Combined Covers and Beth Definability (Extended Version)

Diego Calvanese[1], Silvio Ghilardi[2], Alessandro Gianola[1,3], Marco Montali[1], Andrey Rivkin[1]

[1]Faculty of Computer Science, Free University of Bozen-Bolzano (Italy)
{calvanese, gianola, montali, rivkin}@inf.unibz.it
[2]Dipartimento di Matematica, Università degli Studi di Milano (Italy)
silvio.ghilardi@unimi.it
[3]CSE Department, University of California San Diego (USA)
agianola@eng.ucsd.edu

Abstract. Uniform interpolants were largely studied in non-classical propositional logics since the nineties, and their connection to model completeness was pointed out in the literature. A successive parallel research line inside the automated reasoning community investigated uniform quantifier-free interpolants (sometimes referred to as "covers") in first-order theories. In this paper, we investigate cover transfer to theory combinations in the disjoint signatures case. We prove that, for convex theories, cover algorithms can be transferred to theory combinations under the same hypothesis needed to transfer quantifier-free interpolation (i.e., the equality interpolating property, aka strong amalgamation property). The key feature of our algorithm relies on the extensive usage of the Beth definability property for primitive fragments to convert implicitly defined variables into their explicitly defining terms. In the non-convex case, we show by a counterexample that cover may not exist in the combined theories, even in case combined quantifier-free interpolant do exist.

1 Introduction

Uniform interpolants were originally studied in the context of non-classical logics, starting from the pioneering work by Pitts [26]. We briefly recall what uniform interpolants are; we fix a logic or a theory T and a suitable fragment (propositional, first-order quantifier-free, etc.) of its language L. Given an L-formula $\phi(\underline{x},\underline{y})$ (here $\underline{x},\underline{y}$ are the variables occurring free in ϕ), a *uniform interpolant* of ϕ (w.r.t. $\underline{y}$) is a formula $\phi'(\underline{x})$ where only the $\underline{x}$ occur free, and that satisfies the following two properties: *(i)* $\phi(\underline{x},\underline{y}) \vdash_T \phi'(\underline{x})$; *(ii)* for any further L-formula $\psi(\underline{x},\underline{z})$ such that $\phi(\underline{x},\underline{y}) \vdash_T \psi(\underline{x},\underline{z})$, we have $\phi'(\underline{x}) \vdash_T \psi(\underline{x},\underline{z})$. Whenever uniform interpolants exist, one can compute an interpolant for an entailment like $\phi(\underline{x},\underline{y}) \vdash_T \psi(\underline{x},\underline{z})$ in a way that is *independent* of ψ.

The existence of uniform interpolants is an exceptional phenomenon, which is however not so infrequent; it has been extensively studied in non-classical logics starting from the nineties, as witnessed by a large literature (a non-exhaustive list includes [28,32,17,19,18,11,1,31,23]). The main results from the above papers are that uniform interpolants exist for intuitionistic logic and for some modal systems (like the Gödel-Löb system and the S4.Grz system); they do not exist for instance in $S4$ and $K4$, whereas for the basic modal system K they exist for the local consequence relation but not for the global consequence relation. The connection between uniform interpolants and model completions (for equational theories axiomatizing the varieties corresponding to propositional logics) was first stated in [20] and further developed in [17,31,23].

In the last decade, also the automated reasoning community developed an increasing interest in uniform interpolants, with particular focus on quantifier-free fragments of first-order theories. This is witnessed by various talks and drafts by D. Kapur presented in many conferences and workshops (FloC 2010, ISCAS 2013-14, SCS 2017 [22]), as well as by the paper [21] by Gulwani and Musuvathi in ESOP 2008. In this last paper uniform interpolants were renamed as *covers*, a terminology we shall adopt in this paper too. In these contributions, examples of cover computations were supplied and also some algorithms were sketched. The first formal *proof* about existence of covers in $\mathcal{EUF}$ was however published by the present authors only in [7]; such a proof was equipped with powerful semantic tools (the Cover-by-Extensions Lemma 1 below) coming from the connection to model-completeness, as well as with an algorithm relying on a constrained variant of the Superposition Calculus (two simpler algorithms are studied in [14]). The usefulness of covers in model checking was already stressed in [21] and further motivated by our recent line of research on the verification of data-aware processes [6,5,3,8]. Notably, it is also operationally mirrored in the MCMT [16] implementation since version 2.8. Covers (via quantifier elimination in model completions and hierarchical reasoning) play an important role in symbol elimination problems in theory extensions, as witnesssed in the comprehensive paper [29] and in related papers [25] studying invariant synthesis in model checking applications.

An important question suggested by the applications is the cover transfer problem for combined theories: for instance, when modeling and verifying data-aware processes, it is natural to consider the combination of different theories, such as the theories accounting for the read-write and read-only data storage of the process as well as those for the elements stored therein [6,7,8]. Formally, the cover transfer problem can be stated as follows: *by supposing that covers exist in theories T_1, T_2, under which conditions do they exist also in the combined theory $T_1 \cup T_2$?* In this paper we show that the answer is affirmative in the disjoint signatures convex case, using the same hypothesis (that is, the equality interpolating condition) under which quantifier-free interpolation transfers. Thus, for convex theories we essentially obtain a necessary and sufficient condition, in the precise sense captured by Theorem 6 below. We also prove that if convexity fails, the non-convex equality interpolating property [2] may not be sufficient to ensure the cover transfer property. As a witness for this, we show that $\mathcal{EUF}$ combined with integer difference logic or with linear integer arithmetics constitutes a counterexample.

The main tool employed in our combination result is the *Beth definability theorem for primitive formulae* (this theorem has been shown to be equivalent to the equality interpolating condition in [2]). In order to design a combined cover algorithm, we exploit the equivalence between implicit and explicit definability that is supplied by the Beth theorem. Implicit definability is reformulated, via covers for input theories, at the quantifier-free level. Thus, the combined cover algorithm guesses the implicitly definable variables, then eliminates them via explicit definability, and finally uses the component-wise input cover algorithms to eliminate the remaining (non implicitly definable) variables. The identification and the elimination of the implicitly defined variables via explicitly defining terms is an essential step towards the correctness of the combined cover algorithm: when computing a cover of a formula $\phi(\underline{x}, \underline{y})$ (w.r.t. $\underline{y}$), the variables $\underline{x}$ are (non-eliminable) parameters, and those variables among the $\underline{y}$ that are implicitly definable *need to be discovered and treated in the same way as the parameters* $\underline{x}$. Only after this preliminary step (Lemma 5 below), the input cover algorithms can be suitably exploited (Proposition 1 below).

The combination result we obtain is quite strong, as it is a typical 'black box' combination result: it applies not only to theories used in verification (like the combination of real arithmetics with $\mathcal{EUF}$), but also in other contexts. For instance, since the theory $\mathcal{B}$ of Boolean algebras satisfies our hypotheses (being model completable and strongly amalgamable [13]), we get that uniform interpolants exist in the combination of $\mathcal{B}$ with $\mathcal{EUF}$. The latter is the equational theory algebraizing the basic non-normal classical modal logic system **E** from [27] (extended to n-ary modalities). Notice that this result must be contrasted with the case of many systems of Boolean algebras with operators where existence of uniform interpolation fails [23] (recall that operators on a Boolean algebra are not just arbitrary functions, but are required to be monotonic and also to preserve either joins or meets in each coordinate).

As a last important comment on related work, it is worth mentioning that Gulwani and Musuvathi in [21] also have a combined cover algorithm for convex, signature disjoint theories. Their algorithm looks quite different from ours; apart from the fact that a full correctness and completeness proof for such an algorithm has never been published, we underline that our algorithm is rooted on different hypotheses. In fact, we only need the equality interpolating condition and we show that this hypothesis is not only sufficient, but also necessary for cover transfer in convex theories; consequently, our result is formally stronger. The equality interpolating condition was known to the authors of [21] (but not even mentioned in their paper [21]): in fact, it was introduced by one of them some years before [33]. The equality interpolating condition was then extended to the non convex case in [2], where it was also semantically characterized via the strong amalgamation property.

The paper is organized as follows: after some preliminaries in Section 2, the crucial Covers-by-Extensions Lemma and the relationship between covers and model completions from [7] are recalled in Section 3. In Section 4, we present some preliminary results on interpolation and Beth definability that are instrumental to our machinery. After some useful facts about convex theories in Section 5, we introduce the combined cover algorithms for the convex case and we prove its correctness in Section 6; we also present a detailed example of application of the combined algorithm in case of the combination of $\mathcal{EUF}$ with linear real arithmetic, and we show that the equality interpolating condition is necessary (in some sense) for combining covers. In Section 7 we exhibit a counteraxample to the existence of combined covers in the non-convex case. Finally, in Section 8 we prove that for the 'tame' multi-sorted theory combinations used in our database-driven applications, covers existence transfers to the combined theory under only the stable infiniteness requirement for the shared sorts. Section 9 is devoted to the conclusions and discussion of future work. The current paper is the extended version of [9].

2 Preliminaries

We adopt the usual first-order syntactic notions of signature, term, atom, (ground) formula, and so on; our signatures are always *finite* or *countable* and include equality. To avoid considering limit cases, we assume that signatures always contain at least an individual constant. We compactly represent a tuple $\langle x_1,\ldots,x_n\rangle$ of variables as $\underline{x}$. The notation $t(\underline{x}),\phi(\underline{x})$ means that the term t, the formula ϕ has free variables included in the tuple $\underline{x}$. This tuple is assumed to be formed by *distinct variables*, thus we underline that when we write e.g. $\phi(\underline{x},\underline{y})$, we mean that the tuples $\underline{x},\underline{y}$ are made of distinct variables that are also disjoint from each other.

A formula is said to be *universal* (resp., *existential*) if it has the form $\forall\underline{x}(\phi(\underline{x}))$ (resp., $\exists\underline{x}(\phi(\underline{x}))$), where ϕ is quantifier-free. Formulae with no free variables are called *sentences*. On the semantic side, we use the standard notion of Σ-structure $\mathcal{M}$ and of truth of a formula in a Σ-structure under a free variables assignment. The support of $\mathcal{M}$ is denoted as $|\mathcal{M}|$. The interpretation of a (function, predicate) symbol σ in $\mathcal{M}$ is denoted $\sigma^{\mathcal{M}}$.

A *Σ-theory T* is a set of Σ-sentences; a *model* of T is a Σ-structure $\mathcal{M}$ where all sentences in T are true. We use the standard notation $T\models\phi$ to say that ϕ is true in all models of T for every assignment to the variables occurring free in ϕ. We say that ϕ is *T-satisfiable* iff there is a model $\mathcal{M}$ of T and an assignment to the variables occurring free in ϕ making ϕ true in $\mathcal{M}$.

We now focus on the constraint satisfiability problem and quantifier elimination for a theory T. A Σ-formula ϕ is a *Σ-constraint* (or just a constraint) iff it is a conjunction of literals. The *constraint satisfiability problem* for T is the following: we are given a constraint $\phi(\underline{x})$ and we are asked whether there exist a model $\mathcal{M}$ of T and an assignment $\mathcal{I}$ to the free variables $\underline{x}$ such that $\mathcal{M},\mathcal{I}\models\phi(\underline{x})$. A theory T has *quantifier elimination* iff for every formula $\phi(\underline{x})$ in the signature of T there is a quantifier-free formula $\phi'(\underline{x})$ such that $T\models\phi(\underline{x})\leftrightarrow\phi'(\underline{x})$. Since we are in a computational logic context, when we speak of quantifier elimination, we assume that it is effective, namely that it comes with an algorithm for computing ϕ' out of ϕ. It is well-known that quantifier elimination holds in case we can eliminate quantifiers from *primitive* formulae, i.e., formulae of the kind $\exists\underline{y}\,\phi(\underline{x},\underline{y})$, with ϕ a constraint.

We recall also some further basic notions. Let Σ be a first-order signature. The signature obtained from Σ by adding to it a set $\underline{a}$ of new constants (i.e., 0-ary function symbols) is denoted by $\Sigma^{\underline{a}}$. Analogously, given a Σ-structure $\mathcal{M}$, the signature Σ can be expanded to a new signature $\Sigma^{|\mathcal{M}|}:=\Sigma\cup\{\bar{a}\mid a\in|\mathcal{M}|\}$ by adding a set of new constants $\bar{a}$ (the *name* for a), one for each element a in the support of $\mathcal{M}$, with the convention that two distinct elements are denoted by different "name" constants. $\mathcal{M}$ can be expanded to a $\Sigma^{|\mathcal{M}|}$-structure $\overline{\mathcal{M}}:=(\mathcal{M},a)_{a\in|\mathcal{M}|}$ just interpreting the additional constants over the corresponding elements. From now on, when the meaning is clear from the context, we will freely use the notation $\mathcal{M}$ and $\overline{\mathcal{M}}$ interchangeably: in particular, given a Σ-structure $\mathcal{M}$ and a Σ-formula $\phi(\underline{x})$ with free variables that are all in $\underline{x}$, we will write, by abuse of notation, $\mathcal{M}\models\phi(\underline{a})$ instead of $\overline{\mathcal{M}}\models\phi(\underline{\bar{a}})$.

