UNIVERSITA' DEGLI STUDI DI MILANO

Dipartimento di Scienze dell'Informazione

RAPPORTO INTERNO N° RI 337-12

From Strong Amalgamability to Modularity of Quantifier-Free Interpolation

Roberto Bruttomesso, Silvio Ghilardi, Silvio Ranise

From Strong Amalgamability to Modularity of Quantifier-Free Interpolation

Roberto Bruttomesso[1] and Silvio Ghilardi[1] and Silvio Ranise[2]
[1]Dipartimento di Scienze dell'Informazione, Università degli Studi di Milano (Italy)
[1]FBK (Fondazione Bruno Kessler), Trento, (Italy)

April 24, 2012

Abstract

The use of interpolants in verification is gaining more and more importance. Since theories used in applications are usually obtained as (disjoint) combinations of simpler theories, it is important to modularly re-use interpolation algorithms for the component theories. We show that a sufficient and necessary condition to do this for quantifier-free interpolation is that the component theories have the 'strong (sub-)amalgamation' property. Then, we provide an equivalent syntactic characterization, identify a sufficient condition, and design a combined quantifier-free interpolation algorithm capable of handling both convex and non-convex theories, that subsumes and extends most existing work on combined interpolation.

Contents

1 Introduction 3

2 Formal Preliminaries 4

3 Strong amalgamation and quantifier-free interpolation 6

3.1 Modularity of quantifier-free interpolation 7

4 Equality interpolation and strong amalgamation 8

4.1 Equality interpolation at work 9

4.2 A comparison with the notion of equality interpolation in [32] 10

5 An interpolation algorithm for combinations of theories 12

5.1 Interpolating metarules 13

5.2 A quantifier-free interpolating algorithm 14

6 Conclusion and Related Work 16

A List of Metarules 20

B Proofs 21

B.1 Proofs for Section 3 21

B.2 Proofs for Section 4 24

B.3 Proofs for Section 5 29

C Quantifier elimination through universal axioms 31

C.1 Integer Linear Arithmetic 31

C.2 Unit-Two-Variable-Per-Inequality 32

D Equality interpolating and Beth definability 34

E Interpolation with free fuction symbols 37

F A counterexample: golden cuff links 40

1 Introduction

Algorithms for computing interpolants are more and more used in verification, e.g., in the abstraction-refinement phase of software model checking [16]. Of particular importance in practice are those algorithms capable of computing *quantifier-free* interpolants in presence of some background theory. Since theories commonly used in verification are obtained as combinations of simpler theories, methods to modularly combine available quantifier-free interpolation algorithms are desirable. This paper studies the modularity of quantifier-free interpolation.

Our starting point is the well-known fact [1] that quantifier-free interpolation (for universal theories) is equivalent to the model-theoretic property of *amalgamability.* Intuitively, a theory has the amalgamation property if any two structures $\mathcal{M}_1, \mathcal{M}_2$ in its class of models sharing a common sub-model $\mathcal{M}_0$ can be regarded as sub-structures of a larger model $\mathcal{M}$, called the amalgamated model. Unfortunately, this property is not sufficient to derive a modularity result for quantifier-free interpolation. As shown in this paper, a stronger notion is needed, called *strong amalgamability* [19], that has been thouroughly analyzed in universal algebra and category theory [21,28]. A theory has the strong amalgamation property if in the amalgamated model $\mathcal{M}$, elements from the supports of $\mathcal{M}_1, \mathcal{M}_2$ not belonging to the support of $\mathcal{M}_0$ *cannot be identified.* An example of an amalgamable but not strongly amalgamable theory is the theory of fields: let $\mathcal{M}_0$ be a real field and $\mathcal{M}_1, \mathcal{M}_2$ be two copies of the complex numbers, the imaginary unit in $\mathcal{M}_1$ must be identified with the imaginary unit of $\mathcal{M}_2$ (or with its opposite) in any amalgamating field $\mathcal{M}$ since the polynomial $x^2 + 1$ cannot have more than two roots (more examples will be discussed below, many examples are also supplied in the catalogue of [21]). We show that *strong amalgamability is precisely what is needed for the modularity of quantifier-free interpolation*, in the following sense (here, for simplicity, we assume that theories are universal although in the paper we generalize to arbitrary ones): (a) if T_1 and T_2 are signature disjoint, both stably infinite and strongly amalgamable, then $T_1 \cup T_2$ is also strongly amalgamable and hence quantifier-free interpolating and (b) a theory T is strongly amalgamable iff the disjoint union of T with the theory $\mathcal{EUF}$ of equality with uninterpreted symbols has quantifier-free interpolation (Section 3). The first two requirements of (a) are those for the correctness of the Nelson-Oppen method [26] whose importance for combined satisfiability problems is well-known.

Since the proof of (a) is non-constructive, the result does not provide an algorithm to compute quantifier-free interpolants in combinations of theories. To overcome this problem, we reformulate the notion of *equality interpolating* theory T in terms of the capability of computing some terms that are equal to the variables occurring in disjunctions of equalities entailed (modulo T) by pairs of quantifier-free formulae and show that equality interpolation is *equivalent* to strong amalgamation (Section 4). To put equality interpolation to productive work, we show that *universal* theories admitting elimination of quantifiers are equality interpolating (Section 4.1). This implies that the theories of recursively defined data structures [27], Integer Difference Logic, Unit-Two-Variable-Per-Inequality, and Integer Linear Arithmetic with division-by-n [6] are all equality interpolating. Our notion of equality interpolation is a strict generalization of the one in [32] so that all the theories that are equality interpolating in the sense of [32] are also so according to our definition, e.g., the theory of LISP structures [26] and Linear Arithmetic over the Reals (Section 4.2). Finally, we describe a combination algorithm for the generation of quantifier-free interpolants from finite sets of quantifier-free formulae in unions of signature disjoint, stably infinite, and equality interpolating theories (Section 5). The algorithm uses as sub-modules the interpolation algorithms of the component theories and is based on a sequence of syntactic manipulations organized in groups of syntactic transformations modelled after a non-deterministic version of the Nelson-Oppen combination schema (see, e.g., [31]). All the proofs are in Appendix B. The other Appendixes contain additional information on related topics, in particular Appendix D connects equality interpolation with Beth definability property, Appendix E investigates interpolation in presence of free function symbols and Appendix F supplies a formal counterexample showing that the convex formulation of the equality interpolation property is insufficient to guarantee combined quantifier-free interpolation for non-convex theories.

2 Formal Preliminaries

We assume the usual syntactic and semantic notions of first-order logic (see, e.g., [12]). The equality symbol "$=$" is included in all signatures considered below. For clarity, we shall use "$\equiv$" in the meta-theory to express the syntactic identity between two symbols or two strings of symbols. Notations like $E(\underline{x})$ means that the expression (term, literal, formula, etc.) E contains free variables only from the tuple $\underline{x}$. A 'tuple of variables' is a list of variables without repetitions and a 'tuple of terms' is a list of terms (possibly with repetitions). Finally, whenever we use a notation like $E(\underline{x}, \underline{y})$ we implicitly assume not only that both the $\underline{x}$ and the $\underline{y}$ are pairwise distinct, but also that $\underline{x}$ and $\underline{y}$ are disjoint. A formula is *universal* (*existential*) iff it is obtained from a quantifier-free formula by prefixing it with a string of universal (existential, resp.) quantifiers.

Theories, elimination of quantifiers, and interpolation. A *theory* T is a pair (Σ, Ax_T), where Σ is a signature and Ax_T is a set of Σ-sentences, called the *axioms* of T (we shall sometimes write directly T for Ax_T). The *models* of T are those Σ-structures in which all the sentences in Ax_T are true. A Σ-formula ϕ is *T-satisfiable* if there exists a model $\mathcal{M}$ of T such that ϕ is true in $\mathcal{M}$ under a suitable assignment $\mathtt{a}$ to the free variables of ϕ (in symbols, $(\mathcal{M}, \mathtt{a}) \models \phi$); it is *$T$-valid* (in symbols, $T \vdash \varphi$) if its negation is T-unsatisfiable or, equivalently, φ is provable from the axioms of T in a complete calculus for first-order logic. A theory $T = (\Sigma, Ax_T)$ is *universal* iff there is a theory $T' = (\Sigma, Ax_{T'})$ such that all sentences in $Ax_{T'}$ are universal and the sets of T-valid and T'-valid sentences coincide. A formula φ_1 *T-entails* a formula φ_2 if $\varphi_1 \to \varphi_2$ is *T-valid* (in symbols, $\varphi_1 \vdash_T \varphi_2$ or simply $\varphi_1 \vdash \varphi_2$ when T is clear from the context). The *satisfiability modulo the theory T ($SMT(T)$) problem* amounts to establishing the T-satisfiability of quantifier-free Σ-formulae.

A theory T admits *quantifier-elimination* iff for every formula $\phi(\underline{x})$ there is a quantifier-free formula $\phi'(\underline{x})$ such that $T \vdash \phi \leftrightarrow \phi'$. A theory T *admits quantifier-free interpolation* (or, equivalently, *has quantifier-free interpolants*) iff for every pair of quantifier-free formulae ϕ, ψ such that $\psi \wedge \phi$ is T-unsatisfiable, there exists a quantifier-free formula θ, called an *interpolant*, such that: (i) ψ T-entails θ, (ii) $\theta \wedge \phi$ is T-unsatisfiable, and (iii) only the variables occurring in both ψ and ϕ occur in θ. A theory admitting quantifier elimination also admits quantifier-free interpolantion; the vice versa does not hold. A more general notion of quantifier-free interpolation property, involving free function symbols, is discussed in Appendix E.

Embeddings, sub-structures, and combinations of theories. The support of a structure $\mathcal{M}$ is denoted with $|\mathcal{M}|$. An embedding is a homomorphism that preserves and reflects relations and operations (see, e.g., [10]). Formally, a *Σ-embedding* (or, simply, an embedding) between two Σ-structures $\mathcal{M}$ and $\mathcal{N}$ is any mapping $\mu : |\mathcal{M}| \longrightarrow |\mathcal{N}|$ satisfying the following three conditions: (a) it is a injective function; (b) it is an algebraic homomorphism, that is for every n-ary function symbol f and for every $a_1, \ldots, a_n \in |\mathcal{M}|$, we have $f^{\mathcal{N}}(\mu(a_1), \ldots, \mu(a_n)) = \mu(f^{\mathcal{M}}(a_1, \ldots, a_n))$; (c) it preserves and reflects interpreted predicates, i.e. for every n-ary predicate symbol P, we have $(a_1, \ldots, a_n) \in P^{\mathcal{M}}$ iff $(\mu(a_1), \ldots, \mu(a_n)) \in P^{\mathcal{N}}$. If $|\mathcal{M}| \subseteq |\mathcal{N}|$ and the embedding $\mu : \mathcal{M} \longrightarrow \mathcal{N}$ 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 a *superstructure* of $\mathcal{M}$. As it is well-known, the truth of a universal (resp. existential) sentence is preserved through substructures (resp. superstructures).

A theory T is *stably infinite* iff every T-satisfiable quantifier-free formula (from the signature of T) is satisfiable in an infinite model of T. By compactness, it is possible to show that T is stably infinite iff every model of T embeds into an infinite one (see [14]). A theory T is *convex* iff for every conjunction of literals δ, if $\delta \vdash_T \bigvee_{i=1}^{n} x_i = y_i$ then $\delta \vdash_T x_i = y_i$ holds for some $i \in \{1, ..., n\}$.

Let T_i be a stably-infinite theory over the signature Σ_i such that the $SMT(T_i)$ problem is decidable for $i = 1, 2$ and Σ_1 and Σ_2 are disjoint (i.e. the only shared symbol is equality). Under these assumptions, the Nelson-Oppen combination method [26] tells us that the SMT problem for the combination $T_1 \cup T_2$ of the theories T_1 and T_2 (i.e. the union of their axioms) is decidable.

3 Strong amalgamation and quantifier-free interpolation

We first generalize the notions of amalgamability and strong amalgamability to arbitrary theories.

Definition 3.1. A theory T has the *sub-amalgamation property* iff whenever we are given models $\mathcal{M}_1$ and $\mathcal{M}_2$ of T and a common substructure $\mathcal{A}$ 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 $|\mathcal{A}|$ coincide.[1]

A theory T has the *strong sub-amalgamation property* if the embeddings μ_1, μ_2 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 $|\mathcal{A}|$ such that $m_1 = a = m_2$.

If the theory T is universal, any substructure of a model of T is also a model of T and we can assume that the substructure $\mathcal{A}$ in the definition above is also a model of T. In this sense, Definition 3.1 introduces generalizations of the standard notions of amalgamability and strong amalgamability for universal theories (see, e.g., [21] for a survey). The result of [1] relating universal theories and quantifier-free interpolation can be easily extended.

Theorem 3.2. *A theory T has the sub-amalgamation property iff it admits quantifier-free interpolants.*

A theory admitting quantifier elimination has the sub-amalgamation property: this follows, e.g., from Theorem 3.2 above. On the other hand, quantifier elimination is not sufficient to guarantee the strong sub-amalgamation property. In fact, from Theorem 3.5 below and the counterexample given in [5], it follows that Presburger arithmetic does not have the strong sub-amalgamation property. However, in Section 4, we shall see that it is sufficient to enrich the signature of Presburger Arithmetic with (integer) division-by-n (for every $n \geq 1$) to have strong amalgamability.

Examples. For any signature Σ, let $\mathcal{EUF}(\Sigma)$ be the pure equality theory over Σ. It is easy to see that $\mathcal{EUF}(\Sigma)$ is universal and has the strong amalgamation property by building a model $\mathcal{M}$ of $\mathcal{EUF}(\Sigma)$ from two models $\mathcal{M}_1$ and $\mathcal{M}_2$ sharing a substructure $\mathcal{M}_0$ as follows. Without

[1] For the results of this paper to be correct, the notion of structure (and of course that of substructure) should encompass the case of structures with empty domains. Readers feeling unconfortable with empty domains can assume that signatures always contain an individual constant.

loss of generality, assume that $|\mathcal{M}_0| = |\mathcal{M}_1| \cap |\mathcal{M}_2|$; let $|\mathcal{M}|$ be $|\mathcal{M}_1| \cup |\mathcal{M}_2|$ and arbitrarily extend the interpretation of the function and predicate symbols to make them total on $|\mathcal{M}|$.

