \top]$, hence $\omega \models \phi[x \mapsto \top]$ (as $\phi[x \mapsto \top]$ does not contain any occurrences of $x$).

We thus showed that $\omega \models \phi[x \mapsto \bot] \wedge \phi[x \mapsto \top]$, i. e. $\omega \models \forall x\phi$. □

Proposition 1. *Let $\Phi$ be a finite set of propositional formulae. Assume that $\mathrm{At}(\Phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$. Then the following hold:*

*(1) $\Phi \models \exists x_1 \ldots \exists x_n \Phi$.*
*(2) Let $\psi$ be a propositional formula with $\mathrm{At}(\psi) \subseteq \{y_1, \ldots, y_m\}$ such that $\Phi \models \psi$. Then $\exists x_1 \ldots \exists x_n \Phi \models \psi$.*
*(3) $\forall x_1 \ldots \forall x_n \Phi \models \Phi$ (i. e. $\forall x_1 \ldots \forall x_n \Phi \models \phi_i$ for every formula $\phi_i \in \Phi$).*
*(4) Let $\psi$ be a formula with $\mathrm{At}(\psi) \subseteq \{y_1, \ldots, y_m\}$ such that $\psi \models \Phi$. Then $\psi \models \forall x_1 \ldots \forall x_n \Phi$.*

Proof. Assume that $\Phi = \{\phi_1, \ldots, \phi_n\}$, and let $\phi = \phi_1 \wedge \cdots \wedge \phi_n$. All results follow from Lemma 1 (applied to $\phi$) by induction on the number of variables which are eliminated. □

In what follows, we will assume that variables for which we use different names are different variables.

Corollary 1. *Let $\Phi$ and $\Phi'$ be finite sets of formulae such that $\Phi \cup \Phi' \models \bot$, $\mathrm{At}(\Phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ and $\mathrm{At}(\Phi') = \{y_1, \ldots, y_m, z_1, \ldots, z_k\}$. Then (the propositional formula equivalent to) $\exists x_1 \ldots \exists x_n \Phi$ has the following properties:*

*(1) it contains only propositional variables in $\{y_1, \ldots, y_m\}$;*
*(2) $\Phi \models \exists x_1 \ldots \exists x_n \Phi$;*
*(3) $\Phi' \cup \{\exists x_1 \ldots \exists x_n \Phi\}$ is unsatisfiable;*
*(4) $\exists x_1 \ldots \exists x_n \Phi$ entails any other propositional formula $\psi$ with $\mathrm{At}(\psi) \subseteq \{y_1, \ldots, y_m\}$ such that $\Phi \models \psi$*

*and is therefore a uniform interpolant of $\Phi$ w.r.t. $\{y_1, \ldots, y_m\}$.*

Proof. (1) is obvious. (2) and (4) are easy consequences of Proposition 1. (3) Assume that $\Phi' \cup \{\exists x_1 \ldots \exists x_n \Phi\}$ was satisfiable and let $\omega$ be a model of $\Phi' \cup \{\exists x_1 \ldots \exists x_n \Phi\}$. Then there would exist a model $\omega'$ of $\Phi$, agreeing with $\omega$ on all proposition variables except for $x_1, \ldots, x_n$. Such an interpretation would then be also a model of $\Phi'$, which is not possible because it contradicts the assumption that $\Phi \cup \Phi' \models \bot$. □

In the following, we define some auxiliary notions for interpolants and make some first observations regarding the structure of $\mathbb{I}(\Phi, \Phi')$.

Proposition 2. *Let $\Phi$ and $\Phi'$ be finite sets of formulae.*
*(1) If $\Phi \cup \Phi' \not\models \bot$ then $\mathbb{I}(\Phi, \Phi') = \emptyset$.*
*(2) If $\Phi \models \bot$ and $\Phi' \models \bot$ then $\mathbb{I}(\Phi, \Phi') = \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi'))$.*
*(3) If $\Phi \models \bot$ and $\Phi' \not\models \bot$ then $\mathbb{I}(\Phi, \Phi') = \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')) \cap \{\phi \mid \phi \equiv \neg\phi' \text{ for some } \phi' \in \mathsf{Cn}(\Phi')\}$.*
*(4) If $\Phi \not\models \bot$ and $\Phi' \models \bot$ then $\mathbb{I}(\Phi, \Phi') = \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')) \cap \mathsf{Cn}(\Phi)$.*
*(5) If $\varphi \in \mathbb{I}(\Phi, \Phi')$ then $\neg\varphi \in \mathbb{I}(\Phi', \Phi)$.*

Proof. (1) Assume that $\mathbb{I}(\Phi, \Phi') \neq \emptyset$. Then there exists a formula $\phi$ such that $\Phi \models \phi$ and $\Phi' \cup \{\phi\} \models \bot$. Hence, $\Phi \cup \Phi' \models \bot$.
(2) If $\Phi \models \bot$ and $\Phi' \models \bot$ then (i) $\Phi \models \phi$ for all $\phi \in \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi'))$ and (ii) $\Phi' \cup \{\phi\} \models \bot$ for all $\phi \in \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi'))$.
(3) If $\Phi \models \bot$ then $\Phi \models \phi$ for all $\phi \in \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi'))$. If $\Phi' \not\models \bot$ then $\Phi' \cup \{\phi\} \models \bot$ for all $\phi$ such that (i) $\phi \equiv \neg\phi'$, where $\phi' \in \mathsf{Cn}(\Phi')$ or (ii) $\phi \models \bot$ (and $\phi \equiv \neg\top$, where $\top \in \mathsf{Cn}(\Phi')$).
(4) If $\Phi \not\models \bot$ and $\Phi' \models \bot$ then a formula $\phi$ is an interpolant iff $\Phi \models \phi$ and $\phi \in \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi'))$.
(5) $\varphi \in \mathbb{I}(\Phi, \Phi')$ iff (i) $\Phi \cup \Phi' \models \bot$, (ii) $\mathsf{At}(\varphi) \subseteq \mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')$, (iii) $\Phi \models \varphi$ and $\{\varphi\} \cup \Phi' \models \bot$. It is easy to see that (iii) holds iff (iii)' $\Phi \cup \{\neg\varphi\} \models \bot$ and $\Phi' \models \neg\varphi$. But (i), (ii) and (iii)' hold iff $\neg\varphi \in \mathbb{I}(\Phi', \Phi)$. □

As the notion of an interpolant is a syntactical one, the set $\mathbb{I}(\Phi, \Phi')$ also contains infinitely many syntactical variants for each interpolant (except for the case where $\Phi \cup \Phi'$ is consistent); in particular, if $\phi \in \mathbb{I}(\Phi, \Phi')$ then for every formula $\phi' \in \mathcal{L}(\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi'))$, if $\phi \equiv \phi'$ then $\phi' \in \mathbb{I}(\Phi, \Phi')$. However, we will consider the following finite representation of $\mathbb{I}(\Phi, \Phi')$. For that, for all $A \subseteq \mathsf{At}$, let $[\cdot]_A : \mathcal{L}(A) \to 2^{\mathcal{L}(A)}$ be the function that maps each formula $\varphi \in \mathcal{L}(A)$ to its *equivalence class* w.r.t. $\equiv$, that is, $[\varphi]_A = \{\psi \in \mathcal{L}(A) \mid \psi \equiv \varphi\}$.

**Definition 2.** Let $\Phi$ and $\Phi'$ be finite sets of formulae. We denote by $\mathbb{SI}(\Phi, \Phi')$ the set of equivalence classes of interpolants w.r.t. semantical equivalence, i. e.,

$$\mathbb{SI}(\Phi, \Phi') = \{[\phi]_A \in 2^{\mathcal{L}(A)} \mid \phi \in \mathbb{I}(\Phi, \Phi')\}, \text{ where } A = \mathsf{At}(\Phi) \cap \mathsf{At}(\Phi').$$

In the following, by abuse of notation we refer to $[\phi]$ also as an interpolant if $\phi$ is an interpolant. We will omit the subscript $_A$ from [.] when $A$ is clear from context.

**Remark 3.** It can be easily seen that the elements of $\mathbb{SI}(\Phi, \Phi')$ form a lattice w.r.t. the partial ordering defined by $[\phi| \leq [\psi]$ iff $\phi \models \psi$ and the operations $\wedge$ and $\vee$.

If the sets $\Phi, \Phi'$ of propositional formulae are finite then the set $\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')$ is finite, so the set of all equivalence classes of formulae over this set of atoms is finite and has at most $2^{2^{|\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')|}}$ elements, thus $\mathbb{SI}(\Phi, \Phi')$ is finite, so $(\mathbb{SI}(\Phi, \Phi'), \wedge, \vee)$ is a complete lattice.

Formally, we can make the following observations. In what follows, if $\Phi$ is a finite set of formulae, say $\Phi = \{\phi_1, \ldots, \phi_n\}$ we denote by $\neg(\wedge\Phi)$ the formula $\neg(\phi_1 \wedge \cdots \wedge \phi_n)$.

Lemma 2. *Let $\Phi$ and $\Phi'$ be finite sets of propositional formulae with $\mathsf{At}(\Phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ and $\mathsf{At}(\Phi') = \{y_1, \ldots, y_m, z_1, \ldots, z_k\}$ and such that $\Phi \cup \Phi'$ is unsatisfiable (i. e. $\Phi \models \neg(\wedge\Phi')$).*
*Let $\phi_w := \forall z_1 \ldots \forall z_k \neg(\wedge\Phi')$. Then the following hold:*
*(1) $\Phi \models \phi_w$;*
*(2) $\phi_w \wedge \Phi' \models \bot$;*
*(3) $\phi_w$ contains only propositional variables occurring in both $\Phi$ and $\Phi'$;*
*(4) For every interpolant $\phi$ of $\Phi$ and $\Phi'$, $\phi \models \phi_w$.*

Proof. (1) can be proved by induction on $n$ using Lemma 1(4), using the fact that $\Phi \models \neg(\wedge\Phi')$; (2) can be proved by induction on $n$ using Lemma 1(3). (3) is obvious. To prove (4) let $\phi$ be an interpolant of $\Phi$ and $\Phi'$. Then, in particular, $\Phi' \cup \{\phi\} \models \bot$, so $\phi \models \neg(\wedge\Phi')$. By Lemma 1(4) it follows that $\phi \models \forall z_1 \ldots \forall z_k \neg(\wedge\Phi')$, i. e. $\phi \models \phi_w$. □

Proposition 3. *Let $\Phi$ and $\Phi'$ be finite sets of formulae such that $\Phi \cup \Phi' \models \bot$.*
*(1) If $[\phi], [\phi'] \in \mathbb{SI}(\Phi, \Phi')$ then $[\phi \wedge \phi'] \in \mathbb{SI}(\Phi, \Phi')$.*
*(2) If $[\phi], [\phi'] \in \mathbb{SI}(\Phi, \Phi')$ then $[\phi \vee \phi'] \in \mathbb{SI}(\Phi, \Phi')$.*
*(3) There is a uniquely defined $[\phi_w] \in \mathbb{SI}(\Phi, \Phi')$ with $\phi' \models \phi_w$ for all $[\phi'] \in \mathbb{SI}(\Phi, \Phi')$.*
*(4) There is a uniquely defined $[\phi_s] \in \mathbb{SI}(\Phi, \Phi')$ with $\phi_s \models \phi'$ for all $[\phi'] \in \mathbb{SI}(\Phi, \Phi')$.*

Proof. (1) If $[\phi], [\phi'] \in \mathbb{SI}(\Phi, \Phi')$ then $\phi, \phi' \in \mathbb{I}(\Phi, \Phi')$ so: (i) $\Phi \models \phi$ and $\Phi \models \phi'$, so $\Phi \models \phi \wedge \phi'$; (ii) $\Phi' \cup \{\phi\} \models \bot$, so $\Phi' \cup \{\phi \wedge \phi'\} \models \bot$ and (iii) $\mathrm{At}(\phi), \mathrm{At}(\phi') \subseteq \mathrm{At}(\Phi) \cap \mathrm{At}(\Phi')$, hence this is the case also for $\phi \wedge \phi'$. Thus, $\phi \wedge \phi' \in \mathbb{I}(\Phi, \Phi')$, so $[\phi \wedge \phi'] \in \mathbb{SI}(\Phi, \Phi')$.

(2) If $[\phi], [\phi'] \in \mathbb{SI}(\Phi, \Phi')$ then $\phi, \phi' \in \mathbb{I}(\Phi, \Phi')$, so (i) $\Phi \models \phi$ so $\Phi \models \phi \vee \phi'$; (ii) $\Phi' \cup \{\phi\} \models \bot$ and $\Phi' \cup \{\phi'\} \models \bot$ hence $\Phi' \cup \{\phi \vee \phi'\} \models \bot$ and (iii) $\mathrm{At}(\phi), \mathrm{At}(\phi') \subseteq \mathrm{At}(\Phi) \cap \mathrm{At}(\Phi')$, hence this is the case also for $\phi \vee \phi'$. Thus, $\phi \vee \phi' \in \mathbb{I}(\Phi, \Phi')$, so $[\phi \vee \phi'] \in \mathbb{SI}(\Phi, \Phi')$.

For proving (3) and (4) we assume that the sets of atoms of $\Phi$ and $\Phi'$ are $\mathrm{At}(\Phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ and $\mathrm{At}(\Phi') = \{y_1, \ldots, y_m, z_1, \ldots, z_k\}$.

(3) Let $\phi_w := \forall z_1 \ldots \forall z_k \neg(\wedge \Phi')$. Then, by Lemma 2, the following hold:

- $\Phi \models \phi_w$;
- $\phi_w \wedge \Phi' \models \bot$;
- $\phi_w$ contains only propositional variables occurring in both $\Phi$ and $\Phi'$;
- For every interpolant $\phi$ of $\Phi$ and $\Phi'$, $\phi \models \phi_w$.

Clearly, $[\phi_w]$ is uniquely determined. If $[\phi]$ is a different class with the same property, then $\phi' \models \phi$ for all $[\phi'] \in \mathbb{SI}(\Phi, \Phi')$, hence also $\phi_w \models \phi$. On the other hand, if $\phi \in \mathbb{SI}(\Phi, \Phi')$ then, by Lemma 2 (4), $\phi \models \phi_w$. Thus, $[\phi_w] = [\phi]$.

(4) Let $\phi_s := \exists x_1 \ldots \exists x_n \Phi$. Then, by Proposition 1 and Corollary 1, the following hold:

- $\Phi \models \phi_s$;
- $\phi_s \wedge \Phi' \models \bot$;
- $\phi_s$ contains only propositional variables occurring in both $\Phi$ and $\Phi'$;
- For every interpolant $\phi$ of $\Phi$ and $\Phi'$, $\phi_s \models \phi$.

To show that $[\phi_s]$ is uniquely determined one can use a proof analogous to the one used for $[\phi_w]$. □

Conditions (1) and (2) of the proposition above are clear. We illustrate the intuition of both conditions (3) and (4) with an example.

**Example 2.** Consider the knowledge bases $\Phi_1 = \{a, b\}$ and $\Phi_2 = \{a, \neg b\}$. We have that

$$\mathbb{SI}(\Phi_1, \Phi_2) = \{[a \wedge b], [b], [a \rightarrow b], [a \leftrightarrow b]\}.$$

Note that every formula in $\mathbb{SI}(\Phi_1, \Phi_2)$ implies $a \rightarrow b$ (the $\phi_w$ of condition 3). Similarly, the formula $a \wedge b$ implies all formulae in $\mathbb{SI}(\Phi_1, \Phi_2)$, which makes $a \wedge b$ the formula $\phi_s$ of condition 4.

We call $[\phi_w]$ the *weakest* interpolant and $[\phi_s]$ the *strongest* interpolant, and will refer to any formula in $[\phi_w]$ as a weakest interpolant and to any formula in $[\phi_s]$ as a strongest interpolant. It can be easily seen that

$$[\phi_w] = \left[ \bigvee_{[\phi] \in \mathbb{SI}(\Phi, \Phi')} \phi \right] \qquad [\phi_s] = \left[ \bigwedge_{[\phi] \in \mathbb{SI}(\Phi, \Phi')} \phi \right]$$

We abbreviate the weakest interpolant of $\mathbb{SI}(\Phi, \Phi')$ by $\mathrm{Weakest}(\Phi, \Phi')$ and the strongest interpolant of $\mathbb{SI}(\Phi, \Phi')$ by $\mathrm{Strongest}(\Phi, \Phi')$. If $\mathbb{SI}(\Phi, \Phi') = \emptyset$ both notions are undefined. We conclude this section with some further observations that will be useful in the remainder of the paper.

Proposition 4. *Let $\Phi$, $\Phi'$, and $\Phi''$ be finite sets of formulae.*

*(1)* $\mathbb{I}(\Phi, \Phi') \subseteq \mathbb{I}(\Phi, \Phi' \cup \Phi'')$.
*(2)* $\mathbb{I}(\Phi, \Phi') \subseteq \mathbb{I}(\Phi \cup \Phi'', \Phi')$.
*(3)* $\mathbb{SI}(\Phi, \Phi') \subseteq \mathbb{SI}(\Phi, \Phi' \cup \Phi'')$.
*(4)* $\mathbb{SI}(\Phi, \Phi') \subseteq \mathbb{SI}(\Phi \cup \Phi'', \Phi')$.

*If $\Phi \cup \Phi' \models \bot$ then the following hold:*

*(5)* $\mathsf{Weakest}(\Phi, \Phi') \models \mathsf{Weakest}(\Phi, \Phi' \cup \Phi'')$.
*(6)* $\mathsf{Weakest}(\Phi, \Phi') \models \mathsf{Weakest}(\Phi \cup \Phi'', \Phi')$.
*(7)* $\mathsf{Strongest}(\Phi, \Phi' \cup \Phi'') \models \mathsf{Strongest}(\Phi, \Phi')$.
*(8)* $\mathsf{Strongest}(\Phi \cup \Phi'', \Phi') \models \mathsf{Strongest}(\Phi, \Phi')$.
*(9) If $\varphi \in \mathsf{Weakest}(\Phi, \Phi')$ then $\neg\varphi \in \mathsf{Strongest}(\Phi', \Phi)$.*
*(10) If $\varphi \in \mathsf{Strongest}(\Phi, \Phi')$ then $\neg\varphi \in \mathsf{Weakest}(\Phi', \Phi)$.*

Proof. Considering item (1), for $\phi \in \mathbb{I}(\Phi, \Phi')$, from $\Phi' \cup \{\phi\} \models \bot$ it directly follows $\Phi' \cup \Phi'' \cup \{\phi\} \models \bot$. The other two conditions of Definition 1 remain valid as well. Item (2) is proven analogously and the items (3) to (8) follow directly from (1) and (2).

(9) Let us suppose for contradiction that there is some formula $\varphi$ such that $\varphi \in \mathsf{Weakest}(\Phi, \Phi')$ but $\neg\varphi \notin \mathsf{Strongest}(\Phi', \Phi)$. From Proposition 2 (item 5), $\neg\varphi \in \mathbb{I}(\Phi', \Phi)$. Let $\psi \in \mathsf{Strongest}(\Phi', \Phi)$. Thus, $\psi \models \neg\varphi$ but $\neg\varphi \not\models \psi$, which implies $\mathsf{Mod}(\psi) \subset \mathsf{Mod}(\neg\varphi)$. Therefore, $\mathsf{Mod}(\neg\neg\varphi) \subset \mathsf{Mod}(\neg\psi)$, that is, $\mathsf{Mod}(\varphi) \subset \mathsf{Mod}(\neg\psi)$. Thus, $\neg\psi \not\models \varphi$. From Proposition 2 (item 5), $\neg\psi \in \mathbb{I}(\Phi, \Phi')$ which means $\varphi \notin \mathsf{Weakest}(\Phi, \Phi')$ which contradicts the hypothesis that $\varphi \in \mathsf{Weakest}(\Phi, \Phi')$.

(10) Let us suppose for contradiction that there is some formula $\varphi$ such that $\varphi \in \mathsf{Strongest}(\Phi, \Phi')$ but $\neg\varphi \notin \mathsf{Weakest}(\Phi', \Phi)$. From Proposition 2 (item 5), $\neg\varphi \in \mathbb{I}(\Phi', \Phi)$. Let $\psi \in \mathsf{Weakest}(\Phi', \Phi)$. Thus, $\neg\varphi \models \psi$ but $\psi \not\models \neg\varphi$, which implies $\mathsf{Mod}(\neg\varphi) \subset \mathsf{Mod}(\psi)$. Therefore, $\mathsf{Mod}(\neg\psi) \subset \mathsf{Mod}(\neg\neg\varphi)$, that is, $\mathsf{Mod}(\neg\psi) \subset \mathsf{Mod}(\varphi)$. Thus, $\varphi \not\models \neg\psi$. From Proposition 2 (item 5), $\neg\psi \in \mathbb{I}(\Phi, \Phi')$ which means $\varphi \notin \mathsf{Strongest}(\Phi, \Phi')$ which contradicts the hypothesis that $\varphi \in \mathsf{Strongest}(\Phi, \Phi')$. □

Corollary 2. *Let $\Phi$ and $\Phi'$ be finite sets of formulae, with $\mathsf{At}(\Phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ and $\mathsf{At}(\Phi') = \{y_1, \ldots, y_m, z_1, \ldots, z_k\}$ and such that $\Phi \cup \Phi' \models \bot$, and let $\phi$ (resp. $\phi'$) be the conjunction of all the formulae in $\Phi$ (resp. $\Phi'$).*

*(1)* $\mathsf{Strongest}(\Phi, \Phi') = [\exists x_1 \ldots \exists x_n \phi]$;
*(2)* $\mathsf{Weakest}(\Phi, \Phi') = [\forall z_1 \ldots \forall z_k \neg\phi']$.