A *Σ-homomorphism* (or, simply, a homomorphism) between two Σ-structures $\mathcal{M}$ and $\mathcal{N}$ is a map $\mu:|\mathcal{M}|\longrightarrow|\mathcal{N}|$ among the support sets $|\mathcal{M}|$ of $\mathcal{M}$ and $|\mathcal{N}|$ of $\mathcal{N}$ satisfying the condition $(\mathcal{M}\models\varphi \quad\Rightarrow\quad \mathcal{N}\models\varphi)$ for all $\Sigma^{|\mathcal{M}|}$-atoms φ ($\mathcal{M}$ is regarded as a $\Sigma^{|\mathcal{M}|}$-structure, by interpreting each additional constant $a \in |\mathcal{M}|$ into itself and $\mathcal{N}$ is regarded as a $\Sigma^{|\mathcal{M}|}$-structure by interpreting each additional constant $a \in |\mathcal{M}|$ into $\mu(a)$). In case the last condition holds for all $\Sigma^{|\mathcal{M}|}$-literals, the homomorphism μ is said to be an *embedding* and if it holds for all first order formulae, the embedding μ is said to be *elementary*. If $\mu : \mathcal{M} \longrightarrow \mathcal{N}$ is an embedding which is just the identity inclusion $|\mathcal{M}| \subseteq |\mathcal{N}|$, we say that $\mathcal{M}$ is a *substructure* of $\mathcal{N}$ or that $\mathcal{N}$ is an *extension* of $\mathcal{M}$. Universal theories can be characterized as those theories T having the property that if $\mathcal{M} \models T$ and $\mathcal{N}$ is a substructure of $\mathcal{M}$, then $\mathcal{N} \models T$ (see [10]). If $\mathcal{M}$ is a structure and $X \subseteq |\mathcal{M}|$, then there is the smallest substructure of $\mathcal{M}$ including X in its support; this is called the substructure *generated by* X. If X is the set of elements of a finite tuple $\underline{a}$, then the substructure generated by X has in its support precisely the $b \in |\mathcal{M}|$ such that $\mathcal{M} \models b = t(\underline{a})$ for some term t.

Let $\mathcal{M}$ be a Σ-structure. The *diagram* of $\mathcal{M}$, written $\Delta_\Sigma(\mathcal{M})$ (or just $\Delta(\mathcal{M})$), is the set of ground $\Sigma^{|\mathcal{M}|}$-literals that are true in $\mathcal{M}$. An easy but important result, called *Robinson Diagram Lemma* [10], says that, given any Σ-structure $\mathcal{N}$, the embeddings $\mu : \mathcal{M} \longrightarrow \mathcal{N}$ are in bijective correspondence with expansions of $\mathcal{N}$ to $\Sigma^{|\mathcal{M}|}$-structures which are models of $\Delta_\Sigma(\mathcal{M})$. The expansions and the embeddings are related in the obvious way: $\bar{a}$ is interpreted as $\mu(a)$.

3 Covers and Model Completions

We report the notion of *cover* taken from [21] and also the basic results proved in [7]. Fix a theory T and an existential formula $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$; call a *residue* of $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$ any quantifier-free formula belonging to the set of quantifier-free formulae $Res(\exists \underline{e}\, \phi) = \{\theta(\underline{y}, \underline{z}) \mid T \models \phi(\underline{e}, \underline{y}) \to \theta(\underline{y}, \underline{z})\}$. A quantifier-free formula $\psi(\underline{y})$ is said to be a *T-cover* (or, simply, a *cover*) of $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$ iff $\psi(\underline{y}) \in Res(\exists \underline{e}\, \phi)$ and $\psi(\underline{y})$ implies (modulo T) all the other formulae in $Res(\exists \underline{e}\, \phi)$. The following "cover-by-extensions" Lemma [7] (to be widely used throughout the paper) supplies a semantic counterpart to the notion of a cover:

Lemma 1 (Cover-by-Extensions). *A formula $\psi(\underline{y})$ is a T-cover of $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$ iff it satisfies the following two conditions: (i) $T \models \forall \underline{y}\, (\exists \underline{e}\, \phi(\underline{e}, \underline{y}) \to \psi(\underline{y}))$; (ii) for every model $\mathcal{M}$ of T, for every tuple of elements $\underline{a}$ from the support of $\mathcal{M}$ such that $\mathcal{M} \models \psi(\underline{a})$ it is possible to find another model $\mathcal{N}$ of T such that $\mathcal{M}$ embeds into $\mathcal{N}$ and $\mathcal{N} \models \exists \underline{e}\, \phi(\underline{e}, \underline{a})$.* $\triangleleft$

We underline that, since our language is at most countable, we can assume that the models $\mathcal{M}, \mathcal{N}$ from (ii) above are at most countable too, by a Löwenheim-Skolem argument.

We say that a theory T has *uniform quantifier-free interpolation* iff every existential formula $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$ (equivalently, every primitive formula $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$) has a T-cover.

It is clear that if T has uniform quantifier-free interpolation, then it has ordinary *quantifier-free interpolation* [2], in the sense that if we have $T \models \phi(\underline{e}, \underline{y}) \to \phi'(\underline{y}, \underline{z})$ (for quantifier-free formulae ϕ, ϕ'), then there is a quantifier-free formula $\theta(\underline{y})$ such that $T \models \phi(\underline{e}, \underline{y}) \to \theta(\underline{y})$ and $T \models \theta(\underline{y}) \to \phi'(\underline{y}, \underline{z})$. In fact, if T has uniform quantifier-free interpolation, then the interpolant θ is independent on ϕ' (the same $\theta(\underline{y})$ can be used as interpolant for all entailments $T \models \phi(\underline{e}, \underline{y}) \to \phi'(\underline{y}, \underline{z})$, varying ϕ').

We say that a *universal* theory T has a *model completion* iff there is a stronger theory $T^* \supseteq T$ (still within the same signature Σ of T) such that (i) every Σ-constraint that is satisfiable in a model of T is satisfiable in a model of T^*; (ii) T^* eliminates quantifiers. Other equivalent definitions are possible [10]: for instance, (i) is equivalent to the fact that T and T^* prove the same universal formulae or again to the fact that every model of T can be embedded into a model of T^*. We recall that the model completion, if it exists, is unique and that its existence implies the quantifier-free interpolation property for T [10] (the latter can be seen directly or via the correspondence between quantifier-free interpolation and amalgamability, see [2]).

A close relationship between model completion and uniform interpolation emerged in the area of propositional logic (see the book [17]) and can be formulated roughly as follows. It is well-known that most propositional calculi, via Lindembaum constructions, can be algebraized: the algebraic analogue of classical logic are Boolean algebras, the algebraic analogue of intuitionistic logic are Heyting algebras, the algebraic analogue of modal calculi are suitable variaties of modal agebras, etc. Under suitable hypotheses, it turns out that a propositional logic has uniform interpolation (for the global consequence relation) iff the equational theory axiomatizing the corresponding variety of algebras has a model completion [17]. In the context of first order theories, we prove an even more direct connection:

Theorem 1. *Suppose that T is a universal theory. Then T has a model completion T^* iff T has uniform quantifier-free interpolation. If this happens, T^* is axiomatized by the infinitely many sentences $\forall \underline{y}\,(\psi(\underline{y}) \to \exists \underline{e}\,\phi(\underline{e},\underline{y}))$, where $\exists \underline{e}\,\phi(\underline{e},\underline{y})$ is a primitive formula and ψ is a cover of it.* $\lhd$

The proof (via Lemma 1, by iterating a chain construction) is in [3] (see also [4]).

4 Equality Interpolating Condition and Beth Definability

We report here some definitions and results we need concerning combined quantifier-free interpolation. Most definitions and result come from [2], but are simplified here because we restrict them to the case of universal convex theories. Further information on the semantic side is supplied in Appendix A.

A theory T is *stably infinite* iff every T-satisfiable constraint is satisfiable in an infinite model of T. The following Lemma comes from a compactness argument (see Appendix A for a proof):

Lemma 2. *If T is stably infinite, then every finite or countable model $\mathcal{M}$ of T can be embedded in a model $\mathcal{N}$ of T such that $|\mathcal{N}| \setminus |\mathcal{M}|$ is countable.* $\lhd$

A theory T is *convex* iff for every constraint δ, if $T \vdash \delta \to \bigvee_{i=1}^n x_i = y_i$ then $T \vdash \delta \to x_i = y_i$ holds for some $i \in \{1,...,n\}$. A convex theory T is 'almost' stably infinite in the sense that it can be shown that every constraint which is T-satisfiable in a T-model whose support has at least two elements is satisfiable also in an infinite T-model. The one-element model can be used to build counterexamples, though: e.g., the theory of Boolean algebras is convex (like any other universal Horn theory) but the constraint $x = 0 \wedge x = 1$ is only satisfiable in the degenerate one-element Boolean algebra. Since we take into account these limit cases, we do not assume that convexity implies stable infiniteness.

Definition 1. *A convex universal theory T is* equality interpolating *iff for every pair y_1, y_2 of variables and for every pair of* constraints $\delta_1(\underline{x}, \underline{z}_1, y_1), \delta_2(\underline{x}, \underline{z}_2, y_2)$ *such that*

$$T \vdash \delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \to y_1 = y_2 \tag{1}$$

there exists a term $t(\underline{x})$ such that

$$T \vdash \delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \to y_1 = t(\underline{x}) \wedge y_2 = t(\underline{x}). \tag{2}$$

$\triangleleft$

Theorem 2. *[33,2] Let T_1 and T_2 be two universal, convex, stably infinite theories over disjoint signatures Σ_1 and Σ_2. If both T_1 and T_2 are equality interpolating and have quantifier-free interpolation property, then so does $T_1 \cup T_2$.* $\triangleleft$

There is a converse of the previous result; for a signature Σ, let us call $\mathcal{EUF}(\Sigma)$ the pure equality theory over the signature Σ (this theory is equality interpolating and has the quantifier-free interpolation property).

Theorem 3. *[2] Let T be a stably infinite, universal, convex theory admitting quantifier-free interpolation and let Σ be a signature disjoint from the signature of T containing at least a unary predicate symbol. Then, $T \cup \mathcal{EUF}(\Sigma)$ has quantifier-free interpolation iff T is equality interpolating.* $\triangleleft$

In [2] the above definitions and results are extended to the non-convex case and a long list of universal quantifier-free interpolating and equality interpolating theories is given. The list includes $\mathcal{EUF}(\Sigma)$, recursive data theories, as well as linear arithmetics. For linear arithmetics (and fragments of its), it is essential to make a very careful choice of the signature, see again [2] (especially Subsection 4.1) for details. All the above theories admit a model completion (which coincides with the theory itself in case the theory admits quantifier elimination).

The equality interpolating property in a theory T can be equivalently characterized using Beth definability as follows. Consider a primitive formula $\exists \underline{z} \phi(\underline{x}, \underline{z}, y)$ (here ϕ is a conjunction of literals); we say that $\exists \underline{z}\, \phi(\underline{x}, \underline{z}, y)$ *implicitly defines* y in T iff the formula

$$\forall y \forall y' \, (\exists \underline{z} \phi(\underline{x}, \underline{z}, y) \wedge \exists \underline{z} \phi(\underline{x}, \underline{z}, y') \to y = y') \tag{3}$$

is T-valid. We say that $\exists \underline{z} \phi(\underline{x}, \underline{z}, y)$ *explicitly defines* y in T iff there is a term $t(\underline{x})$ such that the formula

$$\forall y \, (\exists \underline{z} \phi(\underline{x}, \underline{z}, y) \to y = t(\underline{x})) \tag{4}$$

is T-valid.

For future use, we notice that, by trivial logical manipulations, the formulae (3) and (4) are logically equivalent to

$$\forall y \forall \underline{z} \forall y' \forall \underline{z}' (\phi(\underline{x}, \underline{z}, y) \wedge \phi(\underline{x}, \underline{z}', y') \to y = y') \quad . \tag{5}$$

and to

$$\forall y \forall \underline{z} (\phi(\underline{x}, \underline{z}, y) \to y = t(\underline{x})) \tag{6}$$

respectively (we shall use such equivalences without explicit mention).

