Word equations over graph products

Volker Diekert and Markus Lohrey

Universität Stuttgart,
Institut für Formale Methoden der Informatik (FMI)
Universitätsstr. 38, D-70569 Stuttgart, Germany
diekert,lohrey@informatik.uni-stuttgart.de

Abstract. For monoids that satisfy a weak cancellation condition, it is shown that the decidability of the existential theory of word equations is preserved under graph products. Furthermore, it is shown that the positive theory of a graph product of groups can be reduced to the positive theories of those factors, which commute with all other factors, and the existential theories of the remaining factors. Both results also include suitable constraints for the variables. Larger classes of constraints lead in many cases to undecidability results.

1 Introduction

Since the seminal work of Makanin [36] on equations in free monoids, the decidability of various theories of equations in different monoids and groups has been studied, and several new decidability and complexity results have been shown. Let us mention here the results of [50, 59] for free monoids, [12, 26, 31, 37, 38] for free groups, [17] for free partially commutative monoids (trace monoids), [18] for free partially commutative groups (called semifree groups in [2, 3], right-angled Artin groups in [8], and graph groups in [20]), [14] for plain groups (free products of finite and free groups), [11, 54, 60] for (relatively) torsion-free hyperbolic groups, [34] for virtually-free groups and certain HNN-extensions and amalgamated free products, and [30] for groups with a free regular length function.

In this paper, we will continue this stream of research by considering graph products of monoids (Section 2.3). The graph product construction is a well-known construction in mathematics, see e.g. [25, 27], that generalizes both free products and direct products: An independence relation on the factors of the graph product specifies, which monoids are allowed to commute elementwise. Section 3 deals with existential theories of graph products. Using a general closure result for existential theories (Theorem 3.4), we will show in Section 3.3 that under some algebraic restriction on the factors of a graph product ($ab = 1 = ac$ or $ba = 1 = ca$ has to imply $a = c$) the decidability of the existential theory of word equations is preserved under graph products (Theorem 3.12). This transfer theorem remains also valid if we allow constraints for variables, which means that the value of a variable may be restricted to some specified set. More precisely, we will define an operation, which, starting from a class of constraints for each factor monoid of the graph product, constructs a class of constraints for the graph product. This construction is inspired by the notion of bipartite automata, which was introduced by Sakarovitch [57, 58] in order to study rational sets in free products. We will also present an upper bound for the space complexity of the existential theory of the graph product in terms of the space complexities for the existential theories of the factor monoids. Using known results from [34, 54, 60] it follows that the existential theory of word equations of a graph product of finite monoids, free monoids, virtually-free groups, and torsion-free hyperbolic groups is decidable. This result generalizes the decidability result for graph products of finite monoids, free monoids, and free groups groups from [16].

In Section 4 we will investigate positive theories of equations (a sentence is called positive if it is constructed from atomic formulas using only conjunctions, disjunctions, and quantifiers). We prove that the positive theory of word equations of a graph product of *groups* with recognizable constraints can be reduced to

– the positive theories with recognizable constraints of those factors of the graph product that are allowed to commute elementwise with all the other factors and
– the existential theories of the remaining factors.

As a corollary we obtain the decidability of the positive theory of a graph product of finite and free groups with recognizable constraints. This generalizes the well-known result of Makanin for free groups [37, 38]. The technical part relies on a generalization of the techniques introduced by Merzlyakov for free groups [43]. Our decision method leads only to a nonelementary algorithm for the positive theory, but additional restrictions on the graph underlying the graph product give us an elementary upper bound.

Our transfer theorems for graph products should be compared with similar results for existential and positive theories of HNN-extensions and amalgamated free products (see [35] for a definition of these constructions) from [34].

A short version of this paper appeared in [15].

2 Preliminaries

For a binary relation $\to$ on some set, we denote by $\xrightarrow{+}$ ($\xrightarrow{*}$) the transitive (reflexive and transitive) closure of $\to$. Let A be an alphabet (finite or infinite). The empty word over A is denoted by ε. The length of a word $s \in A^*$ is $|s|$, the set of symbols from A that occur in s is $\mathrm{alph}(s)$.

An *involution* ι on A is a function $\iota : A \to A$ such that $\iota(\iota(a)) = a$ for all $a \in A$. The involution may have fixpoints, i.e., $\iota(a) = a$. A *monoid involution* on a monoid $\mathcal{M} = (M, \circ, 1)$ is an involution $\iota : M \to M$ such that $\iota(a \circ b) = \iota(b) \circ \iota(a)$ for all $a, b \in M$ and $\iota(1) = 1$. A *partial monoid involution* on a monoid $\mathcal{M}$ is given by a submonoid $\mathcal{I}$ of $\mathcal{M}$ together with a monoid involution $\iota : \mathcal{I} \to \mathcal{I}$.

We assume some familiarity with computational complexity, see e.g. the textbook [49] for more details.

2.1 Mazurkiewicz traces

For a detailed introduction into trace theory see [19]. An *independence alphabet* is a pair (A, I), where A is a possibly infinite set and $I \subseteq A \times A$ is symmetric and irreflexive. The relation I is known as the *independence relation*, its complement $D = (A \times A) \setminus I$ is the *dependence relation*. The pair (A, D) is called a *dependence alphabet*. For $a \in A$, we let $I(a) = \{b \in A \mid (a, b) \in I\}$ and $D(a) = \{b \in A \mid (a, b) \in D\} = A \setminus I(a)$. An (A, I)-*clique* is a subset $B \subseteq A$ such that $(a, b) \in I$ for all $a, b \in B$ with $a \neq b$. Let $\mathcal{F}(A, I)$ denote the set of all *finite* (A, I)-cliques. Let $\equiv_I$ be the smallest congruence on A^* that contains all pairs of the form (ab, ba) with $(a, b) \in I$. The *trace monoid (free partially commutative monoid)* $\mathbb{M}(A, I)$ associated to (A, I) is the quotient monoid $A^*/\equiv_I$; its elements are called *traces*. Since A may be infinite, we do not restrict to finitely generated trace monoids. Extreme cases are *free monoids* (if $D = A \times A$) and *free commutative monoids* (if $D = \{(a, a) \mid a \in A\}$). Trace monoids were first investigated in [10]. Mazurkiewicz [40] introduced them in computer science.

The trace represented by the word $s \in A^*$ is denoted by $[s]_I$. The neutral element of $\mathbb{M}(A, I)$ is the empty trace $[\varepsilon]_I$, briefly ε. An element $a \in A$ will be identified with the trace $[a]_I$. More generally, for a finite (A, I)-clique C, we can define a unique trace $[C]_I = [a_1 a_2 \cdots a_n]_I$, where $a_1, a_2, \ldots, a_n$ is an arbitrary enumeration of C. We will omit the subscript I if the independence relation is clear from the context.

Let $t = [s]_I \in \mathbb{M}(A, I)$. We define $|t| = |s|$ (the length of t), $\mathrm{alph}(t) = \mathrm{alph}(s)$, $\max(t) = \{a \in A \mid \exists u \in A^* : t = [ua]_I\}$, and $\min(t) = \{a \in A \mid \exists u \in A^* : t = [au]_I\}$. Note that $\min(t)$ and $\max(t)$ are (A, I)-cliques. For two traces $t, u \in \mathbb{M}(A, I)$ we write $(t, u) \in I$ if $\mathrm{alph}(t) \times \mathrm{alph}(u) \subseteq I$.

Let f be a partially defined function on A with $\mathrm{dom}(f) = B \subseteq A$. We say that f is *compatible* with I if $(a, b) \in I \cap (B \times B)$ implies $(f(a), f(b)) \in I$. This allows us to lift f to a partially defined function on $\mathbb{M}(A, I)$ by setting $f([a_1 \cdots a_n]_I) = [f(a_n) \cdots f(a_1)]_I$. The domain of this lifting is $\mathbb{M}(B, I)$. Note that we reverse the order of the symbols in the f-image of a trace. In our applications, f will be always a partial injection on A like for instance an involution $\iota : B \to B$ that is defined on a subset $B \subseteq A$. In this case, the lifting of ι to $\mathbb{M}(A, I)$ is a partial monoid involution on $\mathbb{M}(A, I)$ with domain $\mathbb{M}(B, I)$. The structure $(\mathbb{M}(A, I), \iota)$ is also called a *trace monoid with partial involution*. Assume that $\iota_j : A_j \to A_j$ ($j \in \{1, 2\}$) is a partially defined involution, and $I_j \subseteq A_j \times A_j$ ($j \in \{1, 2\}$) is an independence relation. Moreover, let $g : A_1 \to A_2$. We define $g(I_1) = \{(g(a), g(b)) \mid (a, b) \in I_1\}$. If ι_j is compatible with I_j, $g(I_1) \subseteq I_2$, and $g(\iota_1(a)) = \iota_2(g(a))$ for all $a \in \mathrm{dom}(\iota_1)$, then g can be uniquely lifted to a homomorphism $g : (\mathbb{M}(A_1, I_1), \iota_1) \to (\mathbb{M}(A_2, I_2), \iota_2)$ by setting $g([a_1 \cdots a_n]_{I_1}) = [g(a_1) \cdots g(a_n)]_{I_2}$.

A trace $t \in \mathbb{M}(A, I)$ can be visualized by its *dependence graph* D_t. To define D_t, choose an arbitrary word $w = a_1 a_2 \cdots a_n$, $a_i \in A$, with $t = [w]_I$ and define $D_t = (\{1, \ldots, n\}, E, \lambda)$, where $E = \{(i, j) \mid i < j, (a_i, a_j) \in D\}$ and $\lambda(i) = a_i$. If we identify isomorphic dependence graphs, then this definition is independent of the chosen word representing t. Moreover, the mapping $t \mapsto D_t$ is injective. As a consequence of the representation of traces by dependence graphs, one obtains Levi's lemma for traces, see e.g. [19, p. 74], which is one of the fundamental facts in trace theory. The formal statement is as follows.

Lemma 2.1. *Let $u_1, \dots, u_m, v_1, \dots, v_n \in \mathbb{M}(A, I)$. Then*

$$u_1 u_2 \cdots u_m = v_1 v_2 \cdots v_n$$

if and only if there exist $w_{i,j} \in \mathbb{M}(A, I)$ ($1 \le i \le m$, $1 \le j \le n$) such that

- *$u_i = w_{i,1} w_{i,2} \cdots w_{i,n}$ for every $1 \le i \le m$,*
- *$v_j = w_{1,j} w_{2,j} \cdots w_{m,j}$ for every $1 \le j \le n$, and*
- *$(w_{i,j}, w_{k,\ell}) \in I$ if $1 \le i < k \le m$ and $n \ge j > \ell \ge 1$.*

The situation in the lemma will be visualized by a diagram of the following kind. The i–th column corresponds to u_i, the j–th row corresponds to v_j, and the intersection of the i–th column and the j–th row represents $w_{i,j}$. Furthermore $w_{i,j}$ and $w_{k,\ell}$ are independent if one of them is left-above the other one.

$$\begin{array}{c||c|c|c|c|c|}
\hline
v_n & w_{1,n} & w_{2,n} & w_{3,n} & \dots & w_{m,n} \\ \hline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hline
v_3 & w_{1,3} & w_{2,3} & w_{3,3} & \dots & w_{m,3} \\ \hline
v_2 & w_{1,2} & w_{2,2} & w_{3,2} & \dots & w_{m,2} \\ \hline
v_1 & w_{1,1} & w_{2,1} & w_{3,1} & \dots & w_{m,1} \\ \hline \hline
 & u_1 & u_2 & u_3 & \dots & u_m
\end{array}$$

A consequence of Levi's Lemma is that trace monoids are cancellative, i.e., $usv = utv$ implies $s = t$ for all traces $s, t, u, v \in \mathbb{M}(A, I)$.

We end this section with a brief discussion of *trace rewriting systems*, which generalize semi-Thue systems [7, 29] from words to traces. Formally, a trace rewriting system over $\mathbb{M}(A, I)$ is a subset $R \subseteq \mathbb{M}(A, I) \times \mathbb{M}(A, I)$. We define the one-step rewrite relation $\to_R$ on $\mathbb{M}(A, I)$ as follows: $s \to_R t$ if there exist $u, v \in \mathbb{M}(A, I)$ and $(\ell, r) \in R$ with $s = u\ell v$ and $t = urv$. With $\overset{*}{\leftrightarrow}_R$ we denote the least equivalence relation on $\mathbb{M}(A, I)$ that contains $\to_R$, it is easily seen to be a congruence on $\mathbb{M}(A, I)$. Hence, we can define the quotient monoid $\mathbb{M}(A, I)/\overset{*}{\leftrightarrow}_R$ that will briefly be denoted by $\mathbb{M}(A, I)/R$. Let $\mathrm{RED}(R) = \{t \mid \exists u : t \to_R u\}$ be the set of *reducible traces* and $\mathrm{IRR}(R) = \mathbb{M}(A, I) \setminus \mathrm{RED}(R)$ be the set of *irreducible traces* (with respect to R). The system R is *terminating* if there does not exist an infinite chain $s_1 \to_R s_2 \to_R s_3 \to_R \cdots$ in $\mathbb{M}(A, I)$. We say that R is *length-reducing* if $|s| > |t|$ for all $(s, t) \in R$. The system R is *confluent* if for all $s, t, u \in \mathbb{M}(A, I)$ with $t \mathrel{{}_R{\overset{*}{\leftarrow}}} s \overset{*}{\to}_R u$ there exists $v \in \mathbb{M}(A, I)$ with $t \overset{*}{\to}_R v \mathrel{{}_R{\overset{*}{\leftarrow}}} u$. We say that R is *locally confluent* if for all $s, t, u \in \mathbb{M}(A, I)$ with

$t\ {}_R\!\!\leftarrow s \to_R u$ there exists $v \in \mathbb{M}(A, I)$ with $t \xrightarrow{*}_R v\ {}_R\!\!\xleftarrow{*} u$. If R is terminating, then by Newman's Lemma [47] confluence is equivalent to local confluence. If R is both terminating and confluent, then for every $s \in \mathbb{M}(A, I)$ there exists a unique *normal form* $\mathrm{NF}_R(s) \in \mathrm{IRR}(R)$ such that $s \xrightarrow{*}_R \mathrm{NF}_R(s)$. This normal form is the unique irreducible trace in the equivalence class with respect to $\xleftrightarrow{*}_R$ of the trace s.

In general, it is undecidable whether a finite length-reducing trace rewriting system is confluent, see [46]. This is in sharp contrast to semi-Thue systems, and makes confluence proofs challenging.

2.2 Rational and recognizable sets

Let $\mathcal{M} = (M, \circ, 1)$ be a monoid. The *product* of two sets $L_1, L_2 \subseteq M$ is $L_1 \circ L_2 = \{a_1 \circ a_2 \mid a_1 \in L_1, a_2 \in L_2\}$. The *Kleene star* of $L \subseteq M$ is $L^* = \bigcup_{i \geq 0} L^i$, where $L^0 = \{1\}$ and $L^{i+1} = L \circ L^i$ for $i \geq 0$. The set $\mathrm{RAT}(\mathcal{M})$ of all *rational subsets* of M is the smallest class of subsets that contains every finite subset of M and that is closed under union, product, and Kleene star. A subset $L \subseteq M$ is called *recognizable* if there exists a finite monoid S and a monoid homomorphism $h : \mathcal{M} \to S$, which may be assumed to be surjective, such that $L = h^{-1}(h(L))$. The class of all recognizable subsets of M is denoted by $\mathrm{REC}(\mathcal{M})$.

The classes $\mathrm{REC}(\mathcal{M})$ and $\mathrm{RAT}(\mathcal{M})$ are classical, see e.g. [6]. If $\mathcal{M}$ is a finitely generated monoid, then $\mathrm{REC}(\mathcal{M}) \subseteq \mathrm{RAT}(\mathcal{M})$ [41]. In general, $\mathrm{REC}(\mathcal{M})$ is a proper subset of $\mathrm{RAT}(\mathcal{M})$. For instance, a subgroup of a group G is recognizable if and only if it has finite index in G [6, p. 55]. Hence, a finite subgroup of an infinite group is rational but not recognizable. Another example, where $\mathrm{REC}(\mathcal{M})$ is a proper subset of $\mathrm{RAT}(\mathcal{M})$, is the trace monoid $\mathcal{M} = \mathbb{N} \times \mathbb{N}$ [19, p. 177]. For a free monoid Γ^* we have $\mathrm{REC}(\Gamma^*) = \mathrm{RAT}(\Gamma^*)$ by Kleene's Theorem.

For every monoid $\mathcal{M}$, the class $\mathrm{REC}(\mathcal{M})$ is an effective boolean algebra, but in general $\mathrm{REC}(\mathcal{M})$ is neither closed under products nor Kleene stars. On the other hand $\mathrm{RAT}(\mathcal{M})$ is in general not a boolean algebra, for instance $\mathrm{RAT}(\mathbb{N} \times \{a, b\}^*)$ is not closed under intersection, see e.g. [19, Example 6.1.16].

For a trace monoid $\mathbb{M} = \mathbb{M}(A, I)$ with A finite, it is easy to see that $L \in \mathrm{REC}(\mathbb{M})$ if and only if the language $\{u \in A^* \mid [u]_I \in L\}$ is a regular subset of A^*, whereas $L \in \mathrm{RAT}(\mathbb{M})$ if and only if there is a regular language $K \subseteq A^*$ such that $L = \{[u]_I \mid u \in K\}$. Thus, every finite subset of $\mathbb{M}$ is recognizable. Moreover, $\mathrm{REC}(\mathbb{M})$ is closed under products and *connected Kleene stars* [48].[1] Therefore, for a *finite* trace rewriting system R over a trace monoid $\mathbb{M}$, we have $\mathrm{RED}(R) \in \mathrm{REC}(\mathbb{M})$ and $\mathrm{IRR}(R) \in \mathrm{REC}(\mathbb{M})$.

[1] A Kleene star L^*, where $L \subseteq \mathbb{M}$, is called connected if every $t \in L$ is a connected trace, i.e., we cannot write $t = uv$ with $(u, v) \in I$ and $u \neq \varepsilon \neq v$.

2.3 Graph products

In this section we will introduce graph products of monoids. The graph product construction generalizes both the free product and the direct product. Graph products were introduced in [25].

Let (Σ, I_Σ) be a *finite* independence alphabet, i.e., Σ is finite, and let $\mathcal{M}_\sigma = (M_\sigma, \circ_\sigma, 1_\sigma)$ be a monoid for every $\sigma \in \Sigma$. Let $A_\sigma = M_\sigma \setminus \{1_\sigma\}$, and define an independence alphabet (A, I) by

$$A = \bigcup_{\sigma \in \Sigma} A_\sigma \quad \text{and} \quad I = \bigcup_{(\sigma,\tau) \in I_\Sigma} A_\sigma \times A_\tau,$$

where w.l.o.g. $A_\sigma \cap A_\tau = \emptyset$ for $\sigma \neq \tau$. Let

$$R_\sigma = \{(ab, c) \mid a, b, c \in A_\sigma, a \circ_\sigma b = c\} \cup \{(ab, \varepsilon) \mid a, b \in A_\sigma, a \circ_\sigma b = 1_\sigma\}$$

and the define the trace rewriting system R over $\mathbb{M}(A, I)$ as $R = \bigcup_{\sigma \in \Sigma} R_\sigma$. Then the *graph product* $\mathbb{P}(\Sigma, I_\Sigma, (\mathcal{M}_\sigma)_{\sigma \in \Sigma})$ is the quotient monoid $\mathbb{M}(A, I)/R$. Special cases are the *free product* $*_{\sigma \in \Sigma} \mathcal{M}_\sigma$ (if $I_\Sigma = \emptyset$) and the *direct product* $\prod_{\sigma \in \Sigma} \mathcal{M}_\sigma$ (if $I_\Sigma = (\Sigma \times \Sigma) \setminus \{(\sigma, \sigma) \mid \sigma \in \Sigma\}$). Let us fix a graph product $\mathbb{P} = \mathbb{P}(\Sigma, I_\Sigma, (\mathcal{M}_\sigma)_{\sigma \in \Sigma})$ for the further discussion. The crucial fact for our further investigation is the following, a proof can be found in [32]:

Lemma 2.2. *The trace rewriting system R is confluent.*

Since R is also terminating, the previous lemma implies that $\mathbb{P}$ is in one-to-one correspondence with $\mathrm{IRR}(R) \subseteq \mathbb{M}(A, I)$, which is the set of all traces that do not contain a factor of the form ab with $a, b \in A_\sigma$ for some $\sigma \in \Sigma$.

