Living without Beth and Craig: Definitions and Interpolants in the Guarded and Two-Variable Fragments

Jean Christoph Jung
University of Hildesheim
jungj@uni-hildesheim.de

Frank Wolter
University of Liverpool
wolter@liverpool.ac.uk

***Abstract*—In logics with the Craig interpolation property (CIP) the existence of an interpolant for an implication follows from the validity of the implication. In logics with the projective Beth definability property (PBDP), the existence of an explicit definition of a relation follows from the validity of a formula expressing its implicit definability. The two-variable fragment, $\mathbf{FO}^2$, and the guarded fragment, GF, of first-order logic both fail to have the CIP and the PBDP. We show that nevertheless in both fragments the existence of interpolants and explicit definitions is decidable. In GF, both problems are** 3ExpTime-**complete in general, and** 2ExpTime-**complete if the arity of relation symbols is bounded by a constant $c \geq 3$. In $\mathbf{FO}^2$, we prove a** coN2ExpTime **upper bound and a** 2ExpTime **lower bound for both problems. Thus, both for GF and $\mathbf{FO}^2$ existence of interpolants and explicit definitions are decidable but harder than validity (in case of $\mathbf{FO}^2$ under standard complexity assumptions).**

I. Introduction

A logic enjoys the *Craig Interpolation Property* (CIP) if an implication $\varphi \Rightarrow \psi$ is valid if, and only if, there exists a formula χ using only the common symbols of φ and ψ such that $\varphi \Rightarrow \chi$ and $\chi \Rightarrow \psi$ are both valid. The formula χ is then called an interpolant for $\varphi \Rightarrow \psi$. The CIP is generally regarded as one of the most important and useful results in formal logic, with numerous applications [1], [2], [3], [4], [5]. One particularly interesting consequence of the CIP is the *Projective Beth Definability Property* (PBDP), which states that if a relation is implicitly definable over symbols in a signature τ, then it is explicitly definable over τ.

From an algorithmic viewpoint, the CIP and PBDP are of interest because they reduce existence problems to validity checking: an interpolant *exists* if, and only if, an implication is valid and an explicit definition *exists* if, and only if, a straightforward formula stating implicit definability is valid. The interpolant and explicit definition existence problems are thus not harder than validity.

In this article, we investigate the interpolant and explicit definition existence problem for two fragments of first-order logic (FO) that fail to have the CIP and PBDP: the guarded fragment (GF) and the two-variable fragment (FO^2) of FO.

GF has been introduced as a generalization of modal logic that enjoys many of its attractive algorithmic and model-theoretic properties, including decidability, the finite model property, the tree-like model property, and preservation properties such as the Łoś-Tarski preservation theorem [6], [7]. Since its introduction, the guarded fragment and variants of it have been investigated extensively [8], [9], [10], [11], not only as a natural generalisation of modal logic but also in databases and knowledge representation [12], [13].

While GF is a good generalization of modal logic in many respects, in contrast to modal logic it neither enjoys the CIP [14] nor the PBDP [15]. Note, however, that GF does enjoy the (non-projective) Beth Definability Property (BDP) in which the signature τ of the implicit and explicit definitions contains all symbols except the relation to be defined [14].

Fragments of first-order logic with at most $k \geq 1$ variables have been investigated in a variety of contexts, for example in finite-model theory [16], [17], [18]. The two-variable fragment FO^2 is of particular interest as it is decidable (and any k-variable fragment with $k \geq 3$ is undecidable) and also generalizes modal logic. In fact, satisfiability of FO^2 formulas is NExpTime-complete [19] and FO^2 shares with modal logic and GF the finite model property. In contrast to modal logic and GF, however, it does not enjoy any tree-like model property and is less robust under extensions [20], [21], [22].

Failure of the CIP for FO^2 was first shown in [23], [24] using algebraic techniques. In contrast to GF, FO^2 does not only not enjoy the PBDP but also not the BDP [25], [26].

In this article, we aim to understand better the complexity of deciding the existence of interpolants and explicit definitions for logics that do not enjoy the CIP and PBDP. In addition, our motivation for investigating these existence problems in GF and FO^2 stems from the following applications.