We say that a theory T has the *Beth definability property for primitive formulae* iff whenever a primitive formula $\exists \underline{z}\, \phi(\underline{x}, \underline{z}, y)$ implicitly defines the variable y then it also explicitly defines it.

Theorem 4. *[2] A convex theory T having quantifier-free interpolation is equality interpolating iff it has the Beth definability property for primitive formulae.* $\triangleleft$

Proof. We recall the easy proof of the left-to-right side (this is the only side we need in this paper). Suppose that T is equality interpolating and that

$$T \vdash \phi(\underline{x},\underline{z},y) \wedge \phi(\underline{x},\underline{z}',y') \to y = y' \ \ ;$$

then there is a term $t(\underline{x})$ such that

$$T \vdash \phi(\underline{x},\underline{z},y) \wedge \phi(\underline{x},\underline{z}',y') \to y = t(\underline{x}) \wedge y' = t(\underline{x}) \ \ .$$

Replacing $\underline{z}', y'$ by $\underline{z}, y$ via a substitution, we get precisely (6). $\dashv$

5 Convex Theories

We now collect some useful facts concerning convex theories. We fix for this section a *convex, stably infinite, equality interpolating universal theory T admitting a model completion T^**. We let Σ be the signature of T. We fix also *a Σ-constraint $\phi(\underline{x},\underline{y})$*, where we assume that $\underline{y} = y_1,\dots,y_n$ (recall that the tuple $\underline{x}$ is disjoint from the tuple $\underline{y}$ according to our conventions from Section 2).

For $i = 1,\dots,n$, we let the formula $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$ be the quantifier-free formula equivalent in T^* to the formula

$$\forall \underline{y} \forall \underline{y}' (\phi(\underline{x},\underline{y}) \wedge \phi(\underline{x},\underline{y}') \to y_i = y_i') \qquad (7)$$

where the $\underline{y}'$ are renamed copies of the $\underline{y}$. Notice that the variables occurring free in ϕ are $\underline{x},\underline{y}$, whereas only the $\underline{x}$ occur free in $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$ (the variable y_i is among the $\underline{y}$ and does not occur free in $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$): these facts coming from our notational conventions are crucial and should be kept in mind when reading this and next section. The following semantic technical lemma is proved in Appendix A:

Lemma 3. *Suppose that we are given a model $\mathcal{M}$ of T and elements $\underline{a}$ from the support of $\mathcal{M}$ such that $\mathcal{M} \not\models \mathtt{ImplDef}^T_{\phi,y_i}(\underline{a})$ for all $i = 1,\dots,n$. Then there exists an extension $\mathcal{N}$ of $\mathcal{M}$ such that for some $\underline{b} \in |\mathcal{N}| \setminus |\mathcal{M}|$ we have $\mathcal{N} \models \phi(\underline{a},\underline{b})$.* $\lhd$

The following Lemma supplies terms which will be used as ingredients in our combined covers algorithm:

Lemma 4. *Let $L_{i1}(\underline{x}) \vee \cdots \vee L_{ik_i}(\underline{x})$ be the disjunctive normal form (DNF) of $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$. Then, for every $j = 1,\dots,k_i$, there is a $\Sigma(\underline{x})$-term $t_{ij}(\underline{x})$ such that*

$$T \vdash L_{ij}(\underline{x}) \wedge \phi(\underline{x},\underline{y}) \to y_i = t_{ij} \ \ . \qquad (8)$$

As a consequence, a formula of the kind $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x}) \wedge \exists \underline{y}\,(\phi(\underline{x},\underline{y}) \wedge \psi)$ is equivalent (modulo T) to the formula

$$\bigvee_{j=1}^{k_i} \exists \underline{y}\,(y_i = t_{ij} \wedge L_{ij}(\underline{x}) \wedge \phi(\underline{x},\underline{y}) \wedge \psi) \ \ . \qquad (9)$$

$\lhd$

Proof. We have that $(\bigvee_j L_{ij}) \leftrightarrow \mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$ is a tautology, hence from the definition of $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$, we have that

$$T^* \vdash L_{ij}(\underline{x}) \to \forall \underline{y} \forall \underline{y}' (\phi(\underline{x},\underline{y}) \wedge \phi(\underline{x},\underline{y}') \to y_i = y_i') \ \ ;$$

however this formula is trivially equivalent to a universal formula (L_{ij} does not depend on $\underline{y},\underline{y}'$), hence since T and T^* prove the same universal formulae, we get

$$T \vdash L_{ij}(\underline{x}) \wedge \phi(\underline{x},\underline{y}) \wedge \phi(\underline{x},\underline{y}') \to y_i = y_i' \ \ .$$

Using Beth definability property (Theorem 4), we get (8), as required, for some terms $t_{ij}(\underline{x})$. Finally, the second claim of the lemma follows from (8) by trivial logical manipulations. $\dashv$

In all our concrete examples, the theory T has decidable quantifier-free fragment (namely it is decidable whether a quantifier-free formula is a logical consequence of T or not), thus the terms t_{ij} mentioned in Lemma 4 can be computed just by enumerating all possible $\Sigma(\underline{x})$-terms: the computation terminates, because the above proof shows that the appropriate terms always exist. However, this is terribly inefficient and, from a practical point of view, one needs to have at disposal dedicated algorithms to find the required equality interpolating terms. For some common theories ($\mathcal{EUF}$, Lisp-structures, linear real arithmetic), such algorithms are designed in [33]; in [2] [Lemma 4.3 and Theorem 4.4], the algorithms for computing equality interpolating terms are connected to quantifier elimination algorithms in the case of universal theories admitting quantifier elimination. Still, an extensive investigation on te topic seems to be missed in the SMT literature.

6 The Convex Combined Cover Algorithm

Let us now fix two theories T_1, T_2 over disjoint signatures Σ_1, Σ_2. We assume that both of them satisfy the assumptions from the previous section, meaning that they are convex, stably infinite, equality interpolating, universal and admit model completions T_1^*, T_2^* respectively. We shall supply a cover algorithm for $T_1 \cup T_2$ (thus proving that $T_1 \cup T_2$ has a model completion too).

We need to compute a cover for $\exists \underline{e}\, \phi(\underline{x}, \underline{e})$, where ϕ is a conjunction of $\Sigma_1 \cup \Sigma_2$-literals. By applying rewriting purification steps like

$$\phi \Longrightarrow \exists d\, (d = t \wedge \phi(d/t))$$

(where d is a fresh variable and t is a pure term, i.e. it is either a Σ_1- or a Σ_2-term), we can assume that our formula ϕ is of the kind $\phi_1 \wedge \phi_2$, where ϕ_1 is a Σ_1-formula and ϕ_2 is a Σ_2-formula. Thus we need to compute a cover for a formula of the kind

$$\exists \underline{e}\, (\phi_1(\underline{x}, \underline{e}) \wedge \phi_2(\underline{x}, \underline{e})), \tag{10}$$

where ϕ_i is a conjunction of Σ_i-literals ($i = 1, 2$). We also assume that both ϕ_1 and ϕ_2 contain the literals $e_i \neq e_j$ (for $i \neq j$) as a conjunct: this can be achieved by guessing a partition of the $\underline{e}$ and by replacing each e_i with the representative element of its equivalence class.

Remark 1. It is not clear whether this preliminary guessing step can be avoided. In fact, Nelson-Oppen [24] combined satisfiability for *convex* theories does not need it; however, combining covers algorithms is a more complicated problem than combining mere satisfiability algorithms and for technical reasons related to the correctness and completeness proofs below, we were forced to introduce guessing at this step. $\triangleleft$

To manipulate formulae, our algorithm employs acyclic explicit definitions as follows. When we write `ExplDef`$(\underline{z}, \underline{x})$ (where $\underline{z}, \underline{x}$ are tuples of distinct variables), we mean any formula of the kind (let $\underline{z} := z_1 \ldots, z_m$)

$$\bigwedge_{i=1}^{m} z_i = t_i(z_1, \ldots, z_{i-1}, \underline{x})$$

where the term t_i is pure (i.e. it is a Σ_i-term) and only the variables $z_1,\dots,z_{i-1},\underline{x}$ can occur in it. When we assert a formula like $\exists \underline{z}\,(\texttt{ExplDef}(\underline{z},\underline{x}) \wedge \psi(\underline{z},\underline{x}))$, we are in fact in the condition of recursively eliminating the variables $\underline{z}$ from it via terms containing only the parameters $\underline{x}$ (the 'explicit definitions' $z_i = t_i$ are in fact arranged acyclically).

A *working formula* is a formula of the kind

$$\exists \underline{z}\,(\texttt{ExplDef}(\underline{z},\underline{x}) \wedge \exists \underline{e}\,(\psi_1(\underline{x},\underline{z},\underline{e}) \wedge \psi_2(\underline{x},\underline{z},\underline{e}))) \ , \tag{11}$$

where ψ_1 is a conjunction of Σ_1-literals and ψ_2 is a conjunction of Σ_2-literals. The variables $\underline{x}$ are called *parameters*, the variables $\underline{z}$ are called *defined variables* and the variables $\underline{e}$ *(truly) existential variables*. The parameters do not change during the execution of the algorithm. We assume that ψ_1,ψ_2 in a working formula (11) always contain the literals $e_i \neq e_j$ (for distinct e_i,e_j from $\underline{e}$) as a conjunct.

In our starting formula (10), there are no defined variables. However, if via some syntactic check it happens that some of the existential variables can be recognized as defined, then it is useful to display them as such (this observation may avoid redundant cases - leading to inconsistent disjuncts - in the computations below).

A working formula like (11) is said to be *terminal* iff for every existential variable $e_i \in \underline{e}$ we have that

$$T_1 \vdash \psi_1 \to \neg\texttt{ImplDef}^{T_1}_{\psi_1,e_i}(\underline{x},\underline{z}) \ \text{ and } \ T_2 \vdash \psi_2 \to \neg\texttt{ImplDef}^{T_2}_{\psi_2,e_i}(\underline{x},\underline{z}) \quad . \tag{12}$$

Roughly speaking, we can say that in a terminal working formula, all variables which are not parameters are either explicitly definable or recognized as not implicitly definable by both theories; of course, a working formula with no existential variables is terminal.

Lemma 5. *Every working formula is equivalent (modulo $T_1 \cup T_2$) to a disjunction of terminal working formulae.* $\lhd$

Proof. We only sketch the proof of this Lemma (see the Appendix A for full details), by describing the algorithm underlying it. To compute the required terminal working formulae, it is sufficient to apply the following non-deterministic procedure (the output is the disjunction of all possible outcomes). The non-deterministic procedure applies one of the following alternatives.

(1) Update ψ_1 by adding it a disjunct from the DNF of $\bigwedge_{e_i\in\underline{e}} \neg\texttt{ImplDef}^{T_1}_{\psi_1,e_i}(\underline{x},\underline{z})$ and ψ_2 by adding to it a disjunct from the DNF of $\bigwedge_{e_i\in\underline{e}} \neg\texttt{ImplDef}^{T_2}_{\psi_1,e_i}(\underline{x},\underline{z})$;

(2.i) Select $e_i \in \underline{e}$ and $h \in \{1,2\}$; then update ψ_h by adding to it a disjunct L_{ij} from the DNF of $\texttt{ImplDef}^{T_h}_{\psi_h,e_i}(\underline{x},\underline{z})$; the equality $e_i = t_{ij}$ (where t_{ij} is the term mentioned in Lemma 4)[1] is added to $\texttt{ExplDef}(\underline{z},\underline{x})$; the variable e_i becomes in this way part of the defined variables.

If alternative (1) is chosen, the procedure stops, otherwise it is recursively applied again and again (we have one truly existential variable less after applying alternative (2.i), so we eventually terminate). $\dashv$

Thus we are left to the problem of computing a cover of a terminal working formula; this problem is solved in the following proposition:

[1] Lemma 4 is used taking as $\underline{y}$ the tuple $\underline{e}$, as $\underline{x}$ the tuple $\underline{x},\underline{z}$, as $\phi(\underline{x},\underline{y})$ the formula $\psi_h(\underline{x},\underline{z},\underline{e})$ and as ψ the formula ψ_{3-h}.

