# From Strong Amalgamability to Modularity of Quantifier-Free Interpolation

Roberto Bruttomesso,[1] Silvio Ghilardi,[1] and Silvio Ranise[2]

[1] Università degli Studi di Milano, Milan, Italy
[2] FBK (Fondazione Bruno Kessler), Trento, Italy

**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 handling both convex and non-convex theories, that subsumes and extends most existing work on combined interpolation.

## 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 that are 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 thoroughly 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-$n$ [5] 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]). For proofs and additional information on related topics, see the Appendixes of the Technical Report [8].

## 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 $(\Sigma, Ax_T)$, where $\Sigma$ is a signature and $Ax_T$ is a set of $\Sigma$-sentences, called the *axioms* of $T$ (we shall sometimes write directly $T$ for $Ax_T$). The *models* of $T$ are those $\Sigma$-structures in which all the sentences in $Ax_T$ are true. A $\Sigma$-formula $\phi$ is *$T$-satisfiable* if there exists a model $\mathcal{M}$ of $T$ such that $\phi$ is true in $\mathcal{M}$ under a suitable assignment $\mathtt{a}$ to the free variables of $\phi$ (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, $\varphi$ 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 $\varphi_1$ *$T$-entails* a formula $\varphi_2$ 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 $\Sigma$-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 $\phi, \psi$ such that $\psi \wedge \phi$ is $T$-unsatisfiable, there exists a quantifier-free formula $\theta$, called an *interpolant*, such that: (i) $\psi$ $T$-entails $\theta$, (ii) $\theta \wedge \phi$ is $T$-unsatisfiable, and (iii) only the variables occurring in both $\psi$ and $\phi$ occur in $\theta$. A theory admitting quantifier elimination also admits quantifier-free interpolation. A more general notion of quantifier-free interpolation property, involving also free function symbols, is analyzed in an Appendix of the extended version [8].

**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 *$\Sigma$-embedding* (or, simply, an embedding) between two $\Sigma$-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 $\delta$, 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 $\Sigma_i$ such that the $SMT(T_i)$ problem is decidable for $i = 1, 2$ and $\Sigma_1$ and $\Sigma_2$ 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 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 $\mu_1, \mu_2$ 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 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 1.** *A theory $T$ has the sub-amalgamation property iff it has quantifier-free interpolants.*

A theory admitting quantifier elimination has the sub-amalgamation property: this follows, e.g., from Theorem 1 above. On the other hand, quantifier elimination is not sufficient to guarantee the strong sub-amalgamation property. In fact, from Theorem 3 below and the counterexample given in [4], it follows that Presburger arithmetic does not have the strong sub-amalgamation property, even if we add congruences modulo $n$ to the language. 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 $\Sigma$, let $\mathcal{EUF}(\Sigma)$ be the pure equality theory over $\Sigma$. 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

[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 uncomfortable with empty domains can assume that signatures always contain an individual constant.

$\mathcal{M}_2$ sharing a substructure $\mathcal{M}_0$ as follows. Without 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 $\mathcal{AX}_{\mathtt{diff}}$ of the theory of arrays considered in [7,9]. The signatures of $\mathcal{AX}_{\text{ext}}$ and $\mathcal{AX}_{\mathtt{diff}}$ contain the sort symbols $\mathtt{ARRAY}, \mathtt{ELEM}$, and $\mathtt{INDEX}$, and the function symbols $rd : \mathtt{ARRAY} \times \mathtt{INDEX} \longrightarrow \mathtt{ELEM}$ and $wr : \mathtt{ARRAY} \times \mathtt{INDEX} \times \mathtt{ELEM} \longrightarrow \mathtt{ARRAY}$. The signature of $\mathcal{AX}_{\mathtt{diff}}$ also contains the function symbol $\mathtt{diff} : \mathtt{ARRAY} \times \mathtt{ARRAY} \longrightarrow \mathtt{INDEX}$. 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 $\mathcal{AX}_{\mathtt{diff}}$ is obtained from that of $\mathcal{AX}_{\text{ext}}$ by replacing the third axiom with its Skolemization:

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

In [9], it is shown that $\mathcal{AX}_{\mathtt{diff}}$ 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 literature): 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 1 to $\mathcal{EUF}(\Sigma)$, $\mathcal{AX}_{\mathtt{diff}}$, 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 $\mathcal{AX}_{\mathtt{diff}}$ [7, 9] have 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 admitting quantifier-free interpolation may not admit quantifier-free interpolation. For example, the union of the theory $\mathcal{EUF}(\Sigma)$ and Presburger arithmetic does not admit quantifier-free interpolation [4]. Fortunately, strong sub-amalgamation is modular when combining stably infinite theories.