Proof. Condition (1) is an immediate consequence of Corollary 1. To prove condition (2) note that, by Proposition 4 items (9) and (10), $\varphi \in \mathsf{Weakest}(\Phi, \Phi')$ iff $\neg\varphi \in \mathsf{Strongest}(\Phi', \Phi)$, which, by (1), holds iff $\neg\varphi \equiv \exists z_1 \ldots \exists z_k \phi'$ iff $\varphi \equiv \forall z_1 \ldots \forall z_k \neg\phi'$. □

Thus, the strongest interpolant of $\Phi$ w.r.t. $\Phi'$ is the (equivalence class of the) uniform interpolant of $\Phi$ w.r.t. variables $\{y_1, \ldots, y_m\}$, and the weakest interpolant of $\Phi$ w.r.t. $\Phi'$ is the (equivalence class of the) negation of the uniform interpolant of $\Phi'$ w.r.t. variables $\{y_1, \ldots, y_m\}$.

## 4 Disagreement Measures based on Interpolation

We consider the scenario of measuring *disagreement* between multiple knowledge bases (Potyka 2018) using interpolation. For that, we denote a *knowledge base profile* by $P = \langle \Phi_1, \ldots, \Phi_n \rangle$ with $\Phi_1, \ldots, \Phi_n$ being knowledge bases. Furthermore, for $P$ of this form and another knowledge base $\Phi$, we denote by $P \circ \Phi$ the concatenation of $P$ with $\Phi$, i. e., $P \circ \Phi = \langle \Phi_1, \ldots, \Phi_n, \Phi \rangle$. Furthermore, for $k \in \mathbb{N}$ define $P \circ^k \Phi = (P \circ^{k-1} \Phi) \circ \Phi$ and $P \circ^1 \Phi = P \circ \Phi$.

Whenever the order in the profile does not matter, we also treat $P$ as a set of knowledge bases, e. g., $\bigcup\langle\Phi_1,\ldots,\Phi_n\rangle$ is defined to be $\Phi_1\cup\ldots\cup\Phi_n$.

Let $\mathfrak{K}$ denote the set of all knowledge base profiles. Then Potyka (2018) defines a disagreement measure as follows. Let $\mathbb{R}^{\infty}_{\geq 0}$ be the set of non-negative real values including $\infty$.

**Definition 3.** A disagreement measure $\mathcal{D}$ is a function $\mathcal{D}:\mathfrak{K}\to\mathbb{R}^{\infty}_{\geq 0}$ that satisfies

**Consistency** $\mathcal{D}(P)=0$ iff $\bigcup P$ is consistent.

**Symmetry** $\mathcal{D}(\langle\Phi_1,\ldots,\Phi_n\rangle)=\mathcal{D}(\langle\Phi_{\sigma(1)},\ldots,\Phi_{\sigma(n)}\rangle)$ for each permutation $\sigma$ of $\{1,\ldots,n\}$.

We also write $\mathcal{D}(\Phi_1,\ldots,\Phi_n)$ for $\mathcal{D}(\langle\Phi_1,\ldots,\Phi_n\rangle)$ to ease notation.

The role of a disagreement measure is to assess the degree of disagreement between the knowledge bases in a profile. If no disagreement exists, then there is nothing to be measured (consistency). Symmetry states that the order in which the knowledge bases appear in the profile is irrelevant.

In the following, we define a series of disagreement measures that work on knowledge base profiles with exactly *two* elements. In order to generalise these measures to arbitrary knowledge base profiles, we consider the following constructions.

**Definition 4.** Let $\mathcal{D}$ be a function $\mathcal{D}:\mathbb{K}\times\mathbb{K}\to\mathbb{R}^{\infty}_{\geq 0}$. Then the induced *sum-measure* $\mathcal{D}^{\Sigma}:\mathfrak{K}\to\mathbb{R}^{\infty}_{\geq 0}$ and *max-measure* $\mathcal{D}^{\max}:\mathfrak{K}\to\mathbb{R}^{\infty}_{\geq 0}$ are defined for every profile $P=\langle\Phi_1,\ldots,\Phi_n\rangle$ via

$$\mathcal{D}^{\Sigma}(P)=\sum_{\Phi,\Phi'\in P}\mathcal{D}(\Phi,\Phi')+\mathcal{D}(\bigcup_{i=1}^{n}\Phi_i,\emptyset)$$

$$\mathcal{D}^{\max}(P)=\max(\{\mathcal{D}(\Phi,\Phi')\mid\Phi,\Phi'\in P\}\cup\{\mathcal{D}(\bigcup_{i=1}^{n}\Phi_i,\emptyset)\})$$

The measure $\mathcal{D}^{\Sigma}(P)$ sums up the disagreement values for all pairs of knowledge bases in the profile $P$, while $\mathcal{D}^{\max}(P)$ takes the maximum of those values. Both measures introduce an additional term $\mathcal{D}(\bigcup_{i=1}^{n}\Phi_i,\emptyset)$ to ensure an overall non-zero value if $\bigcup_{i=1}^{n}\Phi_i$ is inconsistent but all pairs of knowledge bases are consistent. Note that this term was not present in the corresponding definitions of the conference version of this paper (Ribeiro et al. 2020); we correct this mistake now by giving the above revised versions.

The two approaches from Definition 4 are just two examples for generalising two-place measures to disagreement measures with an arbitrary number of knowledge bases in the profile. Other possibilities might include taking the product of the pairwise disagreement values or using any other conceivable aggregation function. It should also be noted that the use of aggregation functions such as those given in Definition 4 may not fully account for conflicts that only arise in the presence of three or more knowledge bases. In scenarios where such collective inconsistencies play a significant role, purely pairwise aggregation may underestimate the overall degree of disagreement. However, we will only consider the instances from Definition 4 for the remainder of the paper. In order to obtain valid disagreement measures using these constructions, we only need to require the *Consistency* property from the used two-place functions:

**Consistency**$^2$ $\mathcal{D}(\Phi_1,\Phi_2)=0$ iff $\Phi_1\cup\Phi_2$ is consistent.

Proposition 5. *Let $\mathcal{D}$ be a function $\mathcal{D}:\mathbb{K}\times\mathbb{K}\to\mathbb{R}^{\infty}_{\geq 0}$ that satisfies* Consistency$^2$ *then both $\mathcal{D}^{\Sigma}$ and $\mathcal{D}^{\max}$ are disagreement measures.*

Proof. If $\bigcup P$ is consistent then $\Phi\cup\Phi'$ is consistent as well for all $\Phi,\Phi'\in P$, implying $\mathcal{D}(\Phi,\Phi')=0$. It follows $\mathcal{D}^{\Sigma}(P)=\mathcal{D}^{\max}(P)=0$. Conversely, assume that $\mathcal{D}^{\Sigma}(P)=\mathcal{D}^{\max}(P)=0$. Then $\mathcal{D}(\bigcup_{i=1}^{n}\Phi_i,\emptyset))=0$, so $\bigcup P$ is consistent. This shows that *Consistency* holds. The property *Symmetry* is satisfied by construction. □

Let us now define some concrete measures using interpolants.

**Definition 5.** Let $\Phi_1, \Phi_2$ be finite sets of formulae. Define the measure $\mathcal{D}_{\mathbb{SI}} : \mathbb{K} \times \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ via

$$\mathcal{D}_{\mathbb{SI}}(\Phi_1, \Phi_2) = |\mathbb{SI}(\Phi_1, \Phi_2)|$$

Observe that $\mathcal{D}_{\mathbb{SI}}$ satisfies *Consistency*$^2$, as $\mathbb{SI}(\Phi, \Phi') = \emptyset$ if $\Phi \cup \Phi'$ is consistent. So $\mathcal{D}^{\Sigma}_{\mathbb{SI}}$ and $\mathcal{D}^{\max}_{\mathbb{SI}}$ are disagreement measures according to Proposition 5.

**Example 3.** Consider the three following knowledge bases: $\Phi_1 = \{a, b\}, \Phi_2 = \{b, c\}$ and $\Phi_3 = \{a, \neg b\}$. Going for their interpolants we get:

- $\mathbb{SI}(\Phi_1, \Phi_2) = \mathbb{SI}(\Phi_2, \Phi_1) = \emptyset$, as $\Phi_1$ is consistent with $\Phi_2$;
- $\mathbb{SI}(\Phi_1, \Phi_3) = \{[b], [a \wedge b], [a \to b], [a \leftrightarrow b]\}$ and $\mathbb{SI}(\Phi_3, \Phi_1) = \{[\neg b], [a \to \neg b], [a \wedge \neg b], [\neg(a \leftrightarrow b)]\}$;
- $\mathbb{SI}(\Phi_2, \Phi_3) = \{[b]\}$ and $\mathbb{SI}(\Phi_3, \Phi_2) = \{[\neg b]\}$, as $\mathsf{At}(\{\Phi_2, \Phi_3\}) = \{b\}$.
- $\mathbb{SI}(\Phi_1 \cup \Phi_2 \cup \Phi_3, \emptyset) = \{[\bot]\}$.

Therefore, $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1, \Phi_2, \Phi_3) = 11$ and $\mathcal{D}^{\max}_{\mathbb{SI}}(\Phi_1, \Phi_2, \Phi_3) = 4$.

We consider another measure based on the information content of strongest (resp. weakest) interpolant. We use the following definition of an information measure, similar in spirit to the definition given by Grant and Hunter (2013).

**Definition 6.** An information measure $J$ is a function $J : \mathcal{L}(\mathsf{At}) \to \mathbb{R}^{\infty}_{\geq 0}$ that satisfies the following four properties:

(1) $J(\top) = 0$.
(2) $J(\bot) = \infty$.
(3) If $\phi \models \phi'$ then $J(\phi) \geq J(\phi')$.
(4) If $\phi \not\models \bot$ then $J(\phi) < \infty$.

Here is a simple example of an information measure:

$$J_M(\phi) = \begin{cases} 0 & \text{if } \top \models \phi \\ \infty & \text{if } \phi \models \bot \\ 1/|\mathsf{Mod}(\phi)| & \text{otherwise} \end{cases} \tag{1}$$

It is easy to verify that $J_M$ is indeed an information measure according to Definition 6. Then consider the following measure.

**Definition 7.** Let $\Phi_1, \Phi_2$ be finite sets of formulae and $J$ some information measure. Define the following measure $\mathcal{D}_J : \mathbb{K} \times \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$:

$$\mathcal{D}_J(\Phi_1, \Phi_2) = \begin{cases} 0 & \text{if } \Phi_1 \cup \Phi_2 \text{ is consistent} \\ \infty & \text{if } J(\mathsf{Weakest}(\Phi_1, \Phi_2)) = 0 \\ \frac{J(\mathsf{Strongest}(\Phi_1, \Phi_2))}{J(\mathsf{Weakest}(\Phi_1, \Phi_2))} & \text{otherwise} \end{cases}$$

with $\infty/\infty = 1$.

Again, observe that $\mathcal{D}_J$ satisfies *Consistency*$^2$ (independently of $J$). So $\mathcal{D}^{\Sigma}_J$ and $\mathcal{D}^{\max}_J$ are disagreement measures according to Proposition 5.

**Example 4.** Consider $\mathsf{At} = \{a, b, c\}$ the three knowledge bases from Example 3: $\Phi_1 = \{a, b\}, \Phi_2 = \{b, c\}$ and $\Phi_3 = \{a, \neg b\}$. In the table below, we show their strongest and weakest interpolants as well as their disagreement measure $D_{J_M}$, where $J_M$ is the information measure defined in (1) above.

| | $(\Phi_1, \Phi_3)$ | $(\Phi_3, \Phi_1)$ | $(\Phi_2, \Phi_3)$ | $(\Phi_3, \Phi_2)$ | $(\Phi_1, \Phi_2)$ | $(\Phi_2, \Phi_1)$ |
|---|---|---|---|---|---|---|
| Weakest | $[\neg a \vee b]$ | $[\neg a \vee \neg b]$ | $[b]$ | $[\neg b]$ | - | - |
| Strongest | $[a \wedge b]$ | $[a \wedge \neg b]$ | $[b]$ | $[\neg b]$ | - | - |
| $D_{J_M}$ | 3 | 3 | 1 | 1 | 0 | 0 |

As $\Phi_1$ is consistent with $\Phi_2$ they do not have any interpolant, and therefore their disagreement measure $D_{J_M}$ is zero. Note that each of the weakest interpolants between $\Phi_1$ and $\Phi_3$ have each 6 models, while their strongest interpolants have each 2 models.

We conclude this section by illustrating how the measures can be used in the application scenario from the introduction, i. e., the semi-automatic merging of business rule bases (Corea and Delfmann 2018; Corea and Thimm 2020).

**Example 5.** Three departments of a company maintain rule bases for handling orders, which are to be merged into a single process model. We consider an order of a premium customer placed on a public holiday and use the atoms *prem* (premium customer), *fast* (fast-track processing), *ship* (the order is shipped immediately), *exp* (express delivery), *paid* (the order is prepaid), *cred* (the credit check has been passed), and *hol* (public holiday). The rule bases of the sales department ($\Phi_S$), the finance department ($\Phi_F$), and the logistics department ($\Phi_L$) are

$$\Phi_S = \{prem,\ prem \rightarrow fast,\ fast \rightarrow ship \wedge exp,\ fast \rightarrow \neg paid\}$$
$$\Phi_F = \{ship \rightarrow paid \vee cred,\ \neg cred\}$$
$$\Phi_L = \{hol,\ hol \rightarrow \neg ship,\ exp \rightarrow ship\}$$

Intuitively, the sales department states that the customer is a premium customer, that orders of premium customers are fast-tracked, and that fast-tracked orders are shipped immediately with express delivery and without prepayment. The finance department only allows an order to be shipped if it is prepaid or the credit check has been passed, and the credit check has not been passed for this order. The logistics department states that the order is placed on a public holiday, that no order is shipped immediately on a public holiday, and that express delivery requires immediate shipping. Each rule base is consistent and $\Phi_F \cup \Phi_L$ is consistent as well, but $\Phi_S \cup \Phi_F$ and $\Phi_S \cup \Phi_L$ are inconsistent.

- The shared atoms of $\Phi_S$ and $\Phi_F$ are *ship* and *paid*, and we have $\text{Strongest}(\Phi_S, \Phi_F) = \text{Weakest}(\Phi_S, \Phi_F) = [ship \wedge \neg paid]$. Hence, $\mathbb{SI}(\Phi_S, \Phi_F) = \{[ship \wedge \neg paid]\}$, $\mathcal{D}_{\mathbb{SI}}(\Phi_S, \Phi_F) = 1$, and $\mathcal{D}_{J_M}(\Phi_S, \Phi_F) = 1$. The conflict has a single explanation in terms of the shared vocabulary. The sales rules require the order to be shipped without prepayment, which the finance rules forbid. The atoms *prem*, *fast*, and *cred*, which occur in only one of the two rule bases, do not appear in this explanation.
- The shared atoms of $\Phi_S$ and $\Phi_L$ are *ship* and *exp*, and we have $\text{Strongest}(\Phi_S, \Phi_L) = [ship \wedge exp]$ and $\text{Weakest}(\Phi_S, \Phi_L) = [ship \vee exp]$. Hence, $\mathbb{SI}(\Phi_S, \Phi_L) = \{[ship \wedge exp], [ship], [exp], [ship \vee exp]\}$, $\mathcal{D}_{\mathbb{SI}}(\Phi_S, \Phi_L) = 4$, and $\mathcal{D}_{J_M}(\Phi_S, \Phi_L) = 3$. As both $[ship]$ and $[exp]$ are interpolants, each of the two commitments of the sales department on the shared vocabulary is inconsistent with $\Phi_L$ on its own. Revising only one of them therefore does not resolve the conflict.

The reversed pairs yield the same values here and all remaining pairs of rule bases are consistent. For the profile $P = \langle \Phi_S, \Phi_F, \Phi_L \rangle$ we thus obtain $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(P) = 1 + 1 + 4 + 4 + 1 = 11$, $\mathcal{D}^{\max}_{\mathbb{SI}}(P) = 4$, $\mathcal{D}^{\Sigma}_{J_M}(P) = 1 + 1 + 3 + 3 + 1 = 9$, and $\mathcal{D}^{\max}_{J_M}(P) = 3$.

In such a setting, a modeller can use the pairwise values to decide which conflict to address first, here the one between sales and logistics. The strongest and weakest interpolants then serve as explanations of each conflict that only refer to the vocabulary shared by the two departments. Note also that each shared vocabulary in Example 5 consists of two atoms only, so the strongest and weakest interpolants can be determined with a

few satisfiability checks, one for each interpretation of the shared atoms and each of the two rule bases (cf. the discussion at the end of Section 6).

## 5 Analysis

Potyka (2018) proposes some desirable properties for disagreement measures, inspired by similar properties from inconsistency measurement (Thimm 2018). Let $\mathcal{D}$ be some disagreement measure. For $P = \langle \Phi_1, \ldots, \Phi_n \rangle$, we say that $\Phi_i$, $i \in \{1, \ldots, n\}$, is *involved in a conflict*, if there is $C \subseteq \{1, \ldots, n\}$ such that $\bigcup_{j \in C} \Phi_j$ is consistent but $\Phi_i \cup \bigcup_{j \in C} \Phi_j$ is inconsistent.

**Monotony (MO)** $\mathcal{D}(\Phi_1, \ldots, \Phi_n) \leq \mathcal{D}(\Phi_1 \cup \Phi', \ldots, \Phi_n)$ (for every $\Phi'$).
**Dominance (DO)** For formulae $\phi, \psi$ with $\phi \models \psi$ and $\phi \not\models \bot$, $\mathcal{D}(\Phi_1 \cup \{\phi\}, \ldots, \Phi_n) \geq \mathcal{D}(\Phi_1 \cup \{\psi\}, \ldots, \Phi_n)$.
**Safe Formula Independence (SFI)** For a formula $\phi$ with $\phi \not\models \bot$ and $\mathrm{At}(\phi) \cap \mathrm{At}(\bigcup \Phi_i) = \emptyset$, $\mathcal{D}(\Phi_1 \cup \{\phi\}, \ldots, \Phi_n) = \mathcal{D}(\Phi_1, \ldots, \Phi_n)$.
**Adjunction Invariance (AI)** For formulae $\phi, \psi$, $\mathcal{D}(\Phi_1 \cup \{\phi, \psi\}, \ldots, \Phi_n) = \mathcal{D}(\Phi_1 \cup \{\phi \wedge \psi\}, \ldots, \Phi_n)$.
**Tautology (TA)** If $\Phi$ is a knowledge base with $\top \models \Phi$ then $\mathcal{D}(P) \geq \mathcal{D}(P \circ \Phi)$.
**Contradiction (CO)** If $\Phi$ is inconsistent then $\mathcal{D}(P) \leq \mathcal{D}(P \circ \Phi)$.
**Majority (MAJ)** If $\Phi \in P$ is consistent and involved in a conflict, then there is $k \in \mathbb{N}$ with $\mathcal{D}(P \circ^k \Phi) < \mathcal{D}(P)$.
**Majority Agreement in the Limit (MAJL)** If $M$ is a maximal consistent subset of $\bigcup P$ then $\lim_{k \to \infty} \mathcal{D}(P \circ^k M) = 0$.

*Monotony* states that more information can only increase the disagreement value, as it can add more conflicts. *Dominance* states that, given that $\phi$ entails $\psi$, the formula $\phi$ is at least as involved as $\psi$ in the conflicts. Moreover, if a formula $\phi$ is *safe* (Thimm 2013), i. e., its signature is disjoint from the signature of the rest of the profile and it is not in itself inconsistent, then it does not add anything to the value of disagreement (*Safe Formula Independence*). *Adjunction Invariance* states that disagreement measures are irrelevant to some syntax variances: two formulae contribute as much as their conjunction to the disagreement value. *Tautology* states that adding a tautological knowledge base to the profile cannot increase the disagreement value. Conversely, inconsistent knowledge bases cannot decrease the disagreement value (*Contradiction*). *Majority* states that the disagreement value can be decreased when one consistent part of the inconsistency is sufficiently often duplicated in the profile (thus forming a majority position). In the limit, this should lead to a disagreement value of zero in case that the added information is maximally consistent (*Majority Agreement in the Limit*).

For more discussion on the above properties, as well as some relationships and incompatibilities, we refer to Potyka (2018). Note that the property CO requires the additional premise for $\Phi$ being inconsistent in order to allow for *relative* disagreement measurement, cf. the discussion by Besnard and Grant (2020). In particular, disagreement measures such as the ones by Potyka (2018) and ours, can decrease their value if many consistent knowledge bases that are not in conflict with other knowledge bases are added to the profile.

Note that there is a conceptual difference between adding knowledge $\Phi$ to an existing knowledge base in the profile or appending it as a new knowledge base. More precisely, for $P = \langle \Phi_1, \ldots, \Phi_n \rangle$ and a set of formulae $\Phi$, we do not impose

$$\mathcal{D}(\langle \Phi_1, \ldots, \Phi_n \cup \Phi \rangle) = \mathcal{D}(\langle \Phi_1, \ldots, \Phi_n, \Phi \rangle)$$

In particular, note that our properties above actually require a different behaviour for these scenarios. The property MO requires that adding knowledge to some knowledge base (left-hand side above) cannot decrease the disagreement, while MAJ and MAJL allow for adding elements to profile (right-hand side above) to decrease the disagreement. The reason for this is that the new information added to an existing knowledge base is seen as an extension of a single agent's perspective. This may make that agent's position more specific or even more extreme, which can increase or leave unchanged the overall disagreement, but should not reduce it. On the other

Table 1. Compliance of investigated disagreement measures with postulates.

| | MO | DO | SFI | AI | TA | CO | MAJ | MAJL |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{D}^{\Sigma}_{\mathbb{SI}}$ | ✓ | ✗ | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ |
| $\mathcal{D}^{\max}_{\mathbb{SI}}$ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| $\mathcal{D}^{\Sigma}_{J}$ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| $\mathcal{D}^{\max}_{J}$ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |

hand, adding a separate knowledge base corresponds to adding a new agent to the scenario, which can potentially have a mediating or consensus-building position. In particular, this can reduce the aggregated disagreement in the profile, especially in majority-based procedures, as more agents may share this position.