Let us now consider two variants $\mathcal{AX}_{\text{ext}}$ and $\mathcal{AX}_{\texttt{diff}}$ of the theory of arrays considered in [9]. The signatures of $\mathcal{AX}_{\text{ext}}$ and $\mathcal{AX}_{\texttt{diff}}$ contain the sort symbols $\texttt{ARRAY}, \texttt{ELEM}$, and $\texttt{INDEX}$, and the function symbols $rd : \texttt{ARRAY} \times \texttt{INDEX} \longrightarrow \texttt{ELEM}$ and $wr : \texttt{ARRAY} \times \texttt{INDEX} \times \texttt{ELEM} \longrightarrow \texttt{ARRAY}$. The signature of $\mathcal{AX}_{\texttt{diff}}$ also contains the function symbol $\texttt{diff} : \texttt{ARRAY} \times \texttt{ARRAY} \longrightarrow \texttt{INDEX}$. The set $\mathcal{AX}_{\text{ext}}$ of axioms contains the following three sentences:

$$\forall y, i, j, e.\ i \neq j \Rightarrow rd(wr(y, i, e), j) = rd(y, j), \qquad \forall y, i, e.\ rd(wr(y, i, e), i) = e,$$
$$\forall x, y.\ x \neq y \Rightarrow (\exists i.\ rd(x, i) \neq rd(y, i))$$

whereas the set of axioms for $\mathcal{AX}_{\texttt{diff}}$ is obtained from that of $\mathcal{AX}_{\text{ext}}$ by replacing the third axiom with its Skolemization:

$$\forall x, y. \qquad x \neq y \Rightarrow rd(x, \texttt{diff}(x, y)) \neq rd(y, \texttt{diff}(x, y)) \ .$$

In [7] (the extended version of [9]), it is shown that $\mathcal{AX}_{\texttt{diff}}$ has the strong sub-amalgamation property while $\mathcal{AX}_{\text{ext}}$ does not. However $\mathcal{AX}_{\text{ext}}$ (which is *not* universal) enjoys the following property (this is the standard notion of amalgamability from the literatrure): given two models $\mathcal{M}_1$ and $\mathcal{M}_2$ of $\mathcal{AX}_{\text{ext}}$ sharing a substructure $\mathcal{M}_0$ *which is also a model of* $\mathcal{AX}_{\text{ext}}$, there is a model $\mathcal{M}$ of $\mathcal{AX}_{\text{ext}}$ endowed with embeddings from $\mathcal{M}_1, \mathcal{M}_2$ agreeing on the support of $\mathcal{M}_0$.

The application of Theorem 3.2 to $\mathcal{EUF}(\Sigma)$, $\mathcal{AX}_{\texttt{diff}}$, and $\mathcal{AX}_{\text{ext}}$ allows us to derive in a uniform way results about quantifier-free interpolation that are available in the literature: that $\mathcal{EUF}(\Sigma)$ (see, e.g., [13, 24]) and $\mathcal{AX}_{\texttt{diff}}$ [9] admit quantifier-free interpolants, and that $\mathcal{AX}_{\text{ext}}$ does not [20].

3.1 Modularity of quantifier-free interpolation

Given the importance of combining theories in SMT solving, the next step is to establish whether sub-amalgamation is a modular property. Unfortunately, this is not the case since the combination of two theories having quantifier-free interpolation may not have quantifier-free interpolation. For example, the union of the theory $\mathcal{EUF}(\Sigma)$ and Presburger arithmetic does not have quantifier-free interpolation [5]. Fortunately, strong sub-amalgamation is modular when combining stably infinite theories.

Theorem 3.3. *Let T_1 and T_2 be two stably infinite theories over disjoint signatures Σ_1 and Σ_2. If both T_1 and T_2 have the strong sub-amalgamation property, then so does $T_1 \cup T_2$.*

Theorems 3.2 and 3.3 obviously imply that strong sub-amalgamation is sufficient for the modularity of quantifier-free interpolation for stable infinite theories.

Corollary 3.4. *Let T_1 and T_2 be two stably infinite theories over disjoint signatures Σ_1 and Σ_2. If both T_1 and T_2 have the strong sub-amalgamation property, then $T_1 \cup T_2$ admits quantifier-free interpolation.*

We can also show that strong sub-amalgamation is necessary as explained by the following result.

Theorem 3.5. *Let T be a theory admitting quantifier-free interpolation and 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 has the strong sub-amalgamation property.*

Although Corollary 3.4 is already useful to establish whether combinations of theories admit quantifier-free interpolants, proving the strong sub-amalgamability property can be complex. In the next section, we study an alternative ("syntactic") characterization of strong sub-amalgamability that can be more easily applied to commonly used theories.

4 Equality interpolation and strong amalgamation

There is a tight relationship between the strong sub-amalgamation property of a theory T and the fact that disjunctions of equalities among variables are entailed by T. To state this precisely, we need to introduce some preliminary notions. Given two finite tuples $\underline{t} \equiv t_1, \ldots, t_n$ and $\underline{v} \equiv v_1, \ldots, v_m$ of terms,

$$\text{the notation } \underline{t} \cap \underline{v} \neq \emptyset \text{ stands for the formula } \bigvee_{i=1}^{n} \bigvee_{j=1}^{m} (t_i = v_j).$$

We use $\underline{t}_1\underline{t}_2$ to denote the juxtaposition of the two tuples $\underline{t}_1$ and $\underline{t}_2$ of terms. So, for example, $\underline{t}_1\underline{t}_2 \cap \underline{v} \neq \emptyset$ is equivalent to $(\underline{t}_1 \cap \underline{v} \neq \emptyset) \vee (\underline{t}_2 \cap \underline{v} \neq \emptyset)$.

Definition 4.1. A theory T is *equality interpolating* iff it has the quantifier-free interpolation property and satisfies the following condition:

- for every quintuple $\underline{x}, \underline{y}_1, \underline{z}_1, \underline{y}_2, \underline{z}_2$ of tuples of variables and pair of quantifier-free formulae $\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1)$ and $\delta_2(\underline{x}, \underline{z}_2, \underline{y}_2)$ such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{y}_2 \neq \emptyset \tag{1}$$

there exists a tuple $\underline{v}(\underline{x})$ of terms such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \underline{y}_1\underline{y}_2 \cap \underline{v} \neq \emptyset \ . \tag{2}$$

We are now in the position to formally state the equivalence between strong sub-amalgamation and equality interpolating property.

Theorem 4.2. *A theory T has the strong sub-amalgamation property iff it is equality interpolating.*

4.1 Equality interpolation at work

We now illustrate some interesting applications of Theorem 4.2 so that, by using Corollary 3.4, we can establish when combinations of theories admit quantifier-free interpolation. To ease the application of Theorem 4.2, we first study the relationship between quantifier-elimination and equality interpolation for universal theories.

Theorem 4.3. *A universal theory admitting quantifier elimination is equality interpolating.*

Interestingly, the proof of this theorem (see Appendix B.2) is constructive and shows how an available quantifier elimination algorithm (for a universal theory) can be used *to find the terms $\underline{v}$ satisfying condition* (2) *of Definition 4.1*; this is key to the combined interpolation algorithm presented in Section 5 below.

Examples. The theory $\mathcal{RDS}$ of *recursive data structures* [27] consists of two unary function symbols *car* and *cdr* and a binary function symbol *cons*, and it is axiomatized by the following infinite set of sentences:

$$\begin{aligned} \forall x, y.car(cons(x, y)) = x, \quad & \forall x, y.cdr(cons(x, y)) = y, \qquad \text{(CCC)} \\ \forall x, y.cons(car(x), cdr(x)) = x, \quad & \forall x.x \neq t(x) \end{aligned}$$

where t is a term obtained by finitely many applications of *car* and *cdr* to the variable x (e.g., $car(x) \neq x$, $cdr(cdr(x)) \neq x$, $cdr(car(x)) \neq x$, and so on). Clearly, $\mathcal{RDS}$ is universal; the fact that it admits elimination of quantifiers is known since an old work by Mal'cev [17].

Following [12], we define the theory $\mathcal{IDL}$ of *integer difference logic* to be the theory whose signature contains the constant symbol 0, the unary function symbols *succ* and *pred*, and the binary predicate symbol $<$, and which is axiomatized by adding to the irreflexivity, transitivity and linearity axioms for $<$ the following set of sentences:

$$\begin{aligned} &\forall x.succ(pred(x)) = x, && \forall x.pred(succ(x)) = x, \\ &\forall x, y.x < succ(y) \leftrightarrow (x < y \vee x = y), && \forall x, y.pred(x) < y \leftrightarrow (x < y \vee x = y). \end{aligned}$$

$\mathcal{IDL}$ is universal and the fact that admits elimination of quantifiers can be shown by adapting the procedure for a similar theory of natural numbers with successor and ordering in [12]. The key observation is that the atoms of $\mathcal{IDL}$ are equivalent to formulae of the form $i \bowtie f^n(j)$

(for $n \in \mathbb{Z}$, $\bowtie \in \{=, <\}$) where i, j are variables or the constant 0, $f^0(j)$ is j, $f^k(j)$ abbreviates $succ(succ^{k-1}(j))$ when $k > 0$ or $pred(pred^{k-1}(j))$ when $k < 0$. (Usually, $i \bowtie f^n(j)$ is written as $i - j \bowtie n$ or as $i \bowtie j + n$ from which the name of "integer difference logic.")

The theory $\mathcal{LAI}$ of Linear Arithmetic over the Integers contains the binary predicate symbol $<$, the constant symbols 0 and 1, the unary function symbol $-$, the binary function symbol $+$ and the unary function symbols $div[n]$ (integer division by n, for $n > 1$). As axioms, we take a set of sentences such that all true sentences in the standard model of the integers can be derived. This can be achieved for instance by adding to the axioms for totally ordered Abelian groups the following sentences (below $x\ rem[n]$ abbreviates $x - n(x\ div[n])$, moreover kt denotes the sum $t + \cdots + t$ having k addends all equal to the term t and k stands for $k1$):

$$0 < 1, \quad \forall y.\neg(0 < y \wedge y < 1), \quad \text{and} \quad \forall x.x\ rem[n] = 0 \vee \cdots \vee x\ rem[n] = n - 1\,.$$

$\mathcal{LAI}$ can be seen as a variant of Presburger Arithmetic obtained by adding the functions $div[n]$ instead of the 'congruence modulo n' relations (for $n = 1, 2, 3, ...$), which are needed to have quantifier elimination (see, e.g., [12]). For the application of Theorem 4.3, the problem with adding the 'congruence modulo n' is that the resulting theory is not universal. Instead, $\mathcal{LAI}$ is universal and the fact that admits elimination of quantifiers can be derived by adapting existing quantifier-elimination procedures (e.g., the one in [12]) and observing that x is congruent to y modulo n can be defined as $x\ rem[n] = y\ rem[n]$ (more details can be found in Appendix C.1).

By Theorem 4.3, $\mathcal{RDS}$, $\mathcal{IDL}$, and $\mathcal{LAI}$ are equality interpolating. The theory $\mathcal{UTVPI}$ of Unit-Two-Variable-Per-Inequality (see, e.g., [11]) is also equality interpolating (for lack of space, this is shown in Appendix C.2).

4.2 A comparison with the notion of equality interpolation in [32]

We now show that the notion of equality interpolating theories proposed here reduces to that of [32] when considering convex theories.

Proposition 4.4. *A convex theory T having quantifier-free interpolation is equality interpolating iff for every pair y_1, y_2 of variables and for every pair of conjunctions of literals $\delta_1(\underline{x}, \underline{z}_1, y_1), \delta_2(\underline{x}, \underline{z}_2, y_2)$ such that*

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2 \tag{3}$$

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

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v \wedge y_2 = v. \tag{4}$$

The implication (3) $\Rightarrow$ (4) is exactly the definition of equality interpolation in [32]. In the following, a convex quantifier-free interpolating theory satisfying (3) $\Rightarrow$ (4) will be called *YMc equality interpolating.* By Proposition 4.4, an YMc equality interpolating (convex) theory is also equality interpolating according to Definition 4.1. For example, the theory $\mathcal{LST}$ of list structures [26] contains the function symbols of $\mathcal{RDS}$, a unary predicate symbol *atom*, and it is axiomatized by the axioms of $\mathcal{RDS}$ labelled (CCC) and the sentences:

$$\forall x, y. \neg atom(cons(x, y)), \quad \forall x. \neg atom(x) \to cons(car(x), cdr(x)) = x\,.$$

$\mathcal{LST}$ is a (universal) convex theory [26] that was shown to be YMc equality interpolating in [32]. By Proposition 4.4, we conclude that $\mathcal{LST}$ is equality interpolating in the sense of Definition 4.1. In [32], also Linear Arithmetic over the Reals ($\mathcal{LAR}$) is shown to be YMc equality interpolating (the convexity of $\mathcal{LAR}$ is well-known from linear algebra). By Proposition 4.4, $\mathcal{LAR}$ is equality interpolating in the sense of Definition 4.1. The same result can be obtained from Theorem 4.3 above by identifying a set of universal axioms for the theory and showing that they admit quantifier elimination. For the axioms to be universal, it is essential to include *multiplication by rational coefficients* in the signature of the theory, i.e. the unary function symbols $q * \_$ for every $q \in \mathbb{Q}$. If this is not the case, the theory is not sub-amalgamable and thus not equality interpolating: to see this, consider the embedding of the substructure $\mathbb{Z}$ into two copies of the reals. A direct counterexample to (3) $\Rightarrow$ (4) of Proposition 4.4 can be obtained by taking $\delta_i(x, y_i) \equiv y_i + y_i = x$ for $i = 1, 2$ so that $v(x) \equiv \frac{1}{2} * x$ in (4) and the function symbol $\frac{1}{2} * \_$ is required.

For *non-convex* theories, the notion of equality interpolation in this paper is strictly more general than the one proposed in the extended version of [32]. Such a notion, to be called *YM equality interpolating* below, requires quantifier-free interpolation and the following condition:

– for every tuples $\underline{x}$, $\underline{z}_1$, $\underline{z}_2$ of variables, further tuples $\underline{y}_1 = y_{11}, \ldots, y_{1n}$, $\underline{y}_2 = y_{21}, \ldots, y_{2n}$ of variables, and pairs $\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1), \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2)$ of conjunctions of literals,

$$\text{if } \delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \bigvee_{i=1}^{n} (y_{1i} = y_{2i}) \text{ holds,}$$

then there exists a tuple $\underline{v}(\underline{x}) = v_1, \ldots, v_n$ of terms such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \bigvee_{i=1}^{n} (y_{1i} = v_i \wedge v_i = y_{2i}).$$

We show that the notion of YM equality interpolation implies that of equality interpolation proposed in this paper. Indeed, if a convex theory is YMc equality interpolating, then it is also YM equality interpolating. Since $\mathcal{EUF}(\Sigma)$ is convex and YMc equality interpolating (as shown in [32]), it is YM equality interpolating. By Theorems 3.5 and 4.2 (and the combination result of [32]), if a theory T is YM equality interpolating, it is also equality interpolating in the sense of Definition 4.1. The converse does not hold, i.e. our notion is *strictly weaker* than YM equality interpolation. To prove this, we define a (non-convex) theory T_{cex} that has the strong sub-amalgamation property but is not YM equality interpolating. Let the signature of T_{cex} contain three propositional letters p_1, p_2 and p_3, three constant symbols c_1, c_2, and c_3, and a unary predicate Q. T_{cex} is axiomatized by the following sentences: exactly one among p_1, p_2 and p_3 holds, c_1, c_2, and c_3 are distinct, $Q(x)$ holds for no more than one x, and $p_i \rightarrow Q(c_i)$ for $i = 1, 2, 3$. It is easy to see that T_{cex} is stably infinite and has the strong sub-amalgamation property (T_{cex} is non-convex since $Q(x) \wedge y_1 = c_1 \wedge y_2 = c_2 \wedge y_3 = c_3$ implies the disjunction $x = y_1 \vee x = y_2 \vee x = y_3$ without implying any single disjunct). Now, notice that $Q(x) \wedge Q(y) \vdash_{T_{cex}} x = y$. According to the definition of the YM equality interpolating property (see above), there should be a *single* ground term v such that $Q(x) \wedge Q(y) \vdash_{T_{cex}} x = v \wedge y = v$. This cannot be the case since we must choose among one of the three constants c_1, c_2, c_3 to find such a term v and none of these choices fits our purposes. Hence, T_{cex} is not YM equality interpolating although it has the strong sub-amalgamation property and hence it is equality interpolating according to Definition 4.1.