**Theorem 2.** *Let $T_1$ and $T_2$ be two stably infinite theories over disjoint signatures $\Sigma_1$ and $\Sigma_2$. If both $T_1$ and $T_2$ have the strong sub-amalgamation property, then so does $T_1 \cup T_2$.*

Theorems 1 and 2 obviously imply that strong sub-amalgamation is sufficient for the modularity of quantifier-free interpolation for stably infinite theories.

**Corollary 1.** *Let $T_1$ and $T_2$ be two stably infinite theories over disjoint signatures $\Sigma_1$ and $\Sigma_2$. 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.** *Let $T$ be a theory admitting quantifier-free interpolation and $\Sigma$ be a signature disjoint from the signature of $T$ and containing at least a unary predicate symbol. Then, $T \cup \mathcal{EUF}(\Sigma)$ admits quantifier-free interpolation iff $T$ has the strong sub-amalgamation property.*

Although Corollary 1 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 $\underline{t} \equiv t_1, \dots, t_n$ and $\underline{v} \equiv v_1, \dots, v_m$ of terms,

the notation $\underline{t} \cap \underline{v} \neq \emptyset$ 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 2.** *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*

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

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

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

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

**Theorem 4.** *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 so that, by using Corollary 1, we can establish when combinations of theories admit quantifier-free interpolation. To ease the application of Theorem 4, we first study the relationship between quantifier-elimination and equality interpolation for universal theories.

**Theorem 5.** *A universal theory admitting quantifier elimination is equality interpolating.*

Interestingly, the proof of this theorem (see [8]) 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 2*; 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:

$$\begin{array}{rr} \forall x, y. car(cons(x, y)) = x, & \forall x, y. cdr(cons(x, y)) = y, \\ \forall x, y. cons(car(x), cdr(x)) = x, & \forall x. x \neq t(x) \end{array} \tag{CCC}$$

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:

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

$\mathcal{IDL}$ 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 $n > 1$). The term $x\ div[n]$ (which is new with respect to the language of Presburger arithmetic) represents the unique $q$ such that $x = qn + r$ for some $r = 0, \ldots, n - 1$. 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 5, 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 [8] 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 $x\ rem[n] = y\ rem[n]$.

By Theorem 5, $\mathcal{RDS}$, $\mathcal{IDL}$, and $\mathcal{LAI}$ are equality interpolating. In [8], the theory $\mathcal{UTVPI}$ of Unit-Two-Variable-Per-Inequality (see, e.g., [11]) is shown to be also equality interpolating via Theorem 5.

### 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 1.** *A convex theory $T$ admitting 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) ⇒ (4) is exactly the definition of equality interpolation in [32]. In the following, a convex quantifier-free interpolating theory satisfying (3) ⇒ (4) will be called *YMc equality interpolating.* By Proposition 1, an YMc equality interpolating (convex) theory is also equality interpolating according to Definition 2. 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 1, we conclude that $\mathcal{LST}$ is equality interpolating in the sense of Definition 2. 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 1, $\mathcal{LAR}$ is equality interpolating in the sense of Definition 2. The same result can be obtained from Theorem 5 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) ⇒ (4) of Proposition 1 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 and 4 (and the combination result of [32]), if a theory $T$ is YM equality interpolating, it is also equality interpolating in the sense of Definition 2. 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 $p_i \to Q(c_i)$ 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 2.