We check compliance of the disagreement measures $\mathcal{D}^{\Sigma}_{\mathbb{SI}}, \mathcal{D}^{\max}_{\mathbb{SI}}, \mathcal{D}^{\Sigma}_{J}$, and $\mathcal{D}^{\max}_{J}$ against the disagreement measures postulates mentioned above. For this, we introduce the concept of agreement between interpretations. We say that an interpretation $\omega$ agrees with an interpretation $\omega'$ modulo a set of propositional symbols $X$ iff $\omega(p) = \omega'(p)$, for all $p \in X$. Note that the agreement relation is symmetric, that is, if $\omega$ agrees with $\omega'$ modulo $X$ then $\omega'$ agrees with $\omega$ modulo $X$. We recall the Coincidence Lemma on propositional logic:

Lemma 3 (Coincidence Lemma). *If $\omega$ agrees with $\omega'$ modulo $\mathrm{At}(\alpha)$, then $\omega \models \alpha$ iff $\omega' \models \alpha$.*

Proposition 6. *Let $\Phi$ be knowledge base, and $\varphi$ a consistent formula. If $\mathrm{At}(\Phi) \cap \mathrm{At}(\varphi) = \emptyset$, $\mathrm{At}(\alpha) \subseteq \mathrm{At}(\Phi)$, and $\Phi \cup \{\varphi\} \models \alpha$ then $\Phi \models \alpha$.*

Proof. The case that $\Phi$ is inconsistent is trivial. So we focus on the case that $\Phi$ is consistent. Let then $\omega \in \mathrm{Mod}(\Phi)$ be an interpretation of $\Phi$, we will show that $\omega \in \mathrm{Mod}(\alpha)$. As $\varphi$ is consistent, it has at least one interpretation. Let $\omega' \in \mathrm{Mod}(\varphi)$, and let $\omega_0$ be the following interpretation: $\omega_0(p) = \omega(p)$, if $p \in \mathrm{At}(\Phi)$; and $\omega_0(p) = \omega'(p)$, otherwise. Note that $\omega_0$ agrees with $\omega$ on $\mathrm{At}(\Phi)$, which implies that $\omega_0 \models \Phi$. As $\mathrm{At}(\Phi) \cap \mathrm{At}(\varphi) = \emptyset$, we get that $\omega_0$ agrees with $\omega'$ on $\mathrm{At}(\varphi)$. This implies that $\omega_0 \models \varphi$. Therefore, as $\omega_0 \models \Phi$ and $\omega_0 \models \varphi$, we have that $\omega_0 \models \Phi \cup \{\varphi\}$. Thus, as $\Phi \cup \{\varphi\} \models \alpha$, we get that $\omega_0 \models \alpha$. As, by hypothesis, $\mathrm{At}(\alpha) \subseteq \mathrm{At}(\Phi)$, we get that $\omega$ agrees with $\omega_0$ on $\mathrm{At}(\alpha)$. This implies, from Lemma 3, that $\omega \models \alpha$. □

Theorem 1. *The compliance of the measures $\mathcal{D}^{\Sigma}_{\mathbb{SI}}, \mathcal{D}^{\max}_{\mathbb{SI}}, \mathcal{D}^{\Sigma}_{J}$, and $\mathcal{D}^{\max}_{J}$ is as shown in Table 1 (for any information measure $J$).*

Proof. First, observe that when properties MO, DO, SFI, and AI are satisfied for a function $\mathcal{D} : \mathbb{K} \times \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ then they are also satisfied for $\mathcal{D}^{\Sigma}$ and $\mathcal{D}^{\max}$. We will use this fact below to simplify proofs where this is applicable.

- $\mathcal{D}_{\mathbb{SI}}$

  **MO** Follows from Proposition 4, items 3 and 4.

  **DO** Consider the following counter-example. Let $\Phi_1 = \{\top\}$ and $\Phi_2 = \{\neg p, q\}$. Let $\phi$ stand for the formula $p$, and $\psi$ stand for $(p \vee q) \wedge (p \vee \neg q)$. Note that $\mathbb{SI}(\phi, \Phi_2) = \{[p]\}$ and $\mathbb{SI}(\psi, \Phi_2) = \{[p], [p \vee \neg q]\}$, and $\mathbb{SI}(\Phi_2, \phi) = \{[\neg p]\}$ and $\mathbb{SI}(\Phi_2, \psi) = \{[\neg p], [\neg p \wedge q]\}$. Thus, $\mathcal{D}^{max}_{\mathbb{SI}}(\Phi_1 \cup \{\phi\}, \Phi_2) = 1$ and $\mathcal{D}^{max}_{\mathbb{SI}}(\Phi_1 \cup \{\psi\}, \Phi_2) = 2$, while $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1 \cup \{\phi\}, \Phi_2) = 3$ and $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1 \cup \{\psi\}, \Phi_2) = 5$. Note that $\phi \models \psi$ and $\phi \not\models \bot$, but $\mathcal{D}^{\max}_{\mathbb{SI}}(\Phi_1 \cup \{\phi\}, \Phi_2) \not\geq \mathcal{D}^{max}_{\mathbb{SI}}(\Phi_1 \cup \{\psi\}, \Phi_2)$ and $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1 \cup \{\phi\}, \Phi_2) \not\geq \mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1 \cup \{\psi\}, \Phi_2)$.

  **SFI** Let $P = \langle \Phi_1, \Phi_2 \rangle$ be a knowledge profile, and $\phi$ a consistent formula such that $At(\phi) \cap At(\Phi_1) = At(\phi) \cap At(\Phi_2) = \emptyset$. We have to show $\mathbb{SI}(\Phi_1, \Phi_2) = \mathbb{SI}(\Phi_1 \cup \{\phi\}, \Phi_2)$ and $\mathbb{SI}(\Phi_2, \Phi_1) = \mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi\})$. The case that $\Phi_1$ is inconsistent is trivial. So we focus only on the case that $\Phi_1$ is consistent. From Proposition 4, items 3 and 4, we already have that $\mathbb{SI}(\Phi_1, \Phi_2) \subseteq \mathbb{SI}(\Phi_1 \cup \{\phi\}, \Phi_2)$ and $\mathbb{SI}(\Phi_2, \Phi_1) \subseteq \mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi\})$. So we only need to show that $\mathbb{SI}(\Phi_1 \cup \{\phi\}, \Phi_2) \subseteq \mathbb{SI}(\Phi_1, \Phi_2)$ and $\mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi\}) \subseteq \mathbb{SI}(\Phi_2, \Phi_1)$. Let

$[\psi] \in \mathbb{SI}(\Phi_1 \cup \{\phi\}, \Phi_2)$, we need to show that $[\psi] \in \mathbb{SI}(\Phi_1, \Phi_2)$. Thus, for some formula $\alpha \equiv \psi$, we have $\Phi_1 \cup \{\phi\} \models \alpha$, $\Phi_2 \cup \{\alpha\} \models \bot$ and $At(\alpha) \subseteq At(\Phi_1 \cup \{\phi\}) \cap At(\Phi_2)$ which implies that $At(\alpha) \subseteq At(\Phi_2)$. Thus, as $At(\phi) \cap At(\Phi_2) = \emptyset$, we get $At(\alpha) \cap At(\phi) = \emptyset$ and therefore $At(\alpha) \subseteq At(\Phi_1)$. In summary, we have $\phi$ is consistent, $At(\Phi_1) \cap At(\phi) = \emptyset$, $At(\alpha) \subseteq At(\Phi_1)$ and $\Phi_1 \cup \{\phi\} \models \alpha$. Therefore, from Proposition 6, we get $\Phi_1 \models \alpha$. Therefore, $[\alpha] \in \mathbb{SI}(\Phi_1, \Phi_2)$. Thus, as $\alpha \equiv \psi$, we have that $[\psi] \in \mathbb{SI}(\Phi_1, \Phi_2)$. The proof for $\mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi\}) \subseteq \mathbb{SI}(\Phi_2, \Phi_1)$ is analogous.

**AI** Let $\Phi_1$ and $\Phi_2$ be two knowledge bases, and $\phi$ and $\psi$ be two formulae. First, note that $At(\Phi_1 \cup \{\phi, \psi\}) = At(\Phi_1 \cup \{\phi \wedge \psi\})$, and for every formula $\alpha$, $\Phi_1 \cup \{\phi, \psi\} \models \alpha$ iff $\Phi_1 \cup \{\phi \wedge \psi\} \models \alpha$. Therefore, $\mathbb{SI}(\Phi_1 \cup \{\phi, \psi\}, \Phi_2) = \mathbb{SI}(\Phi_1 \cup \{\phi \wedge \psi\}, \Phi_2)$ and $\mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi, \psi\}) = \mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi \wedge \psi\})$. Thus, $\mathcal{D}_{\mathbb{SI}}(\Phi_1 \cup \{\varphi, \psi\}, \Phi_2) = \mathcal{D}_{\mathbb{SI}}(\Phi_1 \cup \{\phi \wedge \psi\}, \Phi_2)$.

**TA** We will show that **TA** is not satisfied by $\mathcal{D}^{\Sigma}_{\mathbb{SI}}$, but it is satisfied by $\mathcal{D}^{\max}_{\mathbb{SI}}$. For $\mathcal{D}^{\Sigma}_{\mathbb{SI}}$, consider the following counter-example. Let $\Phi_1 = \{p \wedge \neg p\}, \Phi_2 = \{q\}$ and $\Phi_\top = \{p \vee \neg p\}$, and note that $\top \models \Phi_\top$. We will show that $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1, \Phi_2) < \mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1, \Phi_2, \Phi_\top)$. Since $\Phi_1$ is inconsistent, Proposition 2(2) yields $\mathbb{I}(\Phi_1, \Phi_1) = \mathcal{L}(\{p\})$ and hence $|\mathbb{SI}(\Phi_1, \Phi_1)| = 4$, while $\mathbb{SI}(\Phi_2, \Phi_2) = \emptyset$ as $\Phi_2$ is consistent. Moreover, $At(\Phi_1) \cap At(\Phi_2) = \emptyset$, which means that $\mathbb{SI}(\Phi_1, \Phi_2) = \{[\bot]\}$ and $\mathbb{SI}(\Phi_2, \Phi_1) = \{[\top]\}$. As $\mathbb{SI}(\Phi_1 \cup \Phi_2, \emptyset) = \{[\bot]\}$, we obtain $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1, \Phi_2) = 4 + 0 + 1 + 1 + 1 = 7$. For the extended profile, note that $\mathbb{SI}(\Phi_\top, \Phi_\top) = \emptyset$, that $\mathbb{SI}(\Phi_1, \Phi_\top) = \{[\bot]\}$ and $\mathbb{SI}(\Phi_\top, \Phi_1) = \{[\top]\}$, and that $\Phi_2 \cup \Phi_\top$ is consistent, so that $\mathbb{SI}(\Phi_2, \Phi_\top) = \mathbb{SI}(\Phi_\top, \Phi_2) = \emptyset$. Since furthermore $\mathbb{SI}(\Phi_1 \cup \Phi_2 \cup \Phi_\top, \emptyset) = \{[\bot]\}$, we get $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1, \Phi_2, \Phi_\top) = 4 + 0 + 0 + 1 + 1 + 1 + 1 + 1 + 0 + 0 + 1 = 9$. Thus, $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1, \Phi_2) = 7 < 9 = \mathcal{D}^{\Sigma}_{\mathbb{SI}}(\Phi_1, \Phi_2, \Phi_\top)$. For $\mathcal{D}^{\max}_{\mathbb{SI}}$, let us suppose for contradiction that it does not satisfy **TA**. Thus, for some $\Phi_\top$ and some knowledge base profile $P = \langle \Phi_1, \cdots, \Phi_n \rangle$: $\mathcal{D}^{\max}_{\mathbb{SI}}(P) \not\geq \mathcal{D}^{\max}_{\mathbb{SI}}(P \circ \Phi_\top)$, and $\top \models \Phi_\top$. This means that $\mathcal{D}^{\max}_{\mathbb{SI}}(P \circ \Phi_\top) > \mathcal{D}^{\max}_{\mathbb{SI}}(P)$. Let $S = \mathcal{D}^{\max}_{\mathbb{SI}}(P)$, then for all $\Phi_i, \Phi_j \in P$, $|\mathbb{SI}(\Phi_i, \Phi_j)| \leq S$. Moreover, as $\mathcal{D}^{\max}_{\mathbb{SI}}(P \circ \Phi_\top) > \mathcal{D}^{\max}_{\mathbb{SI}}(P)$, we have that for some $m \in \{1, \cdots, n\}$, either $|\mathbb{SI}(\Phi_m, \Phi_\top)| > S$ or $|\mathbb{SI}(\Phi_\top, \Phi_m)| > S$. First note that both $\mathbb{SI}(\Phi_m, \Phi_\top)$ and $\mathbb{SI}(\Phi_\top, \Phi_m)$ have at most one interpolant, which is either $[\bot]$ or $[\top]$. This means that $\mathbb{SI}(\Phi_\top, \Phi_m)$ and $\mathbb{SI}(\Phi_m, \Phi_\top)$ have sizes 0 or 1. As either $|\mathbb{SI}(\Phi_\top, \Phi_m)|$ or $|\mathbb{SI}(\Phi_m, \Phi_\top)|$ is strictly greater than $S$, we get that $S < 1$, and either $|\mathbb{SI}(\Phi_\top, \Phi_m)| = 1$ or $|\mathbb{SI}(\Phi_m, \Phi_\top)| = 1$. In either case, $\Phi_m \models \bot$. But as $S < 1$, we get that $S = 0$ which implies that for all $\Phi_i \in P$, $\Phi_i \not\models \bot$ which contradicts the hypothesis that $\Phi_m \models \bot$. Therefore, $\mathcal{D}^{\max}_{\mathbb{SI}}$ satisfies **TA**.

- $\mathcal{D}_J$

**MO** If $\Phi_1 \cup \Phi_2$ is consistent then $\mathcal{D}_J(\Phi_1, \Phi_2) = 0$ and the property is satisfied. If $J(\text{Weakest}(\Phi_1, \Phi_2)) = 0$ then $J(\text{Weakest}(\Phi_1 \cup \Phi', \Phi_2)) = 0$ as well due to Proposition 4, item 6, and Definition 6, item 3. In that case $\mathcal{D}_J(\Phi_1, \Phi_2) = \mathcal{D}_J(\Phi_1 \cup \Phi', \Phi_2) = \infty$. If $J(\text{Weakest}(\Phi_1, \Phi_2)) \neq 0$ and $J(\text{Weakest}(\Phi_1 \cup \Phi', \Phi_2)) = 0$ then $\mathcal{D}_J(\Phi_1, \Phi_2) < \mathcal{D}_J(\Phi_1 \cup \Phi', \Phi_2) = \infty$. It remains the case that $\Phi_1 \cup \Phi_2$ is inconsistent (and so is $\Phi_1 \cup \Phi' \cup \Phi_2$) and $J(\text{Weakest}(\Phi_1, \Phi_2)) \neq 0$ and $J(\text{Weakest}(\Phi_1 \cup \Phi', \Phi_2)) \neq 0$. Then observe

$$\begin{aligned} \mathcal{D}_J(\Phi_1 \cup \Phi', \Phi_2) &= \frac{J(\text{Strongest}(\Phi_1 \cup \Phi', \Phi_2))}{J(\text{Weakest}(\Phi_1 \cup \Phi', \Phi_2))} \\ &\geq \frac{J(\text{Strongest}(\Phi_1, \Phi_2))}{J(\text{Weakest}(\Phi_1, \Phi_2))} = \mathcal{D}_J(\Phi_1, \Phi_2) \end{aligned}$$

Note that the inequality is due to Proposition 4, items 6 and 8, and Definition 6, item 3.

**DO** We recall the counter-example we used on **DO** for $\mathcal{D}_{\mathbb{SI}}$. Let $\text{At} = \{p, q\}$, $\Phi_1 = \{\top\}$, and $\Phi_2 = \{\neg p, q\}$. Let $\phi$ stand for the formula $p$, and $\psi$ stand for $(p \vee q) \wedge (p \vee \neg q)$, and $J_M$ be the information measure from the example of Definition 6. We summarize below the interpolants, the weakest and strongest interpolants of $\phi$ and $\psi$ modulo $\Phi_2$. We also indicate the value of $\mathcal{D}_{J_M}$ for $\phi$ and $\psi$ modulo $\Phi_2$.

| | $(\phi, \Phi_2)$ | $(\Phi_2, \phi)$ | $(\psi, \Phi_2)$ | $(\Phi_2, \psi)$ |
|---|---|---|---|---|
| $\mathbb{SI}$ | $\{[p]\}$ | $\{[\neg p]\}$ | $\{[p], [p \vee \neg q]\}$ | $\{[\neg p], [\neg p \wedge q]\}$ |
| Weakest | $[p]$ | $[\neg p]$ | $[p \vee \neg q]$ | $[\neg p]$ |
| Strongest | $[p]$ | $[\neg p]$ | $[p]$ | $[\neg p \wedge q]$ |
| $\mathcal{D}_{J_M}$ | 1 | 1 | 3/2 | 2 |

For the remaining summands of Definition 4, note that both $\Phi_1 \cup \{\phi\}$ and $\Phi_2$ (and likewise $\Phi_1 \cup \{\psi\}$) are consistent, so the terms $\mathcal{D}_{J_M}(\Phi, \Phi)$ vanish. Moreover, $\Phi_1 \cup \{\phi\} \cup \Phi_2$ is inconsistent and $At(\emptyset) = \emptyset$, so $\mathbb{SI}(\Phi_1 \cup \{\phi\} \cup \Phi_2, \emptyset) = \{[\bot]\}$ and hence $\text{Weakest}(\Phi_1 \cup \{\phi\} \cup \Phi_2, \emptyset) = \text{Strongest}(\Phi_1 \cup \{\phi\} \cup \Phi_2, \emptyset) = [\bot]$, which gives $\mathcal{D}_{J_M}(\Phi_1 \cup \{\phi\} \cup \Phi_2, \emptyset) = 1$; the same holds for $\psi$. Thus, $\mathcal{D}^{\Sigma}_{J_M}(\Phi_1 \cup \{\phi\}, \Phi_2) = 1 + 1 + 1 = 3$ and $\mathcal{D}^{\Sigma}_{J_M}(\Phi_1 \cup \{\psi\}, \Phi_2) = 3/2 + 2 + 1 = 9/2$, while $\mathcal{D}^{\max}_{J_M}(\Phi_1 \cup \{\phi\}, \Phi_2) = 1$ and $\mathcal{D}^{\max}_{J_M}(\Phi_1 \cup \{\psi\}, \Phi_2) = 2$. Note that $\phi \models \psi$ and $\phi \not\models \bot$, but $\mathcal{D}^{\Sigma}_{J_M}(\Phi_1 \cup \{\phi\}, \Phi_2) \not\geq \mathcal{D}^{\Sigma}_{J_M}(\Phi_1 \cup \{\psi\}, \Phi_2)$ and $\mathcal{D}^{\max}_{J_M}(\Phi_1 \cup \{\phi\}, \Phi_2) \not\geq \mathcal{D}^{\max}_{J_M}(\Phi_1 \cup \{\psi\}, \Phi_2)$.

**SFI** It suffices to show that $\mathbb{SI}(\Phi_1, \Phi_i) = \mathbb{SI}(\Phi_1 \cup \{\phi\}, \Phi_i)$ and $\mathbb{SI}(\Phi_i, \Phi_1) = \mathbb{SI}(\Phi_i, \Phi_1 \cup \{\phi\})$. We have already shown in the proof of $\mathcal{D}_{\mathbb{SI}}$ for **SFI**, that this is indeed the case. Therefore, $\mathcal{D}_J$ satisfies **SFI**.

**AI** Let $\Phi_1$ and $\Phi_2$ be two finite knowledge bases. Fist, note that $\mathbb{SI}(\Phi_1 \cup \{\phi, \psi\}, \Phi_2) = \mathbb{SI}(\Phi_1 \cup \{\phi \wedge \psi\}, \Phi_2)$ and $\mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi, \psi\}) = \mathbb{SI}(\Phi_2, \Phi_1 \cup \{\phi \wedge \psi\})$. This means that $\text{Strongest}(\Phi_1 \cup \{\phi, \psi\}, \Phi_2) = \text{Strongest}(\Phi_1 \cup \{\phi \wedge \psi\}, \Phi_2)$ and $\text{Weakest}(\Phi_1 \cup \{\phi, \psi\}, \Phi_2) = \text{Weakest}(\Phi_1 \cup \{\phi \wedge \psi\}, \Phi_2)$. Thus, $\mathcal{D}_J(\Phi_1 \cup \{\phi, \psi\}, \Phi_2) = \mathcal{D}_J(\Phi_1 \cup \{\phi \wedge \psi\}, \Phi_2)$.