Proposition 1. *A cover of a terminal working formula* (11) *can be obtained just by unravelling the explicit definitions of the variables $\underline{z}$ from the formula*

$$\exists \underline{z}\, (\mathtt{ExplDef}(\underline{z},\underline{x}) \wedge \theta_1(\underline{x},\underline{z}) \wedge \theta_2(\underline{x},\underline{z})) \tag{13}$$

where $\theta_1(\underline{x},\underline{z})$ is the T_1-cover of $\exists \underline{e} \psi_1(\underline{x},\underline{z},\underline{e})$ and $\theta_2(\underline{x},\underline{z}))$ is the T_2-cover of $\exists \underline{e} \psi_2(\underline{x},\underline{z},\underline{e})$. $\lhd$

Proof. In order to show that Formula (13) is the $T_1 \cup T_2$-cover of a terminal working formula (11), we prove, by using the Cover-by-Extensions Lemma 1, that, for every $T_1 \cup T_2$-model $\mathcal{M}$, for every tuple $\underline{a}, \underline{c}$ from $|\mathcal{M}|$ such that $\mathcal{M} \models \theta_1(\underline{a},\underline{c}) \wedge \theta_2(\underline{a},\underline{c})$ there is an extension $\mathcal{N}$ of $\mathcal{M}$ such that $\mathcal{N}$ is still a model of $T_1 \cup T_2$ and $\mathcal{N} \models \exists \underline{e}(\psi_1(\underline{a},\underline{c},\underline{e}) \wedge \psi_2(\underline{a},\underline{c},\underline{e}))$. By a Löwenheim-Skolem argument, since our languages are countable, we can suppose that $\mathcal{M}$ is at most countable and actually that it is countable by stable infiniteness of our theories, see Lemma 2 (the fact that $T_1 \cup T_2$ is stably infinite in case both T_1, T_2 are such, comes from the proof of Nelson-Oppen combination result, see [24],[30], [12]).

According to the conditions (12) and the definition of a cover (notice that the formulae $\neg \mathtt{ImplDef}^{T_h}_{\psi_h,e_i}(\underline{x},\underline{z})$ do not contain the $\underline{e}$ and are quantifier-free) we have that

$$T_1 \vdash \theta_1 \to \neg \mathtt{ImplDef}^{T_1}_{\psi_1,e_i}(\underline{x},\underline{z}) \ \text{ and } \ T_2 \vdash \theta_2 \to \neg \mathtt{ImplDef}^{T_2}_{\psi_2,e_i}(\underline{x},\underline{z})$$

(for every $e_i \in \underline{e}$). Thus, since $\mathcal{M} \not\models \mathtt{ImplDef}^{T_1}_{\psi_1,e_i}(\underline{a},\underline{c})$ and $\mathcal{M} \not\models \mathtt{ImplDef}^{T_2}_{\psi_2,e_i}(\underline{a},\underline{c})$ holds for every $e_i \in \underline{e}$, we can apply Lemma 3 and conclude that there exist a T_1-model $\mathcal{N}_1$ and a T_2-model $\mathcal{N}_2$ such that $\mathcal{N}_1 \models \psi_1(\underline{a},\underline{c},\underline{b}_1)$ and $\mathcal{N}_2 \models \psi_2(\underline{a},\underline{c},\underline{b}_2)$ for tuples $\underline{b}_1 \in |\mathcal{N}_1|$ and $\underline{b}_2 \in |\mathcal{N}_2|$, both disjoint from $|\mathcal{M}|$. By a Löwenheim-Skolem argument, we can suppose that $\mathcal{N}_1, \mathcal{N}_2$ are countable and by Lemma 2 even that they are both countable extensions of $\mathcal{M}$.

The tuples $\underline{b}_1$ and $\underline{b}_2$ have equal length because the ψ_1, ψ_2 from our working formulae entail $e_i \neq e_j$, where e_i, e_j are different existential variables. Thus there is a bijection $\iota : |\mathcal{N}_1| \to |\mathcal{N}_2|$ fixing all elements in $\mathcal{M}$ and mapping component-wise the $\underline{b}_1$ onto the $\underline{b}_2$. But this means that, exactly as it happens in the proof of the completeness of the Nelson-Oppen combination procedure, the Σ_2-structure on $\mathcal{N}_2$ can be moved back via ι^{-1} to $|\mathcal{N}_1|$ in such a way that the Σ_2-substructure from $\mathcal{M}$ is fixed and in such a way that the tuple $\underline{b}_2$ is mapped to the tuple $\underline{b}_1$. In this way, $\mathcal{N}_1$ becomes a $\Sigma_1 \cup \Sigma_2$-structure which is a model of $T_1 \cup T_2$ and which is such that $\mathcal{N}_1 \models \psi_1(\underline{a},\underline{c},\underline{b}_1) \wedge \psi_2(\underline{a},\underline{c},\underline{b}_1)$, as required. $\dashv$

From Lemma 5, Proposition 1 and Theorem 1, we immediately get

Theorem 5. *Let T_1, T_2 be convex, stably infinite, equality interpolating, universal theories over disjoint signatures admitting a model completion. Then $T_1 \cup T_2$ admits a model completion too. Covers in $T_1 \cup T_2$ can be effectively computed as shown above.* $\lhd$

Notice that the input cover algorithms in the above combined cover computation algorithm are used not only in the final step described in Proposition 1, but also every time we need to compute a formula $\mathtt{ImplDef}^{T_h}_{\psi_h,e_i}(\underline{x},\underline{z})$: according to its definition, this formula is obtained by eliminating quantifiers in T_i^* from (7) (this is done via a cover computation, reading $\forall$ as $\neg\exists\neg$). In practice, implicit definability is not very frequent, so that in many concrete cases $\mathtt{ImplDef}^{T_h}_{\psi_h,e_i}(\underline{x},\underline{z})$ is trivially equivalent to $\bot$ (in such cases, Step (2.i) above can obviously be disregarded).

An Example. We now analyze an example in detail. Our results apply for instance to the case where T_1 is $\mathcal{EUF}(\Sigma)$ and T_2 is linear real arithmetic. We recall that covers are computed in real arithmetic by quantifier elimination, whereas for $\mathcal{EUF}(\Sigma)$ one can apply the superposition-based algorithm from [7]. Let us show that the cover of

$$\exists e_1 \cdots \exists e_4 \begin{pmatrix} e_1 = f(x_1) \wedge e_2 = f(x_2) \wedge \\ \wedge f(e_3) = e_3 \wedge f(e_4) = x_1 \wedge \\ \wedge x_1 + e_1 \leq e_3 \wedge e_3 \leq x_2 + e_2 \wedge e_4 = x_2 + e_3 \end{pmatrix} \tag{14}$$

is the following formula

$$\begin{aligned} &[x_2 = 0 \wedge f(x_1) = x_1 \wedge x_1 \leq 0 \wedge x_1 \leq f(0)] \vee \\ &\vee [x_1 + f(x_1) < x_2 + f(x_2) \wedge x_2 \neq 0] \vee \\ &\vee \begin{bmatrix} x_2 \neq 0 \wedge x_1 + f(x_1) = x_2 + f(x_2) \wedge f(2x_2 + f(x_2)) = x_1 \wedge \\ \wedge f(x_1 + f(x_1)) = x_1 + f(x_1) \end{bmatrix} \end{aligned} \tag{15}$$

Formula (14) is already purified. Notice also that the variables e_1, e_2 are in fact already explicitly defined (only e_3, e_4 are truly existential variables).

We first make the partition guessing. There is no need to involve defined variables into the partition guessing, hence we need to consider only two partitions; they are described by the following formulae:

$$\begin{aligned} P_1(e_3, e_4) &\equiv e_3 \neq e_4 \\ P_2(e_3, e_4) &\equiv e_3 = e_4 \end{aligned}$$

We first analyze **the case of P_1**. The formulae ψ_1 and ψ_2 to which we need to apply exhaustively Step (1) and Step (2.i) of our algorithm are:

$$\begin{aligned} \psi_1 &\equiv f(e_3) = e_3 \wedge f(e_4) = x_1 \wedge e_3 \neq e_4 \\ \psi_2 &\equiv x_1 + e_1 \leq e_3 \wedge e_3 \leq x_2 + e_2 \wedge e_4 = x_2 + e_3 \wedge e_3 \neq e_4 \end{aligned}$$

We first compute the implicit definability formulae for the truly existential variables with respect to both T_1 and T_2.

- We first consider $\mathtt{ImplDef}^{T_1}_{\psi_1, e_3}(\underline{x}, \underline{z})$. Here we show that the cover of the negation of formula (7) is equivalent to $\top$ (so that $\mathtt{ImplDef}^{T_1}_{\psi_1, e_3}(\underline{x}, \underline{z})$ is equivalent to $\bot$). We must quantify over truly existential variables and their duplications, thus we need to compute the cover of

  $$f(e_3') = e_3' \wedge f(e_3) = e_3 \wedge f(e_4') = x_1 \wedge f(e_4) = x_1 \wedge e_3 \neq e_4 \wedge e_3' \neq e_4' \wedge e_3' \neq e_3$$

  This is a saturated set according to the superposition based procedure of [7], hence the result is $\top$, as claimed.
- The formula $\mathtt{ImplDef}^{T_1}_{\psi_1, e_4}(\underline{x}, \underline{z})$ is also equivalent to $\bot$, by the same argument as above.
- To compute $\mathtt{ImplDef}^{T_2}_{\psi_2, e_3}(\underline{x}, \underline{z})$ we use Fourier-Motzkin quantifier elimination. We need to eliminate the variables e_3, e_3', e_4, e_4' (intended as existentially quantified variables) from

  $$\begin{aligned} &x_1 + e_1 \leq e_3' \leq x_2 + e_2 \wedge x_1 + e_1 \leq e_3 \leq x_2 + e_2 \wedge e_4' = x_2 + e_3' \wedge \\ &\wedge e_4 = x_2 + e_3 \wedge e_3 \neq e_4 \wedge e_3' \neq e_4' \wedge e_3' \neq e_3 \ . \end{aligned}$$

  This gives $x_1 + e_1 \neq x_2 + e_2 \wedge x_2 \neq 0$, so that $\mathtt{ImplDef}^{T_2}_{\psi_2, e_3}(\underline{x}, \underline{z})$ is $x_1 + e_1 = x_2 + e_2 \wedge x_2 \neq 0$. The corresponding equality interpolating term for e_3 is $x_1 + e_1$.

- The formula $\texttt{ImplDef}^{T_2}_{\psi_2,e_4}(\underline{x},\underline{z})$ is also equivalent to $x_1+e_1=x_2+e_2\wedge x_2\neq 0$ and the equality interpolating term for e_4 is $x_1+e_1+x_2$.

So, if we apply Step 1 we get

$$\exists e_1\cdots\exists e_4 \begin{pmatrix} e_1=f(x_1)\ \wedge\ e_2=f(x_2)\ \wedge \\ \wedge\ f(e_3)=e_3\ \wedge\ f(e_4)=x_1\ \wedge\ e_3\neq e_4\ \wedge \\ \wedge\ x_1+e_1\leq e_3\ \wedge\ e_3\leq x_2+e_2\ \wedge\ e_4=x_2+e_3\ \wedge\ x_1+e_1\neq x_2+e_2 \end{pmatrix} \tag{16}$$

(notice that the literal $x_2\neq 0$ is entailed by ψ_2, so we can simplify it to $\top$ in $\texttt{ImplDef}^{T_2}_{\psi_2,e_3}(\underline{x},\underline{z})$ and $\texttt{ImplDef}^{T_2}_{\psi_2,e_4}(\underline{x},\underline{z})$). If we apply Step (2.i) (for i=3), we get (after removing implied equalities)

$$\exists e_1\cdots\exists e_4 \begin{pmatrix} e_1=f(x_1)\ \wedge\ e_2=f(x_2)\ \wedge\ e_3=x_1+e_1\ \wedge \\ \wedge\ f(e_3)=e_3\ \wedge\ f(e_4)=x_1\ \wedge\ e_3\neq e_4\ \wedge \\ \wedge\ e_4=x_2+e_3\ \wedge\ x_1+e_1=x_2+e_2 \end{pmatrix} \tag{17}$$

Step (2.i) (for i=4) gives a formula logically equivalent to (17). Notice that (17) is terminal too, because all existential variables are now explicitly defined (this is a lucky side-effect of the fact that e_3 has been moved to the defined variables). Thus the exhaustive application of Steps (1) and (2.i) is concluded.