2.4 Relational structures and logic

For more details on first-order logic see e.g. [28]. The notion of a structure (or model) is defined as usual in logic. Formally, we will only consider *relational structures*, but we will feel free to use also constants and (partially defined) operations. They can be encoded by relations in the usual way. Let us fix a relational structure $\mathbb{A} = (A, (R_i)_{i \in J})$, where $R_i \subseteq A^{n_i}$, $i \in J$. The *signature of* $\mathbb{A}$ contains the equality symbol $=$, and for each $i \in J$ it contains a relation symbol of arity n_i that we denote without risk of confusion by R_i as well. Given further relations R_j, $j \in K$, $J \cap K = \emptyset$, we also write $(\mathbb{A}, (R_i)_{i \in K})$ for the structure $(A, (R_i)_{i \in J \cup K})$. It is called an *extension* of $\mathbb{A}$.

Next, let us introduce *first-order logic (FO-logic)*. Let $\mathbb{V}$ be a countable infinite set of *first-order variables*, which range over elements of the universe A. First-order variables are denoted by x, y, z, x', etc. *FO-formulas* over the signature of $\mathbb{A}$ are constructed from the atomic formulas $R_i(x_1, \ldots, x_{n_i})$ and $x = y$ (where $i \in J$ and $x_1, \ldots, x_{n_i}, x, y \in \mathbb{V}$) using boolean connectives and (existential and universal) quantifications over variables from $\mathbb{V}$. The notion of a free variable is defined as usual. A formula without free variables is called a *sentence*. If $\varphi(x_1, \ldots, x_n)$ is an FO-formula with free variables among $x_1, \ldots, x_n$ and $a_1, \ldots, a_n \in A$, then $\mathbb{A} \models \varphi(a_1, \ldots, a_n)$ means that φ evaluates to true in $\mathbb{A}$ if the free variable x_i evaluates to a_i. The *first-order theory* of $\mathbb{A}$, denoted by $\mathrm{FOTh}(\mathbb{A})$, is the set of all first-order sentences φ such that $\mathbb{A} \models \varphi$. The *existential first-order theory* $\exists\mathrm{FOTh}(\mathbb{A})$ of $\mathbb{A}$ is the set of all sentences in $\mathrm{FOTh}(\mathbb{A})$ of the form $\exists x_1 \cdots \exists x_n : \varphi(x_1, \ldots, x_n)$, where $\varphi(x_1, \ldots, x_n)$ is a boolean combination of atomic formulas. The *positive theory* $\mathrm{posTh}(\mathbb{A})$ is the set of all sentences in $\mathrm{FOTh}(\mathbb{A})$ that do not use negations, i.e., that are built from atomic formulas using conjunctions, disjunctions, and existential and universal quantifications.

The length $|\varphi|$ of an FO-formula is the length of the binary encoding of φ. Here, we have to assume that the index set J for the relations is finite or countably infinite. Then, every relation R_i can be encoded by a finite bit string. We do not define the precise binary encoding of formulas because it is not really relevant for the purpose of this paper.

We view a monoid $\mathcal{M} = (M, \circ, 1)$ as a relational structure by considering the multiplication $\circ$ as a ternary relation and the constant 1 as a unary relation. Instead of $\circ(x, y, z)$ we write $x \circ y = z$ or briefly $xy = z$. We will also consider extensions $(\mathcal{M}, (R_i)_{i\in J})$ of the structure $\mathcal{M}$, where R_i is a relation of arbitrary arity over M. In case $\mathcal{C}$ is a class of subsets of M, we also write $(\mathcal{M}, \mathcal{C}, (R_i)_{i\in J})$ instead of $(\mathcal{M}, (L)_{L\in\mathcal{C}}, (R_i)_{i\in J})$ and call formulas of the form $x \in L$ for $L \in \mathcal{C}$ *constraints*. In many cases, a partial monoid involution ι will belong to the R_i (see e.g. Section 3.1). It is viewed as a binary relation on M.

Remark 2.3. Usually the (existential) first-order theory of a monoid is defined by allowing arbitrary equations of the form $u = v$, where u and v are words over the variables, as atomic predicates. But this formulation is easily seen to be equivalent to our definition and we deliberately write down such equations. Moreover, also constants from $\mathcal{M}$ are usually allowed in equations. We can deal with constants by including them as singleton subsets to the additional relations R_i.

Note that if $\mathcal{M}$ is *finitely generated* by Γ, then constants from Γ suffice in order to define all monoid elements of $\mathcal{M}$. In this case, we call $\mathrm{FOTh}(\mathcal{M}, (a)_{a\in\Gamma})$ the *first-order theory of $\mathcal{M}$ with constants*. On the other hand, the further investigations are not restricted to finitely generated monoids.

A well-known example of a decidable theory of equations is Presburger's Arithmetic [51]. Translated into our framework, the results of [5] imply the following statement, where $\mathrm{RAT}(\mathbb{N})$ and $\mathrm{RAT}(\mathbb{Z})$ are the classes of *semi-linear sets* in $\mathbb{N}$ and $\mathbb{Z}$, respectively:

Proposition 2.4 (cf [5]). *If $\mathcal{M} = \mathbb{N}$ or $\mathcal{M} = \mathbb{Z}$, then* $\mathrm{FOTh}(\mathcal{M}, \mathrm{RAT}(\mathcal{M}))$ *belongs to SPACE*$(2^{2^{O(n)}})$.

Remark 2.5. It is known that $\mathrm{FOTh}(\{a,b\}^*, a, b)$ is undecidable [52], in fact already the $\forall\exists^3$-fragment of this theory is undecidable [21, 39]. Together with Presburger's result, it follows that the decidability of the full first-order theory of equations is not preserved under free products. For a restricted class of monoids, we will show such a closure result in Section 3.3 for the existential case, even for general graph products.

The following result, which will be needed later, can be easily deduced from Proposition 2.4, basically because the free product $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$ of two copies of $\mathbb{Z}/2\mathbb{Z}$ is isomorphic to the semi-direct product of $\mathbb{Z}$ by $\mathbb{Z}/2\mathbb{Z}$, see [16] for a proof.

Corollary 2.6. *For $\mathcal{M} = \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$, the theory $\mathrm{FOTh}(\mathcal{M}, \mathrm{RAT}(\mathcal{M}))$ belongs to SPACE($2^{2^{O(n)}}$).*

Another example of a theory that can be easily reduced to Presburger's Arithmetic is the theory of the bicyclic monoid $\{a,b\}^*/_{ab=\varepsilon}$ with the constants a and b and the sets a^* and b^* as constraints:

Corollary 2.7. *The theory $\mathrm{FOTh}(\{a,b\}^*/_{ab=\varepsilon}, a, b, a^*, b^*)$ is in SPACE($2^{2^{O(n)}}$).*

Proof. An element of $\{a,b\}^*/_{ab=\varepsilon}$ can be uniquely written as $b^m a^n$ for $m, n \geq 0$. Moreover, $(b^i a^j)(b^k a^\ell) = b^m a^n$ in $\{a,b\}^*/_{ab=\varepsilon}$ if and only if either $j > k$, $m = i$, and $n = \ell + j - k$ or $j \leq k$, $m = i + k - j$, and $n = \ell$. This is a formula of Presburger's Arithmetic over $\mathbb{Z}$. □

3 Existential theories of graph products

Based on results from [18] for (finitely generated) trace monoids with a partial involution (see Section 3.1), we will prove in Section 3.2 a general preservation theorem for existential theories. In Section 3.3 we will use this result in order to show that for a large class of monoids the decidability of the existential theory is preserved under graph products.

3.1 Trace monoids with a partial involution

All our decidability results in this section are based on the main result from [18]. In order to state this result in its whole generality, we have to introduce the following graph theoretical concept: Let (A, I) be an independence alphabet. We define on A an equivalence relation $\sim_I$ by $a \sim_I b$ if and only if $I(a) = I(b)$ (which is equivalent to $D(a) = D(b)$). Note that $a \sim_I b$ implies $(a, b) \notin I$: if $I(a) = I(b)$ and $(a, b) \in I$, then also $(a, a) \in I$, which contradicts the irreflexivity of I. The following lemma will be needed later.

Lemma 3.1. *For $s, t \in \mathbb{M}(A, I)$ we have $s \neq t$ if and only if there exists an equivalence class $C \subseteq A$ of $\sim_I$ such that one of the following three cases holds:*

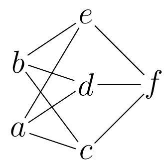
$$\exists u, v, w \in \mathbb{M}(A, I)\, \exists a, b \in C : s = uav \ \wedge\ t = ubw \ \wedge\ a \neq b \tag{1}$$
$$\exists u, v, w \in \mathbb{M}(A, I) : s = uv \ \wedge\ t = uw \ \wedge\ v \in C\,\mathbb{M}(A, I) \ \wedge\ w \notin C\,\mathbb{M}(A, I) \tag{2}$$
$$\exists u, v, w \in \mathbb{M}(A, I) : s = uv \ \wedge\ t = uw \ \wedge\ v \notin C\,\mathbb{M}(A, I) \ \wedge\ w \in C\,\mathbb{M}(A, I) \tag{3}$$

Proof. The if-direction is easy to see using the fact that $\mathbb{M}(A, I)$ is cancellative and that $C \times C \subseteq D$ for every equivalence class C of $\sim_I$. Now assume that $s \neq t$. Since $\mathbb{M}(A, I)$ is cancellative, we can assume by induction that $\min(s) \cap \min(t) = \emptyset$. If either s or t is empty, then (2) or (3) holds for some C (with $u = \varepsilon$). Now assume that $s \neq \varepsilon \neq t$. Moreover, assume that (1) does not hold. Take $a \in \min(s)$ and let C be the unique equivalence class of $\sim_I$ containing a. Since $a \notin \min(t)$ and (1) does not hold, we have $C \cap \min(t) = \emptyset$. Hence, (2) holds (with $u = \varepsilon$). $\square$

An equivalence class B of $\sim_I$ is called a *thin clan* of (A, I), if $I(a) \neq \emptyset$ for some (and hence all) $a \in B$. The cardinality of the set of thin clans of (A, I) is denoted by $\tau(A, I)$ – of course it may be infinite if A is infinite. The following facts are easy to verify:

- $\tau(A, I)$ is bounded by the cardinality of A.
- There exist at most one equivalence class of $\sim_I$, which is not a thin clan. It consists of all the isolated nodes of (A, I).
- The cardinality of a largest (A, I)-clique is at most $\max\{1, \tau(A, I)\}$.
- $\tau(A, I) \neq 1$, and $\tau(A, I) = 0$ if and only if $I = \emptyset$.

For the independence alphabet below, the equivalence classes of $\sim_I$ are $\{a, b\}$, $\{c, d, e\}$, and $\{f\}$, and they are all thin clans.



Now we can state the main result from [18].

Theorem 3.2. *For every $k \geq 0$, the following problem is in PSPACE:*

INPUT: A finite independence alphabet (A, I) with $\tau(A, I) \leq k$, a partial involution ι on A that is compatible with I, and an existential sentence ϕ over the signature of $(\mathbb{M}(A, I), \iota, \mathrm{REC}(\mathbb{M}(A, I)))$ (with ι lifted to $\mathbb{M}(\mathrm{dom}(\iota), I)$).

QUESTION: Does $(\mathbb{M}(A, I), \iota, \mathrm{REC}(\mathbb{M}(A, I))) \models \phi$ hold?

A few remarks should be made on Theorem 3.2.

- A recognizable set $L \in \mathrm{REC}(\mathbb{M}(A, I))$ has to be represented by a finite automaton for the regular language $\{u \in A^* \mid [u]_I \in L\}$. This is crucial. For instance, if recognizable trace languages are represented by loop-connected automata (see e.g. [45]), then already universality is EXPSPACE-complete for some fixed independence alphabet [45].

– Since every singleton subset belongs to $\mathrm{REC}(\mathbb{M}(A,I))$, constants are implicitly allowed in Theorem 3.2.
– In [18], Theorem 3.2 is only stated for a totally defined involution $\iota : A \to A$. But if the involution is only defined on $B \subsetneq A$, then we can introduce a new dummy symbol $\overline{a}$ for every $a \in A \setminus B$, extend the involution by $\iota(a) = \overline{a}$ and $\iota(\overline{a}) = a$, and restrict every variable to the original alphabet A, which is a recognizable constraint.

Theorem 3.2 cannot be extended to the case of rational constraints: For $\mathcal{M} = \{a,b\}^* \times \{c,d\}^*$ it is undecidable whether for given $L_1, L_2 \in \mathrm{RAT}(\mathcal{M})$ it holds $L_1 \cap L_2 = \emptyset$, see [1]. A further investigation leads to the following characterization of Muscholl, see [44, Prop. 2.9.2 and 2.9.3].

Proposition 3.3. *Let $\mathbb{M} = \mathbb{M}(A,I)$ be a trace monoid with A finite. Then $\exists\mathrm{FOTh}(\mathbb{M}, \mathrm{RAT}(\mathbb{M}))$ is decidable if and only if $\mathbb{M}$ is a free product of free commutative monoids, i.e., $\mathbb{M} = *_{i=1}^{n} \mathbb{N}^{k_i}$ for $n, k_1, \ldots, k_n \in \mathbb{N}$.*

3.2 A general preservation theorem

The aim of this section is to prove a general preservation theorem for existential theories. We will apply this result in the next section to existential theories of graph products.

For the further discussion let us fix a set A together with a partial involution ι on A and a countable subset $\mathcal{C} \subseteq 2^A$. Let

$$\mathbb{A} = (A, \iota, (L)_{L \in \mathcal{C}}). \tag{4}$$

Moreover, we have given an independence relation $I \subseteq A \times A$ and additional predicates R_j $(1 \leq j \leq m)$ of arbitrary arity on A such that:

(a) ι is compatible with I,
(b) the set $\{I(a) \mid a \in A\}$ is finite (thus, $\sim_I$ has only finitely many equivalence classes),
(c) $\mathrm{dom}(\iota)$ as well as every equivalence class of $\sim_I$ belong to $\mathcal{C}$, and
(d) $\exists\mathrm{FOTh}(\mathbb{A}, (R_j)_{1 \leq j \leq m})$ is decidable.

Due to (a), we can lift ι to a partial monoid involution on $\mathbb{M}(A,I)$. Moreover, (b) and (c) imply that the independence relation I is definable by a boolean formula over $(A, (L)_{L \in \mathcal{C}})$, because I is a finite union of Cartesian products of equivalence classes of $\sim_I$.

From the unary predicates in $\mathcal{C}$ we construct a set $\mathsf{L}(\mathcal{C}, I) \subseteq 2^{\mathbb{M}(A,I)}$ as follows: A *$\mathcal{C}$-automaton* $\mathcal{A}$ is a finite automaton in the usual sense, except that every edge of $\mathcal{A}$ is labeled with some language $L \in \mathcal{C}$. The language $L(\mathcal{A}) \subseteq A^*$ is defined in the obvious way: $a_1 a_2 \cdots a_n \in L(\mathcal{A})$ $(a_i \in A)$ if and only if there exists a path $q_0 \xrightarrow{L_1} q_1 \xrightarrow{L_2} q_2 \cdots \xrightarrow{L_{n-1}} q_{n-1} \xrightarrow{L_n} q_n$ in $\mathcal{A}$ such that q_0 is the initial state of $\mathcal{A}$,

q_n is a final state of $\mathcal{A}$, and $a_i \in L_i$ for $1 \leq i \leq n$. We say that $\mathcal{A}$ is *I-closed* if $[u]_I = [v]_I$ and $u \in L(\mathcal{A})$ imply $v \in L(\mathcal{A})$. In the following, we will identify $L(\mathcal{A})$ with the set of traces $\{[u]_I \mid u \in L(\mathcal{A})\}$. Then $L \subseteq \mathbb{M}(A, I)$ belongs to the class $\mathsf{L}(\mathcal{C}, I)$ if there exists an I-closed $\mathcal{C}$-automaton $\mathcal{A}$ with $L(\mathcal{A}) = L$. We briefly write $\mathsf{L}(\mathcal{C})$ instead of $\mathsf{L}(\mathcal{C}, I)$. For effectiveness statements, it is necessary that languages in $\mathcal{C}$ have some finite representation. Then, also languages from $\mathsf{L}(\mathcal{C})$ have a canonical finite representation by I-closed $\mathcal{C}$-automata and the size of a $\mathcal{C}$-automaton can be defined in a natural way.

Since $A \subseteq \mathbb{M}(A, I)$, we can view every relation R_j also as a relation over the trace monoid $\mathbb{M}(A, I)$. This is done in the following theorem,[2] which is the main result of this section:

Theorem 3.4. *Let $\mathbb{A}$, I, and $(R_j)_{1 \leq j \leq m}$ satisfy (a)–(d) above. Then*

$$\exists\mathrm{FOTh}(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{C}), (R_j)_{1 \leq j \leq m}) \tag{5}$$

is decidable. Moreover, if $\exists\mathrm{FOTh}(\mathbb{A}, (R_j)_{j \in J})$ is decidable in NSPACE($s(n)$), then the theory (5) can be decided in NSPACE($2^{O(n)} + s(n^{O(1)})$).

Later, the relations R_j will be the multiplication relations of the factors of a given graph product.

Reducing the number of generators In this section we will prove a purely combinatorial lemma that will be the key in order to reduce the infinite set A of generators of $\mathbb{M}(A, I)$ in Theorem 3.4 to a finite set of generators B. This will enable us to apply Theorem 3.2.

In the sequel, we will restrict to some reduct $(A, \iota, (L)_{L \in \mathcal{D}})$ of the structure $\mathbb{A}$ from (4), where $\mathcal{D} \subseteq \mathcal{C}$ is finite and contains $\mathrm{dom}(\iota)$ as well as every equivalence class of $\sim_I$. We denote this reduct by $\mathbb{A}$ as well. For the following consideration it is useful to fix some enumeration $L_0, \ldots, L_k$ of $\mathcal{D}$, where $\mathrm{dom}(\iota) = L_0$ and $L_1, \ldots, L_\ell$ ($\ell \leq k$) is an enumeration of the equivalence classes of $\sim_I$. Thus, $\{L_1, \ldots, L_\ell\}$ is a partition of A. Moreover there exists a fixed independence relation I' on $\{1, \ldots, \ell\}$ such that $I = \bigcup_{(i,j) \in I'} L_i \times L_j$.

Given another structure $\mathbb{B} = (B, \zeta, (K_i)_{0 \leq i \leq k})$ (with ζ a partial involution on B, $K_i \subseteq B$, and $K_0 = \mathrm{dom}(\zeta)$), a mapping $f : A \to B$ is a *strong homomorphism* from $\mathbb{A}$ to $\mathbb{B}$ if for all $a \in A$:

- $a \in L_i$ if and only if $f(a) \in K_i$ for all $0 \leq i \leq k$ and
- $f(\iota(a)) = \zeta(f(a))$ if $a \in \mathrm{dom}(\iota)$.

Lemma 3.5. *We can effectively construct a* finite *structure*

$$\mathbb{B} = (B, \zeta, (K_i)_{0 \leq i \leq k})$$

(with ζ a partial involution on B, $K_i \subseteq B$, and $\mathrm{dom}(\zeta) = K_0$) such that

[2] Recall that in contrast to the R_j, the partial involution ι was lifted from A to the whole trace monoid $\mathbb{M}(A, I)$.

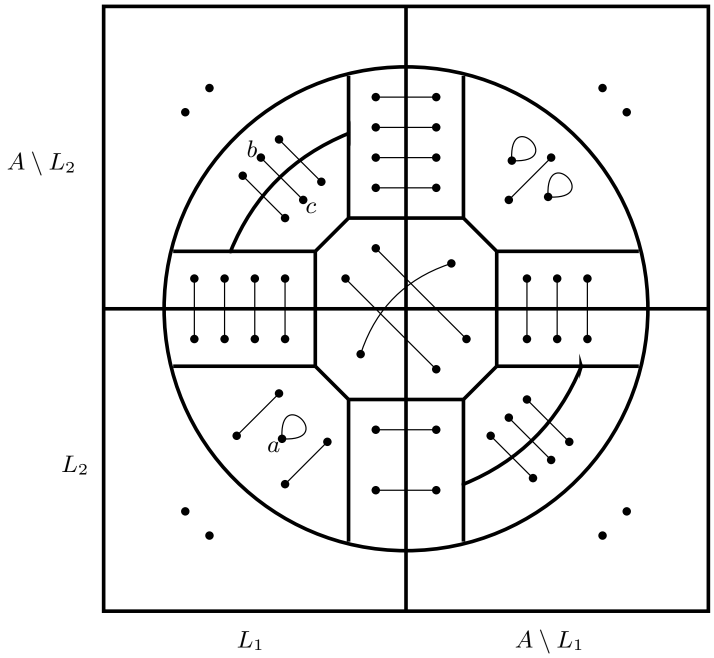


Fig. 1. The construction from the proof of Lemma 3.5

- $|B| \leq 2^{k+1}(2^{k+1}+3)$ *and*
- *there exist strong homomorphisms* $f : \mathbb{A} \to \mathbb{B}$ *and* $g : \mathbb{B} \to \mathbb{A}$ *with* f *surjective.*

Effectiveness in this context means that given a finite set $\mathcal{D} \subseteq \mathcal{C}$*, we can construct the finite structure* $\mathbb{B}$ *effectively.*

Lemma 3.5 is our key lemma. The surjective strong homomorphism $f : \mathbb{A} \to \mathbb{B}$ defines a partition of A into finitely many equivalence classes. Roughly speaking, elements in the same equivalence classes do not have to be distinguished for the purpose of satisfying a boolean formula over $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{C}), (R_j)_{1 \leq j \leq m})$.