Strong separability of labeled data under ontologies. There are several scenarios in which one aims to find a logical formula that separates positive from negative examples given in the form of labeled data items. Examples include concept learning in description logic [27], reverse engineering of database queries, also known as query by example (QBE) [28], and generating referring expression (GRE), where the aim is to find a formula that separates a single positive data item from all other data items [29]. In [30], [31] an attempt is made to provide a unifying framework for these scenarios under the assumption that the data is given by a relational database and additional background information is available in the form of an ontology in first-order logic. A natural version of separability then asks whether for an ontology $\mathcal{O}$, a database $\mathcal{D}$, a signature τ of relation symbols, and sets P (of positive examples) and N (of negative examples) of tuples in $\mathcal{D}$ of the same length whether there exists a formula φ over τ that separates P from N in the sense that $\mathcal{O} \cup \mathcal{D} \models \varphi(\mathbf{a})$ for all $\mathbf{a} \in P$, and $\mathcal{O} \cup \mathcal{D} \models \neg\varphi(\mathbf{b})$, for all $\mathbf{b} \in N$. For the fundamental cases that $\mathcal{O}$ is in GF or FO^2 and one asks for a separating formula in GF or FO^2, respectively, it is not difficult to see that there is a polynomial time reduction of separability to interpolant existence. Moreover, interpolants give rise to separating formulas and vice versa.

Explicit definitions of relation symbols under GF and FO^2-sentences. The existence and computation of explicit definitions of relations has been proposed for ontology engineering [4] and for query rewriting under views and query-reformulation and compilation [32], [33], [5]. Thus, in these applications the focus shifts from interpolants to the existence of explicit definitions over a signature.

The following theorem summarizes our results:

Theorem 1. *(i) The explicit GF-definability and the GF-interpolant existence problems are both* 3EXPTIME*-complete in general, and* 2EXPTIME*-complete if the arity of relation symbols is bounded by a constant* $c \geq 3$.

(ii) The explicit FO^2-definability and the FO^2-interpolant existence problems are in CON2EXPTIME *and* 2EXPTIME*-hard.* 2EXPTIME*-hardness holds already for explicit FO^2-definability using any symbol except the defined one.*

For GF, it follows that interpolant and explicit definition existence are exactly one exponential harder than validity, both in general and if the arity of relation symbols is bounded by a constant $c \geq 3$ [7]. We note that exactly one ternary relation symbol is needed to obtain 2EXPTIME-hardness and that it is known from [34] that the fragment of GF with only unary and binary relation symbols enjoys the CIP and the PBDP and so interpolant and explicit definition existence are EXPTIME-complete in that case. Explicit GF-definability using any symbols except the defined one is polynomial time reducible to validity since GF has the BDP. For FO^2, it follows that all these problems are harder than validity, unless CONEXPTIME = 2EXPTIME. Finding tight complexity bounds remains an open problem in this case.

The proofs start with a straightforward model-theoretic characterization of the non-existence of an interpolant for an implication $\varphi \Rightarrow \psi$ by the existence of appropriate bisimulations between models satisfying φ and $\neg\psi$, respectively. The *guarded bisimulations* used for GF were introduced in [6] to characterize the expressive power of GF within FO, see also [35], [36]. The FO^2-bisimulations used for FO^2 are a straightforward variant of the well known pebble games characterizing finite variable logics [37], [38]. For GF, we then employ a mosaic based approach, using as mosaics sets of types over $\varphi, \neg\psi$ which can be satisfied by tuples that are guarded bisimilar. Constraints for sets of such mosaics characterize when they can be linked together to construct, simultaneously, models of φ and $\neg\psi$ and a guarded bisimulation between them. The triple exponential upper bound then follows from the observation that there are triple exponentially many mosaics. If the arity of relation symbols is bounded by a constant, then there are only double exponentially many mosaics. The lower bounds are proved by a reduction of the word problem for space-bounded alternating Turing machines.

For FO^2 we show, using mosaics that are similar to those introduced for GF, that if there are FO^2-bisimilar models satisfying FO^2-formulas φ, $\neg\psi$, then there are such models of at most double-exponential size. The CON2EXPTIME upper bound follows immediately from this finite model property result. The lower bound is again proved by reduction of the word problem for space-bounded alternating Turing machines.