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 1 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 together with Corollary 1 allow us to establish the quantifier-free interpolation property for all those theories obtained by combining a theory axiomatizing a container data structure (such as $\mathcal{EUF}$, $\mathcal{RDS}$, $\mathcal{LST}$, or $\mathcal{AX}_{\mathtt{diff}}$) 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 $\Sigma$ and set $A$ of formulae, $\Sigma^A$ denotes the signature $\Sigma$ expanded with the free constants occurring in $A$. Let $A$ and $B$ be two finite sets of ground formulae in the signatures $\Sigma^A$ and $\Sigma^B$, respectively, and $\Sigma^C := \Sigma^A \cap \Sigma^B$. Given a term, a literal, or a formula $\varphi$ we call it:

- *AB-common* iff it is defined over $\Sigma^C$;
- *A-local* (resp. *B-local*) if it is defined over $\Sigma^A$ (resp. $\Sigma^B$);
- *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 [7, 9] (to which, the interested reader is pointed for more details) and used also in [6] 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

[2] In [6, 7, 9], metarules manipulate pairs of finite sets of literals instead of ground formulae; the difference is immaterial.

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

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

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

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

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

**Table 1.** Interpolating Metarules (taken from [7, 9]): each rule has a proviso *Prov.* and an instruction *Instr.* for recursively computing the new interpolant $\phi'$ from the old one(s) $\phi, \phi_1, \ldots, \phi_k$. Metarules are applied *bottom-up* and interpolants are computed *top-down*. Notation $\phi(t/a)$ is used for substitution.

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 constant 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 $\phi'$ by eliminating the fresh constant $a$ from the recursively known interpolant $\phi$.

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 Table 1). 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:

**Proposition 2 ([7,9]).** *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 $\phi$ such that $A_0 \vdash \phi$ and $B_0 \wedge \phi \vdash \bot$). The interpolant $\phi$ 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 2.

### 5.2 A quantifier-free interpolating algorithm

Let $T_i$ be a stably-infinite and equality interpolating theory over the signature $\Sigma_i$ 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 2 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 [8] 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 $\Sigma_1$- nor $\Sigma_2$-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 $\phi$ 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 $\Sigma_1$- or of only $\Sigma_2$-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 $\Sigma_1^A$- and $\Sigma_2^A$-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 $\Sigma_1$- and $\Sigma_2$-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 [7]. 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 $B$ instead 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). Intuitively, the reason why Term Sharing works is because, in the end, the new constant will have to be replaced with the $AB$-common term $t$, so the interpolant is not affected by the renaming of $a$ to $c$.

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 $\alpha$ to relevant $A$-atoms satisfying $A$, seen as a set of propositional formulæ (relevant $B$-atoms and $B$-assignments are defined similarly). Below, we use the notation $\alpha$ to denote both the assignment $\alpha$ and the set of literals satisfied by $\alpha$.

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 $\perp \notin A \cup B$, use the interpolation algorithm for $T_i$ to find a ground $AB$-common $\theta$ such that $A_i \vdash_{T_i} \theta$ and $\theta \wedge B_i \vdash_{T_i} \perp$; then add $\theta$ and $\perp$ to $B$.

**Decide$_X$:** if there is no $X$-assignment $\alpha$ such that $\alpha \subseteq X$, pick one of them (if there are none, add $\perp$ 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 $\Sigma_i$-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 $\Sigma_i$-formula $\theta$ 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 $\theta$ 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 unsatisfiable 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 6.** *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 [3].

## 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 $\mathcal{LAI}$ [5], $\mathcal{LST}$ with $\mathcal{LAR}$ [32], and $\mathcal{AX}_{\mathtt{diff}}$ with $\mathcal{IDL}$ [6]. 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 combination method in [15] 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] as suggested in [29]. 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 [3]. Theorem 6 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 [5, 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 developments.

## 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. 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.
4. A. Brillout, D. Kroening, P. Rümmer, and T. Wahl. An Interpolating Sequent Calculus for Quantifier-Free Presburger Arithmetic . In *IJCAR*, 2010.
5. 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.
6. 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.
7. R. Bruttomesso, S. Ghilardi, and S. Ranise. Rewriting-based Quantifier-free Interpolation for a Theory of Arrays. In *RTA*, 2011.

8. R. Bruttomesso, S. Ghilardi, and S. Ranise. From Strong Amalgamation to Modularity of Quantifier-Free Interpolation. Technical Report RI 337-12, Dip. Scienze dell'Informazione, Univ. di Milano, 2012.
9. R. Bruttomesso, S. Ghilardi, and S. Ranise. Quantifier-Free Interpolation of a Theory of Arrays. *Logical Methods in Computer Science*, 2012. to appear.
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.