**TA** Let $P = \langle \Phi_1, \cdots, \Phi_n \rangle$ be a knowledge profile, and $\Phi_\top$ be a knowledge base such that $\top \models \Phi_\top$. So the profile is consistent and this will result in the value 0. So we proceed with the case that $P$ is inconsistent. If $\mathcal{D}^{\Sigma}_J(P) = \infty$ and $\mathcal{D}^{\max}_J(P) = \infty$ then $\mathcal{D}^{\Sigma}_J(P) \geq \mathcal{D}^{\Sigma}_J(P \circ \Phi_\top)$, and $\mathcal{D}^{\max}_J(P) \geq \mathcal{D}^{\max}_J(P \circ \Phi_\top)$. If either $\mathcal{D}^{\Sigma}_J(P) \neq \infty$ or $\mathcal{D}^{\max}_J(P) \neq \infty$, then for all $\Phi_i, \Phi_j \in P$, we get $J(\text{Weakest}(\Phi_i, \Phi_j)) \neq 0$. As $J$ is an information measure, we have that $\text{Weakest}(\Phi_i, \Phi_j) \neq [\top]$, for all $\Phi_i, \Phi_j \in P$. This implies that $[\top] \notin \mathbb{SI}(\Phi_i, \Phi_j)$, for all $\Phi_i, \Phi_j \in P$. Thus, there is no inconsistent knowledge base in $P$, which implies that $\Phi_i \cup \Phi_\top$ is consistent, for all $\Phi_i \in P$. Therefore, $\mathcal{D}_J(\Phi_i, \Phi_\top) = \mathcal{D}_J(\Phi_\top, \Phi_i) = 0$. Thus,

$$\left(\sum_{\Phi_i \in P} \mathcal{D}_J(\Phi_i, \Phi_\top)\right) + \left(\sum_{\Phi_i \in P} \mathcal{D}_J(\Phi_\top, \Phi_i)\right) = 0,$$

and $max\{\mathcal{D}_J(\Phi_i, \Phi_\top) \mid \Phi_i \in P\} = max\{\mathcal{D}_J(\Phi_\top, \Phi_i) \mid \Phi_i \in P\} = 0$. Therefore, $\mathcal{D}^{\Sigma}_J(P) = \mathcal{D}^{\Sigma}_J(P \circ \Phi_\top)$ and $\mathcal{D}^{\max}_J(P) = \mathcal{D}^{\max}_J(P \circ \Phi_\top)$, which means that $\mathcal{D}^{\Sigma}_J(P) \geq \mathcal{D}^{\Sigma}_J(P \circ \Phi_\top)$ and $\mathcal{D}^{\max}_J(P) \geq \mathcal{D}^{\max}_J(P \circ \Phi_\top)$.

As for properties CO, MAJ and MAJL observe that for every function $\mathcal{D} : \mathbb{K} \times \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$, every knowledge base profile $P$, and knowledge base $\Phi$ we have that

$$\mathcal{D}^{\Sigma}(P) \leq \mathcal{D}^{\Sigma}(P \circ \Phi)$$
$$\mathcal{D}^{\max}(P) \leq \mathcal{D}^{\max}(P \circ \Phi)$$

as both sum (given non-negative summands) and max are monotonic. Therefore, CO is satisfied, whereas neither MAJ and MAJL is satisfied for any of the considered functions. □

Potyka (2018) has defined disagreement measures from inconsistency measures. The idea is that the degree of disagreement in a knowledge profile would correspond to measuring the degree of inconsistencies between the knowledge bases in a knowledge profile. Potyka then discusses about some principles that disagreement measures should satisfy, and shows that many of the disagreement measures induced from inconsistency measures fail to satisfy some of these principles. He then proposes the disagreement measure $D_\eta$, based on the $\eta$-inconsistency

measure (Knight 2002), to capture some of these principles. He shows that $D_\eta$ satisfies all properties mentioned in the beginning of this section, except Adjunction Invariance. We recall the definition of $D_\eta$ first. For that, a function $\pi : \Omega(\mathrm{At}) \to [0,1]$ is a *probability function* on $\Omega(\mathrm{At})$ iff

$$\sum_{\omega \in \Omega(\mathrm{At})} \pi(\omega) = 1$$

A probability function $\pi$ is extended to formulas $\phi \in \mathcal{L}(\mathrm{At})$ via

$$\pi(\phi) = \sum_{\omega \models \phi} \pi(\omega)$$

Now the disagreement measure $D_\eta$ by Potyka (2018) can be defined as follows.

**Definition 8.** Let $D_\eta : \mathfrak{K} \to \mathbb{R}^\infty_{\geq 0}$ be defined via

$$D_\eta(P) = 1 - \max\left\{p \mid \exists \pi : \frac{\sum_{i=1}^n \min\{\pi(\phi) | \phi \in \Phi_i\}}{|P|} = p\right\}$$

for every $P = \langle \Phi_1, \ldots, \Phi_n \rangle$.

In other words, $D_\eta$ is defined based on the maximal probability $p$ that can be consistently assigned to all formulas of all knowledge bases, averaged by the size of the profile. In the following, we illustrate two respects in which $D_\eta$ and the disagreement measures based on interpolants behave differently. These observations are based on examples only. A general formal characterisation of these differences, e. g., in terms of rationality postulates, is beyond the scope of this paper and left for future work.

We start by addressing the Adjunction Invariance axiom. Breach of this axiom may be considered unintuitive in some settings. For instance, both $\{\alpha \wedge \neg\alpha\}$ and $\{\alpha, \neg\alpha\}$ arguably should have the same disagreement value, since they are involved in the same conflicts wrt. any profile. Most of the disagreement measures induced from inconsistency measures analysed by Potyka, including his $D_\eta$ measure, are not adjunctive invariant. On the other hand, as shown above, disagreement measures based on interpolants are adjunctive invariant. This is because interpolants are not syntax sensitive, instead they consider formulae that share the same signature. Note, however, that this is a deliberate design choice for $D_\eta$, which assigns the maximal value 1 exactly to those profiles in which every knowledge base contains a contradictory formula (Potyka 2018, Proposition 13).

Potyka's disagreement measure $D_\eta$ plateaus when the profile consists of singleton knowledge bases whose formulae are satisfiable but pairwise inconsistent (Potyka 2018, Example 8). As he himself criticizes, this scenario makes the measurement purely dependable on the size of the knowledge profile. Precisely, in that case $D_\eta(P) = (|P| - 1)/|P|$. In this scenario, it may be desirable that a disagreement measure still distinguishes an increase of conflicts, even if two knowledge profiles have the same size. We illustrate it in the following example.

**Example 6.** Let $P = \langle \Phi_1, \Phi_2, \Phi_3 \rangle$, and $P' = \langle \Phi_1, \Phi_2, \Phi_4 \rangle$ be two knowledge profiles, where $\Phi_1 = \{a \wedge b\}$, $\Phi_2 = \{\neg a \wedge b\}$, $\Phi_3 = \{\neg b\}$, and $\Phi_4 = \{\neg a \wedge \neg b\}$.[3] Note that the four knowledge bases are individually consistent, but their union is inconsistent. For the $D_\eta$ disagreement measure, we would have $D_\eta(P) = D_\eta(P') = 2/3$.

Note that, against $\Phi_1$ and $\Phi_2$, the knowledge base $\Phi_4$ commits to more of the shared vocabulary than $\Phi_3$. It shares both atoms $a$ and $b$ with $\Phi_1$ and $\Phi_2$ and fixes the truth values of both, in conflict with $\Phi_1$ on $a$ and $b$ and with $\Phi_2$ on $b$. In contrast, $\Phi_3$ shares only $b$ with $\Phi_1$ and $\Phi_2$. One may therefore consider $P'$ to be more conflicting than $P$ and, although both profiles have the same size, it may be desirable that a disagreement measure reflects

[3] Representing the knowledge bases as singletons matters here, since $D_\eta$ is not adjunction invariant. For $\Phi'_1 = \{a, b\}, \Phi'_2 = \{\neg a, b\}, \Phi'_3 = \{\neg b\}$, and $\Phi'_4 = \{\neg b, \neg a\}$ we obtain $D_\eta(\langle \Phi'_1, \Phi'_2, \Phi'_3 \rangle) = D_\eta(\langle \Phi'_1, \Phi'_2, \Phi'_3 \rangle) = 1/2$. The values of the interpolant-based measures are the same for both representations.

this difference. The $D_\eta$ measure does not distinguish the two profiles, since for profiles of satisfiable but pairwise inconsistent singleton knowledge bases its value depends only on the size of the profile.

On the other hand, disagreement measures on interpolants can distinguish the two profiles, as illustrated by the following example:

**Example 7.** Consider the $\mathcal{D}^{\Sigma}_{\mathbb{SI}}$ measure. For the knowledge bases above we would get $\mathcal{D}_{\mathbb{SI}}(\Phi_1, \Phi_2) = 4$, $\mathcal{D}_{\mathbb{SI}}(\Phi_1, \Phi_3) = 1$, $\mathcal{D}_{\mathbb{SI}}(\Phi_1, \Phi_4) = 4$, $\mathcal{D}_{\mathbb{SI}}(\Phi_2, \Phi_3) = 1$ and $\mathcal{D}_{\mathbb{SI}}(\Phi_2, \Phi_4) = 4$. Thus, $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(P) = (4 + 1 + 1) * 2 + 1 = 13$ and $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(P') = (4 + 4 + 4) * 2 + 1 = 25$. Similarly, $\mathcal{D}_{J_M}(\Phi_1, \Phi_2) = \mathcal{D}_{J_M}(\Phi_1, \Phi_4) = \mathcal{D}_{J_M}(\Phi_2, \Phi_4) = 3$ and $\mathcal{D}_{J_M}(\Phi_1, \Phi_3) = \mathcal{D}_{J_M}(\Phi_2, \Phi_3) = 1$, with the same values for the reversed pairs, so that $\mathcal{D}^{\Sigma}_{J_M}(P) = 11$ and $\mathcal{D}^{\Sigma}_{J_M}(P') = 19$. In contrast, the max-based variants do not distinguish the two profiles, as $\mathcal{D}^{\max}_{\mathbb{SI}}(P) = \mathcal{D}^{\max}_{\mathbb{SI}}(P') = 4$ and $\mathcal{D}^{\max}_{J_M}(P) = \mathcal{D}^{\max}_{J_M}(P') = 3$.

This illustrates that the sum-based measures $\mathcal{D}^{\Sigma}_{\mathbb{SI}}$ and $\mathcal{D}^{\Sigma}_{J_M}$ can distinguish profiles that $D_\eta$ treats identically. However, the example should not be read as a general statement. The distinction arises because $\Phi_4$ shares more vocabulary with $\Phi_1$ and $\Phi_2$ than $\Phi_3$ and commits to it, and not because the measures count conflicting formulae. For instance, $\mathcal{D}_{\mathbb{SI}}(\Phi_1, \Phi_2) = \mathcal{D}_{\mathbb{SI}}(\Phi_1, \Phi_4) = 4$, although $\Phi_1$ and $\Phi_2$ disagree on $a$ only, while $\Phi_1$ and $\Phi_4$ disagree on both $a$ and $b$. Moreover, $D_\eta$ satisfies MAJ and MAJL, which none of our measures satisfies (see Table 1), so neither approach dominates the other. A formal characterisation of the behaviour illustrated in this section, e. g., in terms of rationality postulates, is left for future work.

## 6 Computational Complexity

We now come to the main contribution of this paper, namely a thorough analysis of the computational complexity of problems related to our measures. The first part of this section establishes the theoretical foundation for analysing the computational hardness of our disagreement measures. In order to bridge the gap between the abstract notions of Craig interpolation and the concrete disagreement measures introduced in Section 4, we first present intermediate definitions, decision problems, functional problems and counting problems that naturally arise when one works with interpolants. These formal building blocks are essential because the measures $\mathcal{D}_{\mathbb{SI}}$ and $\mathcal{D}_{J_M}$ (and their aggregated versions $\mathcal{D}^{\Sigma}$ and $\mathcal{D}^{\max}$) are defined directly in terms of interpolants. In particular, they rely on the existence of a strongest or weakest interpolant, on the size of the set $\mathbb{SI}(\Phi_1, \Phi_2)$, or on the number of models of these interpolants. By determining where each of these auxiliary problems lies, we obtain a precise picture of the computational resources required to evaluate the measures.

The above pre-analysis serves two purposes. First, it clarifies why certain tasks, such as checking whether a formula is a strongest interpolant (Theorem 4) or counting the number of equivalence classes of interpolants (Theorem 7), are intrinsically hard, which in turn explains the complexity bounds that later appear for the disagreement measures themselves. Second, it provides the reader with a self-contained reference that can be consulted when the later results are applied, without having to reconstruct the underlying reductions from scratch.

Most of the decision problems we investigate fall within the polynomial hierarchy. Let us briefly recall its definition. For every complexity class $C$, let $\mathsf{P}^C$ denote the class of decision problems solved in polynomial time by a deterministic Turing machine augmented with an oracle for $C$. The classes $\mathsf{NP}^C$ and $\mathsf{coNP}^C$ are analogously defined. The polynomial hierarchy is defined as:

$$\Delta^p_{i+1} := \mathsf{P}^{\Sigma^p_i}$$
$$\Sigma^p_{i+1} := \mathsf{NP}^{\Sigma^p_i}$$
$$\Pi^p_{i+1} := \mathsf{coNP}^{\Sigma^p_i}$$

for $i \geq 0$ and $\Delta_0^p = \Sigma_0^p = \Pi_0^p := \mathsf{P}$. A decision problem is in $D_i^p$ iff it is the conjunction of a decision problem in $\Sigma_i^p$ and a decision problem in $\Pi_i^p$. We will also consider function problems (where the corresponding algorithms are supposed to determine a value instead of accepting or rejecting the input as for decision problems). The corresponding complexity classes are indicated by the prefix F. For example, the class FP is the class of function problems that can be solved in polynomial time.

The class Pspace (resp. FPspace) is the class of decision problems (resp. function problems) that can be solved with polynomial space. Moreover, the class Exptime (resp. Fexptime) is the class of decision problems (resp. function problems) that can be solved in exponential time.

We will also consider counting problems, in particular the classes $\# \cdot \mathsf{coNP}$ and $\# \cdot \mathsf{NP}$ from the counting complexity class hierarchy defined by Hemaspaandra and Vollmer (1995). Briefly, if C is a complexity class of decision problems, then $\# \cdot \mathsf{C}$ is the class of all counting problems, where deciding whether a candidate solution is indeed a solution (and must be counted) can be decided in C.

The problem of deciding whether a totally quantified QBF is equivalent to $\top$ is known as True QBF (TQBF, for short), and it is a Pspace-complete problem (Papadimitriou 1994).

The $\forall\exists$-TQBF problem corresponds to the special case of TQBF where the input QBFs are restricted to only two alternations of quantifiers: a sequence of universally quantified variables followed by a sequence of existentially quantified variables. Many decisions problems we investigate fall within the second level of the polynomial hierarchy, precisely $\Pi_2^p$. The canonical problem for $\Pi_2^p$ is the $\forall\exists$-TQBF problem, which we present below.

**Definition 9.** Let $x_1, \cdots, x_n \in \mathsf{At}$ and $y_1, \cdots, y_m \in \mathsf{At}$ be $n + m$ distinct atomic propositional variables. A $\forall\exists$-totally quantified QBF ($\forall\exists$-QBF, for short) is a formula $\forall x_1 \cdots x_n \exists y_1 \cdots y_m \psi$ such that $\psi \in \mathcal{L}(\{x_1, \cdots, x_n, y_1, \cdots, y_m\})$ is a propositional formula.

Formally, the $\forall\exists$-TQBF problem is defined as

| | | |
|---|---|---|
| $\forall\exists$-TQBF | **Input**: | A $\forall\exists$-QBF formula $\alpha$, |
| | **Output**: | True iff $\alpha$ is equivalent to $\top$. |

Theorem 2. *(Papadimitriou 1994) The problem $\forall\exists$-TQBF is $\Pi_2^p$-complete.*

The rest of this section is organised as follows. Section 6.1 gives results on general decision problems about interpolants and Section 6.2 is concerned with functional problems. Section 6.3 discusses results for counting problems and Section 6.4 applies all those results for our concrete disagreement measures.

## 6.1 Decision Problems regarding Interpolation

We first present some general observations on interpolants, which seem—to the best of our knowledge—not have been explicitly mentioned in the literature thus far. More precisely, we consider three main decision problems regarding interpolants and characterise their complexities:

| | | |
|---|---|---|
| Interpolant | **Input**: | formula $\alpha \in \mathcal{L}(\mathsf{At})$, and finite sets $\Phi, \Phi' \subseteq \mathcal{L}(\mathsf{At})$ |
| | **Output**: | True iff $\alpha$ is an interpolant of $\Phi$ w.r.t $\Phi'$ |
| Strongest | **Input**: | formula $\alpha \in \mathcal{L}(\mathsf{At})$, and finite sets $\Phi, \Phi' \subseteq \mathcal{L}(\mathsf{At})$ |
| | **Output**: | True iff $\alpha \in \mathsf{Strongest}(\Phi, \Phi')$ |
| Weakest | **Input**: | formula $\alpha \in \mathcal{L}(\mathsf{At})$, and finite sets $\Phi, \Phi' \subseteq \mathcal{L}(\mathsf{At})$ |
| | **Output**: | True iff $\alpha \in \mathsf{Weakest}(\Phi, \Phi')$ |

For Interpolant we obtain the following characterisation.

Theorem 3. *The decision problem* Interpolant *is* coNP*-complete.*

Proof. For coNP-membership, consider the complement problem of deciding whether $\phi$ is *not* an interpolant of $\Phi$ w.r.t. $\Phi'$ and define the following non-deterministic algorithm. On instance $(\phi, \Phi, \Phi')$ we guess a triple $(\omega, \omega', a)$ with interpretations $\omega$, $\omega'$, and an atom $a$. Then $\phi$ is not an interpolant of $\Phi$ w.r.t. $\Phi'$ if either

(1) $\omega \models \Phi$ and $\omega \not\models \phi$,
(2) $\omega' \models \Phi' \cup \{\phi\}$, or
(3) $a \in \mathsf{At}(\phi) \setminus (\mathsf{At}(\Phi) \cap \mathsf{At}(\Phi'))$.

Observe that each of these checks correspond to disproving one item of Definition 1 and that they can be done in polynomial time. It follows coNP-membership of the original problem.

For coNP-hardness, we reduce the problem Taut to our problem. An instance to Taut consists of a formula $\psi$ and it has to be decided whether $\psi$ is tautological. On input $\psi$ we construct the instance $(\phi, \Phi, \Phi')$ for our problem (let $a$ be a fresh atom not appearing in $\psi$) with

$$\phi = a \qquad \Phi = \{\psi \rightarrow a\} \qquad \Phi' = \{\neg a\}$$

It remains to show that $\psi$ is tautological if and only if $\phi$ is an interpolant of $\Phi$ w.r.t. $\Phi'$. Assume that $\psi$ is tautological, then it follows $\Phi \models a$. Obviously, $\Phi' \cup \{a\}$ is inconsistent and $\phi$ only contains atoms of the shared vocabulary. It follows that $\phi$ is an interpolant of $\Phi$ w.r.t. $\Phi'$. The other direction is analogous. □

For the complexity of problems Strongest and Weakest, we first note that, by Proposition 4(9-10), for all formulae $\varphi, \psi, \alpha$:

$$\alpha \in \mathsf{Weakest}(\varphi, \psi) \text{ iff } \neg\alpha \in \mathsf{Strongest}(\psi, \varphi).$$

In what follows, let $\varphi \in \mathcal{L}(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $\psi \in \mathcal{L}(y_1, \ldots, y_m, z_1, \ldots, z_k)$, with $m, n, k \geq 0$.[4] By Corollary 2, we know that $\exists x_1 \ldots \exists x_n \varphi \in \mathsf{Strongest}(\varphi, \psi)$ and $\forall z_1 \ldots \forall z_k \neg\psi \in \mathsf{Weakest}(\varphi, \psi)$ and, conversely, that for every formula $\alpha \in \mathsf{Strongest}(\varphi, \psi)$ we have $\alpha \equiv \exists x_1 \ldots \exists x_n \varphi$ and for every $\beta \in \mathsf{Weakest}(\phi, \psi)$ we have $\beta \equiv \forall z_1 \ldots \forall z_k \neg\psi$.[5]

Proposition 7. *A formula $\alpha$ is a strongest interpolant of $\{\varphi\}$ w.r.t $\{\psi\}$ iff the following hold:*

*(1) $\varphi \wedge \psi$ is inconsistent;*
*(2) $\mathsf{At}(\alpha) \subseteq \mathsf{At}(\varphi) \cap \mathsf{At}(\psi)$, and*
*(3) $\forall y_1, \ldots, y_m, x_1, \ldots, x_n(\neg\varphi \vee \alpha)$ is equivalent to $\top$, and*
*(4) $\forall y_1, \ldots, y_m \exists x_1, \ldots, x_n(\neg\alpha \vee \varphi)$ is equivalent to $\top$,*

*where $\{y_1, \ldots, y_m\} = \mathsf{At}(\psi) \cap \mathsf{At}(\varphi)$, and $\{y_1, \ldots, y_m, x_1, \ldots, x_n\} = \mathsf{At}(\varphi)$.*

Proof. ($\Rightarrow$) Assume that $\alpha$ is the strongest interpolant of $\{\varphi\}$ w.r.t $\{\psi\}$. Then conditions (1) and (2) are satisfied, and $\{\varphi\} \models \alpha$ which implies that $\neg\varphi \vee \alpha$ is valid. Therefore, $\forall y_1, \ldots, y_m, x_1, \ldots, x_n(\neg\varphi \vee \alpha) \equiv \top$, hence condition (3) holds. By Corollary 2, $\alpha \equiv \exists x_1, \ldots, x_n \varphi$, hence, in particular, $\alpha \models \exists x_1, \ldots, x_n \varphi$. Therefore, $\neg\alpha \vee \exists x_1, \ldots, x_n \varphi$ is valid, so $\forall y_1, \ldots, y_m(\neg\alpha \vee \exists x_1, \ldots, x_n \varphi) \equiv \top$. Condition (4) follows from the fact that, since $\mathsf{At}(\alpha) \subseteq \{y_1, \ldots, y_m\}$, we have:

$$\forall y_1, \ldots, y_m(\neg\alpha \vee \exists x_1, \ldots, x_n \varphi) \equiv \forall y_1, \ldots, y_m \exists x_1, \ldots, x_n(\neg\alpha \vee \varphi).$$

[4] All the considerations which follow hold also for the case in which $\varphi$ and $\psi$ do not have common propositional variables, i.e. $m = 0$. In this case, $\varphi \wedge \psi$ is unsatisfiable iff $\varphi$ or $\psi$ is unsatisfiable. The only formulae over $\{y_1, \ldots, y_m\}$ are $\top$ and $\bot$. If both $\varphi$ and $\psi$ are unsatisfiable, then the strongest interpolant is $\bot$ and the weakest is $\top$. If $\varphi$ is unsatisfiable and $\psi$ is satisfiable, the only interpolant is $\bot$. If $\varphi$ is satisfiable and $\psi$ is unsatisfiable then the only interpolant is $\top$. This is completely compatible with all the results presented below, so in what follows we assume that $n, m, k \geq 0$.