Applying the final step of Proposition 1 to (17) is quite easy: it is sufficient to unravel the acyclic definitions. The result, after little simplification, is

$$\begin{aligned} & x_2\neq 0\wedge x_1+f(x_1)=x_2+f(x_2)\wedge \\ & \wedge\ f(x_2+f(x_1+f(x_1)))=x_1\ \wedge\ f(x_1+f(x_1))=x_1+f(x_1); \end{aligned}$$

this can be further simplified to

$$\begin{aligned} & x_2\neq 0\wedge x_1+f(x_1)=x_2+f(x_2)\wedge \\ & \wedge\ f(2x_2+f(x_2))=x_1\ \wedge\ f(x_1+f(x_1))=x_1+f(x_1); \end{aligned} \tag{18}$$

As to formula (16), we need to apply the final cover computations mentioned in Proposition 1. The formulae ψ_1 and ψ_2 are now

$$\begin{aligned} \psi_1' &\equiv\ f(e_3)=e_3\ \wedge\ f(e_4)=x_1\ \wedge\ e_3\neq e_4 \\ \psi_2' &\equiv\ x_1+e_1\leq e_3\leq x_2+e_2\wedge e_4=x_2+e_3\wedge x_1+e_1\neq x_2+e_2\ \wedge\ e_3\neq e_4 \end{aligned}$$

The T_1-cover of ψ_1' is $\top$. For the T_2-cover of ψ_2', eliminating with Fourier-Motzkin the variables e_4 and e_3, we get

$$x_1+e_1<x_2+e_2\wedge x_2\neq 0$$

which becomes

$$x_1+f(x_1)<x_2+f(x_2)\wedge x_2\neq 0 \tag{19}$$

after unravelling the explicit definitions of e_1,e_2. Thus, *the analysis of the case of the partition P_1 gives, as a result, the disjunction of* (18) *and* (19).

We now analyze **the case of** P_2. Before proceeding, we replace e_4 with e_3 (since P_2 precisely asserts that these two variables coincide); our formulae ψ_1 and ψ_2 become

$$\begin{aligned} \psi_1'' &\equiv\ f(e_3)=e_3\ \wedge\ f(e_3)=x_1 \\ \psi_2'' &\equiv\ x_1+e_1\leq e_3\ \wedge\ e_3\leq x_2+e_2\ \wedge\ 0=x_2 \end{aligned}$$

From ψ_1'' we deduce $e_3 = x_1$, thus we can move e_3 to the explicitly defined variables (this avoids useless calculations: the implicit definability condition for variables having an entailed explicit definition is obviously $\top$, so making case split on it produces either tautological consequences or inconsistencies). In this way we get the terminal working formula

$$\exists e_1 \cdots \exists e_3 \begin{pmatrix} e_1 = f(x_1) \wedge e_2 = f(x_2) \wedge e_3 = x_1 \\ \wedge f(e_3) = e_3 \wedge \ f(e_3) = x_1 \wedge \\ \wedge x_1 + e_1 \leq e_3 \wedge e_3 \leq x_2 + e_2 \wedge 0 = x_2 \end{pmatrix} \tag{20}$$

Unravelling the explicit definitions, we get (after exhaustive simplifications)

$$x_2 = 0 \wedge f(x_1) = x_1 \wedge x_1 \leq 0 \wedge x_1 \leq f(0) \tag{21}$$

Now, the disjunction of (18),(19) and (21) is precisely the final result (15) claimed above. This concludes our detailed analysis of our example.

Notice that the example shows that combined cover computations may introduce terms with arbitrary alternations of symbols from both theories (like $f(x_2 + f(x_1 + f(x_1)))$ above). The point is that when a variable becomes explicitly definable via a term in one of the theories, then using such additional variable may in turn cause some other variables to become explicitly definable via terms from the other theory, and so on and so forth; when ultimately the explicit definitions are unraveled, highly nested terms arise with many symbol alternations from both theories.

The Necessity of the Equality Interpolating Condition. The following result shows that equality interpolating is a necessary condition for a transfer result, in the sense that it is already required for minimal combinations with signatures adding uninterpreted symbols:

Theorem 6. *Let T be a convex, stably infinite, universal theory admitting a model completion and let Σ be a signature disjoint from the signature of T containing at least a unary predicate symbol. Then $T \cup \mathcal{EUF}(\Sigma)$ admits a model completion iff T is equality interpolating.* $\lhd$

Proof. The necessity can be shown by using the following argument. By Theorem 1, $T \cup \mathcal{EUF}(\Sigma)$ has uniform quantifier-free interpolation, hence also ordinary quantifier-free interpolation. We can now apply Theorem 3 and get that T must be equality interpolating. Conversely, the sufficiency comes from Theorem 5 together with the fact that $\mathcal{EUF}(\Sigma)$ is trivially universal, convex, stably infinite, has a model completion [7] and is equality interpolating [33],[2]. $\dashv$

7 The Non-Convex Case: a Counterexample

In this section, we show by giving a suitable counterexample that the convexity hypothesis cannot be dropped from Theorems 5, 6. We make use of basic facts about ultrapowers (see [10] for the essential information we need). We take as T_1 integer difference logic $\mathcal{IDL}$, i.e. the theory of integer numbers under the unary operations of successor and predecessor, the constant 0 and the strict order relation $<$. This is stably infinite, universal and has quantifier elimination (thus it coincides with its own model completion). It is not convex, but it satisfies the equality interpolating condition, once the latter is suitably adjusted to non-convex theories, see [2] for the related definition and all the above mentioned facts.

As T_2, we take $\mathcal{EUF}(\Sigma_f)$, where Σ_f has just one unary free function symbol f (this f is supposed not to belong to the signature of T_1).

Proposition 2. *Let T_1, T_2 be as above; the formula*

$$\exists e\,(0 < e \wedge e < x \wedge f(e) = 0) \tag{22}$$

does not have a cover in $T_1 \cup T_2$. $\triangleleft$

Proof. Suppose that (22) has a cover $\phi(x)$. This means (according to Cover-by-Extensions Lemma 1) that for every model $\mathcal{M}$ of $T_1 \cup T_2$ and for every element $a \in |\mathcal{M}|$ such that $\mathcal{M} \models \phi(a)$, there is an extension $\mathcal{N}$ of $\mathcal{M}$ such that $\mathcal{N} \models \exists e\,(0 < e \wedge e < a \wedge f(e) = 0)$.

Consider the model $\mathcal{M}$, so specified: the support of $\mathcal{M}$ is the set of the integers, the symbols from the signature of T_1 are interpreted in the standard way and the symbol f is interpreted so that 0 is not in the image of f. Let a_k be the number $k > 0$ (it is an element from the support of $\mathcal{M}$). Clearly it is not possible to extend $\mathcal{M}$ so that $\exists e\,(0 < e \wedge e < a_k \wedge f(e) = 0)$ becomes true: indeed, we know that all the elements in the interval $(0, k)$ are definable as iterated successors of 0 and, by using the axioms of $\mathcal{IDL}$, no element can be added between a number and its successor, hence this interval cannot be enlarged in a superstructure. We conclude that $\mathcal{M} \models \neg\phi(a_k)$ for every k.

Consider now an ultrapower $\Pi_D\mathcal{M}$ of $\mathcal{M}$ modulo a non-principal ultrafilter D and let a be the equivalence class of the tuple $\langle a_k \rangle_{k\in\mathbb{N}}$; by the fundamental Los theorem [10], $\Pi_D\mathcal{M} \models \neg\phi(a)$. We claim that it is possible to extend $\Pi_D\mathcal{M}$ to a superstructure $\mathcal{N}$ such that $\mathcal{N} \models \exists e\,(0 < e \wedge e < a \wedge f(e) = 0)$: this would entail, by definition of cover, that $\Pi_D\mathcal{M} \models \phi(a)$, contradiction. We now show why the claim is true. Indeed, since $\langle a_k \rangle_{k\in\mathbb{N}}$ has arbitrarily big numbers as its components, we have that, in $\Pi_D\mathcal{M}$, a is bigger than all standard numbers. Thus, if we take a further non-principal ultrapower $\mathcal{N}$ of $\Pi_D\mathcal{M}$, it becomes possible to change in it the evaluation of $f(b)$ for some $b < a$ and set it to 0 (in fact, as it can be easily seen, there are elements $b \in |\mathcal{N}|$ less than a but not in the support of $\Pi_D\mathcal{M}$). $\dashv$

The counterexample still applies when replacing integer difference logic with linear integer arithmetics.

8 Tame Combinations

So far, we only analyzed the mono-sorted case. However, many interesting examples arising in model-checking verification are multi-sorted: this is the case of array-based systems [15] and in particular of the array-based system used in data-aware verification [6],[3].The above examples suggest restrictions on the theories to be combined other than convexity, in particular they suggest restrictions that make sense in a multi-sorted context.

Most definitions we gave in Section 2 have straightforward natural extensions to the multi-sorted case (we leave the reader to formulate them). A little care is needed however for the disjoint signatures requirement. Let T_1, T_2 be multisorted theories in the signatures Σ_1, Σ_2; the disjointness requirement for Σ_1 and Σ_2 can be formulated in this context by saying that the only function or relation symbols in $\Sigma_1 \cap \Sigma_2$ are the equality predicates over the common sorts in $\Sigma_1 \cap \Sigma_2$. We want to strengthen this requirement: we say that the combination $T_1 \cup T_2$ is *tame* iff the sorts in $\Sigma_1 \cap \Sigma_2$ *can only be the codomain sort* (and not a domain sort) of a symbol from Σ_1 other than an equality predicate. In other word, if a relation or a function symbol has as among its domain sorts a sort from $\Sigma_1 \cap \Sigma_2$, then this symbol is from Σ_2 (and not from Σ_1, unless it is the equality predicate).

Tame combinations arise in infinite-state model-checking (in fact, the definition is suggested by this application domain), where signatures can be split into a signature Σ_2 for 'data' and a signature Σ_1 for 'data containers', see [6],[3].

Notice that the notion of a tame combination is not symmetric in T_1 and T_2: to see this, notice that if the sorts of Σ_1 are included in the sorts of Σ_2, then T_1 must be a pure equality theory (but this is not the case if we swap T_1 with T_2). The combination of $\mathcal{IDL}$ and $\mathcal{EUF}(\Sigma)$ used in the counterexample of section 7 is not tame: even if we formulate $\mathcal{EUF}(\Sigma)$ as a two-sorted theory, the unique sort of $\mathcal{IDL}$ must be a sort of $\mathcal{EUF}(\Sigma)$ too, as witnessed by the impure atom $f(e) = 0$ in the formula (22). Because of this, for the combination to be tame, $\mathcal{IDL}$ should play the role of T_2 (the arithmetic operation symbols are defined on a shared sort); however, the unary function symbol $f \in \Sigma$ has a shared sort as domain sort, so the combination is not tame anyway.

In a tame combination, an atomic formula A can only be of two kinds: (1) we say that A is of the *first kind* iff the sorts of its root predicate are from $\Sigma_1 \setminus \Sigma_2$; (2) we say that A is of the *second kind* iff the sorts of its root predicate are from Σ_2. We use the roman letters $e, x, \dots$ for variables ranging over sorts in $\Sigma_1 \setminus \Sigma_2$ and the greek letters $\eta, \xi, \dots$ for variables ranging over sorts in Σ_2. Thus, if we want to display free variables, atoms of the first kind can be represented as $A(e, x, \dots)$, whereas atoms of the second kind can be represented as $A(\eta, \xi, \dots, t(e, x, \dots), \dots)$, where the t are Σ_1-terms.

Suppose not that $T_1 \cup T_2$ is a tame combination and that T_1, T_2 are universal theories admitting model completions T_1^*, T_2^*. We propose the following algorithm to compute the cover of a primitive formula; the latter must be of the kind

$$\exists \underline{e}\, \exists \underline{\eta} (\phi(\underline{e}, \underline{x}) \wedge \psi(\underline{\eta}, \underline{\xi}, \underline{t}(\underline{e}, \underline{x}))) \tag{23}$$

where ϕ is a Σ_1-conjunction of literals, ψ is a conjunction of Σ_2-literals and the $\underline{t}$ are Σ_1-terms. The algorithm has three steps.

(i) We flatten (23) and get

$$\exists \underline{e}\, \exists \underline{\eta}\, \exists \underline{\eta}' \, (\phi(\underline{e}, \underline{x}) \wedge \underline{\eta}' = \underline{t}(\underline{e}, \underline{x}) \wedge \psi(\underline{\eta}, \underline{\xi}, \underline{\eta}'))) \tag{24}$$

where the $\underline{\eta}'$ are fresh variables abstracting out the $\underline{t}$ and $\underline{\eta}' = \underline{t}(\underline{e}, \underline{x})$ is a componentwise conjunction of equalities.