II. Related Work

The problem of deciding the existence of explicit definitions and interpolants has hardly been studied for logics without the PBDP and CIP, respectively. Exceptions are linear temporal logic, LTL, for which the decidability of interpolant existence has been shown in [39], [40], [41] and description logics with nominals and/or role inclusions for which 2EXPTIME-completeness has recently been shown in [42]. Both the upper and lower bound proofs presented in this article are inspired by [42] but the proofs for GF and FO^2 are significantly more involved than for the considered description logics.

The CIP and PBDP are well understood for guarded fragments of FO. The first results on GF itself and various fragments of GF were obtained in [43], [14], [34]. In particular, after proving that GF fails to have the CIP it is shown that a natural modal version of the CIP holds for GF: if $\varphi \Rightarrow \psi$ is valid for guarded formulas φ, ψ, then there exists an interpolant for φ, ψ that may use in addition to the symbols shared by φ and ψ any relation symbol that occurs as a guard in φ or ψ.

More recently, the guarded negation fragment of FO (GNF) has been introduced. GNF extends GF by adding, in a careful way, unions of conjunctive queries [11]. Although GNF extends GF significantly, it is still decidable, has the finite model property, has the tree-like model property, and enjoys various preservation theorems [11], [44]. Importantly, and in contrast to GF, GNF enjoys the CIP and the PBDP [44], [45]. Thus, the existence of Craig interpolants and explicit definitions reduces to validity checking which is 2EXPTIME-complete in GNF and even in EXPTIME if the arity of relation symbols is bounded by a constant. Thus, the existence of interpolants and explicit definitions is one exponential harder in GF than in GNF.

Also related is work on uniform interpolation for GF. As GF does not enjoy the CIP, it also does not enjoy the uniform interpolation property (UIP). However, in [46], the authors consider the same modal-like fragment as [34] and show that the CIP generalizes to the UIP for this fragment. Uniform interpolant existence for GF and FO^2 has been considered in [47]. In contrast to the decidability results obtained in this article, it is shown that uniform interpolant existence is undecidable for both GF and FO^2.

The CIP in finite variable fragments of FO has been investigated from a modal viewpoint in [48]. Variants of the interpolation property and the Beth definability property that hold for finite variable logics are given in [49] using pebble games. Recently, it has been shown in [50] that FO^2 enjoys the weak Beth definability property (wBDP), in contrast to the BDP introduced above. The definition of wBDP is the same as that of BDP except that only those implicit definitions have to be made explicit which also have the existence property, not only the uniqueness property.

Also relevant for this work is the investigation of interpolation and definability in modal logic in general [51] and in hybrid modal logic [52], [53].

III. Preliminaries

Let τ range over relational signatures not containing function or constant symbols. Denote by $\mathrm{FO}(\tau)$ the set of first-order (FO) formulas constructed from atomic formulas $x = y$ and $R(\mathbf{x})$, $R \in \tau$, using conjunction, disjunction, negation, and existential and universal quantification. The signature $\mathrm{sig}(\varphi)$ of an FO-formula φ is the set of relation symbols used in it. As usual, we write $\varphi(\mathbf{x})$ to indicate that the free variables in φ are all from $\mathbf{x}$ and call a formula *sentence* in case it has no free variables. $\mathrm{FO}(\tau)$ is interpreted in τ-structures $\mathfrak{A} = (\mathrm{dom}(\mathfrak{A}), (R^{\mathfrak{A}})_{R\in\tau})$, where $\mathrm{dom}(\mathfrak{A})$ is the non-empty *domain* of $\mathfrak{A}$, and each $R^{\mathfrak{A}}$ is a relation over $\mathrm{dom}(\mathfrak{A})$ whose arity matches that of R. We often drop τ and simply speak of structures $\mathfrak{A}$.