[5] We here use the QBF notation to refer to the propositional formulae equivalent to $\exists x_1 \ldots \exists x_n \varphi$ and $\forall z_1 \ldots \forall z_k \neg\psi$.

($\Leftarrow$) Assume that conditions (1)-(4) hold. From condition (1) we know that $\{\varphi\}$ and $\{\psi\}$ have an interpolant. By condition (2), $\mathsf{At}(\alpha) \subseteq \{y_1, \ldots, y_m\}$, therefore:

$$\begin{array}{lcl} \forall x_1, \ldots, x_n(\neg\varphi \vee \alpha) & \equiv & (\forall x_1, \ldots, x_n \neg\varphi) \vee \alpha \equiv \neg(\exists x_1, \ldots, x_n \varphi) \vee \alpha \\ & \equiv & (\exists x_1, \ldots, x_n \varphi) \rightarrow \alpha \\ \exists x_1, \ldots, x_n(\neg\alpha \vee \varphi) & \equiv & \neg\alpha \vee (\exists x_1, \ldots, x_n \varphi) \\ & \equiv & \alpha \rightarrow (\exists x_1, \ldots, x_n \varphi). \end{array}$$

Thus, conditions (3) and (4) state that $\alpha \equiv (\exists x_1, \ldots, x_n \varphi)$, and by conditions (1) and (2) and by Corollary 2, $\alpha$ is a uniform interpolant of $\varphi$ and $\psi$, hence a strongest interpolant of $\varphi$ w.r.t. $\psi$. □

Lemma 4. *The problem* Strongest *is in* $\Pi_2^p$.

Proof. We construct a polynomial reduction from Strongest to $\forall\exists$-TQBF. Let $\Phi$ and $\Phi'$ be two finite sets of formulae such that $\Phi \subseteq \mathcal{L}(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $\Phi' \subseteq \mathcal{L}(y_1, \ldots, y_m, z_1, \ldots, z_k)$, with $m, n, k \geq 0$, and let $\alpha$ be a formula. Let $\varphi = \bigwedge \Phi$ and $\varphi' = \bigwedge \Phi'$. Observe that $\alpha$ is an interpolant of $\Phi$ w.r.t. $\Phi'$ iff $\alpha$ is an interpolant of $\varphi$ w.r.t. $\varphi'$, and that $\alpha \in \mathsf{Strongest}(\Phi, \Phi')$ iff $\alpha \in \mathsf{Strongest}(\{\varphi\}, \{\varphi'\})$. By Proposition 7, $\alpha$ is a strongest interpolant of $\{\varphi\}$ w.r.t. $\{\varphi'\}$ iff the four conditions below are met:

**C1**: $\varphi \wedge \varphi'$ is inconsistent.
**C2**: $\mathsf{At}(\alpha) \subseteq \mathsf{At}(\varphi) \cap \mathsf{At}(\varphi') \subseteq \{y_1, \ldots, y_m\}$,
**E1**: $\forall y_1, \ldots, y_m, x_1, \ldots, x_n(\neg\varphi \vee \alpha)$ is valid,
**E2**: $\forall y_1, \ldots, y_m \exists x_1, \ldots, x_n(\neg\alpha \vee \varphi)$ is valid.

We can check in polynomial time whether or not $\mathsf{At}(\alpha) \subseteq \mathsf{At}(\varphi) \cap \mathsf{At}(\varphi')$. The conjunction $\varphi \wedge \varphi'$ and the QBFs **E1** and **E2** can also be produced in polynomial time and have size linear in the sum of sizes of $\Phi, \Phi'$ and $\alpha$. Although both QBFs **E1** and **E2** are defined on the same set of variables, in **E1** $x_1, \ldots, x_n$ are universally quantified, while in **E2** $x_1, \ldots, x_n$ are existentially quantified. Therefore, we cannot simply take the conjunction of **E1** and **E2** to produce our final encoding. This can be easily solved by rewriting each $x_i$ in **E2** with new variables $x_i'$, where $i \in \{1, \ldots, n\}$. This yields the equivalent QBF

**E2′**: $\forall y_1, \ldots, y_m \exists x_1', \ldots, x_n'(\neg\alpha \vee \hat{\varphi})$,

where $\hat{\varphi}$ is obtained from $\varphi$ by rewriting each $x_i$ with $x_i'$. Therefore, both **E1** and **E2** are valid QBFs iff

**E**: $\forall y_1, \ldots, y_m, x_1, \ldots, x_n \exists x_1', \ldots, x_n'(\neg\varphi \vee \alpha) \wedge (\neg\alpha \vee \hat{\varphi})$ is valid.

**C1** is equivalent to $\neg\varphi \vee \neg\varphi'$ being valid, i.e. to

**C**: $\forall y_1, \ldots, y_m, x_1, \ldots, x_n, z_1, \ldots, z_k(\neg\varphi \vee \neg\varphi')$ being valid.

Taking the conjunction of **C** with **E** and noting that $\varphi$ does not contain the propositional variables $z_1, \ldots, z_k$ and $\alpha, \hat{\varphi}$ do not contain the propositional variables $x_1, \ldots, x_n, z_1, \ldots, z_k$ we obtain the equivalent QBF:

**EC**: $\forall y_1, \ldots, y_m, x_1, \ldots, x_n, z_1, \ldots, z_k \exists x_1', \ldots, x_k'(\neg\varphi \vee \alpha) \wedge (\neg\alpha \vee \hat{\varphi}) \wedge (\neg\varphi \vee \neg\varphi')$.

This shows that for all sets of formulae $\Phi, \Phi'$ and every formula $\alpha$ we can construct in polynomial time from $(\alpha, \Phi, \Phi')$ the $\forall\exists$-QBF **EC**, such that **EC** $\in \forall\exists$-TQBF iff conditions **C1, E1** and **E2** hold, i.e. iff $(\alpha, \Phi, \Phi') \in$ Strongest. □

Theorem 4. *The problem* Strongest *is* $\Pi_2^p$*-complete.*

Proof. According to Lemma 4, Strongest is in $\Pi_2^p$. So it suffices to show that Strongest is $\Pi_2^p$-hard. For this, we construct a polynomial reduction from $\forall\exists$-TQBF to Strongest, i.e. show that a PTIME computable function $f : \{\forall\overline{y}\exists\overline{x}\,\varphi \mid \varphi \in \mathcal{L}(\mathsf{At})\} \rightarrow \mathcal{L}(\mathsf{At}) \times \mathcal{L}(\mathsf{At}) \times \mathcal{L}(\mathsf{At})$ exists, which maps each $\forall\exists$-QBF $\forall y_1, \ldots, y_m \exists x_1, \ldots, x_n \varphi$ into a triple $(\alpha, \{\psi_1\}, \{\psi_2\})$ such that

$$\forall y_1, \ldots, y_m \exists x_1, \ldots, x_n \varphi \in \forall\exists\text{-TQBF iff } (\alpha, \{\psi_1\}, \{\psi_2\}) \in \textsc{Strongest}.$$

For every $\forall\exists QBF\ \forall y_1, \ldots, y_m \exists x_1, \ldots, x_n \varphi$, let

$$f(\forall y_1, \ldots, y_m \exists x_1, \ldots, x_n \varphi) = (\top, \{\psi_1\}, \{\psi_2\}),$$

where $\psi_1 = \varphi$ and $\psi_2 = \bigwedge_{i=1}^{m}(y_i \wedge \neg y_i) \wedge (z \wedge \neg z)$, with $z \notin \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$.

Clearly, $\psi_1 \wedge \psi_2$ is unsatisfiable, $\psi_1 \models \top$, $\top \wedge \psi_2$ is unsatisfiable and $\mathsf{At}(\top) \subseteq \{y_1, \ldots, y_m\} = \mathsf{At}(\varphi) \cap \mathsf{At}(\psi_2) = \mathsf{At}(\psi_1) \cap \mathsf{At}(\psi_2)$, thus $\top$ is an interpolant of $\psi_1 = \varphi$ and $\psi_2$.

Assume that $\forall y_1, \ldots, y_m \exists x_1, \ldots, x_n \varphi \in \forall\exists$-TQBF, i.e. it evaluates to $\top$. Then $\exists x_1, \ldots, x_n \varphi$ is valid, hence $\exists x_1, \ldots, x_n \varphi \equiv \top$, hence $\top$ is the strongest interpolant of $\psi_1 = \varphi$ w.r.t. $\psi_2$, and so $(\top, \{\psi_1\}, \{\psi_2\}) \in$ Strongest. Assume now that $(\top, \{\psi_1\}, \{\psi_2\}) \in$ Strongest. Then $\top$ is the strongest interpolant of $\psi_1$ w.r.t. $\psi_2$, so $\top \equiv \exists x_1, \ldots, x_n \psi_1$, hence $\exists x_1, \ldots, x_n \psi_1 = \exists x_1, \ldots, x_n \varphi$ is valid, and therefore $\forall y_1 \cdots y_m \exists x_1, \ldots, x_n \varphi \in \forall\exists$-TQBF.

As $f$ can be computed in polynomial time (more precisely, in linear time), we obtain a polynomial reduction from $\forall\exists$-TQBF, which is a $\Pi_2^p$-hard problem, to Strongest. Thus, Strongest is $\Pi_2^p$-hard. Since, by Lemma 4, Strongest is in $\Pi_2^p$, it is $\Pi_2^p$-complete. □

Theorem 5. *The problem Weakest is $\Pi_2^p$-complete.*

Proof. We can inter-reduce in polynomial time Weakest and Strongest. Let $\alpha \in \mathcal{L}(\mathsf{At})$, and $\Phi$ and $\Phi'$ be finite sets of formulae in classical propositional logics. According to Proposition 4 (items 9 - 10), $\alpha$ is a strongest interpolant of $\Phi$ w.r.t. $\Phi'$ iff $\neg\alpha$ is a weakest interpolant of $\Phi'$ w.r.t. $\Phi$. Therefore, by negating $\alpha$ and exchanging $\Phi$ with $\Phi'$ we obtain a polynomial time reduction between the two problems. As Strongest is $\Pi_2^p$-complete, we conclude that Weakest is $\Pi_2^p$-complete. □

## 6.2 Functional Problems regarding Interpolation

We consider some functional problems related to interpolation:

| | | |
|---|---|---|
| FInterpolant | **Input**: | a pair of finite sets $(\Phi, \Phi') \subseteq 2^{\mathcal{L}(\mathsf{At})} \times 2^{\mathcal{L}(\mathsf{At})}$ |
| | **Output**: | an interpolant of $\Phi$ w.r.t $\Phi'$, if $\mathbb{I}(\Phi, \Phi') \neq \emptyset$; reject, otherwise. |
| Fweakest | **Input**: | a pair of finite sets $(\Phi, \Phi') \in 2^{\mathcal{L}(\mathsf{At})} \times 2^{\mathcal{L}(\mathsf{At})}$ |
| | **Output**: | a weakest interpolant of $\Phi$ w.r.t. $\Phi'$, if $\mathbb{I}(\Phi, \Phi') \neq \emptyset$; reject, otherwise. |
| Fstrongest | **Input**: | a pair of finite sets $(\Phi, \Phi') \in \mathcal{L}(\mathsf{At}) \times \mathcal{L}(\mathsf{At})$ |
| | **Output**: | a strongest interpolant of $\Phi$ w.r.t. $\Phi'$, if $\mathbb{I}(\Phi, \Phi') \neq \emptyset$; reject, otherwise. |

We start by showing that Fweakest and Fstrongest are in FPspace, i. e. also in Fexptime (which implies that FInterpolant is also in FPspace and Fexptime, respectively), and discuss methods for computing strongest and weakest interpolants and their complexity.

Proposition 8. *The Fstrongest problem is in FPspace.*

Proof. Let $\Phi$ and $\Phi'$ be sets of formulae. Checking the satisfiabiliy of $\Phi \cup \Phi'$ is in Pspace. We consider the case in which $\Phi \cup \Phi' \models \bot$. Assume that $\mathsf{At}(\Phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ and $\mathsf{At}(\Phi') = \{y_1, \ldots, y_m, z_1, \ldots, z_k\}$.

We generate all models of $\Phi$, one after the other. This can be done reusing the space. For every model $\omega$ of $\Phi$ construct $C_\omega = y_1^{\omega(y_1)} \wedge \cdots \wedge y_m^{\omega(y_m)}$, where $x^{\mathsf{true}} = x$ and $x^{\mathsf{false}} = \neg x$. We write all these conjunctions directly to the output (in the form of a disjunction) while processing the different models $\omega$ of $\Phi$. Let $I = \bigvee_{\omega \in \mathsf{Mod}(\Phi)} C_\omega$. We prove that $I$ is the strongest interpolant of $\Phi$ w.r.t. $\Phi'$:

(i) $I$ contains only shared propositional variables;

(ii) $\Phi \models I$ (for every model $\omega$ of $\Phi$, $C_\omega$ is true);
(iii) $\Phi' \wedge I \models \bot$, otherwise there would be a model $\omega'$ of $\Phi'$ which is also a model of $I$, i.e, $\omega' \models C_\omega$ for some model $\omega$ of $\Phi$. But then $\omega'$ and $\omega$ agree on $\{y_1, \ldots, y_n\}$, so the interpretation $\omega''$ defined by $\omega''(x_i) = \omega(x_i), \omega''(y_j) = \omega(y_j) = \omega'(y_j)$ and $\omega''(z_l) = \omega'(z_l)$ for all $1 \leq i \leq n, 1 \leq j \leq m, 1 \leq l \leq k$ would be a model for both $\Phi$ and $\Phi'$ which is a contradiction.
(iv) Let $\alpha$ be another formula containing only variables $\{y_1, \ldots, y_n\}$ entailed by $\Phi$. Let $\omega'$ be a model of $I$. Then $\omega' \models C_\omega$ for some model $\omega$ of $\Phi$. This means that $\omega$ and $\omega'$ agree on $\{y_1, \ldots, y_m\}$. Since $\Phi \models \alpha$, $\omega \models \alpha$. But since $\omega$ and $\omega'$ agree on $\{y_1, \ldots, y_m\}$, $\omega' \models \alpha$ as well. So $I \models \alpha$.

The problem is in FPSPACE because the method described above uses only polynomial space if we ignore the output. □

COROLLARY 3. *The problems FINTERPOLANT and FWEAKEST are in FPSPACE.*

PROOF. Since every strong interpolant of two formulae is an interpolant of the formulae, Proposition 8 shows that FINTERPOLANT is in FPSPACE. According to Proposition 4 (items 9 - 10), $\alpha$ is a weakest interpolant of $\Phi$ w.r.t. $\Phi'$ iff $\neg\alpha$ is a strongest interpolant of $\Phi'$ w.r.t. $\Phi$. Therefore, we can compute a weakest interpolant of $\Phi$ w.r.t. $\Phi'$ by computing a strongest interpolant of $\Phi'$ w.r.t. $\Phi$ and negating the result, so FWEAKEST is also in FPSPACE. □

One can effectively compute strongest interpolants (hence by Proposition 2 (item 5) and Proposition 4 (items 9–10), also weakest interpolants) using a variant of ordered resolution applied to the CNF of the conjunction of the formulae from $\Phi$ (cf. e.g. (Peuter and Sofronie-Stokkermans 2021) for a more general result or (Koopmann et al. 2025)); using tableau calculi (Koopmann et al. 2025); using methods for quantifier elimination for propositional logic as those implemented in Redlog[6] (Dolzmann and Sturm 1997) (cf. (Seidl and Sturm 2003; Sturm and Zengler 2010)); methods for quantifier elimination in finite fields using Gröbner bases (Gao et al. 2011); or methods for quantifier elimination for OBDDs (Bryant 1986, 1992). The time complexity of such methods is in general exponential, although for certain classes of formulae polynomial complexity can be guaranteed (for instance, methods based on ordered resolution terminate in polynomial time on 2-CNF formulae; Gröbner basis computation is Gaussian elimination if applied to linear polynomials).

## 6.3 Counting Problems regarding Interpolation

We consider the following problems regarding counting the number of interpolants:

| | | |
|---|---|---|
| #INTERPOLANT | **Input**: | a pair of finite sets $(\Phi, \Phi') \in 2^{\mathcal{L}(\mathsf{At})} \times 2^{\mathcal{L}(\mathsf{At})}$ |
| | **Output**: | $\lvert\mathbb{SI}(\Phi, \Phi')\rvert$. |
| #WEAKEST | **Input**: | a pair $(\Phi, \Phi') \in 2^{\mathcal{L}(\mathsf{At})} \times 2^{\mathcal{L}(\mathsf{At})}$, such that $\Phi \cup \Phi'$ is inconsistent |
| | **Output**: | $\lvert\mathsf{Mod}_{\mathsf{At}(\phi)\cap\mathsf{At}(\Phi')}(\phi_w)\rvert$, where $\phi_w \in \mathsf{Weakest}(\Phi, \Phi')$. |
| #STRONGEST | **Input**: | a pair $(\Phi, \Phi') \in 2^{\mathcal{L}(\mathsf{At})} \times 2^{\mathcal{L}(\mathsf{At})}$, such that $\Phi \cup \Phi'$ is inconsistent |
| | **Output**: | $\lvert\mathsf{Mod}_{\mathsf{At}(\phi)\cap\mathsf{At}(\Phi')}(\phi_s)\rvert$, where $\phi_s \in \mathsf{Strongest}(\Phi, \Phi')$. |

The time complexities of these problems are related to two classes of counting problems: # · coNP-complete and # · NP-complete (Durand et al. 2005). We show below two complete problems for these classes.

| | | |
|---|---|---|
| #$\Pi_1$SAT | **Input**: | a QBF $\forall x_1, \ldots, x_n\, \phi$ with $\mathsf{At}(\phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ |
| | **Output**: | $\lvert\{\omega : \{y_1, \ldots, y_m\} \to \{\mathsf{true}, \mathsf{false}\} \mid \forall x_1, \ldots, x_n\, \omega(\phi) \text{ is true}\}\rvert$ |

[6] https://redlog.eu/documentation/tutorial.php?key=tutorial-1.

#$\Sigma_1$SAT **Input**: a QBF $\exists x_1, \ldots, x_n\, \phi$ with $\mathrm{At}(\phi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$
**Output**: $|\{\omega : \{y_1, \ldots, y_m\} \to \{\text{true}, \text{false}\} \mid \exists x_1, \ldots, x_n\, \omega(\phi) \text{ is true}\}|$,
where $\omega(\phi)$ is the (partial) evaluation of $\phi$ under $\omega$ (and as such a formula over $\{x_1, \ldots, x_n\}$).

The problem #$\Pi_1$SAT is # · coNP-complete, while #$\Sigma_1$SAT is # · NP-complete (Durand et al. 2005).

Let $\Phi, \Phi'$ be sets of formulae over At, and let $\mathrm{At}_{\Phi,\Phi'} = \mathrm{At}(\Phi) \cap \mathrm{At}(\Phi')$ be the set of all their common atoms. There is a bijective correspondence between the (equivalence classes of) formulae over $\mathrm{At}_{\Phi,\Phi'}$ and the subsets of the set of maps $\omega : \mathrm{At}_{\Phi,\Phi'} \to \{\text{false}, \text{true}\}$. For every formula over $\mathrm{At}_{\Phi,\Phi'}$ we can associate the set of its models; and for every set of valuations we can find a formula which evaluates to true exactly for those valuations. In what follows, we use this correspondence between formulae and subsets of valuations to find a way of expressing the number of (equivalence classes of) interpolants of $\Phi$ w.r.t. $\Phi'$.

Recall that $\mathrm{Mod}_A(\phi) = \{\omega : A \to \{\text{true}, \text{false}\} \mid \omega \models \phi\}$.

PROPOSITION 9. *Let $\Phi, \Phi'$ be finite sets of formulae over* At, *$\varphi_s$ be a strongest interpolant and $\varphi_w$ a weakest interpolant of $\Phi$ w.r.t. $\Phi'$. Let $A = \mathrm{At}(\Phi) \cap \mathrm{At}(\Phi')$. There is a bijection between the set of all (equivalence classes of) interpolants of $\Phi$ w.r.t. $\Phi'$ and the set of all subsets of $\mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w)$. Therefore, the number of all (equivalence classes of) interpolants of $\Phi$ w.r.t. $\Phi'$ is $2^{|\mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w)|}$, i.e. is equal to $2^{|\mathrm{Mod}_A(\varphi_w)| - |\mathrm{Mod}_A(\varphi_s)|}$.*

PROOF. We first prove that the set of interpolants of $\Phi$ w.r.t. $\Phi'$ is the set of all formulae of the form $\varphi_s \vee F$, where $F$ is a formula over the common atoms of $\Phi$ and $\Phi'$, and $F \models \neg\varphi_s \wedge \varphi_w$.