To conclude the comparison with [32], since the notion of equality interpolation of this paper is *strictly weaker* than that of YM equality interpolation, the scope of applicability of our result about the modularity of theories admitting quantifier-free interpolation (i.e. Corollary 3.4 above) is *broader* than the one in the extended version of [32].

5 An interpolation algorithm for combinations of theories

Although the notion of equality interpolation toghether with Corollary 3.4 allow us to establish the quantifier-free interpolation for all those theories obtained by combining a theory axiomatizing a container data structure (such as $\mathcal{EUF}$, $\mathcal{RDS}$, $\mathcal{LST}$, or $\mathcal{AX}_{\texttt{diff}}$) with relevant fragments of Arithmetics (such as $\mathcal{LAR}$, $\mathcal{IDL}$, $\mathcal{UTVPI}$, or $\mathcal{LAI}$), just knowing that quantifier-free interpolants exist may not be sufficient. It would be desirable to compute interpolants for combinations of theories by modularly reusing the available interpolation algorithms for the component theories. This is the subject of this section.

To simplify the technical development, we work with ground formulae over signatures expanded with free constants instead of quantifier free formulae as done in the previous sections. We use the letters $A, B, \ldots$ to denote finite sets of ground formulae; the logical reading of a set of formulae is the conjunction of its elements. For a signature Σ and set A of formulae, Σ^A denotes the signature Σ expanded with the free constants occurring in A. Let A and B be two finite sets of ground formulae in the signatures Σ^A and Σ^B, respectively, and

$\Sigma^C := \Sigma^A \cap \Sigma^B$. Given a term, a literal, or a formula φ we call it:

- *AB-common* iff it is defined over Σ^C;
- *A-local* (resp. *B-local*) if it is defined over Σ^A (resp. Σ^B);
- *A-strict* (resp. *B-strict*) iff it is A-local (resp. B-local) but not AB-common;
- *AB-mixed* if it contains symbols in both $(\Sigma^A \setminus \Sigma^C)$ and $(\Sigma^B \setminus \Sigma^C)$;
- *AB-pure* if it does not contain symbols in both $(\Sigma^A \setminus \Sigma^C)$ and $(\Sigma^B \setminus \Sigma^C)$.

(Sometimes in the literature about interpolation, "A-local" and "B-local" are used to denote what we call here "A-strict" and "B-strict").

5.1 Interpolating metarules

Our combined interpolation method is based on the abstract framework introduced in [9] (to which, the interested reader is pointed for more details) and used also in [8] that is based on 'metarules.' A metarule applies (bottom-up) to a pair A, B of finite sets of ground formulae[2] producing an equisatisfiable pair of sets of formulae. Each metarule comes with a proviso for its applicability and an instruction for the computation of the interpolant. As an example, consider the metarule (Define0):

$$\frac{A \cup \{a = t\} \mid B \cup \{a = t\}}{A \mid B} \qquad \begin{array}{l} \textit{Proviso}\text{: } t \text{ is } AB\text{-common, } a \text{ is fresh} \\ \textit{Instruction}\text{: } \phi' \equiv \phi(t/a). \end{array}$$

It is not difficult to see that the $A \cup B$ is equisatisfiable to $A \cup B \cup \{a = t\}$ since a is a fresh variable that has been introduced to re-name the AB-common term t according to the proviso of (Define0). The instruction attached to (Define0) allows for the computation of the interpolant ϕ' by eliminating the fresh constant a from the recursively known interpolant ϕ. The idea is to build an *interpolating metarules refutation* for a given unsatisfiable $A_0 \cup B_0$, i.e. a labeled tree having the following properties: (i) nodes are labeled by pairs of finite sets of ground formulae; (ii) the root is labeled by A_0, B_0; (iii) the leaves are labeled by a pair $\tilde{A}, \tilde{B}$ such that $\bot \in \tilde{A} \cup \tilde{B}$; (iv) each non-leaf node is the conclusion of a metarule and its successors are the premises of that metarule (the complete list of metarules is in Appendix A). Once an interpolating metarules refutation has been built, it is possible to recursively compute the interpolant by using (top-down) the instructions attached to the metarules in the tree:

[2] In [8, 9], metarules manipulate pairs of finite sets of literals instead of ground formulae; the difference is immaterial.

Proposition 5.1 ([9]). *If there exists an interpolating metarules refutation for A_0, B_0 then there is a quantifier-free interpolant for A_0, B_0 (i.e., there exists a quantifier-free AB-common sentence ϕ such that $A_0 \vdash \phi$ and $B_0 \wedge \phi \vdash \bot$). The interpolant ϕ is recursively computed by applying the relevant interpolating instructions of the metarules.*

The idea to design the combination algorithm is the following. We design transformations instructions that can be non-deterministically applied to a pair A_0, B_0. Each of the transformation instructions is *justified by metarules*, in the sense that it is just a special sequence of applications of metarules. The instructions are such that, whenever they are applied exhaustively to a pair such that $A_0 \cup B_0$ is unsatisfiable, they produce a tree which is an interpolating metarules refutation for A_0, B_0 from which an interpolant can be extracted according to Proposition 5.1.

5.2 A quantifier-free interpolating algorithm

Let T_i be a stably-infinite and equality interpolating theory over the signature Σ_i such that the $SMT(T_i)$ problem is decidable and $\Sigma_1 \cap \Sigma_2 = \emptyset$ (for $i = 1, 2$). We assume the availability of algorithms for T_1 and T_2 that are able not only to compute quantifier-free interpolants but also the tuples $\underline{v}$ of terms in Definition 4.1 for equality interpolation. Since the $SMT(T_i)$ problem is decidable for $i = 1, 2$, it is always possible to build an equality interpolating algorithm by enumeration; in practice, better algorithms can be designed (see [32] for $\mathcal{EUF}$, $\mathcal{LST}$, $\mathcal{LAR}$ and Appendix B for the possibility to use quantifier elimination to this aim).

Let $\Sigma := \Sigma_1 \cup \Sigma_2$, $T := T_1 \cup T_2$, and A_0, B_0 be a T-unsatisfiable pair of finite sets of ground formulae over the signature $\Sigma^{A_0 \cup B_0}$. Like in the Nelson-Oppen combination method, we have a pre-processing step in which we purify A_0 and B_0 so as to eliminate from them the literals which are neither Σ_1- nor Σ_2-literals. To do this, it is sufficient to repeatedly apply the technique of "renaming terms by constants" described below. Take a term t (occurring in a literal from A_0 or from B_0), add the equality $a = t$ for a fresh constant a and replace all the occurrences of t by a. The transformation can be justified by the following sequence of metarules: Define1, Define2, Redplus1, Redplus2, Redminus1, Redminus2. For example, in the case of the renaming of some term t in A_0, the metarule Define1 is used to add the explicit definition $a = t$ to A_0, the metarule Redplus1 to add the formula $\phi(a/t)$ for each $\phi \in A_0$, and the metarule Redminus1 to remove from A_0 all the formulæ ϕ in which t occurs (except $a = t$).

Because of purification, from now on, *we assume to manipulate pairs A, B of sets of ground formulæ where literals built up of only Σ_1- or of only Σ_2-symbols occur* (besides free constants): this invariant will be in fact maintained during the execution of our algorithm. Given such a pair A, B, we denote by A_1 and A_2 the subsets of Σ_1^A- and Σ_2^A-formulae belonging to A; the sub-sets B_1 and B_2 of B are defined similarly. Notice that *it is false* that $A \equiv A_1 \cup A_2$ and $B \equiv B_1 \cup B_2$, since quantifier-free formulae can mix Σ_1- and Σ_2-symbols even if the literals they are built from do not.

Before presenting our interpolation algorithm for the combination of theories, we need to import a technique, called *Term Sharing*, from [9]. Suppose that A contains a literal $a = t$, where the term t is AB-common and the free constant a is A-strict (a symmetric technique applies to B istead of A). Then it is possible to "make a AB-common" in the following way. First, introduce a fresh AB-common constant c with the explicit definition $c = t$ (to be inserted both in A and in B, as justified by metarule (Define0)); then replace the literal $a = t$ by $a = c$ and replace a by c everywhere else in A; finally, delete $a = c$ too. The result is a pair (A, B) where basically nothing has changed but a has been renamed to an AB-common constant c (the transformation can be easily justified by a suitable subset of the metarules).

An A-*relevant atom* is either an atomic formula occurring in A or it is an A-local equality between free constants; an A-*assignment* is a Boolean assignment α to relevant A-atoms satisfying A, seen as a set of propositional formulæ (relevant B-atoms and B-assignements are defined similarly). Below, we use the notation α to denote both the assignement α and the set of literals $\{L \,|\, \alpha(L) = \mathit{true}\}$.

We are now in the position to present the collection of transformations that should be applied non-deterministically and exhaustively to a pair of purified sets of ground formulæ (all the transformations below can be justified by metarules, the justification is straightforward and left to the reader). In the following, let $i \in \{1, 2\}$ and $X \in \{A, B\}$.

Terminate$_i$: if $A_i \cup B_i$ is T_i-unsatisfiable and $\bot \notin A \cup B$, use the interpolation algorithm for T_i to find a ground AB-common θ such that $A_i \vdash_{T_i} \theta$ and $\theta \wedge B_i \vdash_{T_i} \bot$; then add θ and $\bot$ to B.

Decide$_X$: if there is no X-assignment α such that $\alpha \subseteq X$, pick one of them (if there are none, add $\bot$ to X); then update X to $X \cup \alpha$.

Share$_i$: let $\underline{a} = a_1, \ldots, a_n$ be the tuple of the current A-strict free constants and $\underline{b} = b_1, \ldots, b_m$ be the tuple of the current B-strict free constants. Suppose that $A_i \cup B_i$ is T_i-satisfiable, but $A_i \cup B_i \cup \{\underline{a} \cap \underline{b} = \emptyset\}$ is T_i-unsatisfiable. Since T_i is equality interpolating, there must exist AB-common Σ_i-ground terms $\underline{v} \equiv v_1, \ldots, v_p$ such that

$$A_i \cup B_i \vdash_{T_i} (\underline{a} \cap \underline{v} \neq \emptyset) \vee (\underline{b} \cap \underline{v} \neq \emptyset).$$

Thus the union of $A_i \cup \{\underline{a} \cap \underline{v} = \emptyset\}$ and of $B_i \cup \{\underline{b} \cap \underline{v} = \emptyset\}$ is not T_i-satisfiable and invoking the available interpolation algorithm for T_i, we can compute a ground AB-

common Σ_i-formula θ such that $A \vdash_{T_i} \theta \vee \underline{a} \cap \underline{v} \neq \emptyset$ and $\theta \wedge B \vdash_{T_i} \underline{b} \cap \underline{v} \neq \emptyset$. We choose among $n * p + m * p$ alternatives in order to non-deterministically update A, B. For the first $n * p$ alternatives, we add some $a_i = v_j$ (for $1 \leq i \leq n$, $1 \leq j \leq p$) to A. For the last $m * p$ alternatives, we add θ to A and some $\{\theta, b_i = v_j\}$ to B (for $1 \leq i \leq m$, $1 \leq j \leq p$). Term sharing is finally applied to the updated pair in order to decrease the number of the A-strict or B-strict free constants.

Let $\mathsf{CI}(T_1, T_2)$ be the procedure that, once run on an unsatifiable pair A_0, B_0, first purifies it, then non-deterministically and exhaustively applies the transformation rules above, and finally extracts an interpolant by using the instructions associated to the metarules.

Theorem 5.2. *Let T_1 and T_2 be two signature disjoint, stably-infinite, and equality interpolating theories having decidable SMT problems. Then, $\mathsf{CI}(T_1, T_2)$ is a quantifier-free interpolation algorithm for the combined theory $T_1 \cup T_2$.*

Algorithm $\mathsf{CI}(T_1, T_2)$ paves the way to reuse quantifier-free interpolation algorithms for both conjunctions (see, e.g., [29]) or arbitrary Boolean combinations of literals (see, e.g., [11]). In particular, the capability of reusing interpolation algorithms that can efficiently handle the Boolean structure of formulae seems to be key to enlarge the scope of applicability of verification methods based on interpolants [23]. Indeed, one major issue to address to make $\mathsf{CI}(T_1, T_2)$ practically usable is to eliminate the non-determinism. We believe this is possible by adapting the Delayed Theory Combination approach [4].

6 Conclusion and Related Work

The results of this paper cover several results for the quantifier-free interpolation of combinations of theories that are known from the literature, e.g., $\mathcal{EUF}$ and $\mathcal{LST}$ [32], $\mathcal{EUF}$ and $\mathcal{LAR}$ [11, 25, 29], $\mathcal{EUF}$ and $\mathcal{LAI}$ [6], $\mathcal{LST}$ with $\mathcal{LAR}$ [32], and $\mathcal{AX}_{\mathtt{diff}}$ with $\mathcal{IDL}$ [8]. To the best of our knowledge, the quantifier-free interpolation of the following combinations are new: (a) $\mathcal{RDS}$ with $\mathcal{LAR}$, $\mathcal{IDL}$, $\mathcal{UTVPI}$, $\mathcal{LAI}$, and $\mathcal{AX}_{\mathtt{diff}}$, (b) $\mathcal{LST}$ with $\mathcal{IDL}$, $\mathcal{UTVPI}$, $\mathcal{LAI}$, and $\mathcal{AX}_{\mathtt{diff}}$, and (c) $\mathcal{AX}_{\mathtt{diff}}$ with $\mathcal{LAR}$, $\mathcal{UTVPI}$, and $\mathcal{LAI}$.

In Section 4.2, we have extensively discussed the closely related work of [32], where the authors illustrate a method to derive interpolants in a Nelson-Oppen combination procedure, provided that the component theories satisfy certain hypotheses. The work in [3], among other contributions, recasts the method of [32] in the context of the $DPLL(T)$ paradigm. An alternative combination method is in [15] that has been designed to be efficiently incorporated in state-of-the-art SMT solvers but is complete only for convex theories. An interpolating theorem prover is described in [25], where a sequent-like calculus is used to derive interpolants from proofs in propositional logic, equality with uninterpreted functions, linear rational arithmetic, and their combinations. The "split" prover in [18] applies a sequent calculus for the synthesis of interpolants along the lines of that in [25] and is tuned for predicate abstraction. The "split" prover can handle combinations of theories involving that of arrays without extensionality and fragments of Linear Arithmetic. The CSIsat [2] permits the computation of quantifier-free interpolants over a combination of $\mathcal{EUF}$ and $\mathcal{LAR}$ refining the combination method in [32]. A version of MathSAT [11] features interpolation capabilities for $\mathcal{EUF}$, $\mathcal{LAR}$, $\mathcal{IDL}$, $\mathcal{UTVPI}$ and $\mathcal{EUF} + \mathcal{LAR}$ by extending Delayed Theory Combination [4]. Theorem 5.2 is the key to combine the strength of these tools and to widen the scope of applicability of available interpolation algorithms to richer combinations of theories. Methods [6, 20, 22, 23] for the computation of quantified interpolants in the combination of the theory of arrays and Presburger Arithmetic have been proposed. Our work focus on quantifier-free interpolants by identifying suitable variants of the component theories (e.g., $\mathcal{AX}_{\texttt{diff}}$ instead of $\mathcal{AX}_{\text{ext}}$ and $\mathcal{LAI}$ instead of Presburger Arithmetic). Orthogonal to our approach is the work in [30] where interpolation algorithm are developed for extensions of convex theories admitting quantifier-free interpolation.