In the *guarded fragment*, GF of FO [6], [7], formulas are built from atomic formulas $R(\mathbf{x})$ and $x = y$ by applying the Boolean connectives and *guarded quantifiers* of the form

$$\forall \mathbf{y}(\alpha(\mathbf{x},\mathbf{y}) \rightarrow \varphi(\mathbf{x},\mathbf{y})) \text{ and } \exists \mathbf{y}(\alpha(\mathbf{x},\mathbf{y}) \wedge \varphi(\mathbf{x},\mathbf{y}))$$

where $\varphi(\mathbf{x},\mathbf{y})$ is a guarded formula, and $\alpha(\mathbf{x},\mathbf{y})$ is an atomic formula that contains all variables in $\mathbf{x},\mathbf{y}$. The formula α is called the *guard of the quantifier*. $\mathrm{GF}(\tau)$ denotes the set of all guarded formulas (also called GF-formulas) over signature τ. We regard $\forall \mathbf{y}(\alpha(\mathbf{x},\mathbf{y}) \rightarrow \varphi(\mathbf{x},\mathbf{y}))$ as an abbreviation for $\neg\exists \mathbf{y}(\alpha(\mathbf{x},\mathbf{y}) \wedge \neg\varphi(\mathbf{x},\mathbf{y}))$. The *two-variable fragment*, FO^2, of FO consists of all formulas in FO using two distinct variables.

Let $\mathfrak{A}$ be structure. A pair $\mathfrak{A}, \mathbf{a}$ with $\mathbf{a}$ a tuple in $\mathfrak{A}$ is called a *pointed structure*. It will be convenient to use the notation $[\mathbf{a}] = \{a_1, \ldots, a_n\}$ to denote the set of components of the tuple $\mathbf{a} = (a_1, \ldots, a_n) \in \mathrm{dom}(\mathfrak{A})^n$. Similarly, for a tuple $\mathbf{x} = (x_1, \ldots, x_n)$ of variables we use $[\mathbf{x}]$ to denote the set $\{x_1, \ldots, x_n\}$.

We next recall model-theoretic characterizations of when pointed structures cannot be distinguished in either GF or FO^2. We begin by introducing $\mathrm{GF}(\tau)$-bisimulations (often called guarded τ-bisimulations) [36]. A set $G \subseteq \mathrm{dom}(\mathfrak{A})$ is *guarded* in $\mathfrak{A}$ if G is a singleton or there exists R with $\mathfrak{A} \models R(\mathbf{a})$ such that $G = [\mathbf{a}]$. A tuple $\mathbf{a} \in \mathrm{dom}(\mathfrak{A})^n$ is *guarded* in $\mathfrak{A}$ if $[\mathbf{a}]$ is a subset of some guarded set in $\mathfrak{A}$.

For tuples $\mathbf{a} = (a_1, \ldots, a_n)$ in $\mathfrak{A}$ and $\mathbf{b} = (b_1, \ldots, b_n)$ in $\mathfrak{B}$ we call a mapping p from $[\mathbf{a}]$ to $[\mathbf{b}]$ with $p(a_i) = b_i$ for $1 \leq i \leq n$ (written $p : \mathbf{a} \mapsto \mathbf{b}$) a *partial τ-isomorphism* if p is an isomorphism from the τ-reduct of $\mathfrak{A}_{|[\mathbf{a}]}$ onto $\mathfrak{B}_{|[\mathbf{b}]}$.

A set I of partial τ-isomorphisms $p : \mathbf{a} \mapsto \mathbf{b}$ from guarded tuples $\mathbf{a}$ in $\mathfrak{A}$ to guarded tuples $\mathbf{b}$ in $\mathfrak{B}$ is called a *GF(τ)-bisimulation* if the following hold for all $p : \mathbf{a} \mapsto \mathbf{b} \in I$:

(i) for every guarded tuple $\mathbf{a}'$ in $\mathfrak{A}$ there exists a guarded tuple $\mathbf{b}'$ in $\mathfrak{B}$ and $p' : \mathbf{a}' \mapsto \mathbf{b}' \in I$ such that p' and p coincide on $[\mathbf{a}] \cap [\mathbf{a}']$.

(ii) for every guarded tuple $\mathbf{b}'$ in $\mathfrak{B}$ there exists a guarded tuple $\mathbf{a}'$ in $\mathfrak{A}$ and $p' : \mathbf{a}' \mapsto \mathbf{b}' \in I$ such that p'^{-1} and p^{-1} coincide on $[\mathbf{b}] \cap [\mathbf{b}']$.