(ii) We apply the cover algorithm of T_1 to the formula

$$\exists \underline{e} \, (\phi(\underline{e}, \underline{x}) \wedge \underline{\eta}' = \underline{t}(\underline{e}, \underline{x})) \ ; \tag{25}$$

this gives as a result a formula $\tilde{\phi}(\underline{x}, \underline{\eta}')$ that we put in DNF. A disjunct of ϕ will have the form $\phi_1(\underline{x}) \wedge \phi_2(\underline{\eta}', \underline{t}'(\underline{x}))$ after separation of the literals of the first and of the second kind. We pick such a disjunct $\phi_1(\underline{x}) \wedge \phi_2(\underline{\eta}', \underline{t}'(\underline{x}))$ of the DNF of $\tilde{\phi}(\underline{x}, \underline{\eta}')$ and update our current primitive formula to

$$\exists \underline{\xi}' \, (\underline{\xi}' = \underline{t}'(\underline{x}) \wedge (\exists \underline{\eta}\, \exists \underline{\eta}' \, (\phi_1(\underline{x}) \wedge \phi_2(\underline{\eta}', \underline{\xi}') \wedge \psi(\underline{\eta}, \underline{\xi}, \underline{\eta}')))) \tag{26}$$

(this step is nondeterministic: in the end we shall output the disjunction of all possible outcomes). Here again the $\underline{\xi}'$ are fresh variables abstracting out the terms $\underline{t}'$. Notice that, according to the definition of a tame combination, $\phi_2(\underline{\eta}', \underline{\xi}')$ must be a conjunction of equalities and disequalities between variable terms, because it is a Σ_1-formula (it comes from a T_1-cover computation) and $\underline{\eta}', \underline{\xi}'$ are variables of Σ_2-sorts.

(iii) We apply the cover algorithm of T_2 to the formula

$$\exists \underline{\eta}\, \exists \underline{\eta}'\, (\phi_2(\underline{\eta}', \underline{\xi}') \wedge \psi(\underline{\eta}, \underline{\xi}, \underline{\eta}')) \tag{27}$$

this gives as a result a formula $\psi_1(\underline{\xi}, \underline{\xi}')$. We update our current formula to

$$\exists \underline{\xi}'\, (\underline{\xi}' = \underline{t}'(\underline{x}) \wedge \phi_1(\underline{x}) \wedge \psi_1(\underline{\xi}, \underline{\xi}'))$$

and finally to the equivalent quantifier-free formula

$$\phi_1(\underline{x}) \wedge \psi_1(\underline{\xi}, \underline{t}'(\underline{x})) \ . \tag{28}$$

We now show that the above algorithm is correct under very mild hypotheses. We need some technical facts about stably infinite theories in a multi-sorted context. We say that a multi-sorted theory T is *stably infinite with respect to a set of sorts $\mathcal{S}$ from its signature* iff every T-satisfiable constraint is satisfiable in a model $\mathcal{M}$ where, for every $S \in \mathcal{S}$, the set $S^{\mathcal{M}}$ (namely the interpretation of the sort S in $\mathcal{M}$) is infinite. The next Lemma is a light generalization of Lemma 2 and is proved in the same way (the proof is reported in Appendix A.5):

Lemma 6. *Let T be stably infinite with respect to a subset $\mathcal{S}$ of the set of sorts of the signature of T. Let $\mathcal{M}$ be a model of T and let, for every $S \in \mathcal{S}$, X_S be a superset of $S^{\mathcal{M}}$. Then there is an extension $\mathcal{N}$ of $\mathcal{M}$ such that for all $S \in \mathcal{S}$ we have $S^{\mathcal{N}} \supseteq X_S$.* ◁

Lemma 7. *Let T_1, T_2 be universal signature disjoint theories which are stably infinite with respect to the set of shared sorts (we let Σ_1 be the signature of T_1 and Σ_2 be the signature of T_2). Let $\mathcal{M}_0$ be model of $T_1 \cup T_2$ and let $\mathcal{M}_1$ be a model of T_1 extending the Σ_1-reduct of $\mathcal{M}_0$. Then there exists a model $\mathcal{N}$ of $T_1 \cup T_2$, extending $\mathcal{M}_0$ as a $\Sigma_1 \cup \Sigma_2$-structure and whose Σ_1-reduct extends $\mathcal{M}_1$.* ◁

Proof. Using the previous lemma, build a chain of models $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \mathcal{M}_2 \subseteq \cdots$ such that for all i, $\mathcal{M}_{2i}$ is a model of T_2, $\mathcal{M}_{2i+1}$ is a model of T_1 and $\mathcal{M}_{2i+2}$ is a Σ_2-extension of $\mathcal{M}_{2i}$, whereas $\mathcal{M}_{2i+3}$ is a Σ_2-extension of $\mathcal{M}_{2i+1}$. The union over this chain of models will be the desired $\mathcal{N}$. ⊣

We are now ready for the main result of this section:

Theorem 7. *Let $T_1 \cup T_2$ be a tame combination of two universal theories admitting a model completion. If T_1, T_2 are also stably infinite with repect to their shared sorts, then $T_1 \cup T_2$ has a model completion. Covers in $T_1 \cup T_2$ can be computed as shown in the above three-steps algorithm.* ◁

Proof. Since condition (i) of Lemma 1 is trivially true, we need only to check condition (ii), namely that given a $T_1 \cup T_2$-model $\mathcal{M}$ and elements $\underline{a}, \underline{b}$ from its support such that $\mathcal{M} \models \phi_1(\underline{a}) \wedge \psi_1(\underline{b}, \underline{t}'(\underline{a}))$ as in (28), then there is an extension $\mathcal{N}$ of $\mathcal{M}$ such that (23) is true in $\mathcal{N}$ when evaluating $\underline{x}$ over $\underline{a}$ and $\underline{\xi}$ over $\underline{b}$.

If we let $\underline{b}'$ be the tuple such that $\mathcal{M} \models \underline{b}' = \underline{t}'(\underline{a})$, then we have $\mathcal{M} \models \underline{b}' = \underline{t}'(\underline{a}) \wedge \phi_1(\underline{a}) \wedge \psi_1(\underline{b}, \underline{b}')$. Since $\psi_1(\underline{\xi}, \underline{\xi}')$ is the T_2-cover of (27), the Σ_2-reduct of $\mathcal{M}$ embeds into a T_2-model where (27) is true under the evaluation of the $\underline{\xi}$ as the $\underline{b}$. By Lemma 7, this model can be embedded into a $T_1 \cup T_2$-model $\mathcal{M}'$ in such a way that $\mathcal{M}'$ is an extension of $\mathcal{M}$ and that $\mathcal{M}' \models \underline{b}' = \underline{t}'(\underline{a}) \wedge \phi_1(\underline{a}) \wedge \phi_2(\underline{c}', \underline{b}') \wedge \psi(\underline{c}, \underline{b}, \underline{c}')$ for some $\underline{c}, \underline{c}'$. Since $\phi_1(\underline{x}) \wedge \phi_2(\underline{\eta}', \underline{t}'(\underline{x}))$ implies the T_1-cover of (25) and $\mathcal{M}' \models \phi_1(\underline{a}) \wedge \phi_2(\underline{c}', \underline{t}(\underline{a}))$, then the Σ_1-reduct of $\mathcal{M}'$ can be expanded to a T_1-model where (25) is true when evaluating the $\underline{x}, \underline{\eta}'$ to the $\underline{a}, \underline{c}'$. Again by Lemma 7, this model can be expanded to a $T_1 \cup T_2$-model $\mathcal{N}$ such that $\mathcal{N}$ is an extension of $\mathcal{M}'$ (hence also of $\mathcal{M}$) and $\mathcal{N} \models \phi(\underline{a}', \underline{a}) \wedge \underline{c}' = \underline{t}(\underline{a}', \underline{a}) \wedge \psi(\underline{c}, \underline{b}, \underline{c}')$, that is $\mathcal{N} \models \phi(\underline{a}', \underline{a}) \wedge \psi(\underline{c}, \underline{b}, \underline{t}(\underline{a}', \underline{a}))$. This means that $\mathcal{N} \models \exists \underline{e}\, \exists \underline{\eta} (\phi(\underline{e}, \underline{a}) \wedge \psi(\underline{\eta}, \underline{b}, \underline{t}(\underline{e}, \underline{a})))$, as desired. $\dashv$

9 Conclusions and Future Work

In this paper we showed that covers (aka uniform interpolants) exist in the combination of two convex universal theories over disjoint signatures in case they exist in the component theories and in case the component theories also satisfy the equality interpolating condition - this further condition is nevertheless needed in order to transfer the existence of (ordinary) quantifier-free interpolants. In order to prove that, Beth definability property for primitive fragments turned out to be the crucial ingredient to extensively employ. In case convexity fails, we showed by a counterexample that covers might not exist anymore in the combined theory. The last result raises the following research problem. Even if in general covers do not exist for combination of non-convex theories, it would be interesting to see under what conditions one can decide whether a given cover exists and, in the affirmative case, to compute it.

Applications suggested a different line of investigations, i.e., what we called 'tame combinations'. In database-driven verification [6],[5],[3] one uses tame combinations $T_1 \cup T_2$, where T_1 is a multi-sorted version of $\mathcal{EUF}(\Sigma)$ in a signature Σ containing only unary function symbols and relations of any arity. In this context, quantifier elimination in T_1^* for primitive formulae is quadratic in complexity. Model-checkers like MCMT represent sets of reachable states by using conjunctions of literals and primitive formulae to which quantifier elimination should be applied arise from preimage computations. Now, in this context, if all relation symbols are at most binary, then quantifier elimination in T_1^* produces conjunctions of literals out of primitive formulae, thus step (ii) in the above algorithm becomes deterministic and the only reason why the algorithm may become expensive (i.e. non polynomial) lies in the final quantifier elimination step for T_2^*. The latter might be extremely expensive if substantial arithmetic is involved, but it might still be efficiently handled in practical cases where only very limited arithmetic (e.g. difference bound constraints like $x - y \leq n$ or $x \leq n$, where n is a constant) is involved. This is why we feel that our algorithm for covers in tame combinations can be really useful in the applications.This is confirmed by our first experiments with version 2.9 of MCMT, where such algorithm has been implemented.

A final future research line could consider cover transfer properties to non-disjoint signatures combinations, analogously to similar results obtained in [13] for the transfer of quantifier-free interpolation. Indeed, the main challenge here seems to consist in finding sufficient condition for existence of covers in combination of non-convex theories: in fact, we know from Section 7 that the non-convex version of the equality interpolation property [2] is not enough for this purpose.

References

1. M. Bílková. Uniform interpolation and propositional quantifiers in modal logics. *Studia Logica*, 85(1):1–31, 2007.
2. R. Bruttomesso, S. Ghilardi, and S. Ranise. Quantifier-free interpolation in combinations of equality interpolating theories. *ACM Trans. Comput. Log.*, 15(1):5:1–5:34, 2014.
3. D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. SMT-based verification of data-aware processes: a model-theoretic approach. *Mathematical Structures in Computer Science*. To appear.
4. D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. Quantifier elimination for database driven verification. Technical Report arXiv:1806.09686, arXiv.org, 2018.
5. D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. Formal modeling and SMT-based parameterized verification of data-aware BPMN. In *Proc. of BPM*, volume 11675 of *LNCS*. Springer, 2019.
6. D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. From model completeness to verification of data aware processes. In *Description Logic, Theory Combination, and All That*, volume 11560 of *LNCS*. Springer, 2019.
7. D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. Model completeness, covers and superposition. In *Proc. of CADE*, volume 11716 of *LNCS (LNAI)*. Springer, 2019.
8. D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. Verification of data-aware processes: Challenges and opportunities for automated reasoning. In *Proc. of ARCADE*. EPTCS, 2019.
9. D. Calvanese, S. Ghilardi, A. Gianola, M. Montali, and A. Rivkin. Combined covers and Beth definability. In *Proc. of IJCAR*, LNCS (LNAI). Springer, 2020.
10. C.-C. Chang and J. H. Keisler. *Model Theory*. North-Holland Publishing Co., Amsterdam-London, third edition, 1990.
11. S. Ghilardi. An algebraic theory of normal forms. *Ann. Pure Appl. Logic*, 71(3):189–245, 1995.
12. S. Ghilardi. Model theoretic methods in combined constraint satisfiability. *J. Autom. Reasoning*, 33(3-4):221–249, 2004.
13. S. Ghilardi and A. Gianola. Modularity results for interpolation, amalgamation and superamalgamation. *Annals of Pure and Applied Logic*, 169(8):731–754, 2018.
14. S. Ghilardi, A. Gianola, and D. Kapur. Compactly representing uniform interpolants for EUF using (conditional) DAGS. Technical Report arXiv:2002.09784, arXiv.org, 2020.
15. S. Ghilardi and S. Ranise. Backward reachability of array-based systems by SMT solving: Termination and invariant synthesis. *Logical Methods in Computer Science*, 6(4), 2010.
16. S. Ghilardi and S. Ranise. MCMT: A model checker modulo theories. In *Proc. of IJCAR*, pages 22–29, 2010.
17. S. Ghilardi and M. Zawadowski. *Sheaves, games, and model completions*, volume 14 of *Trends in Logic—Studia Logica Library*. Kluwer Academic Publishers, Dordrecht, 2002. A categorical approach to nonclassical propositional logics.
18. S. Ghilardi and M. W. Zawadowski. A sheaf representation and duality for finitely presenting heyting algebras. *J. Symb. Log.*, 60(3):911–939, 1995.
19. S. Ghilardi and M. W. Zawadowski. Undefinability of propositional quantifiers in the modal system S4. *Studia Logica*, 55(2):259–271, 1995.
20. S. Ghilardi and M. W. Zawadowski. Model completions, r-Heyting categories. *Ann. Pure Appl. Logic*, 88(1):27–46, 1997.
21. S. Gulwani and M. Musuvathi. Cover algorithms and their combination. In *Proc. of ESOP, Held as Part of ETAPS*, pages 193–207, 2008.