Proof of Lemma 3.5. First we will define B and $f : A \to B$ such that every L_i is a finite union of preimages $f^{-1}(c)$ ($c \in B$), i.e., f saturates every L_i. Moreover,

(i) $f(a) = f(a')$ and $a, a' \in \mathrm{dom}(\iota)$ will imply $f(\iota(a)) = f(\iota(a'))$, and
(ii) $f(a) = f(\iota(a))$ will imply $a' = \iota(a')$ for some a' with $f(a) = f(a')$.

Figure 1, where $k = 2$ is assumed, visualizes the construction. The sets L_1 and L_2 are represented by the left half and lower half, respectively, of the whole square, which represents A. The right half (resp. upper half) represents $A \setminus L_1$ (resp. $A \setminus L_2$), the big inner circle represents $\mathrm{dom}(\iota) = L_0$, and the thin lines represent the partial involution ι on A. The 22 regions that are bounded by thick lines represent the preimages $f^{-1}(b)$ ($b \in B$) and hence the elements of B. Of course, the sets $f^{-1}(b)$ will be infinite in general.

Let $[k] = \{0, \ldots, k\}$. We realize B as a subset

$$B \subseteq 2^{[k]} \cup (2^{[k]} \times 2^{[k]}) \cup (2^{[k]} \times \{0, 1\}).^3$$

For a subset $\alpha \subseteq [k]$ define

$$L^\alpha = \bigcap_{i \in \alpha} L_i \cap \bigcap_{i \notin \alpha} A \setminus L_i.$$

If $\alpha \subseteq [k]$ is such that $0 \notin \alpha$ (i.e., $L^\alpha \cap \mathrm{dom}(\iota) = \emptyset$) and $L^\alpha \neq \emptyset$, then we put α into B and define the function f on L^α by $f(L^\alpha) = \alpha$. Note that by assumption (d) we can check effectively whether $L^\alpha \neq \emptyset$, we just have to decide whether $\mathbb{A} \models \exists x : x \in L^\alpha$. For instance the four outer regions in Figure 1 would be represented by $\{1, 2\}$, $\{1\}$, $\{2\}$, and $\emptyset$. If $0 \in \alpha$, i.e., $L^\alpha \subseteq \mathrm{dom}(\iota)$, then L^α has to be split into possibly several (but finitely many) preimages of f in order to satisfy (i) and (ii) above. To represent them in B, take a second subset $\beta \subseteq [k]$ with $0 \in \beta$. In case $\alpha \neq \beta$ we check whether $L^\alpha \cap \iota(L^\beta) \neq \emptyset$, i.e., $\mathbb{A} \models \exists x \in L^\alpha\ \exists y \in L^\beta : x = \iota(y)$. If this is true, then we put (α, β) and (β, α) into B and define $f(L^\alpha \cap \iota(L^\beta)) = (\alpha, \beta)$ and $f(L^\beta \cap \iota(L^\alpha)) = (\beta, \alpha)$. Now assume that $\alpha = \beta$. We proceed with testing whether $\mathbb{A} \models \exists x \in L^\alpha : x = \iota(x)$. If this holds, then we put (α, α) into B and define $f(L^\alpha \cap \iota(L^\alpha)) = (\alpha, \alpha)$. For instance, the region containing a in Figure 1 is represented by $(\{0, 1, 2\}, \{0, 1, 2\})$. On the other hand, if $\mathbb{A} \models \neg\exists x \in L^\alpha : x = \iota(x)$, then we check whether $L^\alpha \cap \iota(L^\alpha) \neq \emptyset$, i.e., $\mathbb{A} \models \exists x, y \in L^\alpha : \iota(x) = y$. If this holds, then due to (ii) the set $L^\alpha \cap \iota(L^\alpha)$ has to be split into precisely two preimages C_0 and C_1 of f, where $\iota(a) \in C_i$ for all $a \in C_{1-i}$. These two classes can be represented by the pairs $(\alpha, 0)$ and $(\alpha, 1)$, which we put into B. We set $f(C_i) = (\alpha, i)$. For instance the two regions containing b and $c = \iota(b)$ in Figure 1 are represented by $(\{0, 1\}, 0)$ and $(\{0, 1\}, 1)$ (it does not matter which of the two possible assignments is chosen). This completes the construction of the alphabet B as well as the definition of the surjection f. The size bound $|B| \leq 2^{k+1}(2^{k+1} + 3)$ follows immediately from the construction.

We define the involution ζ on B as follows: If $\alpha, \beta \in 2^{[k]}$ are such that $(\alpha, \beta) \in B$, then we define $\zeta(\alpha, \beta) = (\beta, \alpha)$. If $\alpha \in 2^{[k]}$ is such that $(\alpha, 0), (\alpha, 1) \in B$, then $\zeta(\alpha, i) = (\alpha, 1 - i)$ for $i \in \{0, 1\}$. We define the set $K_i \subseteq B$ by

$$\begin{aligned} K_i = &\{\alpha \in B \mid \alpha \in 2^{[k]}, i \in \alpha\} \cup \{(\alpha, \beta) \in B \mid \alpha, \beta \in 2^{[k]}, i \in \alpha\} \cup \\ &\{(\alpha, j) \in B \mid \alpha \in 2^{[k]}, j \in \{0, 1\}, i \in \alpha\}. \end{aligned}$$

This finishes the construction of $\mathbb{B}$. Clearly $K_i = f(L_i)$, $B \setminus K_i = f(A \setminus L_i)$, and $\zeta(f(a)) = f(\iota(a))$, i.e., $f : \mathbb{A} \to \mathbb{B}$ is a strong homomorphism.

We have defined $f : A \to B$ such that if $\zeta(b) = b$, then there exists $a \in f^{-1}(b)$ with $\iota(a) = a$ (see (ii)). This allows to select $g(b) \in f^{-1}(b)$ for every $b \in B$ such

[3] The specific representation of B is not really important, we only need some finite representation.

that $\iota(g(b)) = g(\zeta(b))$. Moreover, since $g(b) \in f^{-1}(b)$, we have $b \in K_i$ if and only if $f(g(b)) \in K_i$ if and only if $g(b) \in L_i$. Thus, $g : \mathbb{B} \to \mathbb{A}$ is a strong homomorphism as well. $\square$

Note that since the strong homomorphism f is surjective in the previous lemma and $\{L_1, \dots, L_\ell\}$ is a partition of A, also $\{K_1, \dots, K_\ell\}$ is a partition of B.

Now assume that we have given a third structure $\mathbb{C} = (C, \xi, (\Lambda_i)_{0 \le i \le k})$, where C is finite, ξ is a partial involution on C, $\Lambda_i \subseteq C$ for $0 \le i \le k$, $\mathrm{dom}(\xi) = \Lambda_0$, and $\{\Lambda_1, \dots, \Lambda_\ell\}$ is a partition of C (with $\Lambda_i = \emptyset$ allowed). In the sequel, an *embedding of* $\mathbb{C}$ *in* $\mathbb{A}$ is an injective strong homomorphism $h : \mathbb{C} \to \mathbb{A}$.

Lemma 3.6. *Given* $\mathbb{C}$ *as above, we can effectively construct a* finite *structure* $\mathbb{B} = (B, \zeta, (K_i)_{0 \le i \le k})$ *(with* ζ *a partial involution on* B*,* $K_i \subseteq B$*, and* $\mathrm{dom}(\zeta) = K_0$*) together with an independence relation* $J \subseteq B \times B$ *such that:*

- $C \subseteq B$,
- $|B| \le 2^{k+1}(2^{k+1} + 3) + |C|$,
- ζ *is compatible with* J*, and*
- *for every embedding* $h : \mathbb{C} \to \mathbb{A}$ *there exist strong homomorphisms* $f : \mathbb{A} \to \mathbb{B}$ *and* $g : \mathbb{B} \to \mathbb{A}$ *such that* $f(I) \subseteq J$*,* $g(J) \subseteq I$*, and* $f(h(c)) = c$*,* $g(c) = h(c)$ *for all* $c \in C$*.*

Proof. By Lemma 3.5 we can construct a finite structure

$$\mathbb{B}' = (B', \zeta', (K'_i)_{0 \le i \le k})$$

such that

- $\mathrm{dom}(\zeta') = K'_0$, $|B'| \le 2^{k+1}(2^{k+1} + 3)$, and
- there exist strong homomorphisms $f' : \mathbb{A} \to \mathbb{B}'$ and $g' : \mathbb{B}' \to \mathbb{A}$ with f' surjective.

W.l.o.g. $B' \cap C = \emptyset$. Note that $\{K'_1, \dots, K'_\ell\}$ must be a partition of B'. Now we define the structure

$$\mathbb{B} = (B, \zeta, (K_i)_{0 \le i \le k})$$

by $B = B' \cup C$, $\zeta = \zeta' \cup \xi$, and $K_i = K'_i \cup \Lambda_i$ for $0 \le i \le k$. The given size bound for $|B|$ in the lemma follows from $|B'| \le 2^{k+1}(2^{k+1}+3)$. Since $\{K_1, \dots, K_\ell\}$ is a partition of B, we can define the independence relation J on B by $J = \bigcup_{(i,j) \in I'} K_i \times K_j$. Recall here that I' is an independence relation $\{1, \dots, \ell\}$ such that $I = \bigcup_{(i,j) \in I'} L_i \times L_j$.

Given an embedding $h : \mathbb{C} \to \mathbb{A}$, we define $f : A \to B$ by $f(h(c)) = c$ for $c \in C$ (since h is injective, this is well-defined) and $f(a) = f'(a)$ for $a \in A \setminus h(C)$. We define $g : B \to A$ by $g(b) = g'(b)$ for $b \in B'$ and $g(c) = h(c)$ for $c \in C$. Since $h : \mathbb{C} \to \mathbb{A}$, $f' : \mathbb{A} \to \mathbb{B}'$, and $g' : \mathbb{B}' \to \mathbb{A}$ are strong homomorphisms, the following properties are easy to verify for all $a \in A$ and $b \in B = B' \cup C$:

(i) $a \in L_i$ if and only if $f(a) \in K_i$ and $b \in K_i$ if and only if $g(b) \in L_i$.
(ii) $f(\iota(a)) = \zeta(f(a))$ and $g(\zeta(b)) = \iota(g(b))$ (for the first identity note that $a \in h(C)$ if and only if $\iota(a) \in h(C)$).

Thus, $f : \mathbb{A} \to \mathbb{B}$ and $g : \mathbb{B} \to \mathbb{A}$ are strong homomorphisms with $f(h(c)) = c$ and $g(c) = h(c)$ for all $c \in C$. Moreover, since $I = \bigcup_{(i,j)\in I'} L_i \times L_j$ and $J = \bigcup_{(i,j)\in I'} K_i \times K_j$, (i) implies that $(a, a') \in I$ if and only if $(f(a), f(a')) \in J$ and $(b, b') \in J$ if and only if $(g(b), g(b')) \in I$. In particular, $f(I) \subseteq J$ and $g(J) \subseteq I$.

In order to see that ζ is compatible with J assume that $(b, b') \in J$ and $b, b' \in \mathrm{dom}(\zeta) = K_0$. Then $(g(b), g(b')) \in I$ and $g(b), g(b') \in \mathrm{dom}(\iota) = L_0$. Since ι is compatible with I, we obtain $(\iota(g(b)), \iota(g(b'))) = (g(\zeta(b)), g(\zeta(b'))) \in I$. Hence, $(\zeta(b), \zeta(b')) \in J$. $\square$

Proof of Theorem 3.4 For the proof of Theorem 3.4 let us take a boolean formula θ over the signature of $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{C}), (R_j)_{1\le j\le m})$. We have to decide whether θ is satisfiable in the structure, $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{C}), (R_j)_{1\le j\le m})$. For this, we will present a nondeterministic algorithm that constructs a finitely generated trace monoid $\mathbb{M}(B, J)$ with a partial involution ζ and a boolean formula ϕ' over the signature of $(\mathbb{M}(B, J), \zeta, \mathrm{REC}(\mathbb{M}(B, J)))$ such that θ is satisfiable in $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{C}), (R_j)_{1\le j\le m})$ if and only if for at least one outcome of our nondeterministic algorithm, ϕ' is satisfiable in $(\mathbb{M}(B, J), \zeta, \mathrm{REC}(\mathbb{M}(B, J)))$. This allows to apply Theorem 3.2.

Assume that every $\mathcal{C}$-automaton in θ only uses sets among the finite set $\mathcal{D} \subseteq \mathcal{C}$. Assume that also $\mathrm{dom}(\iota)$ as well as every $\sim_I$-equivalence class belongs to $\mathcal{D}$. Let $\mathcal{D} = \{L_0, \ldots, L_k\}$, where $L_0 = \mathrm{dom}(\iota)$ and $L_1, \ldots, L_\ell$ ($\ell \le k$) is an enumeration of the $\sim_I$-equivalence classes of (A, I). Note that $k \in O(|\theta|)$ for any reasonable encoding of formulas with constraints from $\mathsf{L}(\mathcal{C})$.

Step 1 (pushing negations down and eliminating disjunctions). First we may push negations to the level of atomic subformulas in θ. Moreover, disjunctions may be eliminated by nondeterministically guessing one of the two corresponding disjuncts. Thus, we may assume that θ is a conjunction of atomic predicates and negated atomic predicates. We replace every negated equation $xy \neq z$ by $xy = z' \wedge z \neq z'$, where z' is a new variable. Similarly an equation $\iota(x) \neq y$ is replaced by $\iota(x) = z \wedge z \neq y$. Thus, we may assume that all negated predicates in θ are of the form $x \neq y$, $x \notin L$, and $\neg R_j(x_1, \ldots, x_n)$ for variables $x, y, x_1, \ldots, x_n$ and $L \in \mathsf{L}(\mathcal{D})$.

Step 2 (eliminating disequations). We can write θ as a conjunction $\phi \wedge \psi$, where ψ contains all predicates of the form $(\neg)R_j(x_1, \ldots, x_n)$. Recall that $R_j \subseteq A^{n_j}$, hence the formula ψ is the "$\mathbb{A}$-local" part of θ, which only speaks about the base structure $(\mathbb{A}, (R_j)_{1\le j\le m})$. Let $x \neq y$ be a negated equation in ϕ, where x and y are variables. Since $x \neq y$ is interpreted in the trace monoid $\mathbb{M}(A, I)$, we can by Lemma 3.1 replace $x \neq y$ by either

$$\begin{array}{ll} x = zau \ \wedge \ y = zbv \ \wedge \ a, b \in L_i \ \wedge \ a \neq b & \text{or} \\ x = zu \ \wedge \ y = zv \ \wedge \ u \in L_i \, \mathbb{M}(A, I) \ \wedge \ v \notin L_i \, \mathbb{M}(A, I) & \text{or} \\ x = zu \ \wedge \ y = zv \ \wedge \ u \notin L_i \, \mathbb{M}(A, I) \ \wedge \ v \in L_i \, \mathbb{M}(A, I), & \end{array}$$

where z, u, v, a, b are new variables and $i \in \{1, \ldots, \ell\}$ is guessed nondeterministically. In the first case, we shift $a, b \in L_i \ \wedge \ a \neq b$ to the $\mathbb{A}$-local part ψ. In the second and third case, we have to construct an I-closed $\mathcal{D}$-automaton for $L_i \, \mathbb{M}(A, I)$, which is easy, since all $\sim_I$-equivalence classes belong to $\mathcal{D}$. Thus, in the sequel we may assume that ϕ does not contain negated equations.

Step 3 (eliminating local constraints in the non-local part). So far, we have obtained a conjunction $\phi \ \wedge \ \psi$, where ϕ (the non-local part) is interpreted in $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{D}))$ and ψ (the local part) is interpreted in the base structure $(\mathbb{A}, (R_j)_{1 \leq j \leq m})$. The non-local part ϕ does not contain negated equations. Let $\varXi$ be the set of all variables that occur in $\phi \ \wedge \ \psi$, and let $\varOmega \subseteq \varXi$ contain all variables that occur in the local part ψ. Thus, all variables from $\varOmega$ are implicitly restricted to $A \subseteq \mathbb{M}(A, I)$. Note that variables from $\varOmega$ may of course also occur in ϕ. In case ϕ contains a constraint $x \in L$ with $L \in \mathsf{L}(\mathcal{D})$ and $x \in \varOmega$, then we can guess an $L' \in \mathcal{D}$, which labels a transition from the initial state to a final state of the automaton for L, and replace $x \in L$ by the constraint $x \in L'$. We shift this constraint to ψ. Hence, we may assume that for every constraint $x \in L$ that occurs in ϕ, we have $x \in \varXi \setminus \varOmega$.

Step 4 (saturating the local part ψ). Next, for every variable $x \in \varOmega$ we guess whether $x \in L_0 = \mathrm{dom}(\iota)$ or $x \notin \mathrm{dom}(\iota)$ holds and add the corresponding (negated) constraint to ψ. In case $x \in \mathrm{dom}(\iota)$ was guessed, we add a new variable $\overline{x}$ to $\varOmega$ and add $\iota(x) = \overline{x} \ \wedge \ \overline{x} \in L_0$ to ψ. Next, we guess for all different variables $x, y \in \varOmega$ (here $\varOmega$ refers to the new set of variables including the added copis $\overline{x}$), whether $x = y$ or $x \neq y$. In case $x = y$ is guessed, we can replace y by x everywhere. Thus, we may assume that for all different variables $x, y \in \varOmega$ the negated equation $x \neq y$ belongs to ψ. Finally, for every set L_i with $1 \leq i \leq k$ and every $x \in \varOmega$ we guess whether $x \in L_i$ or $x \notin L_i$ holds and add the corresponding constraint to ψ. We denote the resulting formula by ψ as well.

Most of the guessed local formulas ψ will be not satisfiable in $(\mathbb{A}, (R_j)_{1 \leq j \leq m})$ (e.g., if $L_i \cap L_j = \emptyset$ and the constraints $x \in L_i$ and $x \in L_j$ were guessed). But since $\exists\mathrm{FOTh}(\mathbb{A}, (R_j)_{1 \leq j \leq m})$ is decidable, we can effectively check whether the guessed formula ψ is satisfiable. If it is not satisfiable, then we reject on the corresponding computation path. Let us fix a specific guess, which results in a satisfiable local formula ψ, for the further consideration.

Step 5 (applying Lemma 3.6). Now we construct a finite structure

$$\mathbb{C} = (\widetilde{\varOmega}, \xi, (\varLambda_i)_{0 \leq i \leq k})$$

from ψ as follows: Let $\widetilde{\Omega} = \{\widetilde{x} \mid x \in \Omega\}$ be a disjoint copy of the set of variables Ω. For $0 \leq i \leq k$ let Λ_i be the set of all $\widetilde{x} \in \widetilde{\Omega}$ such that $x \in L_i$ belongs to the local part ψ. Finally, we define the partial involution ξ on $\widetilde{\Omega}$ as follows: The domain of ξ is Λ_0 and $\xi(\widetilde{x}) = \widetilde{y}$ in case $\iota(x) = y$ or $\iota(y) = x$ belongs to the local part ψ. Since ψ is satisfiable and $x \neq y$ belongs to ψ for all pairwise different variables x and y, ξ is indeed a partial involution on $\widetilde{\Omega}$. Moreover, since $\{L_1, \dots, L_\ell\}$ is a partition of A, $\{\Lambda_1 \dots, \Lambda_\ell\}$ must be a partition of $\widetilde{\Omega}$ (with $\Lambda_i = \emptyset$ allowed). Thus, $\mathbb{C}$ satisfies all the requirements preceding Lemma 3.6, and we can apply Lemma 3.6 to the structures $\mathbb{A}$ and $\mathbb{C}$. Hence, from $\mathbb{C}$ we can effectively determine a finite structure

$$\mathbb{B} = (B, \zeta, (K_i)_{0 \leq i \leq k})$$

together with an independence relation $J \subseteq B \times B$ such that

- $\widetilde{\Omega} \subseteq B$,
- $|B| \leq |\widetilde{\Omega}| + 2^{O(k)} \leq 2^{O(|\theta|)}$,
- ζ is compatible with J, and
- for every embedding $h : \mathbb{C} \to \mathbb{A}$ there exist strong homomorphisms $f : \mathbb{A} \to \mathbb{B}$ and $g : \mathbb{B} \to \mathbb{A}$ with $f(I) \subseteq J$, $g(J) \subseteq I$, and $f(h(\widetilde{x})) = \widetilde{x}$, $g(\widetilde{x}) = h(\widetilde{x})$ for every variable $x \in \Omega$.