Let $\gamma$ be an interpolant of $\Phi$ w.r.t. $\Phi'$. Then $\varphi_s \models \gamma$ and $\gamma \models \varphi_w$. Since $\varphi_s \models \gamma$, $\gamma \equiv \varphi_s \vee \gamma \equiv \varphi_s \vee (\neg\varphi_s \wedge \gamma)$. Since $\gamma \models \varphi_w$, it follows that $(\neg\varphi_s \wedge \gamma) \models \varphi_w$. Therefore, $\gamma$ is of the form $\varphi_s \vee F$ (for $F = \neg\varphi_s \wedge \gamma$), where $F \models \neg\varphi_s \wedge \varphi_w$.[7]

Conversely, let $\gamma$ be a formula with atoms in $A$ of the form $\gamma = \varphi_s \vee F$, where $F \models \neg\varphi_s \wedge \varphi_w$. Then $\gamma$ has the properties (i) $\varphi_s \models \gamma$, and (ii) $\gamma \models \varphi_w$ (because $\varphi_s \models \varphi_w$ and $F \models \varphi_w$). From (i) and the fact that $\Phi \models \varphi_s$ it follows that $\Phi \models \gamma$. From (ii) and the fact that $\Phi' \cup \{\varphi_w\}$ is unsatisfiable it follows that $\Phi' \cup \{\gamma\}$ is unsatisfiable. Therefore, $\gamma$ is an interpolant of $\Phi$ w.r.t. $\Phi'$.

Note also that if $\varphi_s \vee F_1 \equiv \varphi_s \vee F_2$ for two formulae $F_1, F_2$ with $F_1 \models \neg\varphi_s \wedge \phi_w$ and $F_2 \models \neg\varphi_s \wedge \phi_w$ then $\neg\varphi_s \wedge (\varphi_s \vee F_1) \equiv \neg\varphi_s \wedge (\varphi_s \vee F_2)$, so $\neg\varphi_s \wedge F_1 \equiv \neg\varphi_s \wedge F_2$, hence, as $F_1 \models \neg\varphi_s$ and $F_2 \models \neg\varphi_s$, $F_1 \equiv \neg\varphi_s \wedge F_1 \equiv \neg\varphi_s \wedge F_2 \equiv F_2$. Therefore,

$$\begin{aligned} \mathbb{I}(\Phi, \Phi') &= \{\varphi_s \vee F \mid F \in \mathcal{L}(A), F \models \neg\varphi_s \wedge \varphi_w\} \\ \mathbb{SI}(\Phi, \Phi') &= \{[\varphi_s \vee F] \mid F \in \mathcal{L}(A), F \models \neg\varphi_s \wedge \varphi_w\}. \end{aligned}$$

By the remarks above, $\{[\varphi_s \vee F] \mid F \in \mathcal{L}(A), F \models \neg\varphi_s \wedge \varphi_w\}$ is in bijection with $\{[F] \mid F \in \mathcal{L}(A), F \models \neg\varphi_s \wedge \varphi_w\}$.

Therefore, in order to determine the number of all the (equivalence classes of) interpolants of $\Phi$ w.r.t. $\Phi'$ it is sufficient to determine the number of all (equivalence classes of) formulae $F$ over the common atoms of $\Phi$ and $\Phi'$ such that $F \models \neg\varphi_s \wedge \varphi_w$.

For every formula $F$ with atoms $A$ such that $F \models \neg\varphi_s \wedge \varphi_w$, the set of all interpretations $\omega : A \to \{\text{true}, \text{false}\}$ with $\omega \models F$, which completely determines $F$ (up to equivalence), is a subset of $\mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w)$. Conversely, every subset of $\mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w)$ defines the equivalence class of a formula $F$ with $F \models \neg\varphi_s \wedge \varphi_w$. Therefore, $|\{[F] \mid F \in \mathcal{L}(A), F \models \neg\phi_s \wedge \phi_w\}| = 2^{|\mathrm{Mod}_A(\neg\phi_s \wedge \phi_w)|}$.

It can be seen that $|\mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w)| = |\mathrm{Mod}_A(\varphi_w)| - |\mathrm{Mod}_A(\varphi_s)|$. Indeed, for every $\omega : A \to \{\text{true}, \text{false}\}$ we have: $\omega \in \mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w)$ iff $\omega \in \mathrm{Mod}_A(\varphi_w)$ and $\omega \notin \mathrm{Mod}_A(\varphi_s)$. Thus, $\mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w) = \mathrm{Mod}_A(\varphi_w) \setminus \mathrm{Mod}_A(\varphi_s)$. Since $\varphi_s \models \varphi_w$, it follows that $\mathrm{Mod}_A(\varphi_s) \subseteq \mathrm{Mod}_A(\varphi_w)$. Hence, $|\mathrm{Mod}_A(\neg\varphi_s \wedge \varphi_w)| = |\mathrm{Mod}_A(\varphi_w)| - |\mathrm{Mod}_A(\varphi_s)|$. □

THEOREM 6. *Let $\Phi$ and $\Phi'$ be two finite sets of formulae with $\Phi \subseteq \mathcal{L}(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $\Phi' \subseteq \mathcal{L}(y_1, \ldots, y_m, z_1, \ldots, z_k)$ such that $\Phi \cup \Phi' \models \bot$. Let $\varphi = \bigwedge \Phi$ and $\varphi' = \bigwedge \Phi'$.*

[7] If $F = \bot$ then $\varphi_s \vee F$ is $\varphi_s$ (the strongest interpolant), if $F = \neg\varphi_s \wedge \varphi_w$ then $\varphi_s \vee F$ is $\varphi_w$ (the weakest interpolant).

*(1) Let $q$ be the number of models (w.r.t. variables $\{y_1, \ldots, y_n\}$) of $\forall x_1 \ldots x_n \neg\varphi \wedge \forall z_1 \ldots z_k \neg\varphi'$. Then the number of (equivalence classes of) interpolants of $\Phi$ w.r.t. $\Phi'$ is $2^q$.*
*(2) Let $m$ and $n$ be the number of models (w.r.t. variables $\{y_1, \ldots, y_n\}$) of $\phi_w$ and $\phi_s$ respectively, where $\phi_s$ is the strongest and $\phi_w$ is the weakest interpolant of $\Phi$ w.r.t. $\Phi'$. Then the number of (equivalence classes of) interpolants of $\Phi$ w.r.t. $\Phi'$ is $2^{m-n}$.*

Proof. This follows from Proposition 9, the fact that $\varphi_s = \exists x_1 \ldots x_n \varphi$ is a strongest interpolant of $\Phi$ w.r.t. $\Phi'$, $\varphi_w = \forall z_1 \ldots z_k \neg\varphi'$ is a weakest interpolant of $\Phi$ w.r.t. $\Phi'$, and $\neg\varphi_s \wedge \varphi_w \equiv \forall x_1 \ldots x_n \neg\varphi \wedge \forall z_1 \ldots z_k \neg\varphi'$. □

The above observation allows us to link the complexity of #Interpolant to the complexity of counting models in quantified Boolean formulae with free variables. More precisely, we can observe a close relationship to the problem of counting the number of models (w.r.t. variables $y_1, \ldots, y_m$) of a formula $\forall x_1, \ldots, x_n\, \phi$, which is a $\# \cdot$ coNP-complete problem (Durand et al. 2005). In order to simplify the reduction, we consider a slight variant of the problem #Interpolant:

| | | |
|---|---|---|
| #Interpolant' | **Input**: | a pair of finite sets $(\Phi, \Phi') \in 2^{\mathcal{L}(\mathrm{At})} \times 2^{\mathcal{L}(\mathrm{At})}$ |
| | **Output**: | $\log_2 \lvert\mathbb{SI}(\Phi, \Phi')\rvert$ |

Note that $x$ is the solution to an instance of #Interpolant' if and only if $2^x$ is the solution to the same instance of #Interpolant.

**Remark 4.** By Theorem 6, $\lvert\mathbb{SI}(\Phi, \Phi')\rvert$ is always of the form $2^q$ with $q \leq 2^{\lvert At(\Phi) \cap At(\Phi')\rvert}$. Written in binary, $q$ requires only polynomially many bits in the size of the input, whereas $2^q$ requires exponentially many. Since every function in $\# \cdot$ coNP is bounded by $2^{p(n)}$ for some polynomial $p$, the problem #Interpolant cannot belong to $\# \cdot$ coNP for reasons of output length alone, independently of the difficulty of the underlying counting task. The two problems are nevertheless equivalent in the sense that each is obtained from the other by the bijective rescaling $x \mapsto 2^x$, which involves no computation beyond writing down the result. We therefore use the exponent $q$ as a succinct representation of $\lvert\mathbb{SI}(\Phi, \Phi')\rvert$ and state all completeness results in terms of #Interpolant'.

For #Interpolant', we obtain the following result.

Theorem 7. *The problem #Interpolant' is $\# \cdot$ coNP-complete.*

Proof. Let $(\Phi, \Phi') \in 2^{\mathcal{L}(\mathrm{At})} \times 2^{\mathcal{L}(\mathrm{At})}$ be the input to #Interpolant' with $\Phi \subseteq \mathcal{L}(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $\Phi' \subseteq \mathcal{L}(y_1, \ldots, y_m, z_1, \ldots, z_k)$. Let $\omega$ be an interpretation $\omega : \{y_1, \ldots, y_m\} \to \{\mathrm{true}, \mathrm{false}\}$. Define $\varphi = \bigwedge \Phi$ and $\varphi' = \bigwedge \Phi'$ and

$$\varphi_\omega := \varphi[y_1 \mapsto \omega(y_1), \ldots, y_m \mapsto \omega(y_m)]$$
$$\varphi'_\omega := \varphi'[y_1 \mapsto \omega(y_1), \ldots, y_m \mapsto \omega(y_m)]$$

$\omega$ is a model of $\forall x_1 \ldots \forall x_n \neg\varphi \wedge \forall z_1 \ldots z_k \neg\varphi'$ iff $\neg\varphi_\omega$ and $\neg\varphi'_\omega$ are both valid. Checking that both $\neg\varphi_\omega$ and $\neg\varphi'_\omega$ are valid is a coNP-complete problem. By Theorem 6, the solution of #Interpolant' is the number of such $\omega$, so it is in $\# \cdot$ coNP.

For an instance $I = \forall x_1, \ldots, x_n\, \varphi$ of $\#\Pi_1$Sat with $\mathrm{At}(\varphi) = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ define

$$\Phi_I = \{\neg\varphi\}$$
$$\Phi'_I = \{y_1, \ldots, y_m, \neg y_1, \ldots, \neg y_m, z, \neg z\}, \text{ where } z \notin \{x_1, \ldots, x_n, y_1, \ldots, y_m\}.$$

By Theorem 6 we know that the number of equivalence classes of interpolants of $\Phi_I$ wrt. $\Phi'_I$ is $2^q$ where $q$ is the number of models (w.r.t. $\{y_1, \ldots, y_n\}$) of

$$\begin{aligned} & \forall x_1, \ldots, x_n \neg\neg\varphi \wedge \forall z \neg(y_1 \wedge \ldots \wedge y_m \wedge \neg y_1 \wedge \ldots \wedge \neg y_m \wedge z \wedge \neg z) \\ \equiv\ & \forall x_1, \ldots, x_n : \varphi \wedge \top \\ \equiv\ & \forall x_1, \ldots, x_n\, \varphi \end{aligned}$$

So an algorithm for #Interpolant' returns $\log_2 2^q = q$, which is the solution of #$\Pi_1$Sat for the instance $I$. Since the construction of $\Phi_I$ and $\Phi'_I$ is polynomial in size, this shows # · coNP-hardness. □

Lemma 5. *Let $\Phi$ and $\Phi'$ be finite jointly inconsistent sets of formulae, and $A = \mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')$. If $\alpha \in \mathsf{Weakest}(\Phi, \Phi')$ and $\omega : A \to \{\mathsf{true}, \mathsf{false}\}$ is an interpretation such that $\omega \not\models \alpha$ then there is an interpretation $\omega' : \mathsf{At}(\Phi') \to \{\mathsf{true}, \mathsf{false}\}$ such that $\omega' \models \Phi'$ and $\omega(x) = \omega'(x)$ for all $x \in A$.*

Proof. Let $\Phi$ and $\Phi'$ be finite jointly inconsistent sets of formulae. Let $\alpha \in \mathsf{Weakest}(\Phi, \Phi')$, and $\omega : A \to \{\mathsf{true}, \mathsf{false}\}$ be an interpretation that does not satisfy $\alpha$. Let

$$\gamma = \bigwedge \Big( \{x \in A \mid \omega(x) = \mathsf{true}\} \cup \{\neg x \mid x \in A, \omega(x) = \mathsf{false}\} \Big).$$
$$\beta = \alpha \vee \gamma.$$

Thus, $\alpha \models \beta$. However, as $\omega \not\models \alpha$, it follows that $\beta \not\models \alpha$. Therefore, as $\alpha$ is a weakest interpolant, it follows that $\beta \notin \mathbb{I}(\Phi, \Phi')$. As $\alpha$ is a weakest interpolant of $\Phi$ w.r.t. $\Phi'$, it follows that $\Phi \models \alpha$. Therefore, as $\alpha \models \beta$, we get that $\Phi \models \beta$. By construction, $\mathsf{At}(\beta) = A$. Therefore, as $\beta \notin \mathbb{I}(\Phi, \Phi')$, it follows that $\beta \wedge \bigwedge \Phi'$ is consistent. So, there is an interpretation $\omega' : \mathsf{At}(\Phi') \to \{\mathsf{true}, \mathsf{false}\}$ that satisfies $\beta \wedge \bigwedge \Phi'$. Therefore, $\omega' \models \Phi'$. By definition, $\beta = \alpha \vee \gamma$. Thus, as $\alpha$ is inconsistent with $\Phi'$, it follows that $\beta \wedge \bigwedge \Phi' \equiv \gamma \wedge \bigwedge \Phi'$. Thus, $\omega' \models \gamma$. Note that, by construction, $\omega' \models \gamma$ iff $\omega'(x) = \omega(x)$, for all $x \in A$. □

Theorem 8. *Let $\Phi$ and $\Phi'$ be jointly inconsistent sets of formulae, $A = \mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')$, and $\alpha \in \mathsf{Weakest}(\Phi, \Phi')$, and let $\varphi' = \bigwedge \Phi'$. An interpretation $\omega : A \to \{\mathsf{true}, \mathsf{false}\}$ satisfies $\alpha$ iff the QBF $\forall x_1, \cdots, x_n \omega(\neg\varphi')$ is valid, where $\{x_1, \cdots, x_n\} = \mathsf{At}(\Phi') \setminus A$.*

Proof. Let $\Phi$ and $\Phi'$ be jointly inconsistent formulae, $\varphi' = \bigwedge \Phi'$, $A = \mathsf{At}(\Phi) \cap \mathsf{At}(\Phi')$, and $\alpha \in \mathsf{Weakest}(\Phi, \Phi')$.

($\Rightarrow$) Let us suppose for contradiction purposes that there is an interpretation $\omega : A \to \{\mathsf{true}, \mathsf{false}\}$ satisfying $\alpha$, but the QBF $\forall x_1, \cdots, x_n\, \omega(\neg\varphi')$ is not valid. Let $\psi = \omega(\neg\varphi')$. Thus, there is some interpretation $\omega' : \mathsf{At}(\Phi') \to \{\mathsf{true}, \mathsf{false}\}$ that does not satisfy $\psi$. Thus, $\omega' \models \neg\psi$. We define the interpretation $\omega'' : \mathsf{At}(\Phi') \to \{\mathsf{true}, \mathsf{false}\}$, where

$$\omega''(x) = \begin{cases} \omega(x), & \text{if } x \in A \\ \omega'(x), & \text{otherwise} \end{cases}$$

As $\neg\psi$ has no symbol from $A$, and $\omega''$ assigns the same truth values as $\omega'$ to the propositional variables not in $A$, it follows that $\omega''$ satisfies $\neg\psi$. As $\psi = \omega(\neg\varphi')$, it follows that $\neg\psi = \neg\omega(\neg\varphi')$ which implies that $\neg\psi \equiv \omega(\varphi')$. Thus, $\omega'' \models \omega(\varphi')$. Hence, $\omega'' \models \varphi'$, so:

$$\omega'' \models \Phi'.$$

Moreover, as $\omega''$ preserves the truth values of $\omega$ w.r.t. the symbols in $A$, and $\mathsf{At}(\alpha) \subseteq A$ and $\omega \models \alpha$, it follows that $\omega'' \models \alpha$. Hence, $\omega'' \models \alpha$ and $\omega'' \models \Phi'$ which contradicts the hypothesis that $\alpha$ is an interpolant of $\Phi$ w.r.t. $\Phi'$.

($\Leftarrow$) Let us suppose for contradiction purposes that $\omega$ is an interpretation such that the QBF $\forall x_1, \cdots, x_n\, \omega(\neg\varphi')$ is valid, but $\omega \not\models \alpha$, where $\{x_1, \cdots, x_n\} = \mathsf{At}(\Phi') \setminus A$. Thus, from Lemma 5, there is an interpretation $\omega' : \mathsf{At}(\Phi') \to \{\mathsf{true}, \mathsf{false}\}$ such that $\omega' \models \Phi'$ and $\omega(x) = \omega'(x)$ for all $x \in A$. Thus, $\omega' \models \varphi'$ which means $\omega' \not\models \neg\varphi'$. Therefore, the QBF $\forall x_1, \cdots, x_n \omega(\neg\varphi')$ is not valid. □

Corollary 4. *Let $\Phi$ and $\Phi'$ be jointly inconsistent sets of formulae, $A = \mathrm{At}(\Phi) \cap \mathrm{At}(\Phi')$, and $\alpha \in \mathrm{Strongest}(\Phi, \Phi')$, and let $\varphi = \bigwedge \Phi$. An interpretation $\omega : A \to \{\mathrm{true}, \mathrm{false}\}$ satisfies $\alpha$ iff the QBF $\exists x_1, \cdots, x_n \omega(\varphi)$ is valid, where $\{x_1, \cdots, x_n\} = \mathrm{At}(\Phi) \setminus A$.*

Proof. Let $\Phi$ and $\Phi'$ be jointly inconsistent sets of formulae, $A = \mathrm{At}(\Phi) \cap \mathrm{At}(\Phi')$, and $\alpha \in \mathrm{Strongest}(\Phi, \Phi')$. Let $\varphi = \bigwedge \Phi$. From Proposition 4, item (10), have that $\neg\alpha \in \mathrm{Weakest}(\Phi', \Phi)$. From Theorem 8, we get that an interpretation $\omega : A \to \{\mathrm{true}, \mathrm{false}\}$ satisfies $\neg\alpha$ iff the QBF $\forall x_1, \cdots, x_n \omega(\neg\varphi)$ is valid, where $\{x_1, \cdots, x_n\} = \mathrm{At}(\Phi) \setminus A$. Therefore, since $\omega(\neg\varphi)$ contains only variables in $\{x_1, \cdots, x_n\}$, it follows that an interpretation $\omega : A \to \{\mathrm{true}, \mathrm{false}\}$ satisfies $\alpha$ iff the QBF $\exists x_1, \cdots, x_n \omega(\varphi)$ is valid, where $\{x_1, \cdots, x_n\} = \mathrm{At}(\Phi) \setminus A$. □

Theorem 9. *#Weakest is # · coNP-complete.*

Proof. Let $\Phi$ and $\Phi'$ be jointly inconsistent sets of formulae. From Theorem 8, the number of models of the weakest interpolant of $\Phi$ w.r.t. $\Phi'$ corresponds to the number of truth assignments of the partial QBF $\forall x_1, \cdots, x_n \neg\varphi'$, where $\varphi' = \bigwedge \Phi'$. Therefore, the solution of #$\Pi_1$Sat on the input $\forall x_1, \cdots, x_n \neg\varphi'$ gives precisely the number of models of the weakest interpolant of $\Phi$ w.r.t. $\Phi'$, so #Weakest is in # · coNP.

For # · coNP-hardness, we reduce #$\Pi_1$Sat to #Weakest. Let $\forall x_1, \cdots, x_n\ \varphi$ be a partial QBF such that $\mathrm{At}(\varphi) = \{x_1, \cdots, x_n, y_1, \cdots, y_m\}$. Let $A = \{y_1, \ldots, y_m\}$, and $\Phi = \{\bigwedge_{y \in A} y_i \wedge \neg y_i \wedge (z \wedge \neg z)\}$, where $z \notin \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$. Clearly, $\Phi$ is inconsistent, and therefore inconsistent with $\Phi' := \{\neg\varphi\}$. Let $\alpha$ be a weakest interpolant of $\Phi$ w.r.t. $\Phi'$. So, $\mathrm{At}(\alpha) \subseteq A$. From Theorem 8, an interpretation $\omega : A \to \{\mathrm{true}, \mathrm{false}\}$ satisfies $\alpha$ iff the QBF $\forall x_1, \cdots, x_n \omega(\neg\neg\varphi)$ is valid. Equivalently, $\omega$ satisfies $\alpha$ iff the QBF $\forall x_1, \cdots, x_n\ \omega(\varphi)$ is valid. Therefore, the number of models of $\alpha$ corresponds to the solution of the #$\Pi_1$Sat problem on the input $\forall x_1, \cdots, x_n \varphi$. Thus, the solution of the problem #$\Pi_1$Sat on $\forall x_1, \cdots, x_n \varphi$ reduces to the solution of #Weakest for instance $(\Phi, \Phi')$. □

Theorem 10. *#Strongest is # · NP-complete.*

Proof. Let $\Phi$ and $\Phi'$ be jointly inconsistent sets of formulae. From Corollary 4, the number of models of the strongest interpolant of $\Phi$ w.r.t. $\Phi'$ corresponds to the number of truth assignments of the partial QBF $\exists x_1, \cdots, x_n \varphi$, where $\varphi = \bigwedge \Phi$. Therefore, the solution of #$\Sigma_1$Sat on the input $\exists x_1, \cdots, x_n \varphi$ corresponds to the number of models of the strongest interpolant of $\Phi$ w.r.t. $\Phi'$, so #Strongest is in # · NP.