The framework proposed in this paper allows us to give a uniform and coherent view of many results available in the literature and we hope that it will be the starting point for new developements.

References

[1] P. D. Bacsich. Amalgamation properties and interpolation theorems for equational theories. *Algebra Universalis*, 5:45–55, 1975.

[2] D. Beyer, D. Zufferey, and R. Majumdar. CSIsat: Interpolation for LA+EUF. In *Proc. of CAV*, volume 5123 of *LNCS*, pages 304–308, 2008.

[3] M. P. Bonacina and M. Johansson. On interpolation in decision procedures. In *Proc. of TABLEAUX'11*, pages 1–16, 2011.

[4] M. Bozzano, R. Bruttomesso, A. Cimatti, T. Junttila, P. Van Rossum, S. Ranise, and R. Sebastiani. Efficient Satisfiability Modulo Theories via Delayed Theory Combination. In *CAV'05*, pages 335–349, 2005.

[5] A. Brillout, D. Kroening, P. Rümmer, and T. Wahl. An Interpolating Sequent Calculus for Quantifier-Free Presburger Arithmetic . In *IJCAR*, 2010.

[6] A. Brillout, D. Kroening, P. Rümmer, and T. Wahl. Beyond quantifier-free interpolation in extensions of Presburger arithmetic. In *Proc. of VMCAI*, pages 88–102. Springer-Verlag, 2011.

[7] R. Bruttomesso, S. Ghilardi, and S. Ranise. Rewriting-based Quantifier-free Interpolation for a Theory of Arrays. Technical Report RI 334-10, Dip. Scienze dell'Informazione, Univ. di Milano, 2010.

[8] R. Bruttomesso, S. Ghilardi, and S. Ranise. A Combination of Rewriting and Constraint Solving for the Quantifier-free Interpolation of Arrays with Integer Difference Constraints. In *FroCoS*, 2011.

[9] R. Bruttomesso, S. Ghilardi, and S. Ranise. Rewriting-based Quantifier-free Interpolation for a Theory of Arrays. In *RTA*, 2011.

[10] C. Chang and J. H. Keisler. *Model Theory.* North-Holland, Amsterdam-London, third edition, 1990.

[11] A. Cimatti, A. Griggio, and R. Sebastiani. Efficient Interpolant Generation in Satisfiability Modulo Theories. In *TACAS*, pages 397–412, 2008.

[12] Herbert B. Enderton. *A Mathematical Introduction to Logic.* Academic Press, New York-London, 1972.

[13] A. Fuchs, A. Goel, J. Grundy, S. Krstić, and C. Tinelli. Ground Interpolation for the Theory of Equality. In *TACAS*, pages 413–427, 2009.

[14] S. Ghilardi. Model theoretic methods in combined constraint satisfiability. *Journal of Automated Reasoning*, 33(3-4):221–249, 2004.

[15] A. Goel, S. Krstić, and C. Tinelli. Ground Interpolation for Combined Theories. In *Proc. of CADE 22*, LNCS, pages 183–198, 2009.

[16] T. Henzinger and K. L. McMillan R. Jhala, R. Majumdar. Abstractions from Proofs. In *POPL*, 2004.

[17] Mal'cev A. I. Axiomatizable classes of locally free algebras of certain types. *Sibirsk. Mat. Ž.*, 3:729–743, 1962.

[18] R. Jhala and K. L. McMillan. A Practical and Complete Approach to Predicate Refinement. In *TACAS*, pages 459–473, 2006.

[19] Bjarni Jónsson. Universal relational systems. *Math. Scand.*, 4:193–208, 1956.

[20] D. Kapur, R. Majumdar, and C. Zarba. Interpolation for Data Structures. In *SIGSOFT'06/FSE-14*, pages 105–116, 2006.

[21] E. W. Kiss, L. Márki, P. Pröhle, and W. Tholen. Categorical algebraic properties. A compendium on amalgamation, congruence extension, epimorphisms, residual smallness, and injectivity. *Studia Sci. Math. Hungar.*, 18(1):79–140, 1982.

[22] L. Kovács and A. Voronkov. Finding Loop Invariants for Programs over Arrays Using a Theorem Prover. In *FASE*, pages 470–485, 2009.

[23] K. McMillan. Interpolants from Z3 proofs. In *Proc. of FMCAD*, 2011.

[24] K. L. McMillan. An Interpolating Theorem Prover. In *TACAS*, pages 16–30, 2004.

[25] K. L. McMillan. An Interpolating Theorem Prover. *Theor. Comput. Sci.*, 345(1):101–121, 2005.

[26] G. Nelson and D. C. Oppen. Simplification by Cooperating Decision Procedures. *ACM Transactions on Programming Languages and Systems*, 1(2):245–57, 1979.

[27] D. C. Oppen. Reasoning about recursively defined data structures. *Journal of the ACM*, 27:403–411, 1980.

[28] Claus Michael Ringel. The intersection property of amalgamations. *J. Pure Appl. Algebra*, 2:341–342, 1972.

[29] A. Rybalchenko and V. Sofronie-Stokkermans. Constraint Solving for Interpolation. *J. of Symbolic Logic*, 45(11):1212–1233, 2010.

[30] V. Sofronie-Stokkermans. Interpolation in Local Theory Extensions. In *IJCAR'06: Int. Conf. on Automated Reasoning*, volume 4130 of *LNCS*, pages 235–250, 2006.

[31] C. Tinelli and M. T. Harandi. A new correctness proof of the Nelson-Oppen combination procedure. In *Proc. FroCoS 1996*, pages 103–119, 1996.

[32] G. Yorsh and M. Musuvathi. A combination method for generating interpolants. In *Automated deduction—CADE-20*, volume 3632 of *LNCS*, pages 353–368. Springer, Berlin, 2005. Extended version available as Technical Report MSR-TR-2004-108, Microsoft Research, October 2004.

A List of Metarules

Close1	Close2	Propagate1	Propagate2
$\dfrac{}{A \mid B}$ *Prov.*: A is unsat. *Instr.*: $\phi' \equiv \bot$.	$\dfrac{}{A \mid B}$ *Prov.*: B is unsat. *Instr.*: $\phi' \equiv \top$.	$\dfrac{A \mid B \cup \{\psi\}}{A \mid B}$ *Prov.*: $A \vdash \psi$ and ψ is AB-common. *Instr.*: $\phi' \equiv \phi \wedge \psi$.	$\dfrac{A \cup \{\psi\} \mid B}{A \mid B}$ *Prov.*: $B \vdash \psi$ and ψ is AB-common. *Instr.*: $\phi' \equiv \psi \to \phi$.

Define0	Define1	Define2
$\dfrac{A \cup \{a = t\} \mid B \cup \{a = t\}}{A \mid B}$ *Prov.*: t is AB-common, a fresh. *Instr.*: $\phi' \equiv \phi(t/a)$.	$\dfrac{A \cup \{a = t\} \mid B}{A \mid B}$ *Prov.*: t is A-local and a is fresh. *Instr.*: $\phi' \equiv \phi$.	$\dfrac{A \mid B \cup \{a = t\}}{A \mid B}$ *Prov.*: t is B-local and a is fresh. *Instr.*: $\phi' \equiv \phi$.

Disjunction1	Disjunction2
$\dfrac{\cdots \quad A \cup \{\psi_k\} \mid B \quad \cdots}{A \mid B}$ *Prov.*: $\bigvee_{k=1}^{n} \psi_k$ is A-local and $A \vdash \bigvee_{k=1}^{n} \psi_k$. *Instr.*: $\phi' \equiv \bigvee_{k=1}^{n} \phi_k$.	$\dfrac{\cdots \quad A \mid B \cup \{\psi_k\} \quad \cdots}{A \mid B}$ *Prov.*: $\bigvee_{k=1}^{n} \psi_k$ is B-local and $B \vdash \bigvee_{k=1}^{n} \psi_k$. *Instr.*: $\phi' \equiv \bigwedge_{k=1}^{n} \phi_k$.

Redplus1	Redplus2	Redminus1	Redminus2
$\dfrac{A \cup \{\psi\} \mid B}{A \mid B}$ *Prov.*: $A \vdash \psi$ and ψ is A-local. *Instr.*: $\phi' \equiv \phi$.	$\dfrac{A \mid B \cup \{\psi\}}{A \mid B}$ *Prov.*: $B \vdash \psi$ and ψ is B-local. *Instr.*: $\phi' \equiv \phi$.	$\dfrac{A \mid B}{A \cup \{\psi\} \mid B}$ *Prov.*: $A \vdash \psi$ and ψ is A-local. *Instr.*: $\phi' \equiv \phi$.	$\dfrac{A \mid B}{A \mid B \cup \{\psi\}}$ *Prov.*: $B \vdash \psi$ and ψ is B-local. *Instr.*: $\phi' \equiv \phi$.

ConstElim1	ConstElim2	ConstElim0
$\dfrac{A \mid B}{A \cup \{a = t\} \mid B}$ *Prov.*: a is A-strict and does not occur in A, t. *Instr.*: $\phi' \equiv \phi$.	$\dfrac{A \mid B}{A \mid B \cup \{b = t\}}$ *Prov.*: b is B-strict and does not occur in B, t. *Instr.*: $\phi' \equiv \phi$.	$\dfrac{A \mid B}{A \cup \{c = t\} \mid B \cup \{c = t\}}$ *Prov.*: c, t are AB-common, c does not occur in A, B, t. *Instr.*: $\phi' \equiv \phi$.

Table 1: Interpolating Metarules: each rule has a proviso *Prov.* and an instruction *Instr.* for recursively computing the new interpolant ϕ' from the old one(s) $\phi, \phi_1, \ldots, \phi_k$. Metarules are applied *bottom-up* and interpolants are computed *top-down.*

B Proofs

We give here all the proofs not included in the text.

B.1 Proofs for Section 3

Lemma B.1. *Let T be a theory in a signature Σ and let $\underline{a}, \underline{b}, \underline{c}$ be tuples of (distinct) free constants; let also Θ_1, Θ_2 be sets of ground formulae having the following properties:*

- *in Θ_1 at most the free constants $\underline{a}, \underline{c}$ occur;*
- *in Θ_2 at most the free constants $\underline{b}, \underline{c}$ occur;*
- *there is no ground formula $\theta(\underline{c})$ such that $\Theta_1 \vdash_T \theta(\underline{c})$ and $\Theta_2 \vdash_T \neg\theta(\underline{c})$.*

Then there are models $\mathcal{M}_1, \mathcal{M}_2$ of T such that $\mathcal{M}_1 \models \Theta_1$, $\mathcal{M}_2 \models \Theta_2$ and such that the intersection of the supports of $\mathcal{M}_1$ and $\mathcal{M}_2$ is precisely the substructure generated by the interpretation of the constants $\underline{c}$.

Proof. Let us call Σ^A the signature Σ expanded with the free constants $\underline{a} \cup \underline{c}$ and Σ^B the signature Σ expanded with the free constants $\underline{b} \cup \underline{c}$ (we put $\Sigma^C := \Sigma^A \cap \Sigma^B = \Sigma \cup \{\underline{c}\}$). As a first step, we build a maximal T-consistent set Γ of ground Σ^A-formulae and a maximal T-consistent set Δ of ground Σ^B-formulae such that $\Theta_1 \subseteq \Gamma$, $\Theta_2 \subseteq \Delta$, and $\Gamma \cap \Sigma^C = \Delta \cap \Sigma^C$.[3] For simplicity[4] let us assume that Σ is at most countable, so that we can fix two enumerations

$$\phi_1, \phi_2, \ldots \qquad \psi_1, \psi_2, \ldots$$

of ground Σ^A- and Σ^B-formulae, respectively. We build inductively Γ_n, Δ_n such that for every n (i) Γ_n contains either ϕ_n or $\neg\phi_n$; (ii) Δ_n contains either ψ_n or $\neg\psi_n$; (iii) there is no ground Σ^C-formula θ such that $\Gamma_n \cup \{\neg\theta\}$ and $\Delta_n \cup \{\theta\}$ are not T-consistent. Once this is done, we can get our Γ, Δ as $\Gamma \equiv \bigcup \Gamma_n$ and $\Delta \equiv \bigcup \Delta_n$.

We let Γ_0 be Θ_1 and Δ_0 be Θ_2 (notice that (iii) holds by assumption). To build Γ_{n+1} we have two possibilities, namely $\Gamma_n \cup \{\phi_n\}$ and $\Gamma_n \cup \{\neg\phi_n\}$. Suppose they are both unsuitable because there are $\theta_1, \theta_2 \in \Sigma^C$ such that the sets

$$\Gamma_n \cup \{\phi_n, \neg\theta_1\}, \quad \Delta_n \cup \{\theta_1\}, \quad \Gamma_n \cup \{\neg\phi_n, \neg\theta_2\}, \quad \Delta_n \cup \{\theta_2\}$$

are all T-inconsistent. If we put $\theta \equiv \theta_1 \vee \theta_2$, we get that $\Gamma_n \cup \{\neg\theta\}$ and $\Delta_n \cup \{\theta\}$ are not T-consistent, contrary to induction hypothesis. A similar argument shows that we can also build Δ_n.

[3] By abuse, we use Σ^C to indicate not only the signature Σ^C but also the set of formulae in the signature Σ^C.

[4] This is just to avoid a (straightforward indeed) transfinite induction argument.

Let now $\mathcal{M}_1$ be a model of Γ and $\mathcal{M}_2$ be a model of Δ. Consider the substructures $\mathcal{A}_1, \mathcal{A}_2$ of $\mathcal{M}_1, \mathcal{M}_2$ generated by the interpretations of the constants from Σ^C: since they satisfy the same literals from Σ^C (because $\Gamma \cap \Sigma^C = \Delta \cap \Sigma^C$), we have that $\mathcal{A}_1$ and $\mathcal{A}_2$ are Σ^C-isomorphic. Up to renaming, we can suppose that $\mathcal{A}_1$ and $\mathcal{A}_2$ are just the same substructure. □

Theorem 3.2 *A theory T admits quantifier-free interpolants iff T has the sub-amalgamation property.*

Proof. Suppose first that T has sub-amalgamation; let ϕ, ψ be quantifier-free formulae such that $\phi \wedge \psi$ is not T-satisfiable. Let us replace variables with free constants in ϕ, ψ; let us call Σ^A the signature Σ expanded with the free constants from ϕ and Σ^B the signature Σ expanded with the free constants from ψ (we put $\Sigma^C := \Sigma^A \cap \Sigma^B$). For reductio, suppose that there is no ground formula θ such that: (a) ϕ T-entails θ; (b) $\theta \wedge \psi$ is T-unsatisfiable; (c) only free constants from Σ^C occur in θ. By Lemma B.1, taking $\Theta_1 := \{\phi\}, \Theta_2 := \{\psi\}$, we know that there are models $\mathcal{M}_1, \mathcal{M}_2$ of T such that $\mathcal{M}_1 \models \phi, \mathcal{M}_2 \models \psi$ and such that the intersection of the supports of $\mathcal{M}_1$ and $\mathcal{M}_2$ is precisely the substructure generated by the interpretation of the constants from Σ^C (let us we call this substructure $\mathcal{A}$ for short). By the sub-amalgamation property, there is a T-amalgam $\mathcal{M}$ of $\mathcal{M}_1$ and $\mathcal{M}_2$ over $\mathcal{A}$. Now ϕ, ψ are ground formulae true in $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, hence they are both true in $\mathcal{M}$, which is impossible because $\phi \wedge \psi$ was assumed to be T-inconsistent.