Since the partial involution $\zeta : B \to B$ is compatible with J, we can lift ζ to a partial involution on $\mathbb{M}(B, J)$. We denote this lifting by ζ as well.

Recall that we have to check whether there exist assignments $\kappa : \Omega \to A$ and $\lambda : \Xi \setminus \Omega \to \mathbb{M}(A, I)$ such that κ satisfies ψ in $(\mathbb{A}, (R_j)_{1 \leq j \leq m})$ and $\kappa \cup \lambda$ satisfies ϕ in $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{D}))$. We have already verified that the conjunction ψ is satisfiable in $(\mathbb{A}, (R_j)_{1 \leq j \leq m})$. For the following consideration let us fix an arbitrary assignment $\kappa : \Omega \to A$ that satisfies ψ in $(\mathbb{A}, (R_j)_{1 \leq j \leq m})$.[4] Since $x \neq y$ belongs to ψ for all different variables $x, y \in \Omega$, κ defines an embedding $h : \mathbb{C} \to \mathbb{A}$ by $h(\widetilde{x}) = \kappa(x)$ for $x \in \Omega$. Therefore, by Lemma 3.6, there exist strong homomorphisms $f : \mathbb{A} \to \mathbb{B}$ and $g : \mathbb{B} \to \mathbb{A}$ with

$$\forall x \in \Omega : f(\kappa(x)) = \widetilde{x} \ \wedge \ g(\widetilde{x}) = \kappa(x). \tag{6}$$

Moreover, $f(\iota(a)) = \zeta(f(a))$ for all $a \in A$, $g(\zeta(b)) = \iota(g(b))$ for all $b \in B$, $f(I) \subseteq J$, and $J \subseteq g(I)$. Hence, by lifting and f and g to $\mathbb{M}(A, I)$ and $\mathbb{M}(B, J)$, respectively, we obtain the following homomorphisms between trace monoids with partial involution:

$$\begin{aligned} f &: (\mathbb{M}(A, I), \iota) \ \to (\mathbb{M}(B, J), \zeta) \\ g &: (\mathbb{M}(B, J), \zeta) \to (\mathbb{M}(A, I), \iota). \end{aligned}$$

Given an I-closed $\mathcal{D}$-automaton $\mathcal{A}$, we define a new automaton $\mathcal{A}'$ by replacing every edge $p \xrightarrow{L_i} q$ in $\mathcal{A}$ by $p \xrightarrow{K_i} q$ (and changing nothing else). Recall that

[4] We do not have to determine this assignment explicitly, only its existence is important.

$K_i \subseteq B$. Since $\mathcal{A}$ is I-closed, $\mathcal{A}'$ is easily seen to be J-closed. Moreover, since B is finite, $L(\mathcal{A}') \subseteq \mathbb{M}(B, J)$ is a recognizable trace language. Recall that for every $0 \leq i \leq k$, we have $a \in L_i$ if and only if $f(a) \in K_i$ and $b \in K_i$ if and only if $g(b) \in L_i$. Thus, the following statement is obvious:

Lemma 3.7. *Let $t \in \mathbb{M}(A, I)$ and $u \in \mathbb{M}(B, J)$:*

- *$t \in L(\mathcal{A})$ if and only if $f(t) \in L(\mathcal{A}')$.*
- *$u \in L(\mathcal{A}')$ if and only if $g(u) \in L(\mathcal{A})$.*

Next, we transform the non-local formula ϕ into a conjunction ϕ', which will be interpreted over $(\mathbb{M}(B, J), \zeta, \mathrm{REC}(\mathbb{M}(B, J)))$, by replacing in ϕ every occurrence of a variable $x \in \Omega$ by the constant $\widetilde{x} \in \widetilde{\Omega} \subseteq B$. Thus, ϕ' contains constants from $\widetilde{\Omega} \subseteq \mathbb{M}(B, J)$ and variables from $\Xi \backslash \Omega$, which range over the trace monoid $\mathbb{M}(B, J)$. Moreover, every constraint $x \in L(\mathcal{A})$ (resp. $x \notin L(\mathcal{A})$) in ϕ is replaced by $x \in L(\mathcal{A}')$ (resp. $x \notin L(\mathcal{A}')$) (note that $x \in \Xi \setminus \Omega$ by Step 3). Thus, all constraint languages in ϕ' are recognizable trace languages.

Lemma 3.8. *The following two statements are equivalent:*

(a) There exists an assignment $\lambda : \Xi \backslash \Omega \to \mathbb{M}(A, I)$ such that $\kappa \cup \lambda$ satisfies the boolean formula ϕ in $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{D}))$.

(b) There exists an assignment $\lambda' : \Xi \backslash \Omega \to \mathbb{M}(B, J)$ that satisfies the boolean formula ϕ' in $(\mathbb{M}(B, J), \zeta, \mathrm{REC}(\mathbb{M}(B, J)))$.

Proof. First, assume that (a) holds. We claim that (b) holds with $\lambda' = f \circ \lambda$. Consider a constraint $x \in L(\mathcal{A}')$ (resp. $x \notin L(\mathcal{A}')$) of ϕ'. Then $x \in \Xi \setminus \Omega$ and $x \in L(\mathcal{A})$ (resp. $x \notin L(\mathcal{A})$) is a constraint of ϕ. Thus, $(\kappa \cup \lambda)(x) = \lambda(x) \in L(\mathcal{A})$ (resp. $\lambda(x) \notin L(\mathcal{A})$), which implies $\lambda'(x) = f(\lambda(x)) \in L(\mathcal{A}')$ (resp. $\lambda'(x) \notin L(\mathcal{A}')$) by Lemma 3.7. Now let $u' = v'$ be an equation of ϕ', which results from the equation $u = v$ of ϕ. The only syntactic difference between $u = v$ and $u' = v'$ is that every occurrence of every variable $x \in \Omega$ in $u = v$ is replaced by the constant $\widetilde{x}$ in $u' = v'$. The assignment $\kappa \cup \lambda$ is a solution of $u = v$ in $(\mathbb{M}(A, I), \iota)$. Since f is a homomorphism between trace monoids with partial involution, $f \circ (\kappa \cup \lambda) = f \circ \kappa \cup f \circ \lambda = f \circ \kappa \cup \lambda'$ is a solution of $u = v$ in $(\mathbb{M}(B, J), \zeta)$. Since $f(\kappa(x)) = \widetilde{x}$ for every $x \in \Omega$ by (6), the mapping λ' is a solution of $u' = v'$ in $(\mathbb{M}(B, J), \zeta)$.

Now assume that (b) holds. We claim that (a) holds with $\lambda = g \circ \lambda'$. Let $x \in L(\mathcal{A})$ (resp. $x \notin L(\mathcal{A})$) be a constraint of ϕ. Then $x \in \Xi \setminus \Omega$ and $x \in L(\mathcal{A}')$ (resp. $x \notin L(\mathcal{A}')$) is a constraint of ϕ'. Hence, $\lambda'(x) \in L(\mathcal{A}')$ (resp. $\lambda'(x) \notin L(\mathcal{A}')$). Lemma 3.7 implies that $\lambda(x) = g(\lambda'(x)) \in L(\mathcal{A})$ (resp. $\lambda(x) \notin L(\mathcal{A})$). Now consider an equation $u = v$ of ϕ and let $u' = v'$ be the corresponding equation of ϕ'. Thus, λ' is a solution of $u' = v'$ in $(\mathbb{M}(B, J), \zeta)$. Let the function π map every variable $x \in \Omega$ to the constant $\widetilde{x} \in \widetilde{\Omega} \subseteq B$. By construction of $u' = v'$, $\lambda' \cup \pi$ is a solution of $u = v$ in $(\mathbb{M}(B, J), \zeta)$. Since g is a homomorphism between trace monoids with partial involution and $g(\pi(x)) = g(\widetilde{x}) = \kappa(x)$ for every $x \in \Omega$ by (6), the mapping $g \circ (\lambda' \cup \pi) = \lambda \cup \kappa$ is a solution of $u = v$ in $(\mathbb{M}(A, I), \iota)$. $\square$

For the previous lemma it is crucial that the conjunction ϕ does not contain negated equations (see Step 2), because the homomorphisms f and g are not injective in general, and therefore do not preserve inequalities.

Since Lemma 3.8 holds for every $\kappa : \Omega \to A$ that satisfies ψ in the structure $(\mathbb{A}, (R_j)_{1\leq j\leq m})$, and we already know that such an assignment exists, it only remains to check whether ϕ' is satisfiable in $(\mathbb{M}(B,J), \zeta, \mathrm{REC}(\mathbb{M}(B,J)))$. By Theorem 3.2 this can be done effectively. This proves the decidability statement in Theorem 3.4.

For the upper complexity bound in Theorem 3.4 one has to notice the following two points:

- The size of the new alphabet B is bounded by $2^{O(|\theta|)}$ and size of the formula ϕ' is bounded by $|\theta|^{O(1)}$, where θ is the initial formula. Moreover, for the number of thin clans we have $\tau(B,J) = \tau(A,I)$, where the latter is a fixed finite constant in Theorem 3.4. This allows to apply the complexity statement from Theorem 3.2 in order to check in NSPACE($2^{O(|\theta|)}$) whether ϕ' is satisfiable in $(\mathbb{M}(B,J), \zeta, \mathrm{REC}(\mathbb{M}(B,J)))$.
- During the construction of B and ϕ', we had to check the validity of existential formulas of size $|\theta|^{O(1)}$ in the structure $(\mathbb{A}, (R_j)_{1\leq j\leq m})$, which can be done in NSPACE($s(|\theta|^{O(1)})$) by assumption.

3.3 Closure under graph products

In this section we will apply Theorem 3.4 in order to show that under some restrictions, the decidability of the existential theory is preserved by graph products. Other closure results for graph products can be found for instance in [24, 27, 33, 42, 63, 64]. Concerning graph products we will use the notation from Section 2.3 in the following.

We fix a graph product $\mathbb{P} = \mathbb{P}(\Sigma, I_\Sigma, (\mathcal{M}_\sigma)_{\sigma\in\Sigma})$ for the further discussion, where $\mathcal{M}_\sigma = (M_\sigma, \circ_\sigma, 1_\sigma)$ is a monoid. Let $A_\sigma = M_\sigma \backslash \{1_\sigma\}$ and define

$$A = \bigcup_{\sigma\in\Sigma} A_\sigma \quad \text{and} \quad I = \bigcup_{(\sigma,\tau)\in I_\Sigma} A_\sigma \times A_\tau,$$

where w.l.o.g. $A_\sigma \cap A_\tau$ for $\sigma \neq \tau$. In Section 2.3 we have defined the trace rewriting system

$$R = \bigcup_{\sigma\in\Sigma} \{ab \to c \mid a,b,c \in A_\sigma, a \circ_\sigma b = c\} \cup \{ab \to \varepsilon \mid a,b \in A_\sigma, a \circ_\sigma b = 1_\sigma\}$$

over $\mathbb{M}(A,I)$. We have stated that R is confluent (Lemma 2.2) and that $\mathbb{P}$ is in one-to-one correspondence with $\mathrm{IRR}(R) \subseteq \mathbb{M}(A,I)$. Define the binary relations

$$\mathrm{inv}_\sigma = \{(a,b) \in A_\sigma \times A_\sigma \mid a \circ_\sigma b = 1_\sigma\} \quad \text{and} \quad \mathrm{inv} = \bigcup_{\sigma\in\Sigma} \mathrm{inv}_\sigma. \tag{7}$$

Let $U_\sigma = \mathrm{dom}(\mathrm{inv}_\sigma)$, $V_\sigma = \mathrm{ran}(\mathrm{inv}_\sigma)$, $U = \bigcup_{\sigma\in\Sigma} U_\sigma = \mathrm{dom}(\mathrm{inv})$, and $V = \bigcup_{\sigma\in\Sigma} V_\sigma = \mathrm{ran}(\mathrm{inv})$.

Constraints Our announced closure result will also include constraints. In this paragraph we present a general construction that defines a class of constraints in the graph product $\mathbb{P}$, starting from a constraint class for every factor monoid $\mathcal{M}_\sigma$.

For every $\sigma \in \Sigma$ let $\mathcal{C}_\sigma \subseteq 2^{M_\sigma}$ be a class of languages and let

$$\mathcal{D}_\sigma = \{L \setminus \{1_\sigma\} \mid L \in \mathcal{C}_\sigma\} \subseteq 2^{A_\sigma}.$$

It is not required that $\mathcal{D}_\sigma \subseteq \mathcal{C}_\sigma$. Let $\mathcal{C} = \bigcup_{\sigma \in \Sigma} \mathcal{C}_\sigma$ and $\mathcal{D} = \bigcup_{\sigma \in \Sigma} \mathcal{D}_\sigma \subseteq 2^A$. Recall the definition of the class $\mathsf{L}(\mathcal{D}, I) \subseteq 2^{\mathbb{M}(A,I)}$ (briefly $\mathsf{L}(\mathcal{D})$) from Section 3.2. We define the class $\mathsf{IL}(\mathcal{C}, I, R) \subseteq 2^{\mathbb{M}(A,I)}$ by

$$\mathsf{IL}(\mathcal{C}, I, R) = \{L \cap \mathrm{IRR}(R) \mid L \in \mathsf{L}(\mathcal{D}, I)\}.$$

In the following, we will briefly write $\mathsf{IL}(\mathcal{C})$ for $\mathsf{IL}(\mathcal{C}, I, R)$. Using the one-to-one correspondence between $\mathbb{P}$ and $\mathrm{IRR}(R)$, we may view $L \cap \mathrm{IRR}(R)$ also as a subset of the graph product $\mathbb{P}$, hence $\mathsf{IL}(\mathcal{C}) \subseteq 2^{\mathbb{P}}$.

Alternatively, we can also define the class $\mathsf{IL}(\mathcal{C})$ by I-closed $\mathcal{D}$-automata which accept subsets of $\mathrm{IRR}(R)$. To see this, let $\mathcal{A}$ be an I-closed $\mathcal{D}$-automaton. The closure properties of recognizable trace languages (Section 2.2) imply that

$$K = \bigcup_{\sigma \in \Sigma} \mathbb{M}(\Sigma, I_\Sigma)\sigma\sigma\mathbb{M}(\Sigma, I_\Sigma) \in \mathrm{REC}(\mathbb{M}(\Sigma, I_\Sigma)).$$

Hence,

$$L = \Sigma^* \setminus \{u \in \Sigma^* \mid [u]_{I_\Sigma} \in K\} \subseteq \Sigma^* \tag{8}$$

is a regular string language. Let $\mathcal{B}$ be a finite automaton for L. An automaton for $L(\mathcal{A}) \cap \mathrm{IRR}(R)$ can be obtained by a product construction from $\mathcal{A}$ and $\mathcal{B}$. The product automaton contains a transition $(p, q) \xrightarrow{D} (p', q')$ if and only if $p \xrightarrow{D} p'$ is a transition of $\mathcal{A}$, $D \subseteq A_\sigma$, and $q \xrightarrow{\sigma} q'$ is a transition of $\mathcal{B}$.

The following lemma will be needed later:

Lemma 3.9. *If $\mathbb{P} = \mathbb{P}(\Sigma, I_\Sigma, (\mathcal{M}_\sigma)_{\sigma \in \Sigma})$ is an arbitrary graph product of monoids $\mathcal{M}_\sigma$, then $\mathrm{REC}(\mathbb{P}) \subseteq \mathsf{IL}(\mathcal{C})$ for $\mathcal{C} = \bigcup_{\sigma \in \Sigma} \mathrm{REC}(\mathcal{M}_\sigma)$.*

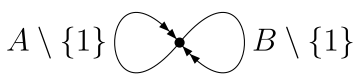
Proof. Let $\mathcal{D} = \bigcup_{\sigma \in \Sigma} \{L \setminus \{1_\sigma\} \mid L \in \mathrm{REC}(\mathcal{M}_\sigma)\} \setminus \{\emptyset\}$. Assume that $L \in \mathrm{REC}(\mathbb{P})$ and let $\varrho : \mathbb{P} \to S$ be a surjective homomorphism onto the finite monoid S such that $L = \varrho^{-1}(F)$ for $F \subseteq S$. We define A_σ, A, I, and R as above. Let

$$\Delta_\sigma = \{A_\sigma \cap \varrho^{-1}(q) \mid q \in S\} \setminus \{\emptyset\} \subseteq 2^{A_\sigma}$$

and $\Delta = \bigcup_{\sigma \in \Sigma} \Delta_\sigma$. Clearly, Δ_σ is a partition of A_σ with finitely many classes. Note that if we restrict ϱ to $\mathcal{M}_\sigma \subseteq \mathbb{P}$, we obtain a homomorphism from $\mathcal{M}_\sigma$ to S. Thus, $\Delta \subseteq \mathcal{D}$. Let $I_\Delta = \bigcup_{(\sigma,\tau) \in I_\Sigma} \Delta_\sigma \times \Delta_\tau$. Hence, (Δ, I_Δ) is a finite independence alphabet. Define the homomorphism $\beta : \mathbb{M}(\Delta, I_\Delta) \to S$ by $\beta(A_\sigma \cap \varrho^{-1}(q)) = q$

if $A_\sigma \cap \varrho^{-1}(q) \neq \emptyset$. Since $(B,C) \in I_\Delta$ implies $\beta(B)\beta(C) = \beta(C)\beta(B)$ in S, this defines indeed a homomorphism. Thus, $\beta^{-1}(F) \in \mathrm{REC}(\mathbb{M}(\Delta, I_\Delta))$. We can also define a homomorphism $\alpha : \mathbb{M}(A,I) \to \mathbb{M}(\Delta, I_\Delta)$ by mapping $a \in A$ to the unique $B \in \Delta$ with $a \in B$. If $h : \mathbb{M}(A,I) \to \mathbb{P}$ denotes the canonical homomorphism that maps a trace $t \in \mathbb{M}(A,I)$ to the element of $\mathbb{P}$ represented by t, then $\beta(\alpha(t)) = \varrho(h(t))$ for all $t \in \mathbb{M}(A,I)$. Let $\mathcal{A}$ be a finite state automaton that accepts $\{w \in \Delta^* \mid [w]_{I_\Delta} \in \beta^{-1}(F)\}$. Since every edge of $\mathcal{A}$ is labeled with a set from $\Delta \subseteq \mathcal{D}$, we can interpret $\mathcal{A}$ also as a $\mathcal{D}$-automaton, which is moreover I-closed. For every $t \in \mathbb{M}(A,I)$ we have: $t \in L(\mathcal{A})$ if and only if $\beta(\alpha(t)) \in F$ if and only if $\varrho(h(t)) \in F$ if and only if $h(t) \in \varrho^{-1}(F) = L$. Hence, $L(\mathcal{A}) = h^{-1}(L)$, i.e., $h(L(\mathcal{A}) \cap \mathrm{IRR}(R)) = L$. Thus, $L \in \mathsf{IL}(\mathcal{C})$. $\square$

The other inclusion $\mathsf{IL}(\mathcal{C}) \subseteq \mathrm{REC}(\mathbb{P})$ for $\mathcal{C} = \bigcup_{\sigma \in \Sigma} \mathrm{REC}(\mathcal{M}_\sigma)$ does not hold in general: Take $\mathbb{P} = \mathbb{Z} * \mathbb{Z}$ and let A (resp. B) be a subgroup of finite index in the first (resp. second) copy of $\mathbb{Z}$ in $\mathbb{P}$. Hence $A, B \in \mathrm{REC}(\mathbb{Z})$ [6]. But the automaton



defines the subgroup $A * B \leq \mathbb{Z} * \mathbb{Z}$, hence $A * B \in \mathsf{IL}(\mathcal{C})$. But since $A * B$ has infinite index in $\mathbb{Z} * \mathbb{Z}$, $A * B \notin \mathrm{REC}(\mathbb{Z} * \mathbb{Z})$.