For # · NP-hardness, we reduce #$\Sigma_1$Sat to #Strongest. Let $\exists x_1, \cdots, x_n \varphi$ be a partial QBF such that $\mathrm{At}(\varphi) = \{x_1, \cdots, x_n, y_1, \cdots, y_m\}$. Let $A = \{y_1, \ldots, y_m\}$, and $\Phi' = \{\bigwedge_{y \in A} y_i \wedge \neg y_i \wedge z \wedge \neg z\}$, where $z \notin \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$. Clearly, $\Phi'$ is inconsistent, and therefore inconsistent with $\Phi := \{\varphi\}$. Let $\alpha$ be a strongest interpolant of $\Phi$ w.r.t. $\Phi'$. So, $\mathrm{At}(\alpha) \subseteq A$. From Corollary 4, an interpretation $\omega : A \to \{\mathrm{true}, \mathrm{false}\}$ satisfies $\alpha$ iff the QBF $\exists x_1, \cdots, x_n \omega(\varphi)$ is valid. Therefore, the number of models of $\alpha$ corresponds to the solution of the #$\Sigma_1$Sat problem on the input $\exists x_1, \cdots, x_n \varphi$. Thus, the solution of the problem #$\Sigma_1$Sat on $\exists x_1, \cdots, x_n \varphi$ reduces to the solution of #Strongest for instance $(\Phi, \Phi')$. □

## 6.4 On the Computational Complexity of Disagreement Measures

We now turn to the complexity of our novel disagreement measures. In order to ease the analysis, we only consider the two-place functions $\mathcal{D}_{\mathbb{SI}}$ and $\mathcal{D}_{J_M}$. We therefore consider the following four problems, adapted from similar problems for inconsistency measures (Thimm and Wallner 2019). In order to avoid numerical issues, we assume that input numbers are rational and represented as two integers encoded in binary. This is not a restriction for our measures, since $\mathcal{D}_{\mathbb{SI}}$ only takes natural numbers and $\mathcal{D}_{J_M}$ only rational numbers (or $\infty$). Note, however, that the values of $\mathcal{D}_{\mathbb{SI}}$ may be doubly exponential in the size of the input, so that for the output of Value$_{\mathcal{D}_{\mathbb{SI}}}$ we use the succinct representation from Remark 4. This is not necessary for $\mathcal{D}_{J_M}$, as the numerators and denominators of its finite values are bounded by $2^{|\mathrm{At}(\Phi_1) \cap \mathrm{At}(\Phi_2)|}$ and can therefore be written in binary with polynomially many bits (see the proofs of Theorems 11 and 12). Let $\mathcal{D} : \mathfrak{K} \to \mathbb{R}^{\infty}_{\geq 0}$ be some function and $\mathbb{Q}^{\geq 0}$ the set of non-negative rational numbers.

| | | |
|---|---|---|
| Exact$_\mathcal{D}$ | **Input**: | $P \in \mathfrak{K}$, $x \in \mathbb{Q}^{\geq 0}$ |
| | **Output**: | true iff $\mathcal{D}(P) = x$ |
| Upper$_\mathcal{D}$ | **Input**: | $P \in \mathfrak{K}$, $x \in \mathbb{Q}^{\geq 0}$ |
| | **Output**: | true iff $\mathcal{D}(P) \leq x$ |
| Lower$_\mathcal{D}$ | **Input**: | $P \in \mathfrak{K}$, $x \in \mathbb{Q}^{\geq 0} \setminus \{0\}$ |
| | **Output**: | true iff $\mathcal{D}(P) \geq x$ |
| Value$_\mathcal{D}$ | **Input**: | $P \in \mathfrak{K}$ |
| | **Output**: | The value of $\mathcal{D}(P)$ |

The three first problems are decision problems, while the last one is a functional problem.

From our observations in the previous sections, we can easily obtain the following insights regarding the concrete measures $\mathcal{D}_{\mathbb{SI}}$ and $\mathcal{D}_{J_M}$ (therefore also only considering the above problems for functions of the form $\mathcal{D} : \mathbb{K} \times \mathbb{K} \to \mathbb{R}^\infty_{\geq 0}$ and profiles $P = \langle \Phi_1, \Phi_2 \rangle$).

Theorem 11. *The computational complexities of Exact$_\mathcal{D}$, Upper$_\mathcal{D}$, Lower$_\mathcal{D}$, and Value$_\mathcal{D}$ for $\mathcal{D} \in \{\mathcal{D}_{\mathbb{SI}}, \mathcal{D}_{J_M}\}$ are as shown in Table 2.*

Proof. Let $P = \langle \Phi_1, \Phi_2 \rangle$ with sets of formulae $\Phi_1 \in \mathcal{L}(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $\Phi_2 \in \mathcal{L}(y_1, \ldots, y_m, z_1, \ldots, z_k)$. We also assume $m > 0$ to avoid trivial cases. Define furthermore $\varphi_1 = \bigwedge \Phi_1$ and $\varphi_2 = \bigwedge \Phi_2$.

- Lower$_{\mathcal{D}_{\mathbb{SI}}}$ is $\Sigma^p_2$-complete: Let $P = \langle \Phi_1, \Phi_2 \rangle$ and $x \in \mathbb{Q}^{\geq 0}$ be the input. Note that $(P, x)$ is a positive instance of Lower$_{\mathcal{D}_{\mathbb{SI}}}$ if $|\mathbb{SI}(\Phi_1, \Phi_2)| \geq x$ which is equivalent to $\log_2 |\mathbb{SI}(\Phi_1, \Phi_2)| \geq \log_2 x$. Let $x' = \lceil \log_2 x \rceil$. Let $\omega : \{y_1, \ldots, y_m\} \to \{\text{true}, \text{false}\}$ be an interpretation. Checking that $\omega$ is a model of $\forall x_1 \ldots x_n \neg\varphi_1 \wedge \forall z_1 \ldots z_k \neg\varphi_2$ is a coNP-complete problem (see also the proof of Theorem 7). By Theorem 6, the number of such models is exactly $\log_2 |\mathbb{SI}(\Phi_1, \Phi_2)|$. So in order to check whether $(P, x)$ is a positive instance, we first check that $\Phi_1 \cup \Phi_2$ is inconsistent (otherwise $\mathbb{SI}(\Phi_1, \Phi_2) = \emptyset$ and, as $x > 0$, $(P, x)$ is a negative instance) using a coNP-oracle, and then guess $x'$ different interpretations $\omega$ and check for each, whether it is a model of $\forall x_1 \ldots x_n \neg\varphi_1 \wedge \forall z_1 \ldots z_k \neg\varphi_2$ (using a coNP-oracle). This shows that Lower$_{\mathcal{D}_{\mathbb{SI}}}$ is in $\mathsf{NP}^{\mathsf{coNP}} = \mathsf{NP}^{\mathsf{NP}} = \Sigma^p_2$. For showing $\Sigma^p_2$-hardness, we consider the canonical $\Sigma^p_2$-complete problem $\exists\forall$-TQBF defined via

  | | | |
  |---|---|---|
  | $\exists\forall$-TQBF | **Input**: | A QBF of the form $\alpha = \exists y_1, \ldots, y_m \forall x_1, \ldots, x_n\, \varphi$ |
  | | **Output**: | True iff $\alpha$ is equivalent to $\top$. |

  We use a similar reduction as in Theorem 7 to map an instance from $\exists\forall$-TQBF to an instance from Lower$_{\mathcal{D}_{\mathbb{SI}}}$. For $\alpha = \exists y_1, \ldots, y_m \forall x_1, \ldots, x_n\, \varphi$ define

  $$\Phi_\alpha = \{\neg\varphi\}$$
  $$\Phi'_\alpha = \{y_1, \ldots, y_m, \neg y_1, \ldots, \neg y_m\}$$

  By Theorem 6 we know that the number of equivalence classes of interpolants of $\Phi_\alpha$ w.r.t. $\Phi'_\alpha$ is $2^q$ where $q$ is the number of models (w.r.t. $\{y_1, \ldots, y_m\}$) of

  $$\begin{aligned} & \forall x_1, \ldots, x_n \neg\neg\varphi \wedge \neg(y_1 \wedge \ldots \wedge y_m \wedge \neg y_1 \wedge \ldots \wedge \neg y_m) \\ \equiv\ & \forall x_1, \ldots, x_n\, \varphi \wedge \top \\ \equiv\ & \forall x_1, \ldots, x_n\, \varphi \end{aligned}$$

  So if $q \geq 1$ (or equivalently $2^q \geq 2^1 = 2$) then $\alpha$ is equivalent to $\top$. It follows that $\alpha$ is a positive instance of $\exists\forall$-TQBF if and only if $(\langle \Phi_\alpha, \Phi'_\alpha \rangle, 2)$ is a positive instance of Lower$_{\mathcal{D}_{\mathbb{SI}}}$.
- Upper$_{\mathcal{D}_{\mathbb{SI}}}$ is $\Pi^p_2$-complete: Since Lower$_{\mathcal{D}_{\mathbb{SI}}}$ is $\Sigma^p_2$-complete (see above), Upper$_{\mathcal{D}_{\mathbb{SI}}}$ is $\Pi^p_2$-complete due to Lemma 6 by Thimm and Wallner (2019).

Table 2. Computational complexities of our disagreement measures; *this pertains to the problem variant where $\log_2 |\mathbb{SI}(\Phi, \Phi')|$ instead of $|\mathbb{SI}(\Phi, \Phi')|$ is determined, cf. Remark 4

| $\mathcal{D}$ | $\textsc{Exact}_{\mathcal{D}}$ | $\textsc{Upper}_{\mathcal{D}}$ | $\textsc{Lower}_{\mathcal{D}}$ | $\textsc{Value}_{\mathcal{D}}$ |
|---|---|---|---|---|
| $\mathcal{D}_{\mathbb{SI}}$ | in $D_2^p$ | $\Pi_2^p$-c | $\Sigma_2^p$-c | $\# \cdot$ coNP-c* |
| $\mathcal{D}_{J_M}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{FP}^{\#\cdot\mathsf{coNP}}$ |

- $\textsc{Exact}_{\mathcal{D}_{\mathbb{SI}}}$ is in $D_2^p$: Since $\textsc{Lower}_{\mathcal{D}_{\mathbb{SI}}}$ is $\Sigma_2^p$-complete (see above), $\textsc{Exact}_{\mathcal{D}_{\mathbb{SI}}}$ is in $D_2^p$ due to Lemma 6 by Thimm and Wallner (2019).
- $\textsc{Value}_{\mathcal{D}_{\mathbb{SI}}}$ is $\# \cdot$ coNP-complete with respect to the succinct representation introduced in Remark 4: under this representation the problem is identical to #Interpolant', which is $\# \cdot$ coNP-complete by Theorem 7.
- $\textsc{Value}_{\mathcal{D}_{J_M}}$ is in $\mathsf{FP}^{\#\cdot\mathsf{coNP}}$: On input $P = \langle \Phi_1, \Phi_2 \rangle$ such that $\Phi_1 \cup \Phi_2$ is inconsistent (which is checked below, as #Weakest is only defined for such inputs), we use one $\# \cdot$ coNP-oracle call to solve #Weakest for $\Phi_1, \Phi_2$ (see Theorem 9), obtaining a value $x$. With another call to the $\# \cdot$ coNP-oracle we solve #Weakest for $\Phi_2, \Phi_1$, obtaining a value $z$. From Proposition 4, items 9/10, we know that for any $\psi, \psi \in \mathsf{Strongest}(\Phi_1, \Phi_2)$ if and only if $\neg\psi \in \mathsf{Weakest}(\Phi_2, \Phi_1)$. Let $A = \mathsf{At}(\Phi_1) \cap \mathsf{At}(\Phi_2)$, then $|\mathsf{Mod}_A(\psi)| = 2^{|A|} - |\mathsf{Mod}_A(\neg\psi)|$, so $y = 2^{|A|} - z$ is the solution of #Strongest for $\Phi_1, \Phi_2$.
  Now, recall that $\mathcal{D}_{J_M}(\Phi_1, \Phi_2) = 0$ if $\Phi_1 \cup \Phi_2$ is consistent. Checking whether $\Phi_1 \cup \Phi_2$ is consistent can be accomplished by one call to the $\# \cdot$ coNP-oracle (more precisely, due to $\#\mathsf{P} \subseteq \# \cdot \mathsf{coNP}$, we can count the models of $\Phi_1 \cup \Phi_2$ with one call and then check whether that number is larger than zero). If this is not the case, recall that $\mathcal{D}_{J_M}(\Phi_1, \Phi_2) = \infty$ if $J_M(\mathsf{Weakest}(\Phi_1, \Phi_2)) = 0$ which is the case if $\top \models \mathsf{Weakest}(\Phi_1, \Phi_2)$, which itself is equivalent to $x = 2^{|A|}$. If $y = 0$, then $\mathsf{Strongest}(\Phi_1, \Phi_2) = [\bot]$ (i. e., $\Phi_1$ is inconsistent) and $J_M(\mathsf{Strongest}(\Phi_1, \Phi_2)) = \infty$, so that $\mathcal{D}_{J_M}(\Phi_1, \Phi_2) = \infty$ if $x > 0$, and $\mathcal{D}_{J_M}(\Phi_1, \Phi_2) = \infty/\infty = 1$ if $x = 0$. In the remaining case we have
  $$\mathcal{D}_{J_M}(\Phi_1, \Phi_2) = \frac{J_M(\mathsf{Strongest}(\Phi_1, \Phi_2))}{J_M(\mathsf{Weakest}(\Phi_1, \Phi_2))} = \frac{|\mathsf{Mod}(\mathsf{Weakest}(\Phi_1, \Phi_2))|}{|\mathsf{Mod}(\mathsf{Strongest}(\Phi_1, \Phi_2))|} = x/y$$
  where the model counts can be taken w.r.t. $A$, as the ratio does not depend on the signature (see Section 4). In this case, $0 < y \leq x < 2^{|A|}$, so both $x$ and $y$ can be written in binary with at most $|A| + 1$ bits. This shows that $\textsc{Value}_{\mathcal{D}_{J_M}}$ is in $\mathsf{FP}^{\#\cdot\mathsf{coNP}}$.
- $\textsc{Lower}_{\mathcal{D}_{J_M}}$, $\textsc{Upper}_{\mathcal{D}_{J_M}}$, and $\textsc{Exact}_{\mathcal{D}_{J_M}}$ are in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$: This follows from the above proof for $\textsc{Value}_{\mathcal{D}_{J_M}}$. After obtaining the value, we just compare it to the input value and return the appropriate answer. □

With the above insights on the complexity of problems related to $\mathcal{D}_{\mathbb{SI}}$ and $\mathcal{D}_{J_M}$ we can also give some upper bounds on the corresponding problems of the induced measures $\mathcal{D}^{\Sigma}$ and $\mathcal{D}^{\max}$ wrt. $\mathcal{D}_{\mathbb{SI}}$ and $\mathcal{D}_{J_M}$. For comparison, we also give the corresponding results for Potyka's measure $D_\eta$ (see Definition 8), which have been missing so far.[8]

Theorem 12. *The computational complexities of $\textsc{Exact}_{\mathcal{D}}$, $\textsc{Upper}_{\mathcal{D}}$, $\textsc{Lower}_{\mathcal{D}}$, and $\textsc{Value}_{\mathcal{D}}$ for $\mathcal{D} \in \{\mathcal{D}^{\Sigma}_{\mathbb{SI}}, \mathcal{D}^{\max}_{\mathbb{SI}}, \mathcal{D}^{\Sigma}_{J_M}, \mathcal{D}^{\max}_{J_M}\}$ are as shown in Table 3.*

Proof. Let $P = \langle \Phi_1, \ldots, \Phi_n \rangle$.

- $\textsc{Exact}_{\mathcal{D}^{\Sigma}_{\mathbb{SI}}}$ is in $\Sigma_3^p$: On input $P$ and $x \in \mathbb{Q}^{\geq 0}$, we guess $n^2 + 1$ numbers $x_{i,j} \in \mathbb{Q}^{\geq 0}$ (for $i, j = 1, \ldots, n$) and $x_0 \in \mathbb{Q}^{\geq 0}$ with $x = \sum_{i,j=1\ldots,n} x_{i,j} + x_0$. Using a $D_2^p$-oracle (since $\textsc{Exact}_{\mathcal{D}_{\mathbb{SI}}}$ is in $D_2^p$), we verify that

[8] Potyka (2018) only gives a straightforward upper bound by stating that computing $D_\eta$ is fixed-parameter tractable when fixing the size of the signature.

$\mathcal{D}_{\mathbb{SI}}(\Phi_i, \Phi_j) = x_{i,j}$ for $i, j = 1, \ldots, n$ and that $\mathcal{D}_{\mathbb{SI}}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) = x_0$. If this is the case, it follows $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(P) = x$. Observe that this requires a polynomial number of calls to the $D_2^p$-oracle, showing $\mathsf{NP}^{D_2^p} = \Sigma_3^p$ membership.

- $\textsc{Upper}_{\mathcal{D}^{\Sigma}_{\mathbb{SI}}}$ is in $\Sigma_3^p$: Analogous to above. On input $P$ and $x \in \mathbb{Q}^{\geq 0}$, we guess $n^2 + 1$ numbers $x_{i,j} \in \mathbb{Q}^{\geq 0}$ (for $i, j = 1, \ldots, n$) and $x_0 \in \mathbb{Q}^{\geq 0}$ with $x = \sum_{i,j=1\ldots,n} x_{i,j} + x_0$. Using a $\Pi_2^p$-oracle (since $\textsc{Upper}_{\mathcal{D}_{\mathbb{SI}}}$ is in $\Pi_2^p$), we verify that $\mathcal{D}_{\mathbb{SI}}(\Phi_i, \Phi_j) \leq x_{i,j}$ for $i, j = 1, \ldots, n$ and that $\mathcal{D}_{\mathbb{SI}}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) \leq x_0$. If this is the case, it follows $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(P) \leq x$. Observe that this requires a polynomial number of calls to the $\Pi_2^p$-oracle, showing $\mathsf{NP}^{\Pi_2^p} = \Sigma_3^p$ membership.
- $\textsc{Lower}_{\mathcal{D}^{\Sigma}_{\mathbb{SI}}}$ is in $\Sigma_3^p$: Analogous to the case for $\textsc{Upper}_{\mathcal{D}^{\Sigma}_{\mathbb{SI}}}$.
- $\textsc{Value}_{\mathcal{D}^{\Sigma}_{\mathbb{SI}}}$ is in $\mathsf{FP}^{\#\cdot\mathsf{coNP}^*}$: For each $i, j = 1, \ldots, n$, we check with one oracle call whether $\Phi_i \cup \Phi_j$ is consistent (in which case $\mathcal{D}_{\mathbb{SI}}(\Phi_i, \Phi_j) = 0$) and otherwise compute $q_{i,j} = \log_2 \mathcal{D}_{\mathbb{SI}}(\Phi_i, \Phi_j)$ with one call to a $\# \cdot \mathsf{coNP}$-oracle for #Interpolant'. The remaining summand is $\mathcal{D}_{\mathbb{SI}}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) = |\{[\bot]\}| = 2^0$ if $\bigcup_{i=1}^{n} \Phi_i$ is inconsistent, and 0 otherwise. As discussed in Remark 4, writing down $2^{q_{i,j}}$ in binary may require exponentially many bits. We therefore represent the output by the list of positions of the non-zero bits in the binary representation of $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(P)$, which generalises the representation from Remark 4. Since $\mathcal{D}^{\Sigma}_{\mathbb{SI}}(P)$ is a sum of at most $n^2 + 1$ powers of two whose exponents can be written with polynomially many bits, this list has at most $n^2 + 1$ entries of polynomially many bits each. It can be computed from the exponents in polynomial time by repeatedly replacing two equal summands $2^q + 2^q$ by $2^{q+1}$.
- $\textsc{Exact}_{\mathcal{D}^{\max}_{\mathbb{SI}}}$ is in $\Delta_3^p$: Using a $D_2^p$-oracle (since $\textsc{Upper}_{\mathcal{D}_{\mathbb{SI}}}$ is in $\Pi_2^p \subseteq D_2^p$), we verify that $\mathcal{D}_{\mathbb{SI}}(\Phi_i, \Phi_j) \leq x$ for $i, j = 1, \ldots, n$ and that $\mathcal{D}_{\mathbb{SI}}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) \leq x$. Then, again using a $D_2^p$-oracle (since $\textsc{Exact}_{\mathcal{D}_{\mathbb{SI}}}$ is in $D_2^p$), we verify that at least one of the following equalities hold: $\mathcal{D}_{\mathbb{SI}}(\Phi_i, \Phi_j) = x$ for some $i, j = 1, \ldots, n$ or $\mathcal{D}_{\mathbb{SI}}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) = x$. If this is the case, it follows $\mathcal{D}^{\max}_{\mathbb{SI}}(P) = x$. Observe that this requires a polynomial number of calls to the $D_2^p$-oracle, showing $\mathsf{P}^{D_2^p} = \Delta_3^p$ membership.
- $\textsc{Upper}_{\mathcal{D}^{\max}_{\mathbb{SI}}}$ is in $\Delta_3^p$: Analogous to above. Using a $\Pi_2^p$-oracle (since $\textsc{Upper}_{\mathcal{D}_{\mathbb{SI}}}$ is in $\Pi_2^p$), we verify that $\mathcal{D}_{\mathbb{SI}}(\Phi_i, \Phi_j) \leq x$ for $i, j = 1, \ldots, n$ and that $\mathcal{D}_{\mathbb{SI}}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) \leq x$. If this is the case, it follows $\mathcal{D}^{\max}_{\mathbb{SI}}(P) \leq x$. Observe that this requires a polynomial number of calls to the $\Pi_2^p$-oracle, showing $\mathsf{P}^{\Pi_2^p} = \mathsf{P}^{\Sigma_2^p} = \Delta_3^p$ membership
- $\textsc{Lower}_{\mathcal{D}^{\max}_{\mathbb{SI}}}$ is in $\Delta_3^p$: Analogous to the case for $\textsc{Upper}_{\mathcal{D}^{\max}_{\mathbb{SI}}}$.
- $\textsc{Value}_{\mathcal{D}^{\max}_{\mathbb{SI}}}$ is in $\mathsf{FP}^{\#\cdot\mathsf{coNP}^*}$: We determine the exponents $q_{i,j}$ and the summands that are 0 as in the previous case. Unless all values are 0, we have $\mathcal{D}^{\max}_{\mathbb{SI}}(P) = 2^q$ for the maximal such exponent $q$. We output $q$, i. e., we use the succinct representation from Remark 4, together with a flag for the case $\mathcal{D}^{\max}_{\mathbb{SI}}(P) = 0$.
- $\textsc{Value}_{\mathcal{D}^{\Sigma}_{J_M}}$ is in $\mathsf{FP}^{\#\cdot\mathsf{coNP}}$: For each $i, j = 1, \ldots, n$, we compute $\mathcal{D}_{J_M}(\Phi_i, \Phi_j)$ as in the proof of Theorem 11, i. e., we check with one oracle call whether $\Phi_i \cup \Phi_j$ is consistent and, if not, use two calls to a $\# \cdot \mathsf{coNP}$-oracle for #Weakest on $(\Phi_i, \Phi_j)$ and $(\Phi_j, \Phi_i)$. This yields either 0, $\infty$, or a rational number $x_{i,j}/y_{i,j}$ with $0 < y_{i,j} \leq x_{i,j} < 2^{|A_{i,j}|}$, where $A_{i,j} = \mathsf{At}(\Phi_i) \cap \mathsf{At}(\Phi_j)$. The remaining summand requires no counting. Since $\mathsf{At}(\emptyset) = \emptyset$, we have $\mathbb{SI}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) = \{[\bot]\}$ if $\bigcup_{i=1}^{n} \Phi_i$ is inconsistent, so that $\mathsf{Weakest}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) = \mathsf{Strongest}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) = [\bot]$ and $\mathcal{D}_{J_M}(\bigcup_{i=1}^{n} \Phi_i, \emptyset) = \infty/\infty = 1$. Otherwise, this summand is 0. This requires one further oracle call. If one of the summands is $\infty$, then $\mathcal{D}^{\Sigma}_{J_M}(P) = \infty$. Otherwise, $\mathcal{D}^{\Sigma}_{J_M}(P)$ is the sum of at most $n^2 + 1$ rational numbers whose numerators and denominators are bounded by $2^{|\mathsf{At}(\bigcup P)|}$. The denominator of this sum is bounded by the product of these denominators and the sum itself is bounded by $(n^2 + 1) \cdot 2^{|\mathsf{At}(\bigcup P)|}$. Hence, the numerator and the denominator of the sum can be written in binary with polynomially many bits and computed in polynomial time. Overall, $O(n^2)$ oracle calls are needed. In contrast to $\mathcal{D}_{\mathbb{SI}}$ (cf. Remark 4), no succinct representation of the output is needed, i. e., the output is given by two integers in binary, or by a special symbol for $\infty$.