Suppose now that T has quantifier-free interpolants. Take two models $\mathcal{M}_1$ and $\mathcal{M}_2$ of T sharing a substructure $\mathcal{A}$; we can freely suppose (up to a renaming) that $|\mathcal{M}_1| \cap |\mathcal{M}_2| = |\mathcal{A}|$ (we use the notation $| - |$ to indicated the support of a structure). In order to show that a T-amalgam of $\mathcal{M}_1, \mathcal{M}_2$ over $\mathcal{A}$ exists, it is sufficient (by Robinson Diagram Lemma [10]) to show that $\Delta_\Sigma(\mathcal{M}_1) \cup \Delta_\Sigma(\mathcal{M}_2)$ is T-consistent, where (for $i = 1, 2$) $\Delta_\Sigma(\mathcal{M}_i)$ is the *diagram* of $\mathcal{M}_i$, namely the set of $\Sigma \cup |\mathcal{M}_i|$-literals true in $\mathcal{M}_i$.

If $\Delta_\Sigma(\mathcal{M}_1) \cup \Delta_\Sigma(\mathcal{M}_2)$ is not T-consistent, by the compactness theorem of first order logic, there exist a $\Sigma \cup |\mathcal{M}_1|$-ground sentence ϕ and a $\Sigma \cup |\mathcal{M}_2|$-ground sentence ψ such that (i) $\phi \wedge \psi$ is T-inconsistent; (ii) ϕ is a conjunction of literals from $\Delta_\Sigma(\mathcal{M}_1)$; (iii) ψ is a conjunction of literals from $\Delta_\Sigma(\mathcal{M}_2)$. By the existence of quantifier-free interpolants, taking free constants instead of variables, we get that there exists a ground $\Sigma \cup |\mathcal{A}|$-sentence θ such that ϕ T-entails θ and $\psi \wedge \theta$ is T-inconsistent. The former fact yields that θ is true in $\mathcal{M}_1$ and hence also in $\mathcal{A}$ and in $\mathcal{M}_2$, because θ is ground. However, the fact that θ is true in $\mathcal{M}_2$ contradicts the fact that $\psi \wedge \theta$ is T-inconsistent. □

The following Lemma is part of the well-known Nelson-Oppen combination results [31], [26]:

Lemma B.2. *Suppose that T_1, T_2 are two stably infinite theories in disjoint signatures Σ_1, Σ_2 and let C be a set of free constants not belonging to $\Sigma_1 \cup \Sigma_2$; let Γ be a partition of C, i.e. a set of ground equalities or inequalities containing the literal $c_1 = c_2$ or the literal $c_1 \neq c_2$, for all pairs of different constants from C. For $i = 1, 2$, let Θ_i be a T_i-consistent set of ground $\Sigma_i \cup C$-formulae containing Γ. Then $\Theta_1 \cup \Theta_2$ is $T_1 \cup T_2$-consistent.*

Proof. Let $\mathcal{M}_1, \mathcal{M}_2$ be two models of $T_1 \cup \Theta_1, T_2 \cup \Theta_2$, respectively. By stable infiniteness and upward Lövenheim-Skolem theorem [10], we can assume that they are both infinite and have the same cardinality (bigger than the cardinality of C). Thus there is a bijection f among their supports and (as equalities of constants from C are interpreted in the same way in $\mathcal{M}_1$ and $\mathcal{M}_2$) we can assume that $f(c^{\mathcal{M}_1}) = c^{\mathcal{M}_2}$. Using this bijection, it is easy to lift the interpretation of the Σ_2-symbols from the support of $\mathcal{M}_2$ to the support of $\mathcal{M}_1$. The lifted model is $\Sigma_2 \cup C$-isomorphic to $\mathcal{M}_2$, thus it is a model of $T_1 \cup T_2 \cup \Theta_1 \cup \Theta_2$. □

Theorem 3.3 *Let T_1, T_2 be two stably infinite theories in disjoint signatures Σ_1, Σ_2. If T_1, T_2 both have the strong sub-amalgamation property, then so does $T_1 \cup T_2$.*

Proof. Consider two models $\mathcal{M}_1, \mathcal{M}_2$ of $T_1 \cup T_2$ together with a common substructure $\mathcal{A}$; we can freely suppose (up to a renaming) that $|\mathcal{M}_1| \cap |\mathcal{M}_2| = |\mathcal{A}|$. By Robinson Diagram Lemma [10], it is sufficient to show the consistency of $T_1 \cup T_2 \cup \Gamma_1 \cup \Gamma_2$, where Γ_i $(i = 1, 2)$ is defined as

$$\Gamma_i \;\equiv\; \Delta_{\Sigma_i}(\mathcal{M}_1) \cup \Delta_{\Sigma_i}(\mathcal{M}_2) \cup \{m_1 \neq m_2 \mid m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|,\ m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|\} \ .$$

By compactness, it is enough to show the $T_1 \cup T_2$-consistency of the subset $T_1 \cup T_2 \cup \Gamma_1^0 \cup \Gamma_2^0$ of $T_1 \cup T_2 \cup \Gamma_1 \cup \Gamma_2$ mentioning just a finite set C of free constants from $|\mathcal{M}_1| \cup |\mathcal{M}_2|$. By the strong amalgamability of T_1 and T_2, we know that $T_1 \cup \Gamma_1^0$ and $T_2 \cup \Gamma_2^0$ are both consistent. Now notice that for every pair c_1, c_2 of distinct constants from C, the set Γ_i (hence also the set Γ_i^0) contains the negative literal $c_1 \neq c_2$: in fact, this inequation is part of the definition of the diagram of a structure or (in case c_1, c_2 are from different supports) it has been added explicitly when building Γ_1, Γ_2. According to Lemma B.2, this is sufficient to infer the consistency of $T_1 \cup T_2 \cup \Gamma_1^0 \cup \Gamma_2^0$, as T_1, T_2 are stably infinite. □

Theorem 3.5 *Let T be a theory admitting quantifier-free interpolation and let Σ be a signature disjoint from the signature of T and containing at least a unary predicate symbol. Then $T \cup \mathcal{EUF}(\Sigma)$ has quantifier-free interpolation iff T has the strong sub-amalgamation property.*

Proof. (Below Σ_T is the signature of T). Let T be strongly amalgamable and let $\mathcal{M}_1, \mathcal{M}_2$ be two models of $T \cup \mathcal{EUF}(\Sigma)$ sharing a submodel $\mathcal{M}_0$ (as usual, we suppose that $|\mathcal{M}_1| \cap |\mathcal{M}_2| =$

$|\mathcal{M}_0|$). To amalgamate them, consider first a model $\mathcal{M}$ of T strongly amalgamating the Σ_T-reducts of $\mathcal{M}_1, \mathcal{M}_2$ over the Σ_T-reduct of $\mathcal{M}_0$. Since the amalgam is strong, up to isomorphism we can consider the support of $\mathcal{M}$ as a superset of $|\mathcal{M}_1| \cup |\mathcal{M}_2|$; thus it is easy to expand $\mathcal{M}$ to a total structure interpreting the symbols of Σ. The expansion is a model of $T \cup \mathcal{EUF}(\Sigma)$ amalgamating $\mathcal{M}_1$ and $\mathcal{M}_2$ over $\mathcal{M}_0$.

Conversely, suppose that T does not have the sub-amalgamation property. Let $\mathcal{M}_1, \mathcal{M}_2$ be models of T_1 and let $\mathcal{A}$ be a substructure of them such that there are no data $\mathcal{M}, \mu_1, \mu_2$ satisfying the conditions for the strong sub-amalgamability property. This means that the set

$$\Gamma \;\equiv\; \Delta_{\Sigma_1}(\mathcal{M}_1) \cup \Delta_{\Sigma_1}(\mathcal{M}_2) \cup \{m_1 \neq m_2 \mid m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|,\ m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|\}$$

is not T-consistent. By compactness, there are $m_1^1, \dots m_1^k \in |\mathcal{M}_1| \setminus |\mathcal{A}|$ and $m_2^1, \dots m_2^k \in |\mathcal{M}_2| \setminus |\mathcal{A}|$ such that

$$T \cup \Delta_{\Sigma_1}(\mathcal{M}_1) \cup \Delta_{\Sigma_1}(\mathcal{M}_2) \models \bigvee_{j=1}^{k} m_1^j = m_2^j \quad . \tag{5}$$

Expand now $\mathcal{M}_1, \mathcal{M}_2$ to $\Sigma_T \cup \Sigma$-structures as follows: the Σ-symbols are interpreted arbitrarily (but in such a way that $\mathcal{A}$ remains a substructure of the expansions) apart from the unary predicate P, which is interpreted as the whole support of $\mathcal{M}_1$ in the expansion of $\mathcal{M}_1$ and as the support of $\mathcal{A}$ in the expansion of $\mathcal{M}_2$. From (5), it is then clear that sub-amalgamation (hence quantifier-free interpolation) fails for $T_1 \cup T_2$: in fact, any $\mathcal{M} \models T$ amalgamating $\mathcal{M}_1, \mathcal{M}_2$ over $\mathcal{A}$, must identify some $m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|$ with some $m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|$, which is impossible as the interpretation of P in $\mathcal{M}$ must agree with the interpretations of P in the expansions of $\mathcal{M}_1$ and $\mathcal{M}_2$. □

B.2 Proofs for Section 4

Theorem 4.2 shows the equivalence between strong amalgamability and equality interpolation; we add one equivalent characterization more in the statement below:

Theorem 4.2 *The following conditions are equivalent for a theory T having quantifier-free interpolation:*

(i) *T is strongly sub-amalgamable;*

(ii) *T is equality interpolating;*

(iii) *for every triple $\underline{x}, \underline{y}_1, \underline{y}_2$ of tuples of variables and for every pair of quantifier-free formulae $\delta_1(\underline{x}, \underline{y}_1), \delta_2(\underline{x}, \underline{y}_2)$ such that*

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{y}_2 \neq \emptyset \tag{6}$$

there is a tuple $\underline{v}(\underline{x})$ *of terms such that*

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1\underline{y}_2 \cap \underline{v} \neq \emptyset \ . \tag{7}$$

Proof. We first show (i) $\Rightarrow$ (ii). Suppose first that T is strongly sub-amalgamable; we show that (1) $\Rightarrow$ (2) holds by contraposition. So, let us fix tuples of fresh free constants $\underline{a}, \underline{m}_1, \underline{n}_1, \underline{m}_2, \underline{n}_2$ and suppose that for every finite tuple $\underline{v}$ of $\Sigma \cup \{\underline{a}\}$-ground terms, the formula

$$\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{n}_2\underline{m}_2) \wedge (\underline{m}_1\underline{m}_2 \cap \underline{v} = \emptyset) \tag{8}$$

is T-consistent (here Σ is the signature of T). We claim that the set

$$\{\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1), \delta_2(\underline{a}, \underline{n}_2, \underline{m}_2)\} \cup \{\underline{m}_1\underline{m}_2 \cap \underline{v} = \emptyset\}_{\underline{v}} \tag{9}$$

is T-consistent, where $\underline{v}$ varies over all possible tuples of such terms. In fact, if (9) were not consistent, by compactness, there would be tuples of $\Sigma \cup \{\underline{a}\}$-ground terms $\underline{v}_1, \dots, \underline{v}_k$ such that

$$\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{n}_2, \underline{m}_2) \wedge \bigwedge_{j=1}^{k} (\underline{m}_1\underline{m}_2 \cap \underline{v}_j = \emptyset)$$

were not T-consistent. Putting $\underline{v}$ equal to the tuple obtained by juxtaposition $\underline{v}_1 \cdots \underline{v}_k$, we would get a $\underline{v}$ contradicting (8).

Let Θ_1 be $\{\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1)\} \cup \{\underline{m}_1 \cap \underline{v} = \emptyset\}_{\underline{v}}$ and let Θ_2 be $\{\delta_2(\underline{a}, \underline{n}_2, \underline{m}_2)\} \cup \{\underline{m}_2 \cap \underline{v} = \emptyset\}_{\underline{v}}$. Since $\Theta_1 \cup \Theta_2$ is equal to (9) which is T-consistent, there is no ground $\Sigma \cup \{\underline{a}\}$-formula $\theta(\underline{a})$ such that $\Theta_1 \vdash_T \theta(\underline{a})$ and such that $\Theta_2 \cup \{\theta(\underline{a})\}$ is not T-consistent. By Lemma B.1, we can then produce models $\mathcal{M}_1, \mathcal{M}_2$ of T such that $\mathcal{M}_1 \models \Theta_1$, $\mathcal{M}_2 \models \Theta_2$ and such that the intersection of their supports is precisely the substructure generated by the interpretation of the constants $\underline{a}$. If we now strongly amalgamate them, we get a model of T in which $\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1), \delta_2(\underline{a}, \underline{n}_2, \underline{m}_2), \underline{m}_1 \cap \underline{m}_2 = \emptyset$ are all true, showing that (1) fails.

The implication (ii) $\Rightarrow$ (iii) is trivial. We prove (iii) $\Rightarrow$ (i). Suppose that we have (6) $\Rightarrow$ (7) and let us prove strong sub-amalgamability. If the latter property fails, by Robinson Diagram Lemma, there are models $\mathcal{M}_1, \mathcal{M}_2$ of T together with a shared substructure $\mathcal{A}$ such that the set of sentences

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

is not T-consistent. By compactness, the sentence

$$\delta_1(\underline{a}, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{m}_2) \to \underline{m}_1 \cap \underline{m}_2 \neq \emptyset$$

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. By the implication (6) ⇒ (7), there exists a finite tuple $\underline{v}(\underline{a})$ of $\Sigma \cup \{\underline{a}\}$-terms such that

$$\delta_1(\underline{a}, \underline{m_1}) \wedge (\underline{m_1} \cap \underline{v}(\underline{a}) = \emptyset) \wedge \delta_2(\underline{a}, \underline{m_2}) \wedge (\underline{m_2} \cap \underline{v}(\underline{a}) = \emptyset)$$

is not T-consistent. Since T has quantifier-free interpolation, there is a ground formula $\theta(\underline{a})$ such that

$$\delta_1(\underline{a}, \underline{m_1}) \wedge (\underline{m_1} \cap \underline{v}(\underline{a}) = \emptyset) \to \theta(\underline{a}) \tag{10}$$

is T-valid and

$$\delta_2(\underline{a}, \underline{m_2}) \wedge (\underline{m_2} \cap \underline{v}(\underline{a}) = \emptyset) \wedge \theta(\underline{a}) \tag{11}$$

is not T-consistent. However this is a contradiction: since $\underline{m_1} \subseteq |\mathcal{M}_1| \setminus |\mathcal{A}|$, the formula $\underline{m_1} \cap \underline{v}(\underline{a}) = \emptyset$ is true in $\mathcal{M}_1$, which entails that $\theta(\underline{a})$ is true in $\mathcal{A}$ and in $\mathcal{M}_2$ too, where (11) consequently holds. □

Notice that (iii) is just the special case of (ii) arising when the tuple $\underline{z}$ is empty; this special case can be enough in the applications (for instance, the combined interpolation algorithm from Section 5 makes use of this special case only).