The main result Throughout this section we will assume that the following two requirements hold:

Assumption 3.10 *For all $\sigma \in \Sigma$ and all $a, b, c \in M_\sigma$, if $a \circ_\sigma b = a \circ_\sigma c = 1_\sigma$ or $b \circ_\sigma a = c \circ_\sigma a = 1_\sigma$, then $b = c$. In other words, the relation* inv_σ *from (7) is a partial injection.*

For example, cancellative monoids (in particular free monoids and groups), the bicyclic monoid $\{a,b\}^*/_{ab=\varepsilon}$, and finite monoids satisfy all this requirement,[5] whereas $\{a,b,c\}^*/_{ab=ac=\varepsilon}$ does not. By Assumption 3.10, $\mathrm{inv} = \bigcup_{\sigma \in \Sigma} \mathrm{inv}_\sigma$ is a partial injection on A with $\mathrm{dom}(\mathrm{inv}) = U$ and $\mathrm{ran}(\mathrm{inv}) = V$. Since inv is compatible with the independence relation I, we can lift inv to $\mathbb{M}(A,I)$ (see Section 2.1). The resulting partial injection inv has domain $\mathbb{M}(U,I)$ and range $\mathbb{M}(V,I)$.

Assumption 3.11 *For all $\sigma \in \Sigma$, the theory $\exists\mathrm{FOTh}(\mathcal{M}_\sigma, \mathcal{C}_\sigma)$ is decidable and $U_\sigma, V_\sigma \in \mathcal{C}_\sigma$, i.e., $U_\sigma, V_\sigma \in \mathcal{D}_\sigma = \{L \setminus \{1_\sigma\} \mid L \in \mathcal{C}_\sigma\}$.*

The following theorem is the main result of this section.

[5] For a finite monoid note that $a \circ b = 1$ implies that the mapping $x \mapsto b \circ x$ is injective, hence it is surjective. Thus, there exists c with $b \circ c = 1$. Clearly $a = c$, i.e., $b \circ a = 1$ and inv_σ is a partial involution.

Theorem 3.12. *Let (Σ, I_Σ) be a finite independence alphabet. Let $\mathcal{M}_\sigma$ be a monoid and $\mathcal{C}_\sigma \subseteq 2^{\mathcal{M}_\sigma}$ be a class of languages such that Assumption 3.10 and Assumption 3.11 hold. Then, for $\mathcal{C} = \bigcup_{\sigma\in\Sigma} \mathcal{C}_\sigma$,*

$$\exists\mathrm{FOTh}(\mathbb{P}(\Sigma, I_\Sigma, (\mathcal{M}_\sigma)_{\sigma\in\Sigma}), \mathsf{IL}(\mathcal{C})) \tag{9}$$

is also decidable. Moreover, if each of the theories $\exists\mathrm{FOTh}(\mathcal{M}_\sigma, \mathcal{C}_\sigma)$ belongs to NSPACE$(s(n))$, then (9) can be decided in NSPACE$(2^{O(n)} + s(n^{O(1)}))$.

Before we go into the details of the proof of Theorem 3.12 let us first present an application. The existential theories with constants and U_σ and V_σ as constraints of the following (classes of) monoids are decidable: finite monoids (trivial), free monoids [38], the bicyclic monoid (Corollary 2.7), virtually-free groups [34], and torsion-free hyperbolic groups [60].[6] Since all these monoids satisfy Assumption 3.10, we obtain the following corollary:

Corollary 3.13. *Let $\mathbb{P}$ be a graph product of finite monoids, free monoids, bicyclic monoids, virtually-free groups, and torsion-free hyperbolic groups, and let Γ be a finite generating set for $\mathbb{P}$. Then $\exists\mathrm{FOTh}(\mathbb{P}, (a)_{a\in\Gamma})$ is decidable.*

The following example shows that already for quite simple monoids, for which Assumption 3.10 fails, the decidability of the existential theory is a difficult problem.

Example 3.14. Let $\mathcal{M} = \{a, b, c\}^*/\{ac = bc = 1\}$. This monoid does not satisfy Assumption 3.10. Clearly, the free monoid $\{a, b\}^*$ is a submonoid of $\mathcal{M}$, and we have $x \in \{a, b\}^*$ if and only if $\exists y : xy = 1$ in $\mathcal{M}$. Moreover, $|x| = |y|$ for $x \in \{a, b\}^*$ if and only if $\exists z : xz = yz = 1$ in $\mathcal{M}$. This shows that the existential theory of a free monoid with length-constraints $|x| = |y|$ can be reduced to the existential theory of $\mathcal{M}$. Whether the former theory is decidable is a longstanding open problem, see e.g. [9].

We begin the proof of Theorem 3.12 with a few simple observations. We have

$$R = \{ab \to \varepsilon \mid (a, b) \in \mathrm{inv}\} \cup \bigcup_{\sigma\in\Sigma} \{ab \to c \mid a, b, c \in A_\sigma, a \circ_\sigma b = c\}.$$

We may assume that $M_\sigma \in \mathcal{C}_\sigma$, i.e., $A_\sigma \in \mathcal{D}$, for every $\sigma \in \Sigma$ without violating Assumption 3.11.[7] Hence, since every equivalence class of $\sim_I$ is a union of some of the A_σ, we may assume that these classes belong to $\mathcal{D}$ as well. Finally, since

[6] Rips and Sela have shown in [54] that it is decidable whether a word equation is solvable over a torsion-free hyperbolic group. In [60], Sela extended the approach of [54] such that also negated equations can be handled.

[7] Note that a constraint of the form $x \in U_\sigma$ could be eliminated by $\exists y : x \circ_\sigma y = 1_\sigma$, but this is not possible for constraints $x \notin U_\sigma$, since we would introduce a universal quantifier in this way. Therefore we assume explicitly that $U_\sigma \in \mathcal{C}_\sigma$.

$U_\sigma, V_\sigma \in \mathcal{D}_\sigma$, we may also assume that $U, V, U \cup V \in \mathcal{D}$. In particular, $\mathbb{M}(U, I)$, which is the domain of the lifting of inv to $\mathbb{M}(A, I)$, belongs to $\mathsf{L}(\mathcal{D})$.

Note that $A_\sigma \in \mathcal{D}$ also implies that $\mathrm{IRR}(R) \in \mathsf{L}(\mathcal{D})$: An I-closed $\mathcal{D}$-automaton for $\mathrm{IRR}(R)$ can be obtained from a finite automaton for the language L in (8) by replacing every label σ by A_σ. It follows that every constraint $x \in \mathsf{IL}(\mathcal{C})$ can be written as $x \in L_1 \wedge x \in L_2$ with $L_1, L_2 \in \mathsf{L}(\mathcal{D})$.

Isolating the structure of the $\mathcal{M}_\sigma$ In this paragraph we finish the proof of Theorem 3.12. Assume that for every $\sigma \in \Sigma$ the theory $\exists\mathrm{FOTh}(\mathcal{M}_\sigma, \mathcal{C}_\sigma)$ is decidable in $\mathrm{NSPACE}(s(n))$. Then the same holds for $\exists\mathrm{FOTh}(A_\sigma, \circ_\sigma, \mathrm{inv}_\sigma, (L)_{L \in \mathcal{D}_\sigma})$, where $\circ_\sigma$ is considered as a ternary relation that is restricted to A_σ. Since by Assumption 3.10, $\mathrm{inv} : U \to V$ is a partial injection, we can define a partial involution ι on A with domain $U \cup V \in \mathcal{D}$ by $\iota(a) = b$ if and only if either $\mathrm{inv}(a, b)$ or $\mathrm{inv}(b, a)$ (note that $\mathrm{inv}(a, b)$ and $\mathrm{inv}(b, c)$ implies $a = c$). This involution on A is compatible with I, hence it can be lifted to a partial monoid involution ι on $\mathbb{M}(A, I)$ with domain $\mathbb{M}(U \cup V, I)$. Let

$$\mathbb{A} = (A, \iota, (L)_{L \in \mathcal{D}}).$$

Since the existential theory of a disjoint union of structures can be reduced (in polynomial time) to the constituent structures (this is a very special case of Feferman-Vaught decomposition [23]), it follows that $\exists\mathrm{FOTh}(\mathbb{A}, (\circ_\sigma)_{\sigma \in \Sigma})$ is decidable in $\mathrm{NSPACE}(s(n))$ as well. Now we apply Theorem 3.4 to the structure $\mathbb{A}$ together with the independence relation I and the additional relations $\circ_\sigma$. Clearly, the structure $(\mathbb{A}, (\circ_\sigma)_{\sigma \in \Sigma})$, the constraint set $\mathcal{D}$, and the independence relation I satisfy the requirements from Section 3.2. It follows from Theorem 3.4 that

$$\exists\mathrm{FOTh}(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{D}), (\circ_\sigma)_{\sigma \in \Sigma})$$

is decidable in $\mathrm{NSPACE}(2^{O(n)} + s(n^{O(1)}))$.

Let θ be a boolean formula with atomic predicates of the form $xy = z$ and $x \in L$ with $L \in \mathsf{IL}(\mathcal{C})$,[8] which is interpreted over $(\mathbb{P}, \mathsf{IL}(\mathcal{C}))$. We have to check, whether there exists an assignment for the variables in θ to elements in $\mathbb{P}$ that satisfies θ.

The rest of the section shows that θ can be transformed in polynomial time into an equivalent existential statement over $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{D}), (\circ_\sigma)_{\sigma \in \Sigma})$. Thus, in some sense we isolate the structure of the factor monoids $\mathcal{M}_\sigma$ into the "$\mathcal{M}_\sigma$-local" multiplication predicates $\circ_\sigma$.

First, we may push negations to the level of atomic subformulas in θ. We replace every negated equation $xy \neq z$ by $xy = z' \wedge z \neq z'$, where z' is a new variable. Thus, we may assume that all negated predicates in θ are of the form $x \neq y$ and $x \notin L$ for variables x and y.

[8] Atomic predicates of the form $x = 1$ are not necessary since $\{1\} \in \mathsf{IL}(\mathcal{C})$.

Recall that $\mathbb{P} \cong \mathbb{M}(A, I)/R$ and that R is confluent and terminating. Hence, if $\widehat{s}$ denotes the unique trace from $\mathrm{IRR}(R)$ that represents $s \in \mathbb{P}$, then for all $s, t, u \in \mathbb{P}$ and $L \in \mathsf{IL}(\mathcal{C})$, we have:

- $s = t$ if and only if $\widehat{s} = \widehat{t}$,
- $st = u$ in $\mathbb{P}$ if and only if $\widehat{s}\,\widehat{t} \xrightarrow{*}_R \widehat{u}$, and
- $s \in L$ if and only if $\widehat{s} \in L$.

For the last point note that every $L \in \mathsf{IL}(\mathcal{C})$, viewed as a subset of $\mathbb{M}(A, I)$, is contained in $\mathrm{IRR}(R)$.

Hence, if we add for every variable x in θ the constraint $x \in \mathrm{IRR}(R)$ (recall that $\mathrm{IRR}(R) \in \mathsf{L}(\mathcal{D})$) and replace every equation $xy = z$ in θ by the rewriting constraint $xy \xrightarrow{*}_R z$, then we obtain a formula, which is satisfiable in the trace monoid $\mathbb{M}(A, I)$ if and only if the original formula θ is satisfiable in $\mathbb{P}$. Using the following lemma, we can replace the rewriting constraints $xy \xrightarrow{*}_R z$ by ordinary equations over $\mathbb{M}(A, I)$ plus A-local $\circ_\sigma$-predicates.

Lemma 3.15. *There exists a fixed positive boolean formula*

$$\psi(x, y, z, x_1, \ldots, x_m)$$

over the signature of $(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{D}), (\circ_\sigma)_{\sigma \in \Sigma})$ *such that for all* $x, y, z \in \mathrm{IRR}(R)$ *we have* $xy \xrightarrow{*}_R z$ *in* $\mathbb{M}(A, I)$ *if and only if*

$$(\mathbb{M}(A, I), \iota, \mathsf{L}(\mathcal{D}), (\circ_\sigma)_{\sigma \in \Sigma}) \models \exists x_1 \cdots \exists x_m : \psi(x, y, z, x_1, \ldots, x_m). \qquad (10)$$

Proof. Recall that $\mathcal{F}(A, I)$ is the set of all independence cliques in (A, I). For the further reasoning it is important to note that $a, b \in A_\sigma$ and $(a, c) \in I$ implies $(b, c) \in I$.

First we show that for all $x, y, z \in \mathrm{IRR}(R)$, $xy \xrightarrow{*}_R z$ in $\mathbb{M}(A, I)$ if and only if there exist $p, r, s, t, u \in \mathrm{IRR}(R)$ and $C_1, C_2 \in \mathcal{F}(A, I)$ such that in $(\mathbb{M}(A, I), \iota)$

$$[C_1][C_2] \xrightarrow{*}_R u, \quad \mathrm{inv}(p, r), \quad x = s\,[C_1]\,p, \quad y = r\,[C_2]\,t, \quad z = s\,u\,t. \qquad (11)$$

If (11) holds, then $xy \xrightarrow{*}_R z$ follows immediately. Now assume that $xy \xrightarrow{*}_R z$. We can choose $p \in \mathbb{M}(A, I)$ of maximal length such that $x = x'p$, $y = ry'$, and $\mathrm{inv}(p, r)$. Let $C_1 = \max(x') \in \mathcal{F}(A, I)$, $C_2 = \min(y') \in \mathcal{F}(A, I)$, and $[C_1][C_2] \xrightarrow{*}_R u \in \mathrm{IRR}(R)$. Hence, there are s and t with $x = s[C_1]p$, $y = r[C_2]t$, and $xy \xrightarrow{*}_R sut \xrightarrow{*}_R z$. Note that $p, r, s, t, u, [C_1], [C_2] \in \mathrm{IRR}(R)$. Since the length of p was chosen maximal, only rules of the form $(ab, c) \in R$, where $a \in C_1$, $b \in C_2$, and $a, b, c \in A_\sigma$ for some $\sigma \in \Sigma$, can be applied to the trace $[C_1][C_2]$. Thus, $\mathrm{alph}(u) = C_1 \cup C_2$, and if $(a, u) \in I$ for $a \in A$, then also $(a, C_1) \in I$. We claim that $sut \in \mathrm{IRR}(R)$, which implies $z = sut$ and hence (11). Assume that there exist $ab \in \mathrm{dom}(R)$ and traces q_1, q_2 such that $sut = q_1abq_2$. By Levi's

Lemma 2.1 we obtain up to symmetry one of the following two diagrams (recall that $s, u, t \in \mathrm{IRR}(R)$):

$$
\begin{array}{c||c|c|c}
q_2 & s_2 & u_2 & t_2 \\ \hline
ab & a & \varepsilon & b \\ \hline
q_1 & s_1 & u_1 & t_1 \\ \hline
 & s & u & t
\end{array}
\qquad\qquad
\begin{array}{c||c|c|c}
q_2 & s_2 & u_2 & t_2 \\ \hline
ab & a & b & \varepsilon \\ \hline
q_1 & s_1 & u_1 & t_1 \\ \hline
 & s & u & t
\end{array}
$$

Assume that $a, b \in A_\sigma$ ($\sigma \in \Sigma$). Let us first consider the left diagram. Since $(a, u_1) \in I$, $(b, u_2) \in I$, and $u = u_1 u_2$, we obtain $(a, u) \in I$ and thus $(a, C_1) \in I$. Furthermore, from the diagram we obtain also $(b, s_2) \in I$. Thus, $(a, s_2) \in I$, which implies $a \in \max(s)$. Together with $(a, C_1) \in I$ it follows that $a \in \max(s[C_1]) = \max(x') = C_1$, which contradicts $(a, C_1) \in I$.

Now let us consider the right diagram. Again we have $a \in \max(s)$. Furthermore, $(a, u_1) \in I$, i.e., $(b, u_1) \in I$. Hence, $b \in \min(u) \cap A_\sigma$. There are now two possibilities: either there exists $a' \in C_1 \cap A_\sigma$ or $b \in C_2$ and $(b, C_1) \in I$. If $a' \in C_1 \cap A_\sigma$, then $s[C_1]$ would contain the factor $aa' \in \mathrm{dom}(R_\sigma)$, which contradicts $x = s[C_1]p \in \mathrm{IRR}(R)$. If $b \in C_2$ and $(b, C_1) \in I$, then also $(a, C_1) \in I$, which implies $a \in \max(s[C_1]) = C_1$; the same contradiction as in the previous paragraph.

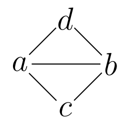
Thus, $xy \xrightarrow{*}_R z$ is equivalent to (11). Next, note that (11) is equivalent to

$$[C_1][C_2] \xrightarrow{*}_R u, \quad x = s\,[C_1]\,p, \quad y = \iota(p)\,[C_2]\,t, \quad p \in \mathbb{M}(U, I), \quad z = s\,u\,t. \quad (12)$$

Recall that $\mathbb{M}(U, I)$ belongs to $\mathsf{L}(\mathcal{D})$. It remains to replace the additional rewriting constraints of the form $[C_1][C_2] \xrightarrow{*}_R u$, where $C_1, C_2 \in \mathcal{F}(A, I)$, by local equations of the form $x' \circ_\sigma y' = z'$. Since $C_i \in \mathcal{F}(A, I)$ we can write down a disjunction over all independence cliques C_1' and C_2' in (Σ, I_Σ), with the meaning that $C_i' = \{\sigma \in \Sigma \mid C_i \cap A_\sigma \neq \emptyset\}$, and replace C_i in (12) by $x_{i,1} x_{i,2} \cdots x_{i,n_i}$, where $n_i = |C_i'| \leq |\Sigma|$ and $x_{i,j}$ is a new variable. Moreover, we add the constraints $x_{i,j} \in A_{\sigma(i,j)}$, where $C_i' = \{\sigma(i,j) \mid 1 \leq j \leq n_i\}$. Since there are at most $|\Sigma|^{\tau(\Sigma, I_\Sigma)+1}$ many cliques in (Σ, I_Σ), this results in a disjunction of $|\Sigma|^{O(\tau(\Sigma, I_\Sigma))}$ many conjunctions of size $O(|\Sigma|)$. Finally the rewriting constraint $x_{1,1} \cdots x_{1,n_1} x_{2,1} \cdots x_{2,n_2} \xrightarrow{*}_R u$ is equivalent to a conjunction of at most $|\Sigma|$ many local equations of the form $x' \circ_\sigma y' = z'$ with $x', y', z' \in A_\sigma$ and a single equation over $\mathbb{M}(A, I)$. $\square$

Let us illustrate the last step in the previous proof with an example:

Example 3.16. Assume that $\Sigma = \{a, b, c, d\}$ and the independence relation I_Σ looks as follows:



Then the rewriting constraint

$$x_a x_b x_c \, x'_a x'_b x_d \xrightarrow{*}_R u,$$

where $x_a, x'_a \in A_a$, $x_b, x'_b \in A_b$, $x_c \in A_c$, and $x_d \in A_d$, is equivalent to

$$x_a \circ_a x'_a = y_a \ \wedge \ x_b \circ_b x'_b = y_b \ \wedge \ u = y_a y_b x_c x_d \ \wedge \ y_a \in A_a \ \wedge \ y_b \in A_b.$$

Here, the equation $u = y_a y_b x_c x_d$ is interpreted in the trace monoid $\mathbb{M}(A, I)$.

By applying Lemma 3.15 to every rewriting constraint $xy \xrightarrow{*}_R z$, we obtain an equivalent formula over $(\mathbb{M}(A, I), \iota, (\circ_\sigma)_{\sigma \in \Sigma}, \mathsf{L}(\mathcal{D}))$. Since (Σ, I_Σ) is assumed to be fixed in Theorem 3.12, the size of the resulting conjunction increased only by a constant factor. This concludes the proof of Theorem 3.12.

4 Positive theories of graph products

In this section we consider positive theories of graph products of *finitely generated groups*. Assume that $\mathbb{P} = \mathbb{P}(\Sigma, I_\Sigma, (\mathcal{G}_\sigma)_{\sigma \in \Sigma})$ is a graph product such that every $\mathcal{G}_\sigma$ is a *finitely generated group*. Let Γ_σ be a finite generating set for $\mathcal{G}_\sigma$. Then $\mathbb{P}$ is generated by $\Gamma = \bigcup_{\sigma \in \Sigma} \Gamma_\sigma$. Let $D_\Sigma = (\Sigma \times \Sigma) \setminus I_\Sigma$ the dependence relation corresponding to I_Σ. A node $\sigma \in \Sigma$ is called an *isolated node of the dependence alphabet* (Σ, D_Σ) if $D_\Sigma(\sigma) = \{\sigma\}$. Throughout this section, we make the following two assumptions:

Assumption 4.1 *For every isolated node σ of the dependence alphabet (Σ, D_Σ), the positive theory* $\mathrm{posTh}(\mathcal{G}_\sigma, (a)_{a \in \Gamma_\sigma}, \mathrm{REC}(\mathcal{G}_\sigma))$ *is decidable.*

Assumption 4.2 *For every nonisolated node σ of (Σ, D_Σ), the existential theory* $\exists\mathrm{FOTh}(\mathcal{G}_\sigma, (a)_{a \in \Gamma_\sigma}, \mathrm{REC}(\mathcal{G}_\sigma))$ *is decidable.*

Since we restrict to finitely generated groups, we obtain finite representations for recognizable constraints. More precisely, since $\mathbb{P}$ is a group, it follows that $L \in \mathrm{REC}(\mathbb{P})$ if and only if there exists a surjective group homomorphism $\rho : \mathbb{P} \to S$ onto a finite group S such that $L = \rho^{-1}(\rho(L))$. Thus, L can be represented by the finite group S, the homomorphism ρ and $F \subseteq S$ with $L = \rho^{-1}(F)$. To represent ρ, it suffices to specify its value $\rho(a)$ for every generator $a \in \Gamma$.