Assume that $\mathbf{a}$ and $\mathbf{b}$ are (possibly not guarded) tuples in $\mathfrak{A}$ and $\mathfrak{B}$. Then we say that the pointed structures $\mathfrak{A}, \mathbf{a}$ and $\mathfrak{B}, \mathbf{b}$ are *GF(τ)-bisimilar*, in symbols $\mathfrak{A}, \mathbf{a} \sim_{\mathrm{GF},\tau} \mathfrak{B}, \mathbf{b}$, if there exists a partial τ-isomorphism $p : \mathbf{a} \mapsto \mathbf{b}$ and a $\mathrm{GF}(\tau)$-bisimulation I such that Conditions (i) and (ii) hold for p.

Next we introduce appropriate bisimulations for FO^2, which are essentially a relational variant of the infinite 2-pebble games which have been used to characterize the expressive power of FO^2, see e.g. [37], [38]. Given structures $\mathfrak{A}, \mathfrak{B}$, a relation $S \subseteq \mathrm{dom}(\mathfrak{A}) \times \mathrm{dom}(\mathfrak{B})$ is an *$FO^2(\tau)$-bisimulation between $\mathfrak{A}$ and $\mathfrak{B}$* if S is *global*, that is, $\mathrm{dom}(\mathfrak{A}) \subseteq \{a \mid (a,b) \in S\}$ and $\mathrm{dom}(\mathfrak{B}) \subseteq \{b \mid (a,b) \in S\}$ and, for every $(a,b) \in S$ the following conditions are satisfied:

(i) for every $a' \in \mathrm{dom}(\mathfrak{A})$, there is a $b' \in \mathrm{dom}(\mathfrak{B})$ such that $(a,a') \mapsto (b,b')$ is a partial τ-isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ and $(a',b') \in S$;

(ii) for every $b' \in \mathrm{dom}(\mathfrak{B})$, there is a $a' \in \mathrm{dom}(\mathfrak{A})$ such that $(a,a') \mapsto (b,b')$ is a partial τ-isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ and $(a',b') \in S$.

For tuples $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n)$ of equal length $n = 0, 1, 2$, we write $\mathfrak{A}, \mathbf{a} \sim_{\mathrm{FO}^2,\tau} \mathfrak{B}, \mathbf{b}$ iff $\mathbf{a} \mapsto \mathbf{b}$ is a partial τ-isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ and there is an $\mathrm{FO}^2(\tau)$-bisimulation S between $\mathfrak{A}$ and $\mathfrak{B}$ such that $(a_i, b_i) \in S$, for all $i \leq n$.

Now, let $\mathcal{L}$ be either GF or FO^2. We write $\mathfrak{A}, \mathbf{a} \equiv_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b}$ and call $\mathfrak{A}, \mathbf{a}$ and $\mathfrak{B}, \mathbf{b}$ *$\mathcal{L}(\tau)$-equivalent* if $\mathfrak{A} \models \varphi(\mathbf{a})$ iff $\mathfrak{B} \models \varphi(\mathbf{b})$ holds for all formulas φ in $\mathcal{L}(\tau)$. The following equivalences are well-known [35], [36].

Lemma 1. *Let $\mathcal{L}$ be either GF or FO^2. Let $\mathfrak{A}, \mathbf{a}$ and $\mathfrak{B}, \mathbf{b}$ be pointed structures and τ a signature. Then*

$$\mathfrak{A}, \mathbf{a} \sim_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b} \quad \textit{implies} \quad \mathfrak{A}, \mathbf{a} \equiv_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b}$$

and, conversely, if $\mathfrak{A}$ and $\mathfrak{B}$ are ω-saturated, then

$$\mathfrak{A}, \mathbf{a} \equiv_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b} \quad \textit{implies} \quad \mathfrak{A}, \mathbf{a} \sim_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b}$$

IV. Interpolants and Explicit Definitions

Let $\mathcal{L}$ be either GF or FO^2. We introduce $\mathcal{L}$-interpolants and explicit $\mathcal{L}$-definitions and provide model-theoretic characterizations of the existence of $\mathcal{L}$-interpolants and explicit $\mathcal{L}$-definitions using $\mathcal{L}$-bisimulations.