We now come to the results concerning equality interpolation and quantifier elimination.

Lemma B.3. *Let T be a theory admitting quantifier elimination; T is universal iff for every quantifier-free formula $\phi(\underline{x}, \underline{y})$, there exists tuples $\underline{t}_1(\underline{x}), \dots, \underline{t}_n(\underline{x})$ of tuples of terms such that*

$$T \vdash \exists \underline{y}\, \phi(\underline{x}, \underline{y}) \leftrightarrow \bigvee_{i=1}^{n} \phi(\underline{x}, t_i(\underline{x})) \quad . \tag{12}$$

Proof. If the condition of the Lemma is true for every $\phi(\underline{x}, \underline{y})$, one can find an equivalent universal set of axioms for T as follows. Notice that the right-to-left side of (12) is a logical validity and the left-to-right side is equivalent to a universal formula. Thus, we can take as axioms for T the universal closures of the left-to-right sides of (12), together with the ground formulae which are logical consequences of T. In fact, axioms (12) are sufficient to find for every sentence a ground formula T-equivalent to it.

Conversely, suppose that T is universal and that there is $\phi(\underline{x}, y_1, \dots, y_m)$ such that (12) does not hold (for all possible tuples of m-tuples of terms). Then, by compactness, we have that the set of sentences

$$\Gamma \;\equiv\; \{\phi(\underline{a}, \underline{b})\} \cup \{\neg\phi(\underline{a}, \underline{t}(\underline{a}))\}_{\underline{t}}$$

is T-consistent (here Σ is the signature of T, $\underline{a}, \underline{b} := b_1, \dots, b_m$ are tuples of fresh constants and $\underline{t}$ vary on the set of m-tuples of $\Sigma \cup \{\underline{a}\}$-terms). Let $\mathcal{M}$ be a T-model of Γ and let $\mathcal{N}$ be the substructure of $\mathcal{M}$ generated by the $\underline{a}$. Since T is universal and truth of universal sentences is preserved under taking substructures, $\mathcal{N}$ is also a model of T and since T has quantifier-elimination, $\exists \underline{y}\, \phi(\underline{a}, \underline{y})$ - being T-equivalent to a quantifier-free $\Sigma \cup \{\underline{a}\}$-sentence - is true in $\mathcal{N}$ too. This is a contradiction because from $\mathcal{N} \models \exists \underline{y}\, \phi(\underline{a}, \underline{y})$ it follows that $\mathcal{N} \models \phi(\underline{a}, \underline{t}(\underline{a}))$ holds for some $\underline{t}$, contrary to the fact that $\mathcal{M} \not\models \phi(\underline{a}, \underline{t}(\underline{a}))$ and to the fact that $\mathcal{N}$ is a substructure of $\mathcal{M}$. □

Theorem 4.3 *A universal theory admitting quantifier elimination is equality interpolating.*

Proof. We show that a universal and quantifier eliminable theory T satisfies the implication (6) $\Rightarrow$ (7). Suppose that (6) holds; by the previous Lemma, there exists tuples of terms $\underline{t}_1(\underline{x}), \dots, \underline{t}_k(\underline{x})$ such that

$$\exists \underline{y}_2\, \delta_2(\underline{x}, \underline{y}_2) \leftrightarrow \bigvee_{j=1}^{k} \delta_2(\underline{x}, \underline{t}_j(\underline{x})) \tag{13}$$

is T-valid. For every $j = 1, \dots, k$, if we replace $\underline{y}_2$ with $\underline{t}_j$ in (6), we get

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{t}_j) \vdash_T \underline{y}_1 \cap \underline{t}_j \neq \emptyset$$

hence also

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \bigvee_{j=1}^{k} \delta_2(\underline{x}, \underline{t}_j) \vdash_T \bigvee_{j=1}^{k} (\underline{y}_1 \cap \underline{t}_j \neq \emptyset) \quad .$$

Taking into account (13) and letting $\underline{v}$ be the tuple $\underline{t}_1 \cdots \underline{t}_k$ obtained by juxtaposition, we get

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \exists \underline{y}_2 \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{v} \neq \emptyset \quad .$$

Removing the existential quantifier in the antecedent of the implication, we obtain

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{v} \neq \emptyset$$

and a fortiori (7), as desired. □

We point out that the obvious converse of Theorem 4.3 is not true: the theory of dense linear orders without endpoints has quantifier elimination, is equality interpolating (because it can be checked it has the strong sub-amalgamation property), but does not admit a universal set of axioms (because it is not closed under substructures).

The proof of Theorem 4.3 is important also from the applications point of view. In fact, in the combined interpolation algorithm designed in Section 5, one is given formulæ δ_1, δ_2 satisfying (6) and is asked to compute terms $\underline{v}(\underline{x})$ satisfying (7). In case our equality interpolating theory is universal and has quantifier elimination, one way to do this is to run the quantifier elimination algorithm over $\exists \underline{y}_2\, \delta_2(\underline{x}, \underline{y}_2)$ and to let $\underline{v}$ be the tuple $\underline{t}_1 \cdots \underline{t}_k$ obtained by juxtaposition from the tuples in the right member of (13).

Lemma B.3 is also interesting in itself. According to Theorem 4.3, a sufficient condition for a theory T to be equality interpolating is to have quantifier elimination via a universal set of axioms. The Lemma gives the possibility of checking the existence of such a set of axioms just by inspecting the quantifier elimination algorithm. Sometimes, this procedure is easy. As an example, we can take the case of linear real arithmetic and Fourier-Motzkin algorithm. It is not difficult to see that Fourier-Motzkin algorithm satisfies the condition of Lemma B.3 in the sense that it always 'eliminates existential quantifiers via tuples of terms'. For instance, when eliminating $\exists x$ from $\exists x\,(x < y_1 \wedge x < y_2 \wedge y_3 < x)$ one gets

$$(t_1 < y_1 \wedge t_1 < y_2 \wedge y_3 < t_1) \vee (t_2 < y_1 \wedge t_2 < y_2 \wedge y_3 < t_2)$$

where $t_1 := y_3 + (y_1 - y_3)/2$ and $t_2 := y_3 + (y_2 - y_3)/2$.

We now show that in the convex case, our notion of an equality interpolating theory coincides with the one given in [32].

Proposition 4.4 *A convex theory T having quantifier-free interpolation is equality interpolating iff for every pair y_1, y_2 of variables and for every pair of conjunctions of literals $\delta_1(\underline{x}, \underline{z}_1, y_1), \delta_2(\underline{x}, \underline{z}_2, y_2)$ such that*

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2 \tag{3}$$

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

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v \wedge y_2 = v. \tag{4}$$

Proof. If $\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2$ holds and T is equality interpolating, it follows that there are terms $\underline{v}(\underline{x}) := v_1(\underline{x}), \ldots, v_n(\underline{x})$ such that

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T \bigvee_{i=1}^{n} (y_1 = v_i) \vee \bigvee_{i=1}^{n} (y_2 = v_i). \tag{14}$$

Let $w_1, \ldots, w_n$ be fresh variables; from (14) it follows that

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \wedge \bigwedge_{i=1}^{n} (w_i = v_i) \vdash_T \bigvee_{i=1}^{n} (y_1 = w_i) \vee \bigvee_{i=1}^{n} (y_2 = w_i).$$

Applying convexity, we obtain that there is some i such that either

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \wedge \bigwedge_{i=1}^{n} (w_i = v_i) \vdash_T y_1 = w_i$$

or

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \wedge \bigwedge_{i=1}^{n} (w_i = v_i) \vdash_T y_2 = w_i$$

holds. Replacing the w's with the v's, this gives either

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v_i$$

or

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_2 = v_i.$$

In both cases (taking into consideration (3)), we get $\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v_i \wedge y_2 = v_i$, as required by (4).

Vice versa, when assuming the implication (3) $\Rightarrow$ (4), it is very easy to show (by applying convexity) that T is equality interpolating.[5] □

B.3 Proofs for Section 5

In this Subsection we prove the relevant properties (soundness, completeness, termination) of our combined interpolation algorithm $\mathsf{CI}(T_1, T_2)$, where T_1, T_2 are two signature-disjoint, stably infinite and equality interpolating theories whose SMT problems are decidable.

Lemma B.4. *If rules **Decide**$_X$, **Share**$_i$ and **Terminate**$_i$ do not apply to a pair A, B, then $A \cup B$ is $T_1 \cup T_2$-satisfiable, unless $\bot \in A \cup B$.*

Proof. Let $\underline{a}, \underline{c}$ the free constants occurring in A and $\underline{b}, \underline{c}$ be the free constants occurring in B. If the above rules do not apply and $\bot \notin A \cup B$, then $A_i \cup B_i \cup \{\underline{a} \cap \underline{b} = \emptyset\}$ is T_i-satisfiable for $i = 1, 2$; moreover A contains an A-assignment α and B contains a B-assignement β. This means that $A_1 \cup A_2$ entails A and $B_1 \cup B_2$ entails B, so that it is sufficient to show the $T_1 \cup T_2$-satisfiability of $A_1 \cup A_2 \cup B_1 \cup B_2$ only. The latter follows from Lemma B.2, because the sets

$$\Theta_i \equiv A_i \cup B_i \cup \{\underline{a} \cap \underline{b} = \emptyset\}$$

satisfy the hypothesis of the Lemma. Pick in fact a pair of constants d_1, d_2 from $\underline{a}, \underline{b}, \underline{c}$: if they are both from $\underline{a}, \underline{c}$ or both from $\underline{b}, \underline{c}$, then either $d_1 = d_2$ or $d_1 \neq d_2$ belongs to Θ_i, as $\alpha \cup \beta$ has assigned a truth value to $d_1 = d_2$. If one of them is in $\underline{a}$ and the other is in $\underline{b}$, then $d_1 \neq d_2 \in \Theta_i$ by construction. □

[5] Notice that in Definition 4.1, we can restrict δ_1, δ_2 to be conjunctions of literals, getting anyway an equivalent definition. In fact, if (1) holds, $\delta_1 \equiv \bigvee_j \theta_{1j}$ and $\delta_2 \equiv \bigvee_k \theta_{2k}$, then we can find tuples $\underline{v}_{jk}$ satisfying $\theta_{1j} \wedge \theta_{2k} \vdash_T \underline{y}_1\underline{y}_2 \cap \underline{v}_{jk} \neq \emptyset$ and finally get by juxtaposition a tuple $\underline{v}$ satisfying (2).

Theorem 5.2 *Let T_1 and T_2 be two signature disjoint, stably-infinite, and equality interpolating theories having decidable SMT problems. Then, $\mathsf{CI}(T_1, T_2)$ is a quantifier-free interpolation algorithm for the combined theory $T_1 \cup T_2$.*

Proof. Let A_0, B_0 be our input $T_1 \cup T_2$-unsatisfiable pair. By repeatedly applying our transformations $\textbf{Decide}_X$, $\textbf{Share}_i$ and $\textbf{Terminate}_i$ to it, we produce a tree τ (the pairs labeling the successors of a node are the possible outcomes of our transformations, which are non deterministic). Clearly $\textbf{Decide}_X$, $\textbf{Share}_i$ and $\textbf{Terminate}_i$ are satisfiability-preserving, in the sense that a pair to which they are applied is $T_1 \cup T_2$-satisfiable iff one of the outcomes is. As a consequence, by Lemma B.4, $\bot$ must belongs to all pairs labeling the leaves. Thus, since $\textbf{Decide}_X$, $\textbf{Share}_i$ and $\textbf{Terminate}_i$ can all be justified by metarules, our tree τ is an interpolating metarules refutation (and we are done by Proposition 5.1), *provided we show that τ is finite.* Finiteness of τ is also needed to prove the termination of our algorithm.

We apply König Lemma and show that all branches of τ are finite. Notice that the transformation $\textbf{Decide}_X$ can be applied many times in a branch: this is because $\textbf{Share}_i$ introduces a new ground formula θ and alters the definition of an A-relevant and a B-relevant atom (it introduces new AB-common constants by Term Sharing). However, $\textbf{Share}_i$ can be applied only finitely many times, as it decreases the number of A-strict or B-strict constants. Once $\textbf{Share}_i$ is no more applied, just single applications of $\textbf{Decide}_A$, $\textbf{Decide}_B$, $\textbf{Terminate}_i$ are possible. □

C Quantifier elimination through universal axioms

In this Appendix we give details concerning a couple of applications of Theorem 4.3.

C.1 Integer Linear Arithmetic

Presburger Arithmetic $\mathcal{PRA}$ is the theory so specified. Its signature consists of the symbols $0, 1, +, -, <$ in addition to the infinite predicates P_n (one for every $n > 0$). A set of axioms for $\mathcal{PRA}$ is the following one

$$\begin{aligned}
&\forall x, y, z.\ x + (y + z) = (x + y) + z \\
&\forall x, y.\ x + y = y + x \\
&\forall x.\ x + 0 = x \\
&\forall x.\ x + (-x) = 0 \\
&\forall x.\ x \not< x \\
&\forall x, y, z.\ (x < y \wedge y < z \to x < z) \\
&\forall x, y.\ x < y \vee x = y \vee y < x \\
&\forall x, y, z.\ x < y \to x + z < y + z, \\
&0 < 1, \\
&\forall y. \neg(0 < y \wedge y < 1), \\
&\forall x\, \exists y.\ \textstyle\bigvee_{0 \leq r < n} x = ny + r \\
&\forall x.\ P_n(x) \leftrightarrow \exists y\ (ny = x)
\end{aligned}$$

(we used the abbreviations nt for the sum of n-copies of t and n for $n1$). Presburger arithmetic enjoys quantifier-elimination: a detailed proof can be found e.g. in [12] or also in the online available notes[6] L. Van Der Dries "Mathematical Logic Lecture Notes" (where we took the above axiomatization from). However, $\mathcal{PRA}$ is not equality interpolating because $\mathcal{PRA} \cup \mathcal{EUF}$ does not enjoys quantifier-free interpolation [5].