22. D. Kapur. Nonlinear polynomials, interpolants and invariant generation for system analysis. In *Proc. of the 2nd International Workshop on Satisfiability Checking and Symbolic Computation co-located with ISSAC*, 2017.
23. T. Kowalski and G. Metcalfe. Uniform interpolation and coherence. *Ann. Pure Appl. Logic*, 170(7):825–841, 2019.
24. G. Nelson and D. C. Oppen. Simplification by cooperating decision procedures. *ACM Trans. Program. Lang. Syst.*, 1(2):245–257, 1979.
25. D. Peuter and V. Sofronie-Stokkermans. On invariant synthesis for parametric systems. In *Automated Deduction - CADE 27 - 27th International Conference on Automated Deduction, Natal, Brazil, August 27-30, 2019, Proceedings*, pages 385–405, 2019.
26. A. M. Pitts. On an interpretation of second order quantification in first order intuitionistic propositional logic. *J. Symb. Log.*, 57(1):33–52, 1992.
27. K. Segerberg. *An Essay in Classical Modal Logic*, volume 13 of *Filosofiska Studier*. Uppsala Universitet, 1971.
28. V. Shavrukov. Subalgebras of diagonalizable algebras of theories containing arithmetic. *Dissertationes Mathematicae*, CCCXXIII, 1993.
29. V. Sofronie-Stokkermans. On interpolation and symbol elimination in theory extensions. *Logical Methods in Computer Science*, 14(3), 2018.
30. C. Tinelli and M. T. Harandi. A new correctness proof of the Nelson-Oppen combination procedure. In *Proc. of FroCoS 1996*, pages 103–119, 1996.
31. S. J. van Gool, G. Metcalfe, and C. Tsinakis. Uniform interpolation and compact congruences. *Ann. Pure Appl. Logic*, 168(10):1927–1948, 2017.
32. A. Visser. Uniform interpolation and layered bisimulation. In P. Hájek, editor, *Gödel 96: Logical foundations on mathematics, computer science and physics – Kurt Gödel's legacy*. Springer Verlag, 1996.
33. G. Yorsh and M. Musuvathi. A combination method for generating interpolants. In *Proc. of CADE-20*, LNCS, pages 353–368. 2005.

A Appendix

In this Appendix we report the proof of the Cover-by-Extensions Lemma 1, of the technical Lemmas 2 and 3 and fill the missing details of the proof of Lemma 5.

A.1 Proof of Lemma 1

The Cover-by-Extension Lemma is not an original result of this paper: the proof is reported here from [7] just for the sake of completeness (the Lemma is crucial in the present paper too).

Lemma 1 [Cover-by-Extensions] *A formula $\psi(\underline{y})$ is a T-cover of $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$ iff it satisfies the following two conditions: (i) $T \models \forall \underline{y}\, (\exists \underline{e}\, \phi(\underline{e}, \underline{y}) \to \psi(\underline{y}))$; (ii) for every model $\mathcal{M}$ of T, for every tuple of elements $\underline{a}$ from the support of $\mathcal{M}$ such that $\mathcal{M} \models \psi(\underline{a})$ it is possible to find another model $\mathcal{N}$ of T such that $\mathcal{M}$ embeds into $\mathcal{N}$ and $\mathcal{N} \models \exists \underline{e}\, \phi(\underline{e}, \underline{a})$.*

Proof. Suppose that $\psi(\underline{y})$ satisfies conditions (i) and (ii) above. Condition (i) says that $\psi(\underline{y}) \in Res(\exists \underline{e}\, \phi)$, so ψ is a residue. In order to show that ψ is also a cover, we have to prove that $T \models \forall \underline{y}, \underline{z}(\psi(\underline{y}) \to \theta(\underline{y}, \underline{z}))$, for every $\theta(\underline{y}, \underline{z})$ that is a residue for $\exists \underline{e}\, \phi(\underline{e}, \underline{y})$. Given a model $\mathcal{M}$ of T, take a pair of tuples $\underline{a}, \underline{b}$ of elements from $|\mathcal{M}|$ and suppose that $\mathcal{M} \models \psi(\underline{a})$. By condition (ii), there is a model $\mathcal{N}$ of T such that $\mathcal{M}$ embeds into $\mathcal{N}$ and $\mathcal{N} \models \exists \underline{e} \phi(\underline{e}, \underline{a})$. Using the definition of $Res(\exists \underline{e}\, \phi)$, we have $\mathcal{N} \models \theta(\underline{a}, \underline{b})$, since $\theta(\underline{y}, \underline{z}) \in Res(\exists \underline{x}\, \phi)$. Since $\mathcal{M}$ is a substructure of $\mathcal{N}$ and θ is quantifier-free, $\mathcal{M} \models \theta(\underline{a}, \underline{b})$ as well, as required.

Suppose that $\psi(\underline{y})$ is a cover. The definition of residue implies condition (i). To show condition (ii) we have to prove that, given a model $\mathcal{M}$ of T, for every tuple $\underline{a}$ of elements from $|\mathcal{M}|$, if $\mathcal{M} \models \psi(\underline{a})$, then there exists a model $\mathcal{N}$ of T such that $\mathcal{M}$ embeds into $\mathcal{N}$ and $\mathcal{N} \models \exists \underline{e} \phi(\underline{e}, \underline{a})$. By reduction to absurdity, suppose that this is not the case: this is equivalent (by using Robinson Diagram Lemma) to the fact that $\Delta(\mathcal{M}) \cup \{\phi(\underline{e}, \underline{a})\}$ is a T-inconsistent $\Sigma^{|\mathcal{M}| \cup \{\underline{e}\}}$-theory. By compactness, there is a finite number of literals $\ell_1(\underline{a}, \underline{b}), ..., \ell_m(\underline{a}, \underline{b})$ (for some tuple $\underline{b}$ of elements from $|\mathcal{M}|$) such that $\mathcal{M} \models \ell_i$ (for all $i = 1, \dots, m$) and $T \models \phi(\underline{e}, \underline{a}) \to \neg(\ell_1(\underline{a}, \underline{b}) \wedge \cdots \wedge \ell_m(\underline{a}, \underline{b}))$, which means that $T \models \phi(\underline{e}, \underline{y}) \to (\neg \ell_1(\underline{y}, \underline{z}) \vee \cdots \vee \neg \ell_m(\underline{y}, \underline{z}))$, i.e. that $T \models \exists \underline{e}\, \phi(\underline{e}, \underline{y}) \to (\neg \ell_1(\underline{y}, \underline{z}) \vee \cdots \vee \neg \ell_m(\underline{y}, \underline{z}))$. By definition of residue, clearly $(\neg \ell_1(\underline{y}, \underline{z}) \vee \cdots \vee \neg \ell_m(\underline{y}, \underline{z})) \in Res(\exists \underline{x}\, \phi)$; then, since $\psi(\underline{y})$ is a cover, $T \models \psi(\underline{y}) \to (\neg \ell_1(\underline{y}, \underline{z}) \vee \cdots \vee \neg \ell_m(\underline{y}, \underline{z}))$, which implies that $\mathcal{M} \models \neg \ell_j(\underline{a}, \underline{b})$ for some $j = 1, \dots, m$, which is a contradiction. Thus, $\psi(\underline{y})$ satisfies conditions (ii) too. $\dashv$

A.2 Proof of Lemma 2

Lemma 2 *If T is stably infinite, then every finite or countable model $\mathcal{M}$ of T can be embedded in a model $\mathcal{N}$ of T such that $|\mathcal{N}| \setminus |\mathcal{M}|$ is countable.*

Proof. Consider $T \cup \Delta(\mathcal{M}) \cup \{c_i \neq a \mid a \in |\mathcal{M}|\}_i \cup \{c_i \neq c_j\}_{i \neq j}$, where $\{c_i\}_i$ is a countable set of fresh constants: by the Diagram Lemma and the downward Löwenheim-Skolem theorem [10], it is sufficient to show that this set is consistent. Suppose not; then by compactness $T \cup \Delta_0 \cup \Delta_1 \cup \{c_i \neq c_j\}_{i \neq j}$ is not satisfiable, for a finite subset Δ_0

of $\Delta(\mathcal{M})$ and a finite subset Δ_1 of $\{c_i \neq a \mid a \in |\mathcal{M}|\}_i$. However, this is a contradiction because by stable infiniteness Δ_0 (being satisfiable in $\mathcal{M}$) is satisfiable in an infinite model of T. $\dashv$

A.3 Proof of Lemma 3

In order to prove Lemma 3, we need further background from [2] concerning amalgamation and strong amalgamation.

Definition 2. *A universal theory T has the* amalgamation property *iff whenever we are given models $\mathcal{M}_1$ and $\mathcal{M}_2$ of T and a common substructure $\mathcal{M}_0$ of them, there exists a further model $\mathcal{M}$ of T endowed with embeddings $\mu_1 : \mathcal{M}_1 \longrightarrow \mathcal{M}$ and $\mu_2 : \mathcal{M}_2 \longrightarrow \mathcal{M}$ whose restrictions to $|\mathcal{M}_0|$ coincide.*

A universal theory T has the strong amalgamation property *if the above embeddings μ_1, μ_2 and the above model $\mathcal{M}$ can be chosen so as to satisfy the following additional condition: if for some m_1, m_2 we have $\mu_1(m_1) = \mu_2(m_2)$, then there exists an element a in $|\mathcal{M}_0|$ such that $m_1 = a = m_2$.* $\triangleleft$

Amalgamation and strong amalgamation are strictly related to quantifier-free interpolation and to combined quantifier-free interpolation, as the result below show:

Theorem 8. *[2] The following two conditions are equivalent for a convex universal theory T: (i) T is equality interpolating and has quantifier-free interpolation; (ii) T has the strong amalgamation property.* $\triangleleft$

Proof. For the sake of completeness, we report the proof of the implication (i) $\Rightarrow$ (ii) (this is the only fact used in the paper). Suppose that T is equality interpolating and has quantifier-free interpolation; we prove that it is strongly amalgamable. If the latter property fails, by Robinson Diagram Lemma, there exist models $\mathcal{M}_1, \mathcal{M}_2$ of T together with a shared submodel $\mathcal{A}$ such that the set of sentences

$$\Delta_\Sigma(\mathcal{M}_1) \cup \Delta_\Sigma(\mathcal{M}_2) \cup \{m_1 \neq m_2 \mid m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|,\ m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|\}$$

is not T-consistent. By compactness, the sentence

$$\delta_1(\underline{a}, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{m}_2) \to \bigvee_{n_1 \in \underline{m}_1, n_2 \in \underline{m}_2} n_1 = n_2$$

is T-valid, for some tuples $\underline{a} \subseteq |\mathcal{A}|$, $\underline{m}_1 \subseteq (|\mathcal{M}_1| \setminus |\mathcal{A}|)$, $\underline{m}_2 \subseteq (|\mathcal{M}_2| \setminus |\mathcal{A}|)$ and for some ground formulae $\delta_1(\underline{a}, \underline{m}_1), \delta_2(\underline{a}, \underline{m}_2)$ true in $\mathcal{M}_1, \mathcal{M}_2$, respectively. If the disjunction is empty, we get $T \models \delta_1(\underline{a}, \underline{m}_1) \to \neg\delta_2(\underline{a}, \underline{m}_2)$ and then we get a contradiction by the quantifier-free interpolation property (the argument is the same as below). Otherwise, by convexity, there are $n_1 \in \underline{m}_1, n_2 \in \underline{m}_2$ such that

$$\delta_1(\underline{a}, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{m}_2) \to n_1 = n_2$$

is T-valid. By the equality interpolating property, there is a term $t(\underline{a})$ such that

$$\delta_1(\underline{a}, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{m}_2) \to n_1 = t(\underline{a})$$

is T-valid. By the quantifier-free interpolation property, there is a quantifier-free formula $\theta(\underline{a})$ such that

$$\delta_1(\underline{a}, \underline{m}_1) \wedge n_1 \neq t(\underline{a}) \to \theta(\underline{a})$$

and

$$\theta(\underline{a}) \to \neg\delta_2(\underline{a}, \underline{m}_2)$$

are both T-valid. Since $n_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|$, we have that $n_1 \neq t(\underline{a})$ is true in $\mathcal{M}_1$. But then we have a contradiction because $\theta(\underline{a})$ is true in $\mathcal{M}_1$, $\mathcal{A}$ and in $\mathcal{M}_2$ as well (truth of quantifier-free formulae moves back and forth via substructures). $\dashv$

We underline that Theorem 8 extends also to the non convex case provided the notion of an equality interpolating theory is suitably adjusted [2].