Let $\varphi(\mathbf{x}), \psi(\mathbf{x})$ be $\mathcal{L}$-formulas with the same free variables $\mathbf{x}$. We call an $\mathcal{L}$-formula $\theta(\mathbf{x})$ an *$\mathcal{L}$-interpolant for φ, ψ* if $\mathrm{sig}(\theta) \subseteq \mathrm{sig}(\varphi) \cap \mathrm{sig}(\psi)$, $\varphi(\mathbf{x}) \models \theta(\mathbf{x})$, and $\theta(\mathbf{x}) \models \psi(\mathbf{x})$.

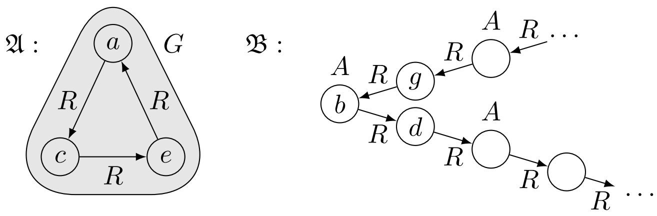

Fig. 1. Models for formulas in Example 1

We are interested in *$\mathcal{L}$-interpolant existence*, the problem to decide for given $\varphi(\mathbf{x}), \psi(\mathbf{x})$ in $\mathcal{L}$ whether an $\mathcal{L}$-interpolant for $\varphi(\mathbf{x}), \psi(\mathbf{x})$ exists. Recall from the introduction that neither GF nor FO^2 enjoy the *Craig Interpolation Property* (CIP) according to which an $\mathcal{L}$-interpolant for $\mathcal{L}$-formulas $\varphi(\mathbf{x}), \psi(\mathbf{x})$ exists iff $\varphi(\mathbf{x}) \models \psi(\mathbf{x})$.

We call $\mathcal{L}$-formulas $\varphi(\mathbf{x}), \psi(\mathbf{x})$ *jointly $\mathcal{L}(\tau)$-consistent* if there exist pointed structures $\mathfrak{A}, \mathbf{a}$ and $\mathfrak{B}, \mathbf{b}$ with $\mathfrak{A} \models \varphi(\mathbf{a})$ and $\mathfrak{B} \models \psi(\mathbf{b})$ such that $\mathfrak{A}, \mathbf{a} \sim_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b}$. The notion of joint consistency has been (implicitely) used to show the lack of the CIP for GF [34]. Using Lemma 1 we show that interpolant existence can in fact be characterized via joint consistency.

Lemma 2. *Let $\mathcal{L}$ be either FO^2 or GF. Let $\varphi(\mathbf{x}), \psi(\mathbf{x})$ be $\mathcal{L}$-formulas and let $\tau = sig(\varphi) \cap sig(\psi)$. Then the following conditions are equivalent:*

1) *there does not exist an $\mathcal{L}$-interpolant for $\varphi(\mathbf{x}), \psi(\mathbf{x})$;*
2) *$\varphi(\mathbf{x}), \neg\psi(\mathbf{x})$ are jointly $\mathcal{L}(\tau)$-consistent.*

The following example illustrates the introduced notions.

Example 1. Consider the GF-formulas $\varphi(x), \psi(x)$ given by

$$\varphi(x) = \exists yz\,(G(x,y,z) \wedge R(x,y) \wedge R(y,z) \wedge R(z,x))$$
$$\psi(x) = A(x) \wedge \forall y \forall z\,(R(y,z) \rightarrow (A(y) \leftrightarrow \neg A(z)))$$

Clearly, we have $\varphi(x) \models \neg\psi(x)$. Moreover, the models $\mathfrak{A}, a$ of $\varphi(x)$ and $\mathfrak{B}, b$ of $\psi(x)$ depicted in Fig. 1 witness that $\varphi(x)$ and $\psi(x)$ are jointly GF($\{R\}$)-consistent; in fact, the witnessing GF-bisimulation contains $n \mapsto m$ for every $n \in \mathrm{dom}(\mathfrak{A})$, $m \in \mathrm{dom}(\mathfrak{B})$. Lemma 2 implies that there is no GF-interpolant for $\varphi(x), \neg\psi(x)$. $\dashv$