The aim of this section is to prove the following result:

Theorem 4.3. *Let $\mathbb{P} = \mathbb{P}(\Sigma, I_\Sigma, (\mathcal{G}_\sigma)_{\sigma \in \Sigma})$ by a graph product such that Assumption 4.1 and Assumption 4.2 hold. Then* $\mathrm{posTh}(\mathbb{P}, (a)_{a \in \Gamma}, \mathrm{REC}(\mathbb{P}))$ *is decidable.*

Since the theory of a finite group is of course decidable, and the same holds for the theory of $\mathbb{Z}$ with rational constraints (Proposition 2.4), we obtain the following corollary, which was already stated in [16]:

Corollary 4.4. *Let $\mathbb{P}$ be a graph product of finite groups and free groups. Then* $\mathrm{posTh}(\mathbb{P}, (a)_{a\in\Gamma}, \mathrm{REC}(\mathbb{P}))$ *is decidable.*

Remark 4.5. Note that Corollary 4.4 cannot be extended by allowing monoids for the factors of the graph product. Already the positive $\forall\exists^3$-theory of the free monoid $\{a, b\}^*$ is undecidable [21, 39]. Similarly, Corollary 4.4 cannot be extended by replacing $\mathrm{REC}(\mathbb{P})$ by $\mathrm{RAT}(\mathbb{P})$, since the latter class contains a free monoid $\{a, b\}^*$ in case $\mathbb{P} = F_2$ is the free group of rank 2.

The proof of Theorem 4.3 follows the arguments from the proof of Corollary 18 in [16]:

- In a first step, we will reduce $\mathrm{posTh}(\mathbb{P}, (a)_{a\in\Gamma}, \mathrm{REC}(\mathbb{P}))$ to the positive theories $\mathrm{posTh}(\mathbb{P}_i, (a)_{a\in\Gamma_i}, \mathrm{REC}(\mathbb{P}_i))$, $1 \le i \le n$, where the $\mathbb{P}_i$ result from the connected components of the dependence alphabet (Σ, D_Σ). Thus, $\mathbb{P} = \prod_{i=1}^n \mathbb{P}_i$. After this step, we may assume that (Σ, D_Σ) is connected and (by Assumption 4.1) contains at least two nodes.
- Next, we will reduce $\mathrm{posTh}(\mathbb{P}, (a)_{a\in\Gamma}, \mathrm{REC}(\mathbb{P}))$ (where the underlying dependence alphabet (Σ, D_Σ) is connected and contains at least two nodes) to $\exists\mathrm{FOTh}(\mathbb{P} * F, (a)_{a\in\Gamma\cup K}, \mathrm{REC}(\mathbb{P} * F) \cup \mathcal{C})$. Here $F = F(K)$ is the free group generated by the finite set K, and the additional constraint class $\mathcal{C}$ contains all subgroups of $\mathbb{P} * F(K)$ of the form $\mathbb{P} * F(K')$ for $K' \subseteq K$.[9] This second step is inspired by techniques of Makanin and Merzlyakov [38, 43] developed for free groups. The proof of the main technical lemma is shifted into Section 4.3.
- The last step consists of an application of Theorem 3.12. In order to apply this theorem to $\exists\mathrm{FOTh}(\mathbb{P} * F, (a)_{a\in\Gamma\cup K}, \mathrm{REC}(\mathbb{P} * F) \cup \mathcal{C})$, we have to "decompose" the constraints using Lemma 3.9.

4.1 Simplifying the graph product $\mathbb{P}$

In a first step we may assume that no finite group $\mathcal{G}_\sigma$, $\sigma \in \Sigma$, is a direct product of two finite nontrivial groups, since otherwise we could replace σ by two independent nodes. In particular, if $\mathcal{G}_\sigma$ is not $\mathbb{Z}/2\mathbb{Z}$, then there must exist $a \in \mathcal{G}_\sigma$ such that $a^2 \neq 1_\sigma$, i.e., $a \neq a^{-1}$ in $\mathcal{G}_\sigma$. Next, assume that the dependence alphabet (Σ, D_Σ) consists of two nonempty disjoint components (Σ_1, D_1) and (Σ_2, D_2), which define graph products $\mathbb{P}_1$ and $\mathbb{P}_2$, respectively. Then $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$. Furthermore by Mezei's Theorem, see e.g. [6], every $L \in \mathrm{REC}(\mathbb{P})$ is effectively a finite union of sets of the form $L_1 \times L_2$ with $L_i \in \mathrm{REC}(\mathbb{P}_i)$. Since the corresponding statement for singleton sets (i.e., constants from Γ) holds as well, we may apply the following Proposition 4.6, which is a decomposition lemma in the style of the Feferman Vaught Theorem [23], see [16] for a proof.

[9] A subgroup $\mathbb{P} * F(K')$ with $K' \subsetneq K$ is not a recognizable subset of $\mathbb{P} * F(K)$, since $\mathbb{P} * F(K')$ has infinite index in $\mathbb{P} * F(K)$.

Proposition 4.6. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be monoids with classes $\mathcal{C}_1 \subseteq 2^{\mathcal{M}_1}$ and $\mathcal{C}_2 \subseteq 2^{\mathcal{M}_2}$. Let $\mathcal{C} \subseteq 2^{\mathcal{M}_1 \times \mathcal{M}_2}$ such that every $L \in \mathcal{C}$ is effectively a finite union of sets of the form $L_1 \times L_2$ with $L_1 \in \mathcal{C}_1$ and $L_2 \in \mathcal{C}_2$. If both* (pos)FOTh$(\mathcal{M}_1, \mathcal{C}_1)$ *and* (pos)FOTh$(\mathcal{M}_2, \mathcal{C}_2)$ *are decidable, then* (pos)FOTh$(\mathcal{M}_1 \times \mathcal{M}_2, \mathcal{C})$ *is decidable, too.*

The construction in our proof of Proposition 4.6 may lead to a nonelementary blow-up with respect to formula size. This will be the main bottle neck in our proof of Theorem 4.3.

By Proposition 4.6 and Assumption 4.1 we may assume that the dependence alphabet (Σ, D_Σ) is connected and contains at least two nodes. By Corollary 2.6 we can also exclude the case that Σ contains exactly two D_Σ-adjacent nodes, which are both labeled by $\mathbb{Z}/2\mathbb{Z}$. Thus, we may assume that either the graph (Σ, D_Σ) contains a path consisting of three different nodes or one of the groups $\mathcal{G}_\sigma$ has a generator $g \in \mathcal{G}_\sigma$ with $g^{-1} \neq g \neq 1_\sigma$. Hence, there exist three different generators $a \in \mathcal{G}_{\sigma(a)} \setminus \{1_{\sigma(a)}\}$, $b \in \mathcal{G}_{\sigma(b)} \setminus \{1_{\sigma(b)}\}$, and $c \in \mathcal{G}_{\sigma(c)} \setminus \{1_{\sigma(c)}\}$ such that

- $\sigma(a) \neq \sigma(b)$ and $(\sigma(a), \sigma(b)) \in D_\Sigma$,
- $\sigma(b) \neq \sigma(c)$ and $(\sigma(b), \sigma(c)) \in D_\Sigma$, and finally
- either $\sigma(a) \neq \sigma(c)$ or $a \neq a^{-1} = c$ in $\mathcal{G}_{\sigma(a)}$.

Thus, the dependency between a, b, and c being used is

$$a \text{ — } b \text{ — } c.$$

In Section 4.3, a, b, and c will always refer to these three elements.

4.2 Reducing to the existential theory

Our strategy for reducing the positive theory of $\mathbb{P}$ to an existential theory is based on [38, 43], but the presence of partial commutation and recognizable constraints makes the construction more involved: Given a positive sentence θ, which is interpreted over $\mathbb{P}$, we construct an *existential sentence* θ', which is interpreted over a free product $\mathbb{P} * F$ of $\mathbb{P}$ and a free group F, such that θ is true in $\mathbb{P}$ if and only if θ' is true in $\mathbb{P} * F$. Roughly speaking, θ' results from θ by replacing the universally quantified variables by the generators of the free group F.

Assume that we have given a positive boolean combination ϕ of equations with constants and recognizable constraints $x_i \in L_i$ $(1 \leq i \leq n)$, where the latter are represented via surjective homomorphisms $\rho_i : \mathbb{P} \to S_i$ such that $L_i = \rho_i^{-1}(\rho_i(L_i))$. Let $S = \prod_{i=1}^n S_i$ and define $\rho(x) = (\rho_1(x), \ldots, \rho_n(x))$ for $x \in \mathbb{P}$. Now we can replace every constraint $x_i \in L_i$ by constraints of the form $\rho(x_i) = q$ for $q \in S$. Note that the number of these constraints is bounded exponentially in the size of the description of ϕ. Thus, we may assume that all recognizable constraints in our initial positive formula are given in the form $\rho(x) = q$ for $q \in S$ and a fixed surjective homomorphism $\rho : \mathbb{P} \to S$ onto a finite group S.

Let K be a finite set of new constants, $K \cap \Gamma = \emptyset$. Recall that $F(K)$ is the free group generated by K. For the free product $\mathbb{P} * F(K)$ we write $\mathbb{P}[K]$ in the following. Instead of $\mathbb{P}[\{k_1, \ldots, k_n\}]$, we write $\mathbb{P}[k_1, \ldots, k_n]$. Similarly, instead of $\mathbb{P}[K \cup \{k\}]$ we write $\mathbb{P}[K, k]$. In the sequel we also have to deal with formulas, where the constraints are given by different extensions of our basic homomorphism $\rho : \mathbb{P} \to S$ to $\mathbb{P}[K]$. For this we introduce the following notation: Let $\mathcal{G}$ be an arbitrary group, and let $\varrho : \mathcal{G} \to S$ be a group homomorphism onto some finite group S. Let $K = \{k_1, \ldots, k_n\}$ and $q_1, \ldots, q_n \in S$. Then $\varrho^{k_1,\ldots,k_n}_{q_1,\ldots,q_n} : \mathcal{G}[K] \to S$ denotes the unique extension of ϱ, defined by $\varrho^{k_1,\ldots,k_n}_{q_1,\ldots,q_n}(k_i) = q_i$. Similarly, if ϕ is some boolean combination of equations and constraints of the form $\varrho(x) = q$, then $\phi^{k_1,\ldots,k_n}_{q_1,\ldots,q_n}$ denotes the formula that results from ϕ by replacing every constraint $\varrho(x) = q$ by $\varrho^{k_1,\ldots,k_n}_{q_1,\ldots,q_n}(x) = q$. Let us now fix a formula[10]

$$\theta(\widetilde{z}) \;\equiv\; \forall x_1 \exists y_1 \cdots \forall x_n \exists y_n \, \phi(x_1, \ldots, y_n, y_1, \ldots, y_n, \widetilde{z}),$$

with ϕ a positive boolean formula over the signature of $(\mathbb{P}, (a)_{a\in\Gamma}, \mathrm{REC}(\mathbb{P}))$ such that all recognizable constraints are given in the form $\rho(x) = q \in S$ for our fixed homomorphism $\rho : \mathbb{P} \to S$. Choose for every universally quantified variable x_i in θ a new constant k_i and let $K = \{k_1, \ldots, k_n\}$. The following theorem yields the reduction from the positive to the existential theory.

Theorem 4.7. *Let $\theta(\widetilde{z}) \equiv \forall x_1 \exists y_1 \cdots \forall x_n \exists y_n \, \phi(x_1, \ldots, x_n, y_1, \ldots, y_n, \widetilde{z})$ be a formula as above. For all $\widetilde{u} \in \mathbb{P}$ we have $\theta(\widetilde{u})$ in $\mathbb{P}$ if and only if*

$$\bigwedge_{q_1 \in S} \exists y_1 \cdots \bigwedge_{q_n \in S} \exists y_n \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le n} y_i \in \mathbb{P}[k_1, \ldots, k_i] \wedge \\ \phi^{k_1,\ldots,k_n}_{q_1,\ldots,q_n}(k_1, \ldots, k_n, y_1, \ldots, y_n, \widetilde{u}) \end{array} \right\} \text{ in } \mathbb{P}[K]. \tag{13}$$

Proof. Using Lemmas 4.9 and 4.10 below, the proof is the same as in [16, Theorem 17]. □

To complete the proof of Theorem 4.3, we apply Theorem 3.12 to the group $\mathbb{P}[K]$, which is a graph product as well: Add every $k \in K$ as an isolated node to the independence alphabet (Σ, I_Σ) and label it with $F(k) \cong \mathbb{Z}$. For every $\sigma \in \Sigma$ let $\mathcal{C}_\sigma = \mathrm{REC}(\mathcal{G}_\sigma) \cup \{\{a\} \mid a \in \Gamma_\sigma\}$ and for every $k \in K$ let $\mathcal{C}_k = \mathrm{RAT}(F(k))$, which contains $\mathrm{REC}(F(k))$ and every singleton subset. Let $\mathcal{C} = \bigcup_{\tau \in \Sigma \cup K} \mathcal{C}_\tau$. By Assumption 4.2 (note that (Σ, D_Σ) does not contain isolated nodes by the simplifications from the previous section), $\exists\mathrm{FOTh}(\mathcal{G}_\sigma, \mathcal{C}_\sigma)$ is decidable for every $\sigma \in \Sigma$. By Proposition 2.4, for every $k \in K$, $\exists\mathrm{FOTh}(F(k), \mathcal{C}_k)$ is decidable as well. Thus, in order to apply Theorem 3.12, it suffices to show that all constraint sets and constants (viewed as singleton sets) in (13) belong to $\mathsf{IL}(\mathcal{C})$. For the constants this is clear – they all belong to $\Gamma \cup K$. Also $\mathbb{P}[k_1, \ldots, k_i] \in \mathsf{IL}(\mathcal{C})$ is easy to see. Finally, $\mathrm{REC}(\mathbb{P}[K]) \subseteq \mathsf{IL}(\mathcal{C})$ by Lemma 3.9.

[10] In the following symbols with a tilde like $\widetilde{x}$ will denote sequences of arbitrary length over some set that will be always clear form the context. If say $\widetilde{a} = (a_1, \ldots, a_m)$, then $\widetilde{a} \in A$ means $a_1 \in A, \ldots, a_m \in A$.

Remark 4.8. Concerning the complexity, it can be shown that our proof of Theorem 4.3 leads to a nonelementary blow-up due to the construction in our proof of Proposition 4.6. On the other hand, if we restrict to connected graphs (Σ, D_Σ), then Proposition 4.6 becomes superfluous. Due to Corollary 2.6 and the complexity statement in Theorem 3.12, we obtain an elementary reduction from the positive theory to the theories in Assumption 4.2.

For the further consideration let us fix a set of constants K and a further constant $k \notin K$. Moreover, let $K_i \subseteq K$ for $1 \le i \le m$. Fix also $q \in S$ and a sequence

$$\widetilde{u} = (u_1, \ldots, u_N) \tag{14}$$

of elements $u_i \in \mathbb{P}$. The simple proof of the following lemma is the same as for [16, Lemma 19].

Lemma 4.9. *Let $\phi(x, y_1, \ldots, y_m, \widetilde{z})$ be a positive boolean formula with constraints of the form $\varrho(y) = p$ for $p \in S$ and (possibly different) extensions $\varrho : \mathbb{P}[K] \to S$ of our fixed homomorphism $\rho : \mathbb{P} \to S$. If*

$$\exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{P}[K_i, k] \wedge \\ \phi_q^k(k, y_1, \ldots, y_m, \widetilde{u}) \end{array} \right\} \text{ in } \mathbb{P}[K, k],$$

then

$$\forall x \in \mathbb{P} \cap \rho^{-1}(q)\ \exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{P}[K_i] \wedge \\ \phi(x, y_1, \ldots, y_m, \widetilde{u}) \end{array} \right\} \text{ in } \mathbb{P}[K].$$

Note that the assertion of Lemma 4.9 does not hold in general if ϕ involves negations. For example $\forall x : x \neq 1$ is false, but $k \neq 1$ is true. On the other hand, the converse implication of Lemma 4.9 is true for arbitrary formulas:

Lemma 4.10. *Let $\phi(x, y_1, \ldots, y_m, \widetilde{z})$ be a not necessarily positive boolean formula with constraints of the form $\varrho(y) = p$ for $p \in S$ and (possibly different) extensions $\varrho : \mathbb{P}[K] \to S$ of our fixed homomorphism $\rho : \mathbb{P} \to S$. If*

$$\forall x \in \mathbb{P} \cap \rho^{-1}(q)\ \exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{P}[K_i] \wedge \\ \phi(x, y_1, \ldots, y_m, \widetilde{u}) \end{array} \right\} \text{ in } \mathbb{P}[K],$$

then

$$\exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{P}[K_i, k] \wedge \\ \phi_q^k(k, y_1, \ldots, y_m, \widetilde{u}) \end{array} \right\} \text{ in } \mathbb{P}[K, k].$$

The statement of Lemma 4.10 will be shown by a reduction to the underlying trace monoid with involution. For this, we need one more lemma. First, we have to introduce a few notations.

In Lemma 4.10, the groups $\mathbb{P} = \mathbb{P}(\Sigma, I_\Sigma, (\mathcal{G}_\sigma)_{\sigma\in\Sigma})$, $\mathbb{P}[K_i]$, $\mathbb{P}[K_i, k]$, $\mathbb{P}[K]$, and $\mathbb{P}[K, k]$ appear ($1 \le i \le m$). Similarly to Section 3.3 we define $A_\sigma = \mathcal{G}_\sigma \setminus \{1_\sigma\}$ for $\sigma \in \Sigma$, $A = \bigcup_{\sigma\in\Sigma} A_\sigma$, and $I = \bigcup_{(\sigma,\tau)\in I_\Sigma} A_\sigma \times A_\tau$. Let $\mathbb{M} = \mathbb{M}(A, I)$. Since every $\mathcal{G}_\sigma$ is a group, we can define a total involution ι on A by taking the inverse in each group $\mathcal{G}_\sigma$ and lift this involution to $\mathbb{M}$ in the standard way. Next, take for each constant $\kappa \in K \cup \{k\}$ a new copy $\overline{\kappa}$. Let $\overline{K} = \{\overline{\kappa} \mid \kappa \in K\}$ and similarly for $\overline{K_i}$. We extend the involution ι on A to $A \cup K \cup \overline{K} \cup \{k, \overline{k}\}$ by setting $\iota(\kappa) = \overline{\kappa}$ and $\iota(\overline{\kappa}) = \kappa$ for $\kappa \in K \cup \{k\}$. Then ι can be also lifted to the free product $\mathbb{M} * (K \cup \overline{K} \cup \{k, \overline{k}\})^*$, which will be the largest trace monoid in our further investigation. We will use the following abbreviations in the sequel: $\mathbb{M}[K_i] = \mathbb{M}*(K_i \cup \overline{K}_i)^*$, $\mathbb{M}[K_i, k] = \mathbb{M}*(K_i \cup \overline{K}_i \cup \{k, \overline{k}\})^*$, $\mathbb{M}[K] = \mathbb{M}*(K \cup \overline{K})^*$, and $\mathbb{M}[K, k] = \mathbb{M}*(K \cup \overline{K} \cup \{k, \overline{k}\})^*$. Finally let R be the trace rewriting system on $\mathbb{M}[K, k]$ defined by

$$R = \bigcup_{\sigma\in\Sigma} \{ab \to c \mid a, b, c \in A_\sigma,\ a \circ_\sigma b = c\} \cup \{\iota(a)a \to \varepsilon \mid a \in A_\sigma\} \cup \bigcup_{\kappa\in K\cup\{k\}} \{\kappa\overline{\kappa} \to \varepsilon, \overline{\kappa}\kappa \to \varepsilon\}. \tag{15}$$

Then R is confluent and $\mathbb{M}[K, k]/\overset{*}{\leftrightarrow}_R \cong \mathbb{P}[K, k]$. Similarly, if we restrict R to traces from $\mathbb{M}[K]$ (resp. $\mathbb{M}$), then $\mathbb{M}[K]/\overset{*}{\leftrightarrow}_R \cong \mathbb{P}[K]$ (resp. $\mathbb{M}/\overset{*}{\leftrightarrow}_R \cong \mathbb{P}$).