Let us now come back to the proof of Lemma 3. For proving it, we fixed a convex, stably infinite, equality interpolating, universal theory T admitting a model completion T^* in a signature Σ. We fixed also *a Σ-constraint* $\phi(\underline{x}, \underline{y})$, where we assumed that $\underline{y} = y_1, \dots, y_n$.

Since T has a model completion, it has uniform quantifier-free interpolants by Theorem1, hence it has also (ordinary) quantifier-free interpolants. By Theorem 8 it is strongly amalgamable because it is equality interpolating. In conclusion, *we are allowed to use strong amalgamation in the proof below*.

Recall that for $i = 1, \dots, n$, the formula $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$ was defined as the quantifier-free formula equivalent in T^* to the formula

$$\forall \underline{y} \forall \underline{y}' (\phi(\underline{x}, \underline{y}) \wedge \phi(\underline{x}, \underline{y}') \to y_i = y'_i)$$

where the $\underline{y}'$ are renamed copies of the $\underline{y}$.

Lemma 3 *Suppose that we are given a model $\mathcal{M}$ of T and elements $\underline{a}$ from the support of $\mathcal{M}$ such that $\mathcal{M} \not\models \mathtt{ImplDef}^T_{\phi,y_i}(\underline{a})$ for all $i = 1, \dots, n$. Then there exists an extension $\mathcal{N}$ of $\mathcal{M}$ such that for some $\underline{b} \in |\mathcal{N}| \setminus |\mathcal{M}|$ we have $\mathcal{N} \models \phi(\underline{a}, \underline{b})$.*

Proof. By strong amalgamability, we can freely assume that $\mathcal{M}$ is generated, as a Σ-structure, by the $\underline{a}$: in fact, if we prove the statement for the substructure generated by the $\underline{a}$, then strong amalgamability will provide the model we want.

By using the Robinson Diagram Lemma, what we need is to prove the consistency of $T \cup \Delta(\mathcal{M})$ with the set of ground sentences

$$\{\phi(\underline{a}, \underline{b})\} \cup \{b_i \neq t(\underline{a})\}_{t, b_i}$$

where $t(\underline{x})$ varies over $\Sigma(\underline{x})$-terms, the $\underline{b} = b_1, \dots, b_n$ are fresh constants and i vary over $1, \dots, n$. By convexity,[2] this set is inconsistent iff there exist a term $t(\underline{x})$ and $i = 1, \dots, n$ such that

$$T \cup \Delta(\mathcal{M}) \vdash \phi(\underline{a}, \underline{y}) \to y_i = t \quad .$$

This however implies that $T \cup \Delta(\mathcal{M})$ has the formula

$$\forall \underline{y} \forall \underline{y}' (\phi(\underline{a}, \underline{y}) \wedge \phi(\underline{a}, \underline{y}') \to y_i = y'_i)$$

as a logical consequence. If we now embed $\mathcal{M}$ into a model $\mathcal{N}$ of T^*, we have that $\mathcal{N} \models \mathtt{ImplDef}^T_{\phi,y_i}(\underline{a})$, which is in contrast to $\mathcal{M} \not\models \mathtt{ImplDef}^T_{\phi,y_i}(\underline{a})$ (because $\mathcal{M}$ is a substructure of $\mathcal{N}$ and $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{a})$ is quantifier-free). $\dashv$

The following Lemma will be useful in the next Subsection:

[2] Strictly speaking, convexity says that if, for a set of literals ϕ and for a non empty disjunction of *variables* $\bigvee_{i=1}^n x_i = y_i$, we have $T \models \phi \to \bigvee_{i=1}^n x_i = y_i$, then we have also $T \models \phi \to x_i = y_i$ for some $i = 1, \dots, n$. If, instead of variables, we have *terms*, the same property nevertheless applies: if we have $T \models \phi \to \bigvee_{i=1}^n t_i = u_i$, then for fresh variables x_i, y_i we get $T \models \phi \wedge \bigwedge_{i=1}^n (x_i = t_i \wedge y_i = u_i) \to \bigvee_{i=1}^n x_i = y_i$, etc.

Lemma 8. *Let T have a model completion T^* and let the constraint $\phi(\underline{x},\underline{y})$ be of the kind $\alpha(\underline{x}) \wedge \phi'(\underline{x},\underline{y})$, where $\underline{y} = y_1,\dots,y_n$. Then for every $i = 1,\dots,n$, the formula $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$ is T-equivalent to $\alpha(\underline{x}) \to \mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$.* $\lhd$

Proof. According to (7), the formula $\mathtt{ImplDef}^T_{\phi,y_i}(\underline{x})$ is obtained by eliminating quantifiers in T^* from

$$\forall \underline{y} \forall \underline{y}' (\alpha(\underline{x}) \wedge \phi'(\underline{x},\underline{y}) \wedge \alpha(\underline{x}) \wedge \phi'(\underline{x},\underline{y}') \to y_i = y'_i) \tag{29}$$

The latter is equivalent, modulo logical manipulations, to

$$\alpha(\underline{x}) \to \forall \underline{y} \forall \underline{y}' (\phi'(\underline{x},\underline{y}) \wedge \phi(\underline{x},\underline{y}') \to y_i = y'_i) \tag{30}$$

whence the claim (eliminating quantifiers in T^* from (29) and (30) gives quantifiers-free T^*-equivalent formulae, hence also T-equivalent formulae because T and T^* prove the same quantifier-free formulae). $\dashv$

A.4 Detailed Proof of Lemma 5

Lemma 5 *Every working formula is equivalent (modulo $T_1 \cup T_2$) to a disjunction of terminal working formulae.*

Proof. To compute the required terminal working formulae, it is sufficient to apply the following non-deterministic procedure (the output is the disjunction of all possible outcomes). The non-deterministic procedure applies one of the following alternatives.

(1) Update ψ_1 by adding it a disjunct from the DNF of $\bigwedge_{e_i \in \underline{e}} \neg \mathtt{ImplDef}^{T_1}_{\psi_1,e_i}(\underline{x},\underline{z})$ and ψ_2 by adding to it a disjunct from the DNF of $\bigwedge_{e_i \in \underline{e}} \neg \mathtt{ImplDef}^{T_2}_{\psi_1,e_i}(\underline{x},\underline{z})$;

(2.i) Select $e_i \in \underline{e}$ and $h \in \{1,2\}$; then update ψ_h by adding to it a disjunct L_{ij} from the DNF of $\mathtt{ImplDef}^{T_h}_{\psi_h,e_i}(\underline{x},\underline{z})$; the equality $e_i = t_{ij}$ (where t_{ij} is the term mentioned in Lemma 4)[3] is added to $\mathtt{ExplDef}(\underline{z},\underline{x})$; the variable e_i becomes in this way part of the defined variables.

If alternative (1) is chosen, the procedure stops, otherwise it is recursively applied again and again: we have one truly existential variable less after applying alternative (2.i), so the procedure terminates, since eventually either no truly existential variable remains or alternative (1) is applied. The correctness of the procedure is due to the fact that the following formula is trivially a tautology:

$$\Big(\bigwedge_{e_i \in \underline{e}} \neg \mathtt{ImplDef}^{T_1}_{\psi_1,e_i}(\underline{x},\underline{z}) \wedge \bigwedge_{e_i \in \underline{e}} \neg \mathtt{ImplDef}^{T_2}_{\psi_2,e_i}(\underline{x},\underline{z})\Big) \vee$$
$$\vee \bigvee_{e_i \in \underline{e}} \mathtt{ImplDef}^{T_1}_{\psi_1,e_i}(\underline{x},\underline{z}) \vee \bigvee_{e_i \in \underline{e}} \mathtt{ImplDef}^{T_2}_{\psi_2,e_i}(\underline{x},\underline{z})$$

The first disjunct is used in alternative (1), the other disjuncts in alternative (2.i). At the end of the procedure, we get a terminal working formula. Indeed, if no truly existential variable remains, then the working formula is trivially terminal. It remains to prove that the working formula obtained after applying alternative (1) is indeed terminal. Let ψ'_k (for $k = 1,2$) be the formula obtained from ψ_k after applying alternative (1). We have that ψ'_k is $\alpha(\underline{x},\underline{z}) \wedge \psi_k(\underline{x},\underline{z},\underline{e})$, where α is a disjunct of the DNF of

[3] Lemma 4 is used taking as $\underline{y}$ the tuple $\underline{e}$, as $\underline{x}$ the tuple $\underline{x},\underline{z}$, as $\phi(\underline{x},\underline{y})$ the formula $\psi_h(\underline{x},\underline{z},\underline{e})$ and as ψ the formula ψ_{3-h}.

$\bigwedge_{e_i\in\underline{e}} \neg\mathtt{ImplDef}^{T_k}_{\psi_k,e_i}(\underline{x},\underline{z})$. We need to show that $T_k \vdash \psi'_k \to \neg\mathtt{ImplDef}^{T_k}_{\psi'_k,e_j}(\underline{x},\underline{z})$ for every j. Fix such a j; according to Lemma 8, we must show that

$$T_k \vdash \alpha(\underline{x},\underline{z}) \wedge \psi_k(\underline{x},\underline{z},\underline{e}) \to \neg(\alpha(\underline{x},\underline{z}) \to \mathtt{ImplDef}^{T_k}_{\psi_k,e_j}(\underline{x},\underline{z}))$$

which is indeed the case because $\alpha(\underline{x},\underline{z})$ logically implies $\neg\mathtt{ImplDef}^{T_k}_{\psi'_k,e_j}(\underline{x},\underline{z})$ being a disjunct of the DNF of $\bigwedge_{e_i\in\underline{e}} \neg\mathtt{ImplDef}^{T_k}_{\psi_k,e_i}(\underline{x},\underline{z})$. $\dashv$

A.5 Proof of Lemma 6

Lemma 6 *Let T be stably infinite with respect to a subset $\mathcal{S}$ of the set of sorts of the signature of T. Let $\mathcal{M}$ be a model of T and let, for every $S \in \mathcal{S}$, X_S be a superset of $S^{\mathcal{M}}$. Then there is an extension $\mathcal{N}$ of $\mathcal{M}$ such that for all $S \in \mathcal{S}$ we have $S^{\mathcal{N}} \supseteq X_S$.*

Proof. Let us expand the signature of T with the set C of fresh constants (we take one constant for every $c \in X_S \setminus S^{\mathcal{M}}$). We need to prove the T-consistency of $\Delta(\mathcal{M})$ with a the set D of disequalities asserting that all $c \in C$ are different from each other and from the names of the elements of the support of $\mathcal{M}$. By compactness, it is sufficient to ensure the T-consistency of $\Delta_0 \cup D_0$, where Δ_0 and D_0 are finite subsets of $\Delta(\mathcal{M})$ and D, respectively. Since $\mathcal{M} \models \Delta_0$, this set is T-consistent and hence it is satisfied in a T-model $\mathcal{M}'$ where all the sorts in $\mathcal{S}$ are interpreted as infinite sets; in such $\mathcal{M}'$, it is trivially seen that we can interpret also the constants occurring in D_0 so as to make D_0 true too. $\dashv$