Let φ be an $\mathcal{L}$-sentence, $\theta(\mathbf{x})$ an $\mathcal{L}$-formula, and τ a signature. An $\mathcal{L}(\tau)$-formula $\psi(\mathbf{x})$ is an *explicit $\mathcal{L}(\tau)$-definition of θ under φ* if $\varphi \models \forall\mathbf{x}(\theta(\mathbf{x}) \leftrightarrow \psi(\mathbf{x}))$. We call θ *explicitly $\mathcal{L}(\tau)$-definable under* φ if such an explicit $\mathcal{L}(\tau)$-definition of θ under φ exists. We call θ *implicitly $\mathcal{L}(\tau)$-definable under* φ if $\varphi \wedge \varphi' \models \forall\mathbf{x}(\theta(\mathbf{x}) \leftrightarrow \theta'(\mathbf{x}))$, where φ' and θ' are obtained from φ and θ, respectively, by renaming all non-τ symbols R to fresh R' of the same arity. Obviously, explicit $\mathcal{L}(\tau)$-definability implies implicit $\mathcal{L}(\tau)$-definability. Recall from the introduction that neither GF nor FO^2 enjoy the *projective Beth definability property* (PBDP) according to which the converse implication holds.

We consider the problem of *explicit $\mathcal{L}$-definability*, that is, the problem to decide for given $\varphi, \theta(\mathbf{x}), \tau$ whether there is an explicit $\mathcal{L}(\tau)$-definition of $\theta(\mathbf{x})$ under φ. We first observe that explicit definition existence reduces to interpolant existence.

Lemma 3. *Let $\mathcal{L}$ be either FO^2 or GF. There is a polynomial time reduction of explicit $\mathcal{L}$-definability to $\mathcal{L}$-interpolant existence.*

Lemma 3 suggests that there is a characterization of explicit definability in terms of joint $\mathcal{L}(\tau)$-consistency as well. Indeed, we give this characterization next.

Lemma 4. *Let $\mathcal{L}$ be either FO^2 or GF. For every $\mathcal{L}$-sentence φ, every $\mathcal{L}$-formula $\theta(\mathbf{x})$, and signature τ, the following conditions are equivalent:*

1) *there does not exist an explicit $\mathcal{L}(\tau)$-definition of $\theta(\mathbf{x})$ under φ;*
2) *$\varphi \wedge \theta(\mathbf{x})$ and $\varphi \wedge \neg\theta(\mathbf{x})$ are jointly $\mathcal{L}(\tau)$-consistent.*

Let us also illustrate the failure of the projective Beth definability property in FO^2 using Lemma 4.

Example 2. Consider the FO^2-sentence φ given by

$$\begin{aligned}
\varphi = {} & \forall xy\,((Y(x) \wedge Y(y)) \rightarrow x = y) \;\wedge \\
& \forall x\,(Z(x) \rightarrow \textstyle\bigvee_{i=0}^{3}(\varphi_i(x) \wedge \varphi'_{3-i}(x))) \wedge \\
& \forall xy\,((R(x,y) \wedge \neg Z(x)) \rightarrow I(x)) \wedge \\
& \forall xy\,(R(x,y) \rightarrow (I(x) \leftrightarrow I(y))) \wedge \\
& \forall xy\,(R(x,y) \rightarrow (I(x) \rightarrow (A(x) \leftrightarrow \neg A(y))))
\end{aligned}$$

where $\varphi_i(x)$ (resp., $\varphi'_i(x)$) is an FO^2-formula expressing that there is an R-path of length i to (resp., from) an element satisfying Y. Observe that $Z(x)$ is implicitely $\mathrm{FO}^2(\{R\})$-definable under φ since it is explicitely $\mathrm{FO}(\{R\})$-definable under φ: $Z(x)$ is true at $\mathfrak{A}, a$ iff a lies on a cycle of length three. In particular, the last three conjuncts of φ imply that $\neg Z(x)$ cannot be satisfied on any node of a cycle of odd length. To see the lack of an explicit $\mathrm{FO}^2(\{R\})$-definition, consider structures $\mathfrak{A}'$ and $\mathfrak{B}'$ obtained from $\mathfrak{A}, \mathfrak{B}$ in Figure 1:

- $\mathfrak{B}'$ is the extension of $\mathfrak{B}$ in which every node satisfies I;
- $\mathfrak{A}'$ is the disjoint union of $\mathfrak{B}'$ and the extension of $\mathfrak{A}$ in which every node satisfies Z and a satisfies Y.