Let $\widetilde{w} = (w_1, \ldots, w_N)$, where $w_i \in \mathbb{M} \cap \mathrm{IRR}(R)$ is the unique irreducible trace representing the fixed group element $u_i \in \mathbb{P}$ from (14). In the following, we identify a homomorphism $\varrho : \mathbb{P}[K] \to S$ with $h \circ \varrho : \mathbb{M}[K] \to S$, where $h : \mathbb{M}[K] \to \mathbb{P}[K]$ is the canonical homomorphism that maps a trace t to the group element represented by t. Moreover, for $\varrho : \mathbb{M}[K] \to S$ we denote with $\varrho^k_q : \mathbb{M}[K, k] \to S$ the unique extension of ϱ, defined by $\varrho^k_q(k) = q$ and $\varrho^k_q(\overline{k}) = q^{-1}$.

As in Section 3.3, in the following lemma $\circ_\sigma$ denotes the ternary relation $\{(a, b, c) \mid a, b, c \in A_\sigma, a \circ_\sigma b = c\} \subseteq \mathbb{M}^3$ for $\sigma \in \Sigma$.

Lemma 4.11. *Let $\chi(x, y_1, \ldots, y_m, \widetilde{z})$ be a not necessarily positive boolean formula over the signature of $(\mathbb{M}[K], \iota, (a)_{a\in\Gamma\cup K}, \mathrm{REC}(\mathbb{M}[K]), (\circ_\sigma)_{\sigma\in\Sigma})$ such that all recognizable constraints in χ have the form $\varrho(y) = p$ for $p \in S$ and (possibly different) extensions $\varrho : \mathbb{M}[K] \to S$ of our fixed homomorphism $\rho : \mathbb{M} \to S$. If*

$$\forall x \in \mathbb{M} \cap \mathrm{IRR}(R) \cap \rho^{-1}(q)\ \exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge_{1\le i\le m} y_i \in \mathbb{M}[K_i] \cap \mathrm{IRR}(R) \\ \wedge\ \chi(x, y_1, \ldots, y_m, \widetilde{w}) \end{array} \right\} \text{ in } \mathbb{M}[K],$$

then there are $s_1, s_2 \in \mathbb{M} \cap \mathrm{IRR}(R)$ *with* $\rho(s_1) q \rho(s_2) = q$ *in the finite group* S *and*

$$\exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{M}[K_i, k] \cap \mathrm{IRR}(R) \\ \wedge\ \chi_q^k(s_1 k s_2, y_1, \ldots, y_m, \widetilde{w}) \end{array} \right\} \text{ in } \mathbb{M}[K, k].$$

The proof of Lemma 4.11 is the main technical difficulty and shifted to the next section. Using Lemma 4.11, we can finish the proof of Lemma 4.10: Assume that

$$\forall x \in \mathbb{P} \cap \rho^{-1}(q)\ \exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{P}[K_i] \wedge \\ \phi(x, y_1, \ldots, y_m, \widetilde{u}) \end{array} \right\} \text{ in } \mathbb{P}[K].$$

By restricting every variable in ϕ to $\mathbb{M}[K] \cap \mathrm{IRR}(R)$ and replacing every equation $x'y' = z'$ by $x'y' \xrightarrow{*}_R z'$, we obtain a true statement over $\mathbb{M}[K]$. As in the proof of Lemma 3.15, we can replace every rewriting constraint $x'y' \xrightarrow{*}_R z'$ by a formula $\psi(x', y', z', \ldots)$ over the signature of $(\mathbb{M}[K], \iota, (\circ_\sigma)_{\sigma \in \Sigma \cup K})$.[11] This transformation introduces only new existentially quantified variables ($\widetilde{y}$ below). We obtain a formula χ over the signature of $(\mathbb{M}[K], \iota, \mathrm{REC}(\mathbb{M}[K]), (\circ_\sigma)_{\sigma \in \Sigma \cup K})$ such that

$$\forall x \in \mathbb{M} \cap \mathrm{IRR}(R) \cap \rho^{-1}(q)\ \exists y_1 \cdots \exists y_m\ \exists \widetilde{y} \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{M}[K_i] \cap \mathrm{IRR}(R) \\ \wedge\ \widetilde{y} \in \mathbb{M}[K] \cap \mathrm{IRR}(R) \\ \wedge\ \chi(x, y_1, \ldots, y_m, \widetilde{y}, \widetilde{w}) \end{array} \right\}$$

is true in $\mathbb{M}[K]$. Thus, by Lemma 4.11 there exist $s_1, s_2 \in \mathbb{M} \cap \mathrm{IRR}(R)$ such that $\rho(s_1) q \rho(s_2) = q$ in the finite group S and

$$\exists y_1 \cdots \exists y_m \exists \widetilde{y} \left\{ \begin{array}{l} \bigwedge\limits_{1 \le i \le m} y_i \in \mathbb{M}[K_i, k] \cap \mathrm{IRR}(R) \\ \wedge\ \widetilde{y} \in \mathbb{M}[K, k] \cap \mathrm{IRR}(R) \\ \wedge\ \chi_q^k(s_1 k s_2, y_1, \ldots, y_m, \widetilde{y}, \widetilde{w}) \end{array} \right\} \text{ in } \mathbb{M}[K, k].$$

By doing the above transformation from $\mathbb{P}[K]$ to $\mathbb{M}[K]$ via Lemma 3.15 backwards, i.e., from $\mathbb{M}[K, k]$ to $\mathbb{P}[K, k]$, it follows that

$$\exists y_1 \cdots \exists y_m \left\{ \bigwedge_{1 \le i \le m} y_i \in \mathbb{P}[K_i, k]\ \wedge\ \phi_q^k(s_1 k s_2, y_1, \ldots, y_m, \widetilde{u}) \right\} \text{ in } \mathbb{P}[K, k], \quad (16)$$

where $s_i \in \mathbb{M} \cap \mathrm{IRR}(R)$ is identified with the group element from $\mathbb{P}$ it represents. Let us define a group homomorphism $f : \mathbb{P}[K, k] \to \mathbb{P}[K, k]$ by $f(k) = s_1^{-1} k s_2^{-1}$

[11] The formula ψ, constructed in the proof of Lemma 3.15 contains constraints, which we want to avoid here. The constraint $p \in \mathbb{M}(U, I)$ in (12) can be omitted here, because in our situation the involution ι is completely defined. Furthermore, constraints of the form $x \in A_\sigma$ are also used in the proof of Lemma 3.15. Such a constraint is equivalent to $x \circ_\sigma \iota(x) = 1$.

and $f(x) = x$ for $x \in \mathbb{P}[K]$. First, note that f is injective (the homomorphism defined by $g(k) = s_1 k s_2$ defines an inverse). Thus, the truth value of all (negated) equations is preserved by f. Moreover, $f(\widetilde{u}) = \widetilde{u}$ (since $\widetilde{u} \in \mathbb{P}$) and $\rho(s_1) q \rho(s_2) = q$ in S. Thus, $\varrho_q^k(s_1^{-1} k s_2^{-1}) = \rho(s_1)^{-1} q \rho(s_2)^{-1} = q = \varrho_q^k(k)$ for every extension ϱ of ρ, i.e., all recognizable constraints are also preserved by f. Finally, $f(s_1 k s_2) = s_1 s_1^{-1} k s_2^{-1} s_2 = k$ in $\mathbb{P}[K, k]$. Hence, applying f to (16) yields

$$\exists y_1 \cdots \exists y_m \left\{ \bigwedge_{1 \le i \le m} y_i \in \mathbb{P}[K_i, k] \ \wedge \ \phi_q^k(k, y_1, \ldots, y_m, \widetilde{u}) \right\} \text{ in } \mathbb{P}[K, k].$$

4.3 Proof of Lemma 4.11

Recall that $\mathbb{M} \subseteq \mathbb{M}[K_i] \subseteq \mathbb{M}[K] \subseteq \mathbb{M}[K, k]$, $q \in S$, and $\widetilde{w} = (w_1, \ldots, w_N)$ with $w_i \in \mathbb{M} \cap \mathrm{IRR}(R)$ are already fixed. On $\mathbb{M}[K, k]$ we defined the confluent trace rewriting system R by (15). Let $D = (A \cup K \cup \overline{K} \cup \{k, \overline{k}\})^2 \setminus I$ be the dependence relation corresponding to $\mathbb{M}[K, k]$. The involution ι is totally defined on $\mathbb{M}[K, k]$. In the following we will write $\overline{t}$ instead of $\iota(t)$. Recall that (Σ, D_Σ) is assumed to be connected with $|\Sigma| > 1$. Let $\chi(x, y_1, \ldots, y_m, \widetilde{z})$ be an arbitrary boolean formula with atomic predicates of the form $xy = z$, $x = \overline{y}$, $x = t$, $x \circ_\sigma y = z$, and $\varrho(x) = p$, where x, y, and z are variables, $t \in \mathbb{M}[K]$ is a constant (w.l.o.g. $|t| \le 1$), $p \in S$, and $\varrho : \mathbb{M}[K] \to S$ is some extension of our basic homomorphism $\rho : \mathbb{M} \to S$. Since ρ was derived from a corresponding group homomorphism on $\mathbb{P}$, $s \overset{*}{\leftrightarrow}_R t$ implies $\rho(s) = \rho(t)$ for $s, t \in \mathbb{M}$. Let

$$W = \{w_1, \overline{w}_1, \ldots, w_N, \overline{w}_N\} \tag{17}$$

and let d be the number of equations of the form $xy = z$ that occur in χ. Choose a number $\lambda \in \mathbb{N}$ such that $|S|$ divides $\lambda - 1$ and $\lambda \ge 2d + 1$.

We start with the definition of some specific traces. A *chain* is a trace $t = a_1 a_2 \cdots a_\kappa$ such that $a_i \in A_{\sigma_i}$ $(1 \le i \le \kappa)$ and $[\sigma_1, \sigma_2, \ldots, \sigma_\kappa]$ is a path in the dependence graph (Σ, D_Σ) with $\sigma_i \ne \sigma_{i+1}$ for $1 \le i \le \kappa - 1$. Thus, $t \in \mathbb{M} \cap \mathrm{IRR}(R)$ and $(a_i, a_{i+1}) \in D$ for $1 \le i \le \kappa - 1$.

Recall that we have fixed symbols $a, b, c \in A$ at the end of Section 4.1 such that $(a, b), (b, c) \in D$ and either a, b, and c belong to pairwise different A_σ or $a \ne \overline{a} = c$. It is possible that $(a, c) \in I$. If a, b, and c belong to pairwise different A_σ, then let

$$e_b = (b\, a)^{|S|} (c\, b)^{|S|}.$$

Otherwise, we have $a \ne \overline{a}$, i.e., $a^2 = a'$ for some $a' \in A$. Then let

$$e_b = (b\, a)^{|S|-1}\, b\, a'\, b\, (a\, b)^{|S|-1};$$

note that in $\mathbb{P}$ this trace equals $(b\, a)^{|S|} (a\, b)^{|S|}$. In both cases, $e_b \in \mathbb{M} \cap \mathrm{IRR}(R)$ is a trace with $\min(e_b) = \max(e_b) = \{b\}$ and $\rho(e_b) = 1$.

Lemma 4.12. *There is a trace $\ell \in \mathbb{M} \cap \mathrm{IRR}(R)$ such that $\rho(\ell) = q$, $(\ell, t) \in D$ for every $t \neq \varepsilon$, and* $\min(\ell) = \max(\ell) = \{a\}$.

Proof. First, for every $x \in A$ we construct a trace $t(x) \in \mathbb{M} \cap \mathrm{IRR}(R)$ such that $\min(t(x)) = \{x\}$, $\max(t(x)) = \{\overline{x}\}$, and $\rho(t(x)) = 1$. Let s be a chain such that $x\,s\,b$ is a chain, which exists since (Σ, D_Σ) is connected. Then set $t(x) = x\,s\,e_b\,\overline{s}\,\overline{x}$. Now we construct ℓ as follows:

- Select a trace $s = b_1 b_2 \cdots b_\kappa \in \mathrm{IRR}(R)$, $b_i \in A$, with $\rho(b_1 \cdots b_\kappa) = q$. Recall that ρ was assumed to be surjective, hence s exists.
- Let $u_1, \ldots, u_{\kappa+1} \in \mathbb{M} \cap \mathrm{IRR}(R)$ be chains, visiting every subset A_σ ($\sigma \in \Sigma$), such that the trace $au_1b_1u_2b_2 \cdots u_\kappa b_\kappa u_{\kappa+1}a$ is also a chain. These u_i exist, since (Σ, D_Σ) is connected.
- If $u_i = c_1 \cdots c_{\kappa_i}$ with $c_j \in A$, then define $v_i = t(c_1) \cdots t(c_{\kappa_i})$; thus $\rho(v_i) = 1$.
- Finally, let $\ell = t(a)v_1b_1v_2b_2 \cdots v_\kappa b_\kappa v_{\kappa+1}t(\bar{a})$.

The construction implies that ℓ has indeed the desired properties. $\square$

For the rest of the section let $\ell \in \mathbb{M}$ be some trace satisfying the properties from the previous lemma.

A *trace system of degree n* is a tuple $\mathcal{R} = (r_0, \ldots, r_\lambda)$ of $\lambda + 1$ traces $r_i = t_i e_b$ with $t_i \in \{(ba)^{|S|}, (bc)^{|S|}\}^n$ for some n large enough. The value of n will be made more precise later. Note that $\rho(r_i) = 1$ and that the traces r_i are irreducible and almost chains; only the single factor ac (in case $(a, c) \in I$) in e_b leads to commutation. There are $2^{n(\lambda+1)}$ trace systems of degree n. We append the trace e_b to every t_i in order to assure that every r_i starts and ends with b.

An *overlapping of two traces $u, v \in \mathbb{M}$* is a trace s with $u = ts$ and $v = st'$ for some $t, t' \in \mathbb{M}$. The trace system $\mathcal{R} = (r_0, \ldots, r_\lambda)$ has *no long overlapping*, if

- the traces $r_0, \overline{r}_0, \ldots, r_\lambda, \overline{r}_\lambda$ are pairwise different, and
- for all $0 \leq i, j \leq \lambda$, $u \in \{r_i, \overline{r}_i\}$, and $v \in \{r_j, \overline{r}_j\}$ we have: if s is an overlapping of u and v with $|s| \geq \frac{|r_i| - |\ell|}{2} = (n+2)|S| - \frac{|\ell|}{2}$, then $s = u = v$.

Note that this implies in particular that if $r_i \ell r_{i+1} = urv$ with $r \in \{r_j, \overline{r}_j\}$, then either $u = \varepsilon$ and $r_i = r$ or $v = \varepsilon$ and $r_j = r$, i.e., r cannot be properly contained in $r_i \ell r_{i+1}$.

The following lemma can be derived by standard techniques that random strings are incompressible, the formal proof is therefore omitted. The idea is that if the trace system $\mathcal{R}$ has a long overlapping, then, in case n is large enough, the description of $\mathcal{R}$ can be compressed to less than $n(\lambda + 1)$ bits. But this cannot happen for all systems $\mathcal{R}$.

Lemma 4.13. *There exists n_0 (depending only on λ and $|S|$) such that for all $n \geq n_0$ there exists a trace system of degree n without long overlapping.*

Remark 4.14. Later, we will use $\mathcal{R}$ to construct a trace s, which can be replaced by the trace $s_1 k s_2$ in Lemma 4.11. An explicit construction of s without using the notion of random strings is sketched in [13].

Let us fix a trace system $\mathcal{R} = (r_0, \ldots, r_\lambda)$ of degree n without long overlapping, where

$$2|r_i| + |\ell| = 4(n+2)|S| + |\ell| > |w| \tag{18}$$

for all $w \in W$ from (17). For every $1 \le i \le \lambda$ define the length-reducing trace rewriting system

$$T_i = \{r_{i-1}\,\ell\,r_i \to r_{i-1}\,k\,r_i,\ \overline{r_i}\,\overline{\ell}\,\overline{r_{i-1}} \to \overline{r_i}\,\overline{k}\,\overline{r_{i-1}}\}.$$

We consider T_i as a trace rewriting system over our largest trace monoid $\mathbb{M}[K,k]$. Note that $W \cup A \subseteq \mathrm{IRR}(T_i)$ by (18) and that $s \to_{T_i} t$ implies also $\overline{s} \to_{T_i} \overline{t}$.

Lemma 4.15. *Every trace rewriting system T_i is confluent.*

Proof. Since T_i is terminating, we have to verify that T_i is locally confluent. Assume that $t \;{}_{T_i}\!\!\leftarrow s \to_{T_i} u$, where t and u both result from s by an application of the rule $r_{i-1}\ell r_i \to r_{i-1}kr_i$, the other two cases can be dealt analogously. Thus, there exist traces $t_1, t_2, u_1, u_2 \in \mathbb{M}[K,k]$ such that

$$s = t_1 r_{i-1}\,\ell\,r_i t_2 = u_1 r_{i-1}\,\ell\,r_i u_2 \text{ and } t = t_1 r_{i-1}\,k\,r_i t_2,\ u = u_1 r_{i-1}\,k\,r_i u_2.$$

Now we apply Levi's Lemma 2.1 to the identity $t_1 r_{i-1}\,\ell\,r_i t_2 = u_1 r_{i-1}\,\ell\,r_i u_2$. Recall that every r_j starts and ends with b. Hence, nonempty prefixes (resp. suffixes) of r_{i-1} (resp. r_i) are dependent. Moreover, by Lemma 4.12 the trace ℓ is dependent from every nonempty trace. Thus, we obtain up to symmetry one of the following two diagrams:

u_2	ε	s_2	t_2
$r_{i-1}\,\ell\,r_i$	ε	$r_{i-1}\,\ell\,r_i$	ε
u_1	t_1	s_1	ε
	t_1	$r_{i-1}\,\ell\,r_i$	t_2

u_2	ε	ε	u_2
$r_{i-1}\,\ell\,r_i$	ε	s	s_2
u_1	t_1	s_1	v
	t_1	$r_{i-1}\,\ell\,r_i$	t_2

In the first case, $s_1 = \varepsilon = s_2$ and thus $t = u$. In the second case, we may assume that $s_1 \neq \varepsilon \neq s_2$, since otherwise we obtain a special case of the first diagram. Furthermore, if $s = \varepsilon$, then

$$t \to_{T_i} t_1\,r_{i-1}\,k\,r_i\,v\,r_{i-1}\,k\,r_i\,u_2 \;{}_{T_i}\!\!\leftarrow u.$$

Thus, assume that also $s \neq \varepsilon$. Since $r_{i-1}\,\ell\,r_i = s_1\,s = s\,s_2$ with $s_1 \neq \varepsilon \neq s_2$, and $\mathcal{R}$ has no long overlapping, there exist traces r and r' such that $s_1 = r_{i-1}\,\ell\,r$,

$s_2 = r'\,\ell\, r_i$, $r_i = r\, s$, $r_{i-1} = s\, r'$. Since $(v, s) \in I$, we obtain

$$
\begin{aligned}
t = t_1 r_{i-1}\, k\, r_i t_2 \;&=\; t_1\, r_{i-1}\, k\, r\, s\, v\, r'\, \ell\, r_i\, u_2 \\
&=\; t_1\, r_{i-1}\, k\, r\, v\, s\, r'\, \ell\, r_i\, u_2 \\
&\to_{T_i} t_1\, r_{i-1}\, k\, r\, v\, s\, r'\, k\, r_i\, u_2 \\
&=\; t_1\, r_{i-1}\, k\, r\, s\, v\, r'\, k\, r_i\, u_2 \\
{}_{T_i}\!\!&\leftarrow t_1\, r_{i-1}\, \ell\, r\, s\, v\, r'\, k\, r_i\, u_2 \\
&=\; t_1\, r_{i-1}\, \ell\, r\, v\, s\, r'\, k\, r_i\, u_2 = u_1 r_{i-1}\, k\, r_i u_2 = u.
\end{aligned}
$$

Thus, T_i is confluent. $\square$

The previous lemma implies that for every $1 \leq i \leq \lambda$, every trace $s \in \mathbb{M}[K, k]$ has a unique normal form $\mathrm{NF}_{T_i}(s) \in \mathrm{IRR}(T_i)$. In the following, we briefly write $\mathrm{NF}_i(s)$ for $\mathrm{NF}_{T_i}(s)$. The following lemma is easy to verify. For the last point note that $\rho(\ell) = q = \rho_q^k(k)$.