Table 3. Computational complexities of the induced disagreement measures $\mathcal{D}^{\Sigma}$ and $\mathcal{D}^{\max}$ wrt. $\mathcal{D}_{\mathbb{SI}}$ and $\mathcal{D}_{J_M}$, and of $D_\eta$. *w.r.t. the succinct representations of the output described in the proof of Theorem 12, cf. Remark 4.

| $\mathcal{D}$ | $\textsc{Exact}_{\mathcal{D}}$ | $\textsc{Upper}_{\mathcal{D}}$ | $\textsc{Lower}_{\mathcal{D}}$ | $\textsc{Value}_{\mathcal{D}}$ |
|---|---|---|---|---|
| $\mathcal{D}^{\Sigma}_{\mathbb{SI}}$ | in $\Sigma_3^p$ | in $\Sigma_3^p$ | in $\Sigma_3^p$ | in $\mathsf{FP}^{\#\cdot\mathsf{coNP}*}$ |
| $\mathcal{D}^{\max}_{\mathbb{SI}}$ | in $\Delta_3^p$ | in $\Delta_3^p$ | in $\Delta_3^p$ | in $\mathsf{FP}^{\#\cdot\mathsf{coNP}*}$ |
| $\mathcal{D}^{\Sigma}_{J_M}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{FP}^{\#\cdot\mathsf{coNP}}$ |
| $\mathcal{D}^{\max}_{J_M}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$ | in $\mathsf{FP}^{\#\cdot\mathsf{coNP}}$ |
| $D_\eta$ | in $D_1^p$ | NP-c | coNP-c | in $\mathsf{FP}^{\mathsf{NP}}$ |

- $\textsc{Exact}_{\mathcal{D}^{\Sigma}_{J_M}}$, $\textsc{Upper}_{\mathcal{D}^{\Sigma}_{J_M}}$, and $\textsc{Lower}_{\mathcal{D}^{\Sigma}_{J_M}}$ are in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$: We compute $\mathcal{D}^{\Sigma}_{J_M}(P)$ as above and compare it with the input value $x$. If $\mathcal{D}^{\Sigma}_{J_M}(P) = \infty$, the answer is TRUE for $\textsc{Lower}_{\mathcal{D}^{\Sigma}_{J_M}}$ and FALSE for the other two problems. Otherwise, two rational numbers given by integers in binary are compared by cross-multiplication in polynomial time.
- $\textsc{Value}_{\mathcal{D}^{\max}_{J_M}}$ is in $\mathsf{FP}^{\#\cdot\mathsf{coNP}}$: We compute the at most $n^2 + 1$ values $\mathcal{D}_{J_M}(\Phi_i, \Phi_j)$, for $i, j = 1, \ldots, n$, and $\mathcal{D}_{J_M}(\bigcup_{i=1}^{n} \Phi_i, \emptyset)$ as in the previous case and determine their maximum by pairwise comparison, which can be done in polynomial time. The output is represented as in the previous case.
- $\textsc{Exact}_{\mathcal{D}^{\max}_{J_M}}$, $\textsc{Upper}_{\mathcal{D}^{\max}_{J_M}}$, and $\textsc{Lower}_{\mathcal{D}^{\max}_{J_M}}$ are in $\mathsf{P}^{\#\cdot\mathsf{coNP}}$: Analogous to the case for $\mathcal{D}^{\Sigma}_{J_M}$.
- $\textsc{Upper}_{D_\eta}$ is NP-complete: Proposition 9 of Potyka (2018) states that $D_\eta$ can be computed by solving a linear optimisation problem (over an exponential number of variables). Analogous to the proof of determining the upper bound of the *inconsistency measure* $I_\eta$ by Thimm and Wallner (2019) (Proposition 5), we note that this linear optimisation problem has an optimal solution, where only a polynomial number of variables get a non-zero variable (Chvátal 1983). We can therefore guess a set of polynomial many variables, set the objective function to the given upper bound $x$, and solve the corresponding program using a polynomial-time algorithm (as linear programming is in P). If it is feasible, $x$ is indeed an upper bound.
  NP-completeness follows from a straightforward reduction from SAT: note that a formula $\alpha$ is satisfiable if and only if $D_\eta(\langle\{\alpha\}\rangle) = 0$.
- $\textsc{Lower}_{D_\eta}$ is coNP-complete: Since $\textsc{Upper}_{D_\eta}$ is NP-complete (see above), $\textsc{Lower}_{D_\eta}$ is coNP-complete due to Lemma 6 by Thimm and Wallner (2019).
- $\textsc{Exact}_{D_\eta}$ is in $D_1^p$: Since $\textsc{Upper}_{D_\eta}$ is NP-complete (see above), $\textsc{Exact}_{D_\eta}$ is in $D_1^p$ due to Lemma 6 by Thimm and Wallner (2019).
- $\textsc{Value}_{D_\eta}$ is in $\mathsf{FP}^{\mathsf{NP}}$: Since $\textsc{Upper}_{D_\eta}$ is NP-complete (see above), $\textsc{Value}_{D_\eta}$ is in $\mathsf{FP}^{\mathsf{NP}}$ due to Lemma 4 by Thimm and Wallner (2019), analogously to the case for $I_\eta$. □

Since most of computational problems in our framework are intractable, it is important to highlight potential parameters under which the proposed measures may become tractable and thus more practically relevant. For instance, restricting the input to certain propositional classes, such as Horn formulas, Krom formulas, or formulas of bounded treewidth, can lead to significant reductions in complexity, as many reasoning tasks become tractable in these settings (Thimm and Wallner 2019). Additionally, if the shared vocabulary between knowledge bases is bounded, the number of possible interpolants is exponentially reduced, which can make the computation of disagreement measures feasible in practice. Exploring such tractable fragments and parameterized approaches is an important direction for future work, as it would enhance the practical applicability of our disagreement measures in real-world knowledge representation scenarios.

## 7 Summary and Discussion

We investigated the problem of measuring disagreement in belief merging scenarios. For that, we made use of the concept of Craig interpolants and defined disagreement measures that consider the number of semantically equivalent interpolants between two knowledge bases and the information content in the strongest and weakest interpolants. We showed that our measures satisfy a number of desirable properties. We then focused on an analysis of the computational complexity of various problems in this context. We characterised several problems in that manner, in particular related to weakest and strongest interpolants. We specifically considered counting problems as well and applied these insights in analysing the computational complexity of our new measures.

While our framework is restricted to propositional logic, this choice is motivated by the use of propositional logic in other areas of inconsistency measurement (Thimm and Wallner 2019) and the well-understood existence and computational properties of Craig interpolants in this setting, which enable the development of concrete disagreement measures with precise complexity characterisations. Although central results on Craig interpolation also hold in first-order logic, extending our disagreement measures and the associated complexity results to full first-order logic is not straightforward. In general, the set of interpolants is not recursively enumerable, and key decision problems such as interpolant existence and equivalence become undecidable. As such, the complexity bounds we establish for the propositional case do not directly transfer to the full first-order setting. Nevertheless, meaningful extensions to restricted fragments of first-order logic which are decidable and have the interpolation property—such as the monadic fragment, finite variable fragments of the guarded fragment that have the Craig interpolation property (Hoogland and Marx 2002), fragments with bounded quantifier alternation, or other decidable fragments with the Craig interpolation property such as the unary negation fragment (Cate and Segoufin 2013) and the guarded negation fragment (Cate and Comer 2025)—are conceivable. An extension of the results we established for the propositional case to logical theories in which ground interpolants for sets of ground formulae exist, and which allow uniform interpolation for ground formulae seems conceivable under additional conditions (for instance if the theories are universal and have a purely relational signature or, more generally, are locally finite[9] (Calvanese et al. 2021; Sofronie-Stokkermans 2025)). Investigating disagreement measures and their computational properties in such fragments represents a promising avenue for future research, potentially bridging the gap between propositional and more expressive logical frameworks.

Interpolants provide a suitable tool for analysing inconsistency and disagreement because they isolate exactly those parts of two knowledge bases that are jointly responsible for a conflict while abstracting away all symbols on which the bases do not interact, thereby focusing the analysis on the genuine points of disagreement rather than on their internal structure. In contrast to syntax-sensitive constructions, this locality and signature-based restriction makes the resulting disagreement measures invariant under harmless reformulations such as $\{\alpha \wedge \neg\alpha\}$ and $\{\alpha, \neg\alpha\}$ that intuitively encode the same conflict. Moreover, by ranging over all interpolants (up to semantic equivalence) and by exploiting the complete lattice structure of this set, our measures can take into account both the number of distinct explanations for conflicts and their information content. As illustrated in Section 5, this allows the sum-based variants of our measures to distinguish some profiles that $D_\eta$ treats identically. A general formal characterisation of this behaviour is left for future work.

For future work, we will investigate the possibility of defining further measures based on interpolation and investigate their properties. Moreover, a deeper analysis of the differences of our measures with the measures proposed by Potyka (2018) is needed. In particular, the comparison with $D_\eta$ in Section 5 is based on examples, and a formal characterisation of the differences, e. g., by means of rationality postulates, remains open. Precisely, our approach based on interpolants has a semantic perspective in assessing the culpability degree of the inconsistencies between two knowledge bases. We shall investigate what properties this semantic perspective brings upon these

[9] A theory $\mathcal{T}$ is locally finite iff for every finite tuple of variables $\overline{x}$ there are only finitely many equivalence classes wrt $\mathcal{T}$ of atoms $A(\overline{x})$ involving only the variables $\overline{x}$ (cf. (Calvanese et al. 2021)).

inconsistency measures. We will also explore algorithmic approaches to compute our measures and investigate applying our ideas to the area of inconsistency measurement (Thimm 2019).

## Acknowledgments

The research reported here was partially supported by the Deutsche Forschungsgemeinschaft (project numbers 424710479 and 465447331).

## References

K. Arrow. 1963. *Social choice and individual values*. Wiley, New York.

C. Baral, J. Minker, and S. Kraus. 1991. "Combining Multiple Knowledge Bases." *IEEE Transactions on Knowledge and Data Engineering*, 3, 2, 208–221.

P. Besnard and J. Grant. 2020. "Relative inconsistency measures." *Artificial Intelligence*, 280, 103231.

R. E. Bryant. 1986. "Graph-Based Algorithms for Boolean Function Manipulation." *IEEE Trans. Computers*, 35, 8, 677–691. doi:10.1109/TC.1986.1676819.

R. E. Bryant. 1992. "Symbolic Boolean Manipulation with Ordered Binary-Decision Diagrams." *ACM Comput. Surv.*, 24, 3, 293–318. doi:10.1145/136035.136043.

D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. 2021. "Model Completeness, Uniform Interpolants and Superposition Calculus." *J. Autom. Reason.*, 65, 7, 941–969. doi:10.1007/S10817-021-09596-X.

B. ten Cate and J. Comer. 2025. "Craig Interpolation for Decidable First-Order Fragments." *Log. Methods Comput. Sci.*, 21, 3. doi:10.46298/LMCS-21(3:22)2025.

B. ten Cate and L. Segoufin. 2013. "Unary negation." *Log. Methods Comput. Sci.*, 9, 3. doi:10.2168/LMCS-9(3:25)2013.

L. Cholvy and A. Hunter. 1997. "Information fusion in logic: A brief overview." In: *Qualitative and Quantitative Practical Reasoning (ECSQARU'97/FAPR'97)* (Lecture Notes in Computer Science). Vol. 1244. Springer, 86–95.

V. Chvátal. 1983. *Linear Programming*. Freeman, San Francisco.

C. Corea and P. Delfmann. 2018. "Supporting Business Rule Management with Inconsistency Analysis." In: *Proceedings of the Dissertation Award, Demonstration, and Industrial Track at BPM 2018 co-located with 16th International Conference on Business Process Management (BPM 2018), Sydney, Australia, September 9-14, 2018*, 141–147.

C. Corea and M. Thimm. 2020. "Towards Inconsistency Measurement in Business Rule Bases." In: *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI'20)*.

W. Craig. 1957. "Linear Reasoning. A New Form of the Herbrand-Gentzen Theorem." *J. Symb. Log.*, 22, 3, 250–268.

A. M. Díaz and R. P. Pérez. 2017. "Impossibility in Belief Merging." *Artificial Intelligence*, 251, 1–34.

A. Dolzmann and T. Sturm. 1997. "REDLOG: computer algebra meets computer logic." *SIGSAM Bull.*, 31, 2, 2–9. doi:10.1145/261320.261324.

A. Durand, M. Hermann, and P. G. Kolaitis. 2005. "Subtractive reductions and complete problems for counting complexity classes." *Theor. Comput. Sci.*, 340, 3, 496–513. doi:10.1016/J.TCS.2005.03.012.

P. Everaere, S. Konieczny, and P. Marquis. 2015. "Belief Merging versus Judgment Aggregation." In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*.

S. Gao, A. Platzer, and E. M. Clarke. 2011. "Quantifier Elimination over Finite Fields Using Gröbner Bases." In: *Algebraic Informatics - 4th International Conference, CAI 2011, Linz, Austria, June 21-24, 2011. Proceedings* (Lecture Notes in Computer Science). Ed. by F. Winkler. Vol. 6742. Springer, 140–157. doi:10.1007/978-3-642-21493-6_9.

J. Grant. 1978. "Classifications for Inconsistent Theories." *Notre Dame Journal of Formal Logic*, 19, 3, 435–444.

J. Grant and A. Hunter. 2011. "Measuring Consistency Gain and Information Loss in Stepwise Inconsistency Resolution." In: *Proceedings of the 11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2011)* (Lecture Notes in Artificial Intelligence) 6717. Ed. by W. Liu. Springer-Verlag, 362–373.

J. Grant and A. Hunter. 2013. "Measuring the Good and the Bad in Inconsistent Information." In: *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*.

L. A. Hemaspaandra and H. Vollmer. 1995. "The satanic notations: counting classes beyond #P and other definitional adventures." *SIGACT News*, 26, 1, 2–13.

E. Hoogland and M. Marx. 2002. "Interpolation and Definability in Guarded Fragments." *Stud Logica*, 70, 3, 373–409. doi:10.1023/A:1015154431342.

A. Hunter and S. Konieczny. 2008. "Measuring Inconsistency through Minimal Inconsistent Sets." In: *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'2008)*. Ed. by G. Brewka and J. Lang. AAAI Press, Menlo Park, California, Sydney, Australia, 358–366.

K. Knight. 2002. "Measuring Inconsistency." *J. Philosophical Logic*, 31, 1, 77–98.

S. Konieczny and R. P. Pérez. 1998. "On the logic of merging." In: *Proceedings of the Sixth International Conference on Principles of Knowledge Representationa and Reasoning (KR'98)*. Morgan Kaufmann.

P. Koopmann, C. Wernhard, and F. Wolter. 2025. "Interpolation in Classical Propositional Logic." *CoRR*, abs/2508.11449. arXiv: 2508.11449. doi:10.48550/ARXIV.2508.11449.

L. Kovács and A. Voronkov. 2009. "Interpolation and Symbol Elimination." In: *Automated Deduction - CADE-22, 22nd International Conference on Automated Deduction, Montreal, Canada, August 2-7, 2009. Proceedings* (Lecture Notes in Computer Science). Ed. by R. A. Schmidt. Vol. 5663. Springer, 199–213. doi:10.1007/978-3-642-02959-2_17.

K. Mu, W. Liu, and Z. Jin. 2011. "A general framework for measuring inconsistency through minimal inconsistent sets." *Knowledge and Information Systems*, 27, 85–114.

C. H. Papadimitriou. 1994. *Computational complexity*. Addison-Wesley.

D. Peuter and V. Sofronie-Stokkermans. 2021. "Symbol Elimination and Applications to Parametric Entailment Problems." In: *Frontiers of Combining Systems - 13th International Symposium, FroCoS 2021, Birmingham, UK, September 8-10, 2021, Proceedings* (Lecture Notes in Computer Science). Ed. by B. Konev and G. Reger. Vol. 12941. Springer, 43–62. doi:10.1007/978-3-030-86205-3_3.

N. Potyka. 2018. "Measuring Disagreement among Knowledge Bases." In: *Proceedings of the 12th International Conference on Scalable Uncertainty Management (SUM'18)*.

J. S. Ribeiro, V. Sofronie-Stokkermans, and M. Thimm. 2020. "Measuring Disagreement with Interpolants." In: *Proceedings of the 14th International Conference on Scalable Uncertainty Management (SUM'20)* (Lecture Notes in Artificial Intelligence). Ed. by J. Davis and K. Tabia. Vol. 12322. Springer International Publishing, 84–97.

A. Seidl and T. Sturm. 2003. *Boolean quantification in a first-order context*. Fakultät für Mathematik und Informatik, Universität Passau.

V. Sofronie-Stokkermans. 2025. "On Symbol Elimination and Uniform Interpolation in Theory Extensions." In: *Automated Deduction - CADE 30 - 30th International Conference on Automated Deduction, Stuttgart, Germany, July 28-31, 2025, Proceedings* (Lecture Notes in Computer Science). Ed. by C. W. Barrett and U. Waldmann. Springer, 35–55. doi:10.1007/978-3-031-99984-0_3.

T. Sturm and C. Zengler. 2010. "Parametric quantified SAT solving." In: *Symbolic and Algebraic Computation, International Symposium, ISSAC 2010, Munich, Germany, July 25-28, 2010, Proceedings*. Ed. by W. Koepf. ACM, 77–84. doi:10.1145/1837934.1837954.

M. Thimm. 2019. "Inconsistency Measurement." In: *Proceedings of the 13th International Conference on Scalable Uncertainty Management (SUM'19)*.

M. Thimm. 2013. "Inconsistency Measures for Probabilistic Logics." *Artificial Intelligence*, 197, 1–24.

M. Thimm. 2018. "On the Evaluation of Inconsistency Measures." In: *Measuring Inconsistency in Information*. Studies in Logic. Vol. 73. Ed. by J. Grant and M. V. Martinez. College Publications.

M. Thimm and J. P. Wallner. 2019. "On the Complexity of Inconsistency Measurement." *Artificial Intelligence*, 275, 411–456.

J. Williamson. 2009. "Aggregating Judgements by Merging Evidence." *Journal of Logic and Computation*, 19, 3, 461–473.

G. Xiao and Y. Ma. 2012. "Inconsistency Measurement based on Variables in Minimal Unsatisfiable Subsets." In: *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*.

Received 18 May 2026; accepted 28 October 2026