In Subsection 4.1, we proposed the theory $\mathcal{LIA}$, comprising in its language also the unary function symbols $div[n]$ (representing integer division by n, for $n > 1$). In $\mathcal{LIA}$, one can define $P_n(x)$ as $x\ rem[n] = 0$ (recall that $x\ rem[n]$ abbreviate $x - n(x\ div[n])$). Using this definition, we can view $\mathcal{LIA}$ as a supertheory of $\mathcal{PRA}$, because all the axioms of $\mathcal{PRA}$ are derivable in $\mathcal{LIA}$.[7] We are ready to show that Theorem 4.3 applies to $\mathcal{LIA}$:

Proposition C.1. *$\mathcal{LIA}$ is equality interpolating.*

[6] `http://www.math.uiuc.edu/~vddries/`

[7] For the last one, show that the following universal sentences are derivable in $\mathcal{PRA}$ for every $n > 0$:

$$\forall x.\ nx = 0 \to x = 0 \qquad \forall x.\ \bigwedge_{0<r<n} nx \neq r\ .$$

Proof. In view of Theorem 4.3, since $\mathcal{LIA}$ is universal, we only need to show that $\mathcal{LIA}$ has elimination of quantifier. Let $\phi(\underline{x})$ be an abritrary formula of $\mathcal{LIA}$; consider an atom L occurring in ϕ containing an occurrence of a term u of the kind $t\ div[n]$. Modulo $\mathcal{LIA}$, the atom L is equivalent to

$$\exists y \bigvee_{0\leq r<n} (t = ny + r \wedge L[y/u]) \tag{15}$$

(this is because $\bigvee_{0\leq r<n}(t = ny + r) \leftrightarrow y = t\ div[n]$ follows from the axioms of $\mathcal{LIA}$). We can then replace L by (15) in ϕ and get an equivalent formula. If we do this exhaustively, we obtain a formula ϕ' such that $\mathcal{LIA} \vdash \phi \leftrightarrow \phi'$. Since, as we observed above, $\mathcal{LIA}$ is a supertheory of $\mathcal{PRA}$ and the latter enjoys quantifier elimination, we can find a quantifier-free $\phi''(\underline{x})$ such that $\mathcal{LIA} \vdash \phi \leftrightarrow \phi''$. □

C.2 Unit-Two-Variable-Per-Inequality

This theory (called $\mathcal{UTVPI}$ in the literature) is another interesting fragment of integer linear arithmetic, slightly more expressive than $\mathcal{IDL}$. If can be defined as the theory whose axioms are the sentences true in $\mathbb{Z}$ in the signature comprising predecessor *pred*, successor *succ*, $0, <$ and $-$ (the latter is viewed as a unary symbol). We shall exhibit here a set of universal quantifier eliminating axioms for $\mathcal{UTVPI}$ (thus showing that $\mathcal{UTVPI}$ is equality interpolating too, thanks to Theorem 4.3).

Like in the case of $\mathcal{IDL}$, let us examine the shape of the atoms of $\mathcal{UTVPI}$. They are equivalent to formulae having the form $\pm i \bowtie f^n(j)$ (for $n \in \mathbb{Z}$, $\bowtie \in \{=, <, >\}$)[8] where i, j are variables or the constant 0, $f^0(j)$ is j, $f^k(j)$ abbreviates $succ(succ^{k-1}(j))$ when $k > 0$ or $pred(pred^{k-1}(j))$ when $k < 0$. (Usually, $\pm i \bowtie f^n(j)$ is written as $i \pm j \bowtie n$ or as $i \bowtie n \pm j$).

Proposition C.2. *$\mathcal{UTVPI}$ is equality interpolating.*

Proof. We take inspiration from Lemma B.3, that is we directly supply a quantifier elimination algorithm for $\mathcal{UTVPI}$ satisfying (12) (thus, the left to right sides of formulæ (12) will be the relevant axiomatization for $\mathcal{UTVPI}$, once joined with the universal sentences in the signature of $\mathcal{UTVPI}$ which are true in $\mathbb{Z}$).[9] As usual, it is sufficient to eliminate single existentially quantified variables from primitive formulæ [10]. This means that, since negation can be eliminated, we must consider formulae $\exists x\, \phi$ where ϕ is a conjunction of atoms of the following kinds:

$$x = m_i \pm t_i, \qquad x < m_j \pm u_j, \qquad x > m_k \pm v_k,$$

[8] We use $>$ as a defined symbol ($t > u$ stands for $u < t$).

[9] The latter are needed to normalize all atoms to the form $i \bowtie n \pm j$.

where x does not occur in the t_i, u_j, v_k (otherwise either ϕ is inconsistent or the atom is redundant or it simplifies to an atom of the above kinds). If there are literals of the first kind, the quantifier $\exists x$ can be eliminated by substitution (this schema fits (12)), so suppose there are none. If there are no literals of the second kind or no literals of the third kind, $\exists x\, \phi$ is equivalent to $\top$ (use the terms $pred(m_j \pm u_j), succ(m_k \pm v_k)$ to fit (12)). If there are both literals of the second and of the third kind, $\exists x\, \phi$ is equivalent to $\bigvee_k \phi(succ(m_k \pm v_k))$. $\Box$

D Equality interpolating and Beth definability

In this Section we discuss the connection of the notions introduced in this paper with standard topics in mathematical logic and universal algebra. This complementary material is included here for the sake of completeness.

Beth definability theorem [10] is a classical result in model theory; we show that in the convex case equality interpolating can be interpreted as a 'modulo theory' version of a Beth definability property. We find in the non-convex case too a 'Beth-like' formulation of equality interpolation. In the end, we use the Beth definability formulation of equality interpolation in order to briefly discuss the relationship between our results and well-known results concerning strong amalgamation from the algebraic literature.

To begin with, we add a further equivalent characterization to the list (i)-(iii) of Theorem 4.2:

Theorem 4.2 *The following conditions are equivalent for a theory T having quantifier-free interpolation:*

(i) *T is strongly sub-amalgamable;*

(ii) *T is equality interpolating;*

(iii) *T satisfies the implication* (6) ⇒ (7) *(for every δ_1, δ_2);*

(iv) *for every quantifier-free formula $\delta(\underline{x}, \underline{z}, \underline{y})$ such that*

$$\delta(\underline{x}, \underline{z}', \underline{y}') \wedge \delta(\underline{x}, \underline{z}'', \underline{y}'') \vdash_T \underline{y}' \cap \underline{y}'' \neq \emptyset \tag{16}$$

there are terms $\underline{v}(\underline{x})$ such that

$$\delta(\underline{x}, \underline{z}, \underline{y}) \vdash_T \underline{y} \cap \underline{v} \neq \emptyset. \tag{17}$$

Proof. We already proved (in the previous formulation of Theorem 4.2 in Appendix B) that conditions (i)-(ii)-(iii) are all equivalent to each other.

Assume (iv) and (6). Take $\underline{y} := \underline{y}_1, \underline{y}_2$ and put $\delta(\underline{x}, \underline{y}) := \delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{y}_2)$. Now notice that $\delta(\underline{x}, \underline{y}_1', \underline{y}_2') \wedge \delta(\underline{x}, \underline{y}_1'', \underline{y}_2'')$ is

$$\delta_1(\underline{x}, \underline{y}_1') \wedge \delta_2(\underline{x}, \underline{y}_2') \wedge \delta_1(\underline{x}, \underline{y}_1'') \wedge \delta_2(\underline{x}, \underline{y}_2'');$$

since by (6) we have

$$\delta_1(\underline{x}, \underline{y}_1') \wedge \delta_2(\underline{x}, \underline{y}_2'') \vdash_T \underline{y}_1' \cap \underline{y}_2'' \neq \emptyset$$

a fortiori we get

$$\delta(\underline{x}, \underline{y}_1', \underline{y}_2') \wedge \delta(\underline{x}, \underline{y}_1'', \underline{y}_2'') \vdash_T \underline{y}_1'\underline{y}_2' \cap \underline{y}_1''\underline{y}_2'' \neq \emptyset, \tag{18}$$

By (iv), there are terms $\underline{v}(\underline{x})$ such that $\delta(\underline{x}, \underline{y}_1, \underline{y}_2) \vdash_T \underline{y}_1\underline{y}_2 \cap \underline{v} \neq \emptyset$, which is the same as (7).

For the vice versa, we suppose that (1) $\Rightarrow$ (2) holds. Consider $\delta(\underline{x}, \underline{z}, \underline{y})$ such that (16) holds. Then, we can find $\underline{v}(\underline{x})$ such that

$$\delta(\underline{x}, \underline{z}', \underline{y}') \wedge \delta(\underline{x}, \underline{z}'', \underline{y}'') \vdash_T (\underline{y}' \cap \underline{v} \neq \emptyset) \vee (\underline{y}'' \cap \underline{v} \neq \emptyset) \tag{19}$$

holds. Making the substitutions $\underline{z}' \mapsto \underline{z}, \underline{z}'' \mapsto \underline{z}, \underline{y}' \mapsto \underline{y}, \underline{y}'' \mapsto \underline{y}$, this gives precisely (17). □

Condition (iv) above can be interpreted as a 'generalized Beth property'. The situation becomes clearer in the simplified convex case; we first restate Proposition 4.4:

Proposition 4.4 *The following conditions are equivalent for a convex theory T having quantifier-free interpolation:*

(i) *T is equality interpolating;*

(ii) *T satisfies the implication* (3) $\Rightarrow$ (4) *(for every conjunctions of literals δ_1, δ_2);*

(iii) *for every pair $\underline{x}, \underline{z}$ of tuples of variables, for every further variable y and for every conjunction of literals $\delta(\underline{x}, \underline{z}, y)$ such that*

$$\delta(\underline{x}, \underline{z}', y') \wedge \delta(\underline{x}, \underline{z}'', y'') \vdash_T y' = y'' \ ,$$

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

$$\delta(\underline{x}, \underline{z}, y) \vdash_T y = v \ .$$

Proof. Again, we already know from Appendix B that (i) and (ii) are equivalent.

Assume that (iii) holds and consider $\delta_1(\underline{x}, \underline{z}_1, y_1), \delta_2(\underline{x}, \underline{z}_2, y_2)$ satisfying (3). Take $\delta(\underline{x}, \underline{z}_1, \underline{z}_2, y) := \delta_1(\underline{x}, \underline{z}_1, y) \wedge \delta_2(\underline{x}, \underline{z}_2, y)$. Now $\delta(\underline{x}, \underline{z}_1', \underline{z}_2', y') \wedge \delta(\underline{x}, \underline{z}_1'', \underline{z}_2'', y'')$ is

$$\delta_1(\underline{x}, \underline{z}_1', y') \wedge \delta_2(\underline{x}, \underline{z}_2', y') \wedge \delta_1(\underline{x}, \underline{z}_1'', y'') \wedge \delta_2(\underline{x}, \underline{z}_2'', y''),$$

hence (considering the first and the fourth conjunct) from (3) we get

$$\delta(\underline{x}, \underline{z}_1', \underline{z}_2', y') \wedge \delta(\underline{x}, \underline{z}_1'', \underline{z}_2'', y'') \vdash_T y' = y''.$$

By (iii), there is a term $v(\underline{x})$ such that

$$\delta_1(\underline{x}, \underline{z}_1, y) \wedge \delta_2(\underline{x}, \underline{z}_2, y) \vdash_T y = v(\underline{x}). \tag{20}$$

Again by (3), we obtain (after renamings)

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2 \wedge \delta_2(\underline{x}, \underline{z}_2, y_1) \ ;$$

thus (taking into account (20)) also

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2 \wedge y_1 = v(\underline{x})$$

and finally (4).

Vice versa, if (ii) holds and we have $\delta(\underline{x}, \underline{z}', y') \wedge \delta(\underline{x}, \underline{z}'', y'') \vdash_T y' = y''$, we can find $v(\underline{x})$ such that

$$\delta(\underline{x}, \underline{z}', y') \wedge \delta(\underline{x}, \underline{z}'', y'') \vdash_T y' = v \wedge y'' = v \ ;$$

applying the substitution $\underline{z}' \mapsto \underline{z}, \underline{z}'' \mapsto \underline{z}, \underline{y}' \mapsto \underline{y}, \underline{y}'' \mapsto \underline{y}$, this gives our claim $\delta(\underline{x}, \underline{z}, y) \vdash_T y = v$. $\square$

A *primitive* formula is obtained from a conjunction of literals by prefixing to it a string of existential quantifiers. We can reformulate the condition (iii) from Proposition 4.4 above as follows:

(iii)' for every tuple of variables $\underline{x}$, for every further variable y and for every *primitive* formula $\theta(\underline{x}, y)$ such that $\theta(\underline{x}, y') \wedge \theta(\underline{x}, y'') \vdash_T y' = y''$, there is a term $v(\underline{x})$ such that $\theta(\underline{x}, y) \vdash_T y = v$.

This is precisely *Beth definability property* [10], modulo T, for primitive formulæ. Hence equality interpolating coincides with this 'primitive Beth definability property' in the convex case.

To conclude, for the interested reader, we make some observations connecting the above result with the algebraically oriented literature (see [21] for a survey and for pointers to relevant papers). In an appropriate context from universal algebra, strong amalgamability is shown to be equivalent to the conjunction of amalgamability and of regularity of epimorphisms (alternatively: and of regularity of monomorphisms). In the same context, unravelling the definitions and using presentations of algebras as quotient of free ones, it is not difficult to realize that the primitive Beth definability property above is equivalent to regularity of monomorphisms. Thus, *our results perfectly match with the algebraic characterization of strong amalgamability.* Our approach, however, is *orthogonal* to algebraic and category-theoretic approaches: such approaches are able in fact to prove characterizations of strong amalgamability that work in abstract sufficiently complete/cocomplete categories, including consequently categories having nothing to do with models of first order theories. On the other hand, existence of minimal categorical structure fails in our context as soon as we go beyond the universal Horn case. Thus, the two approaches are incomparable and this is reflected by the different techniques employed (we mostly rely on diagrams and compactness, whereas the category-theoretic approach mostly exploit universal properties).

E Interpolation with free fuction symbols

In this paper, we treated quantifier free interpolation only with respect to variables, in the sense that we always considered all non variable symbols as shared symbols. This is not the notion of interpolation commonly used in verification, where also free functions and predicate symbols are not allowed to apper in the interpolants in case they do not occur in both the formulæ to be interpolated. We show here that this more general notion of quantifier free interpolation can be reduced to combined interpolation and thus that it is equivalent to strong sub-amalgamability too.

Definition E.1. Let T be a theory in a signature Σ; we say that T has the *general quantifier-free interpolation property* iff for every signature Σ' (disjoint from Σ) and for every finite sets of ground $\Sigma \cup \Sigma'$-formulæ A, B such that $A \wedge B$ is T-unsatisfiable,[10] there is a ground formula θ such that: (i) A T-entails θ; (ii) $\theta \wedge B$ is T-unsatisfiable; (iv) all predicate, constants and function symbols from Σ' occurring in θ occur also in A and in B.

Notice that the above definition becomes equivalent to the definition of quantifier free interpolation property introduced in Section 2 if we restrict it to the signatures Σ' containing only constant function symbols. One may wonder whether Definition E.1 is the same as asking for quantifier free interpolation for all combined theoris $T \cup \mathcal{EUF}(\Sigma')$; at a first glance, it does not seem to be so because in Definition E.1 we require also that the function and the predicate symbols from Σ' not occurring in both A, B do not occur in θ either. We shall see however that such symbols are immaterial because they can be removed.

Let us fix a theory T in a signature Σ and let Σ' be a further signature (disjoint from Σ). A finite set A of ground $\Sigma \cup \Sigma'$-formulæ is said to be Σ_0-*flat* (for some $\Sigma_0 \subseteq \Sigma'$) iff A is of the kind $A_0 \cup A_1$, where A_1 does not contain Σ_0-symbols and A_0 is a set of literals of the kind

$$f(a_1, \dots, a_n) = b, \quad P(a_1, \dots, a_n), \quad \neg P(a_1, \dots, a_n)$$

where $f, P \in \Sigma_0$ and $a_1, \dots, a_n, b$ are constants not in Σ_0.