Lemma 4.16. *For every $1 \leq i \leq \lambda$ and $s \in \mathbb{M}[K]$ we have:*

- $\mathrm{NF}_i(s) = s$ *if $|s| \leq 1$ or $s \in W$, in particular, if $a \circ_\sigma b = c$ for $a, b, c \in A_\sigma$ ($\sigma \in \Sigma$), then also $\mathrm{NF}_i(a) \circ_\sigma \mathrm{NF}_i(b) = \mathrm{NF}_i(c)$,*
- $\overline{\mathrm{NF}_i(s)} = \mathrm{NF}_i(\overline{s})$*, and*
- $\varrho(s) = \varrho_q^k(\mathrm{NF}_i(s))$ *for every extension $\varrho : \mathbb{M}[K] \to S$ of $\rho : \mathbb{M} \to S$.*

Thus, every normal form mapping NF_i preserves constants, the involution $\overline{}$, and recognizable constraints. On the other hand, concatenation in $\mathbb{M}[K]$ is in general not preserved, but the following statement will suffice:

Lemma 4.17. *Let $u, v \in \mathbb{M}[K]$. There are at most two $i \in \{1, \ldots, \lambda\}$ such that $\mathrm{NF}_i(u)\mathrm{NF}_i(v) \neq \mathrm{NF}_i(uv)$.*

Proof. Assume that $1 \leq i \leq \lambda$ is such that $\mathrm{NF}_i(u)\mathrm{NF}_i(v) \in \mathrm{RED}(T_i)$. We only consider the case that $\mathrm{NF}_i(u)\mathrm{NF}_i(v) = s r_i \ell r_{i+1} t$ for some $s, t \in \mathbb{M}[K]$. Due to the dependencies between nonempty suffixes and prefixes of r_i, ℓ, and r_{i+1}, we obtain one of the following three diagrams (where $r'_j \neq \varepsilon \neq r''_j$ for $j \in \{i-1, i\}$):

$$
\begin{array}{r||c|c|c|c|c|}
\mathrm{NF}_i(v) & s_2 & r''_{i-1} & \ell & r_i & t \\ \hline
\mathrm{NF}_i(u) & s_1 & r'_{i-1} & \varepsilon & \varepsilon & \varepsilon \\ \hline\hline
 & s & r_{i-1} & \ell & r_i & t
\end{array}
\qquad
\begin{array}{r||c|c|c|c|c|}
\mathrm{NF}_i(v) & \varepsilon & \varepsilon & \varepsilon & r''_i & t_2 \\ \hline
\mathrm{NF}_i(u) & s & r_{i-1} & \ell & r'_i & t_1 \\ \hline\hline
 & s & r_{i-1} & \ell & r_i & t
\end{array}
$$

$$
\begin{array}{r||c|c|c|c|c|}
\mathrm{NF}_i(v) & s_2 & \varepsilon & \ell_2 & r_i & t_2 \\ \hline
\mathrm{NF}_i(u) & s_1 & r_{i-1} & \ell_1 & \varepsilon & t_1 \\ \hline\hline
 & s & r_{i-1} & \ell & r_i & t
\end{array}
$$

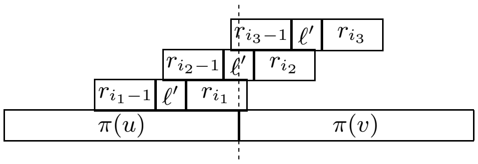


Fig. 2.

Since every r_j starts and ends with b, it follows that $(s_2, b) \in I$ (resp. $(t_1, b) \in I$) in the first and third (resp. second and third) diagram. Let π denote the homomorphism on $\mathbb{M}[K, k]$ that projects onto the subalphabet $\{a, \overline{a}, b, \overline{b}, c, \overline{c}, k, \overline{k}\}$. Thus, $\pi(s_2) = \pi(t_1) = \varepsilon$. It follows that one of the following three cases holds, where $x, y \in \{a, \overline{a}, b, \overline{b}, c, \overline{c}, k, \overline{k}\}^*$ and $\ell' = \pi(\ell)$:

- $\pi(\mathrm{NF}_i(u)) = xr$ and $\pi(\mathrm{NF}_i(v)) = r'\ell' r_i y$ where $r_{i-1} = rr'$
- $\pi(\mathrm{NF}_i(u)) = xr_{i-1}\ell' r$ and $\pi(\mathrm{NF}_i(v)) = r'y$, where $r_i = rr'$
- $\pi(\mathrm{NF}_i(u)) = xr_{i-1}\ell'_1$ and $\pi(\mathrm{NF}_i(v)) = \ell'_2 r_i y$, where $\ell' = \ell'_1\ell'_2$

But then there are also $x', y' \in \{a, \overline{a}, b, \overline{b}, c, \overline{c}\}^*$ with

- $\pi(u) = x'r$ and $\pi(v) = r'\ell' r_i y'$ where $r_{i-1} = rr'$ or
- $\pi(u) = x'r_{i-1}\ell' r$ and $\pi(v) = r'y'$, where $r_i = rr'$ or
- $\pi(u) = x'r_{i-1}\ell'_1$ and $\pi(v) = \ell'_2 r_i y'$, where $\ell' = \ell'_1\ell'_2$

The traces x' and y' result from x and y, respectively, by replacing every occurrence of k and $\overline{k}$, respectively, by ℓ' and $\overline{\ell'}$, respectively. Thus $\pi(u) = x'z_1$, $\pi(v) = z_2 y'$, $z_1 \neq \varepsilon \neq z_2$, and $z_1 z_2 = r_{i-1}\ell' r_i$. Now assume that this holds for three different i_1, i_2, and i_3. Then it is easy to see that two of the three traces $r_{j-1}\ell' r_j$ ($j \in \{i_1, i_2, i_3\}$) have a "long" overlapping, contradicting the fact that $\mathcal{R}$ has no long overlapping. See Figure 2 for a typical constellation, where there is a long overlapping between r_{i_1} and r_{i_2-1} as well as between r_{i_2} and r_{i_3-1}. $\square$

Since moreover $\mathrm{NF}_i(u) = \mathrm{NF}_i(v)$ implies $u = v$ for all $u, v \in \mathbb{M}[K]$, we obtain the following lemma – recall that $\lambda \geq 2d + 1$, where d is the number of equations in the formula χ.

Lemma 4.18. *Let $x_j, y_j, z_j \in \mathbb{M}[K]$ for $1 \leq j \leq d$. Then there exists $1 \leq i \leq \lambda$ such that for all $1 \leq j \leq d$ we have $x_j y_j = z_j$ if and only if $\mathrm{NF}_i(x_j)\mathrm{NF}_i(y_j) = \mathrm{NF}_i(z_j)$.*

Now we are able to prove Lemma 4.11: Assume that

$$\forall x \in \mathbb{M} \cap \mathrm{IRR}(R) \cap \rho^{-1}(q)\ \exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \bigwedge\limits_{1 \leq i \leq m} y_i \in \mathbb{M}[K_i] \cap \mathrm{IRR}(R) \\ \wedge\ \chi(x, y_1, \ldots, y_m, \widetilde{w}) \end{array} \right\} \text{ in } \mathbb{M}[K].$$

Let $s = r_0\,\ell\, r_1\,\ell \cdots r_{\lambda-1}\,\ell\, r_\lambda \in \mathbb{M} \cap \mathrm{IRR}(R)$. Since $\rho(r_i) = 1$ and λ was chosen such that $|S|$ is a divisor of $\lambda - 1$, we have $\rho(s) = \rho(\ell^\lambda) = q^\lambda = q$. Thus, there exist traces $t_i \in \mathbb{M}[K_i] \cap \mathrm{IRR}(R)$, $1 \leq i \leq m$, with $\chi(s, t_1, \ldots, t_m, \widetilde{w})$ in $\mathbb{M}[K]$. By Lemma 4.16 and Lemma 4.18 there exists $1 \leq j \leq \lambda$ such that $\chi^k_q(\mathrm{NF}_j(s), \mathrm{NF}_j(t_1), \ldots, \mathrm{NF}_j(t_m), \widetilde{w})$ in $\mathbb{M}[K, k]$. Since $\mathcal{R}$ has no long overlapping, there exists only a single occurrence of $r_{j-1}\ell r_j$ in s. Thus, we can write $\mathrm{NF}_j(s) = s_1\, k\, s_2$ for $s_1, s_2 \in \mathbb{M} \cap \mathrm{IRR}(R)$ such that $\rho(s_1)q\rho(s_2) = \rho(s) = q$. Thus,

$$\exists y_1 \cdots \exists y_m \left\{ \begin{array}{l} \displaystyle\bigwedge_{1 \leq i \leq m} y_i \in \mathbb{M}[K_i, k] \cap \mathrm{IRR}(R) \\ \wedge\ \chi^k_q(s_1 k s_2, y_1, \ldots, y_m, \widetilde{w}) \end{array} \right\} \text{ in } \mathbb{M}[K, k].$$

5 Open problems

Concerning existential theories, the following problems might deserve further investigations:

- Is the additional exponential summand in Theorem 3.12 unavoidable?
- Is Assumption 3.10 necessary in Theorem 3.12?
- Is the existential theory of a hyperbolic group decidable? Recall that the existential theory of a torsion-free hyperbolic group is decidable [54, 60].
- Is the existential theory of an automatic group [22] undecidable? At least for asynchronous automatic groups [22] this is the case, in fact already conjugacy is in general undecidable for asynchronous automatic groups [4].

Further results concerning undecidable existential theories for groups and monoids can be found in [53, 56].

For positive theories it remains open whether an elementary reduction is possible in Theorem 4.3 in case (Σ, D_Σ) is not connected. One might also investigate, whether the positive theory of a torsion-free hyperbolic group is decidable. Further results on positive theories can be found in [30, 34, 55, 62].

Finally, one may hope to get decidability results for full first-order theories of restricted graph products, like for instance graph groups. Kharlampovich and Myasnikov proved in a series of papers the decidability of the full first-order theory of a free group (this problem was known as Tarski's problem) [31]. One approach might be to generalize the techniques developed by Kharlampovich and Myasnikov to broader classes of groups. We should mention here that elementary decision procedures cannot be expected for full first-order theories. By a result of Semenov [61], the full first-order theory of a free group of rank 2 is nonelementary.

References

1. I. J. Aalbersberg and H. J. Hoogeboom. Characterizations of the decidability of some problems for regular trace languages. *Mathematical Systems Theory*, 22:1–19, 1989.
2. A. Baudisch. Kommutationsgleichungen in semifreien Gruppen. *Acta Mathematica Academiae Scientiarum Hungaricae*, 29:235–249, 1977. In German.

3. A. Baudisch. Subgroups of semifree groups. *Acta Mathematica Academiae Scientiarum Hungaricae*, 38:19–28, 1981.
4. G. Baumslag, S. Gersten, M. Shapiro, and H. Short. Automatic groups and amalgams. *Journal of Pure Applied Algebra*, 76(3):229–316, 1991.
5. L. Berman. The complexity of logical theories. *Theoretical Computer Science*, 11:71–77, 1980.
6. J. Berstel. *Transductions and context–free languages*. Teubner Studienbücher, Stuttgart, 1979.
7. R. V. Book and F. Otto. *String–Rewriting Systems*. Springer, 1993.
8. N. Brady and J. Meier. Connectivity at infinity for right angled Artin groups. *Transactions of the American Mathematical Society*, 353:117–132, 2001.
9. J. R. Büchi and S. Senger. Definability in the existential theory of concatenation and undecidable extensions of this theory. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 34(4):337–342, 1988.
10. P. Cartier and D. Foata. *Problèmes combinatoires de commutation et réarrangements*. Number 85 in Lecture Notes in Mathematics. Springer, 1969.
11. F. Dahmani. Existential questions in (relatively) hyperbolic groups *and* finding relative hyperbolic structures. Technical report, arXiv.org, 2006. `http://arxiv.org/abs/math.GR/0505345`.
12. V. Diekert, C. Gutiérrez, and C. Hagenah. The existential theory of equations with rational constraints in free groups is PSPACE-complete. *Information and Computation*, 202(2):105–140, 2005.
13. V. Diekert and M. Lohrey. Existential and positive theories of equations in graph products. In H. Alt and A. Ferreira, editors, *Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science (STACS 2002), Juan les Pins (France)*, number 2285 in Lecture Notes in Computer Science, pages 501–512. Springer, 2002.
14. V. Diekert and M. Lohrey. A note on the existential theory of equations in plain groups. *International Journal of Algebra and Computation*, 12(1 & 2):1–7, 2002.
15. V. Diekert and M. Lohrey. Word equations over graph products. In *Proceedings of the 23rd Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2003), Mumbai (India)*, number 2914 in Lecture Notes in Computer Science, pages 156–167. Springer, 2003.
16. V. Diekert and M. Lohrey. Existential and positive theories of equations in graph products. *Theory of Computing Systems*, 37(1):133–156, 2004.
17. V. Diekert, Y. Matiyasevich, and A. Muscholl. Solving word equations modulo partial commutations. *Theoretical Computer Science*, 224(1–2):215–235, 1999.
18. V. Diekert and A. Muscholl. Solvability of equations in free partially commutative groups is decidable. *International Journal of Algebra and Computation*, 16(6):1047–1069, 2006.
19. V. Diekert and G. Rozenberg, editors. *The Book of Traces*. World Scientific, 1995.
20. C. Droms. Graph groups, coherence and three-manifolds. *Journal of Algebra*, 106(2):484–489, 1985.
21. V. G. Durnev. Undecidability of the positive $\forall\exists^3$-theory of a free semi-group. *Sibirsky Matematicheskie Jurnal*, 36(5):1067–1080, 1995. English translation.
22. D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston. *Word processing in groups*. Jones and Bartlett, Boston, 1992.
23. S. Feferman and R. L. Vaught. The first order properties of products of algebraic systems. *Fundamenta Mathematicae*, 47:57–103, 1959.
24. E. Fohry and D. Kuske. On graph products of automatic and biautomatic monoids. *Semigroup Forum*, 72(3):337–352, 2006.
25. E. R. Green. *Graph Products of Groups*. PhD thesis, The University of Leeds, 1990.
26. C. Gutiérrez. Satisfiability of equations in free groups is in PSPACE. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC'2000)*, pages 21–27. ACM Press, 2000.
27. S. Hermiller and J. Meier. Algorithms and geometry for graph products of groups. *Journal of Algebra*, 171:230–257, 1995.
28. W. Hodges. *Model Theory*. Cambridge University Press, 1993.
29. M. Jantzen. Confluent string rewriting. In *EATCS Monographs on Theoretical Computer Science*, volume 14. Springer, 1988.
30. B. Khan, A. G. Myasnikov, and D. E. Serbin. On positive theories of groups with regular free length function. to appear in *International Journal of Algebra and Computation*, 2005.

31. O. Kharlampovich and A. Myasnikov. Elementary theory of free non-abelian groups. *Journal of Algebra*, 302(2):451–552, 2006.
32. D. Kuske and M. Lohrey. Logical aspects of Cayley-graphs: the monoid case. *International Journal of Algebra and Computation*, 16(2):307–340, 2006.
33. J. Loeffler, J. Meier, and J. Worthington. Graph products and Cannon pairs. *International Journal of Algebra and Computation*, 12(6):747–754, 2002.
34. M. Lohrey and G. Sénizergues. Theories of HNN-extensions and amalgamated products. In M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, editors, *Proceedings of the 33st International Colloquium on Automata, Languages and Programming (ICALP 2006), Venice (Italy)*, number 4052 in Lecture Notes in Computer Science, pages 681–692. Springer, 2006.
35. R. C. Lyndon and P. E. Schupp. *Combinatorial Group Theory*. Springer, 1977.
36. G. S. Makanin. The problem of solvability of equations in a free semigroup. *Math. Sbornik*, 103:147–236, 1977. In Russian; English translation in *Math. USSR Sbornik 32, 1977*.
37. G. S. Makanin. Equations in a free group. *Izv. Akad. Nauk SSR*, Ser. Math. 46:1199–1273, 1983. In Russian; English translation in *Math. USSR Izvestija 21, 1983*.
38. G. S. Makanin. Decidability of the universal and positive theories of a free group. *Izv. Akad. Nauk SSSR*, Ser. Mat. 48:735–749, 1984. In Russian; English translation in *Math. USSR Izvestija, 25*, 75–88, 1985.
39. S. S. Marchenkov. Unsolvability of the positive $\forall\exists$-theory of a free semi-group. *Sibirsky Matematicheskie Jurnal*, 23(1):196–198, 1982.
40. A. Mazurkiewicz. Concurrent program schemes and their interpretations. DAIMI Rep. PB 78, Aarhus University, Aarhus, 1977.
41. J. D. McKnight. Kleene quotient theorems. *Pacific Journal of Mathematics*, 14:1343–1352, 1964.
42. J. Meier. When is the graph product of hyperbolic groups hyperbolic? *Geometriae Dedicata*, 61:29–41, 1996.
43. Y. I. Merzlyakov. Positive formulas on free groups. *Algebra i Logika Sem.*, 5(4):25–42, 1966. In Russian.
44. A. Muscholl. *Decision and Complexity Issues on Concurrent Systems*. Habilitation thesis, Universität Stuttgart, 1999.
45. A. Muscholl and D. Peled. Message sequence graphs and decision problems on Mazurkiewicz traces. In M. Kutylowski, L. Pacholski, and T. Wierzbicki, editors, *Proceedings of the 24th International Symposium on Mathematical Foundations of Computer Science (MFCS'99), Szklarska Poreba (Poland)*, number 1672 in Lecture Notes in Computer Science, pages 81–91. Springer, 1999.
46. P. Narendran and F. Otto. Preperfectness is undecidable for Thue systems containing only length-reducing rules and a single commutation rule. *Information Processing Letters*, 29:125–130, 1988.
47. M. H. A. Newman. On theories with a combinatorial definition of "equivalence". *Annals of Mathematics*, 43:223–243, 1943.
48. E. Ochmański. Regular behaviour of concurrent systems. *Bulletin of the European Association for Theoretical Computer Science (EATCS)*, 27:56–67, 1985.
49. C. H. Papadimitriou. *Computational Complexity*. Addison Wesley, 1994.
50. W. Plandowski. Satisfiability of word equations with constants is in PSPACE. *Journal of the Association for Computing Machinery*, 51(3):483–496, 2004.
51. M. Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Comptes Rendus du Premier Congrès des Mathématicienes des Pays Slaves*, pages 92–101, Warsaw, 1929.
52. W. V. O. Quine. Concatenation as a basis for arithmetic. *Journal of Symbol Logic*, 11(4):105–114, 1946.
53. N. N. Repin. Some simply presented groups for which an algorithm recognizing solvability of equations is impossible. *Voprosy Kibernetiki (Moskva)*, 134:167–175, 1988. In Russian.
54. E. Rips and Z. Sela. Canonical representatives and equations in hyperbolic groups. *Inventiones Mathematicae*, 120:489–512, 1995.
55. B. V. Rozenblat. Positive theories of free inverse semigroups. *Siberian Mathematical Journal*, 20:910–918, 1980. English translation.
56. B. V. Rozenblat. Diophantine theories of free inverse semigroups. *Siberian Mathematical Journal*, 26:860–865, 1985. English translation.
57. J. Sakarovitch. On regular trace languages. *Theoretical Computer Science*, 52:59–75, 1987.

58. J. Sakarovitch. The “last” decision problem for rational trace languages. In I. Simon, editor, *Proceedings of the 1st Latin American Symposium on Theoretical Informatics (LATIN'92)*, number 583 in Lecture Notes in Computer Science, pages 460–473. Springer, 1992.
59. K. U. Schulz. Makanin's algorithm for word equations — Two improvements and a generalization. In K. U. Schulz, editor, *Word Equations and Related Topics*, number 572 in Lecture Notes in Computer Science, pages 85–150. Springer, 1991.
60. Z. Sela. Diophantine geometry over groups VIII: The elementary theory of a hyperbolic group. Available via `http://www.ma.huji.ac.il/zlil/`, 2002.
61. A. L. Semenov. An interpretation of free algebras in free groups. *Soviet Math. Dokl.*, 21:952–955, 1980.
62. Y. M. Vazhenin and B. V. Rozenblat. Decidability of the positive theory of a free countably generated semigroup. *Mathematics of USSR Sbornik.*, 44(1):109–116, 1983. English translation.
63. A. Veloso da Costa. Graph products of monoids. *Semigroup Forum*, 63(2):247–277, 2001.
64. A. Veloso da Costa. On graph products of automatic monoids. *R.A.I.R.O. — Informatique Théorique et Applications*, 35(5):403–417, 2001.