Lemma E.2. *Let T, Σ, Σ' be as above and let the finite set of ground $\Sigma \cup \Sigma'$-formulæ A be Σ_0-flat (for some $\Sigma_0 \subseteq \Sigma'$). Then it is possible to find a finite set of ground formulæ $A^{-\Sigma_0}$ such that: (i) $A^{-\Sigma_0}$ does not contain Σ_0-symbols; (ii) A T-entails $A^{-\Sigma_0}$; (iii) $A^{-\Sigma_0}$ is T-satisfiable iff A is T-satisfiable.*

[10] By this (and similar notions) we mean that $A \wedge B$ is unsatisfiable in all Σ'-structures whose Σ-reduct is a model of T. We use the same convention as in Section 5 and indicate with the letters A, B both a finite set of ground formulæ and its conjunction.

Proof. Let A be $A_0 \cup A_1$ as prescribed in the definition of Σ_0-flatness. We take as $A^{-\Sigma_0}$ the set of ground formulæ $A'_0 \cup A_1$ where A'_0 is built as follows. For every function symbol $f \in \Sigma_0$ and for every pair of atoms $f(a_1, \dots, a_n) = b, f(a'_1, \dots, a'_n) = b'$ belonging to A_0 we include in A'_0 the ground clause

$$a_1 = a'_1 \wedge \cdots \wedge a_n = a'_n \rightarrow b = b'; \tag{21}$$

similarly, for every predicate symbol $P \in \Sigma_0$ and for every pair of literals $P(a_1, \dots, a_n)$, $\neg P(a'_1, \dots, a'_n)$ belonging to A_0 we include in A'_0 the ground clause

$$a_1 = a'_1 \wedge \cdots \wedge a_n = a'_n \rightarrow \bot. \tag{22}$$

That $A'_0 \cup A_1$ enjoys properties (i)-(ii) is clear; it remains to show that if it is T-satisfiable, so is $A_0 \cup A_1$.[11] Suppose indeed that $\mathcal{M}$ is a $\Sigma \cup (\Sigma' \setminus \Sigma_0)$-model of T in which $A'_0 \cup A_1$ is true. We expand $\mathcal{M}$ to a $\Sigma \cup \Sigma'$-structure as follows. Let $f \in \Sigma_0$ have arity n and let $c_1, \dots, c_n$ be elements from the support of $\mathcal{M}$; then $f^{\mathcal{M}}(c_1, \dots, c_n)$ is arbitrary, unless there are $f(a_1, \dots, a_n) = b \in A_0$ such that $c_1 = a_1^{\mathcal{M}}, \dots, c_n = a_n^{\mathcal{M}}$: in this case, we put $f^{\mathcal{M}}(c_1, \dots, c_n)$ to be equal to $b^{\mathcal{M}}$. Since $\mathcal{M}$ is a model of the clauses (21), the definition is correct. Similarly, if $P \in \Sigma_0$ has arity n, then $P^{\mathcal{M}}$ is the set of n-tuples $c_1, \dots, c_n$ of elements from the support of $\mathcal{M}$ such that there exists $P(a_1, \dots, a_n) \in A_0$ such that $c_1 = a_1^{\mathcal{M}}, \dots, c_n = a_n^{\mathcal{M}}$. The literals from A_0 turns out to be all true by construction and because in $\mathcal{M}$ the clauses (22) hold. □

Theorem E.3. *T has the general quantifier free interpolation property iff it is strongly sub-amalgamable iff it is equality interpolating.*

Proof. Since the general quantifier free interpolation property for T implies the (ordinary) quantifier free interpolation property for all the theories $T \cup \mathcal{EUF}(\Sigma')$, it is clear from Theorem 3.5 that the general quantifier free interpolation property implies strong sub-amalgamability. To show the vice versa, we use our metarules and Lemma E.2 above.

Let Σ be the signature of T and let Σ' be disjoint from Σ; fix also finite sets of ground $\Sigma \cup \Sigma'$-formulae A, B such that $A \wedge B$ is T-unsatisfiable. Let Σ_A be the set of predicate and (non constant) function symbols from Σ' that occur in A but not in B; similarly, let Σ_B be the set of predicate and (non constant) function symbols from Σ' that occur in B but not in A. We show how to transform A into a Σ_A-flat $\tilde{A}$ by using metarules (a similar transformation is applied to B to get a Σ_B-flat $\tilde{B}$). Using metarules (Define1), (Redplus1), (Redminus1) we can add 'defining atoms' $f(a_1, \dots, a_n) = a$ (with fresh a) and replace all occurrences of the term $f(a_1, \dots, a_n)$ in A by a; if we do it repeatedly, A gets flattened, in the sense that function and predicate symbols (different from identity) in A are always applied to constants. With the

[11] The right-to-left side of (iii) is a consequence of (ii).

same technique, we can transform A into a conjunction of defining atoms and ground formulæ in which function symbols from Σ_A do not occur. To take care of predicate symbols $P \in \Sigma_A$, we need guessings and metarule (Disjunction1): for every atom $P(a_1, \dots, a_n)$ occurring in A, we add either $P(a_1, \dots, a_n)$ or $\neg P(a_1, \dots, a_n)$ to A and replace $P(a_1, \dots, a_n)$ with $\top$ or $\perp$, respectively (notice that because of such guessings the transformation from A, B to $\tilde{A}, \tilde{B}$ may be non-deterministic). Since metarules are satisfiability-preserving and are endowed with recursive instructions for computation of interpolants, it will be sufficient to find a desired interpolant θ for $\tilde{A}$ and $\tilde{B}$.

If we apply the tranformations of Lemma E.2 to $\tilde{A} \cup \tilde{B}$ we can get $(\tilde{A} \cup \tilde{B})^{-\Sigma_A} \equiv \tilde{A}^{-\Sigma_A} \cup \tilde{B}$ with the properties (i)-(iii) stated in that Lemma: in particular, function and predicate symbols from Σ_A do not occur anymore in $\tilde{A}^{-\Sigma_A}$. We do the same for $\tilde{B}$ and eventually we get $\bar{A}, \bar{B}$ such that (a) $\bar{A} \cup \bar{B}$ is T-unsatisfiable; (b) $\tilde{A}$ T-entails $\bar{A}$, $\tilde{B}$ T-entails $\bar{B}$; (c) all predicate and (non constant) functions symbols occurring in $\bar{A}$ occur also in $\bar{B}$ and vice versa. Let Σ_C be the set of predicate and (non constant) function symbols occurring in both $\bar{A}$ and $\bar{B}$. Since T is strongly amalgamable, by Theorem 3.5, $T \cup \mathcal{EUF}(\Sigma_C)$ has the quantifier-free interpolation property.[12] Thus, there exists a ground formula θ containing, besides interpreted symbols from Σ, only predicate and function symbols from Σ_C, as well as individual free constants occurring both in $\bar{A}$ and in $\bar{B}$, such that $\bar{A}$ T-entails θ and $\bar{B} \wedge \theta$ is T-inconsistent. By (b) above, we get that $\tilde{A}$ T-entails θ and $\tilde{B} \wedge \theta$ is T-inconsistent, thus θ is the desired interpolant.

The equivalence between strong sub-amalgamability and equality interpolating property comes from Theorem 4.2. □

[12] The proof of the right-to-left side of that Theorem does not need the requirement that Σ_C has at least a unary predicate symbol.

F A counterexample: golden cuff links

Here we show by exhibiting a formal counterexample that the 'convex' formulation of the equality interpolating property is not sufficient to guarantee the modularity of quantifier-free interpolation for non-convex theories. Intuitively, the reason is that disjunctions of equalities must be propagated in the non convex case and the convex formulation of the equality interpolation property does not say anything about them. This Appendix can be read independently on the remaining part of the Technical Report.

We say that a theory T has the *YMc property* ('convex Yorsh-Musuvathi property') iff it has quantifier-free interpolation property and moreover the implication (3) $\Rightarrow$ (4) holds, i.e. for every pair y_1, y_2 of variables and for every pair of conjunctions of literals $\delta_1(\underline{x}, \underline{z}_1, y_1), \delta_2(\underline{x}, \underline{z}_2, y_2)$ such that

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2$$

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

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v \wedge y_2 = v.$$

To build our counterexample, we introduce a theory CL which is meant to describe a set of *cuff links*, containing at most one pair of *golden* cuff links. Formally, in the signature Σ_{CL} of CL we have a unary function symbol $(-)'$ and a unary predicate G.[13] The axioms of CL say that $(-)'$ denotes the 'twin' cuff link

$$\forall x.\ x'' = x, \quad \forall x.\ x \neq x'$$

that twin cuff links are both golden or not

$$\forall x.\ G(x) \leftrightarrow G(x')$$

and that there is at most one pair of golden cuff links:

$$\forall x \forall y.\ G(x) \wedge G(y) \rightarrow x = y \vee x' = y. \tag{23}$$

Lemma F.1. *CL has the quantifier free interpolation property, because it has the sub-amalgamation (but not the strong sub-amalgamation) property.*

Proof. That the sub-amalgamation property holds is quite clear: suppose we are given models $\mathcal{M}_1, \mathcal{M}_2$ of CL sharing the substructure $\mathcal{A}$ (as a side remark, notice that $\mathcal{A}$ is also a model of CL because CL is universal). As usual, we assume that the intersection of the supports

[13] A free constant c_0 is added to the signature Σ_{CL} to prevent it from being empty.

of $\mathcal{M}_1$ and $\mathcal{M}_2$ is the support of $\mathcal{A}$. To amalgamate $\mathcal{M}_1, \mathcal{M}_2$ over $\mathcal{A}$, it is sufficient to take the union of the supports of $\mathcal{M}_1$ and $\mathcal{M}_2$, with just one proviso: if $\mathcal{M}_1, \mathcal{M}_2$ both contain a pair of golden cuff links that is not from $\mathcal{A}$, then such pairs must be merged (the need of such merging is precisely what shows that strong sub-amalgamation fails). □

Proposition F.2. *CL has the YMc property.*

Proof. We shall work with free constants (instead of with variables). Consider finite sets of ground literals A, B in the signature Σ_{CL} enriched with additional free constants (let Σ_A be the signature of A and Σ_B be the signature of B). We call AB-common the ground terms built up from free constants occurring both in A and in B; ground terms built up from constants occurring in A but not in B are called A-strict (B-strict ground terms are defined symmetrically). We call a ground term or literal *pure* iff it is either from Σ_A or from Σ_B. We argue by contraposition. Suppose that, for an A-strict constant a and a B-strict constant b, there is no AB-common ground term t such that $A \cup B \vdash_{CL} t = a \wedge t = b$; we show that $A \cup B \not\vdash_{CL} a = b$ by exhibiting a $\Sigma_A \cup \Sigma_B$-model $\mathcal{M}$ of CL such that $\mathcal{M} \models A, \mathcal{M} \models B$ and $\mathcal{M} \not\models a = b$.

We can freely make further assumptions on our A, B: first, we can assume that there is at least one AB-common ground term.[14], that terms like d'' do not occur in $A \cup B$, [15] and that if a term occurs in $A \cup B$, so does its twin term (here the twin of a constant d is d' and the twin of d' is d).[16] Second, since the number of Σ_A-ground literals is finite (modulo the identification of a term like t'' with t), we can assume that if a Σ_A-ground literal is entailed (modulo CL) by $A \cup B$, then it actually occurs in A (and similarly for B): the addition of such entailed literals does not in fact compromize our claim. So, let us make the above assumptions. Notice that (since there is at least one ground AB-common term), our hypotheses imply that $A \cup B$ is CL-consistent, so no pair of contradictory literals can be there.

We can divide the ground terms occurring in $\Sigma_A \cup \Sigma_B$ into equivalence classes (similarly to what happens in congruence closure algorithms), according to the equivalence relation that holds among d_1 and d_2 iff (i) either they both occur in A and $d_1 = d_2 \in A$, or (ii) they both occur in B and $d_1 = d_2 \in B$, or (iii) d_1 is A-strict, d_2 is B-strict and there exists an AB-common t such that $d_1 = t \in A$, $d_2 = t \in B$, or (iv) d_1 is B-strict, d_2 is A-strict and there exists an AB-common t such that $d_1 = t \in B$, $d_2 = t \in A$. Notice that, because of our assumptions, the equivalence class of a is different from the equivalence class of b.

Since there are no contradictory literals in $A \cup B$, we can build a $\Sigma_A \cup \Sigma_B$-structure $\mathcal{A}$ in which all literals from $A \cup B$ are true: the support of $\mathcal{A}$ is formed by the above equivalence

[14] Because Σ_{CL} has one.

[15] Because they simplify to d.

[16] To ensure the latter, we can just add literals like $d' = d'$ to A or B, if needed.

classes, a free constant is interpreted as the equivalence class it belongs to, the twin C' of an equivalence class C is the equivalence class formed by the twin terms of the terms belonging to C; moreover, C is a golden cuff link in $\mathcal{A}$ iff $G(t) \in A \cup B$ (here t is any term belonging to C). Notice that $\mathcal{A} \not\models a = b$. However, we are not done, because $\mathcal{A}$ may not be a model of CL: the reason is that there might be more than one golden pair of cuff links. We now show how to merge all golden pairs of cuff links of $\mathcal{A}$ and get a model $\mathcal{M}$ of CL having the required properties, namely such that $\mathcal{M} \models A, \mathcal{M} \models B$ and $\mathcal{M} \not\models a = b$.

Consider two different pairs of golden cuff links C, C' and D, D' (when we say that they are different as pairs of cuff links, we mean that C is different from both D and D'). We claim that if we merge C with D and C' with D' as equivalence classes (i.e. if we identify them as elements from the support of $\mathcal{A}$), we still have that the literals from A and B are true. In fact, this could possibly be not the case if there are $t \in C, u \in D$ such that $t \neq u \in A \cup B$. However, literals in $A \cup B$ are all pure, so that either $t, u \in \Sigma_A$ or $t, u \in \Sigma_B$. Suppose $t, u \in \Sigma_A$ (the other case is symmetric); by the construction of $\mathcal{A}$ and since C, D are golden, we have that $G(t), G(u) \in A \cup B$ and hence (by (23)) the entailed literal $t = u'$ belongs to A, so that $C = D'$ which means that C, C' and D, D' are not different pairs of cuff links.

In conclusion, whenever we pick two different pairs of golden cuff links C, C' and D, D' from the support of $\mathcal{A}$, we can merge C with D and D with D', without compromizing the truth of $A \cup B$; notice, however, that we can make the symmetric operation and merge C with D' and C' with D, again keeping the literals in $A \cup B$ true. In the end, we can merge all golden pairs of cuff links into a single one; if a and b belong to C and D, respectively, and if C, D are both golden, we can choose the appropriate merging among the two possible ones, so that in the end we have that D is equal to C', which implies that a and b remains interpreted as different elements in the support of the final model. □

From the above results and Theorem 3.5 , we obtain:

Corollary F.3. *CL has the quantifier-free interpolation property and the YMc property, but $CL \cup \mathcal{EUF}$ does not have the quantifier free interpolation property (if the signature of $\mathcal{EUF}$ has at least a unary predicate symbol).*

A direct counterexample to the quantifier-free interpolation property for the combined theory $CL \cup \mathcal{EUF}$ can be easily obtained by considering the following mutually unsatisfiable sets of ground literals

$$A := \{G(a), P(a), P(a')\}, \qquad B := \{G(b), \neg P(b), \neg P(b')\}$$

(here P is the extra free predicate).