It can be verified that $\mathfrak{A}', a$ is a model of $\varphi \wedge Z(x)$, that $\mathfrak{B}', b$ is a model of $\varphi \wedge \neg Z(x)$, and that $\mathfrak{A}', a \sim_{\mathrm{FO}^2,\{R\}} \mathfrak{B}', b$. By Lemma 4, $Z(x)$ is not explicitely $\mathrm{FO}^2(\{R\})$-definable under φ. $\dashv$

V. Deciding Joint GF(τ)-Consistency

We prove Theorem 1 (i). As Lemma 2 provides a reduction of the complement of GF-interpolant existence to joint GF(τ)-consistency, that is, the problem of deciding whether given $\varphi(\mathbf{x}), \psi(\mathbf{x})$ are jointly GF(τ)-consistent, we will prove the complexity upper bound for the latter problem. For the complexity lower bounds, we will also consider joint GF(τ)-consistency, but for an input of the form given in Lemma 4. This yields the respective lower bounds for explicit definition existence; by Lemma 3, they lift to interpolant existence.

A. Upper Bounds

To decide joint GF(τ)-consistency we pursue a mosaic approach based on types. Throughout the section, let $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ be the input to joint GF(τ)-consistency, for some signature τ. Let $\Xi = \{\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)\}$.

We begin by defining an appropriate notion of type. Let width(Ξ) denote the maximal arity of any relation symbol used in Ξ and let fv(Ξ) be the number of variables in $\mathbf{x}_0$. Let $x_1, \dots, x_{2n}$ be fresh variables, where $n := \max\{\text{width}(\Xi), \text{fv}(\Xi)\}$. We use $\mathsf{cl}(\Xi)$ to denote the smallest set of GF-formulas that is closed under taking subformulas and single negation, and contains:

- Ξ,
- all formulae $x = y$ for distinct variables x, y;
- all formulae $\exists \mathbf{x} R(\mathbf{x}\mathbf{y})$, where R is a relation symbol that occurs in Ξ and $\mathbf{x}\mathbf{y}$ is a tuple of variables.

Let $\mathfrak{A}$ be a structure, $\mathbf{a}$ a tuple of distinct elements from the domain of $\mathfrak{A}$, and $\mathbf{x}$ a tuple of distinct variables in $\{x_1, \dots, x_{2n}\}$ of the same length as $\mathbf{a}$. Consider the bijection $v : \mathbf{x} \mapsto \mathbf{a}$. Then the *$\Xi$-type of* $\mathbf{a}$ *in* $\mathfrak{A}$ *defined through* v is

$$\text{tp}(\mathfrak{A}, v : \mathbf{x} \mapsto \mathbf{a}) = \{\theta \mid \mathfrak{A} \models_v \theta, \theta \in \mathsf{cl}(\Xi)[\mathbf{x}]\},$$

where $\mathsf{cl}(\Xi)[\mathbf{x}]$ is obtained from $\mathsf{cl}(\Xi)$ by substituting in any formula $\theta \in \mathsf{cl}(\Xi)$ the free variables of θ by variables in $[\mathbf{x}]$ in all possible ways. Note that the assumption that v is bijective entails that $\neg(x = y) \in \text{tp}(\mathfrak{A}, v : \mathbf{x} \mapsto \mathbf{a})$ for any two distinct $x, y \in [\mathbf{x}]$. We drop v (and both v and $\mathbf{x}$) and write $\text{tp}(\mathfrak{A}, \mathbf{x} \mapsto \mathbf{a})$ (and $\text{tp}(\mathfrak{A}, \mathbf{a})$, respectively), whenever they are obvious from the context. Any Ξ-type of some $\mathbf{a}$ through some $v : \mathbf{x} \mapsto \mathbf{a}$ is called a *Ξ-type* and simply denoted $t(\mathbf{x})$. The set of all Ξ-types is denoted $T(\Xi)$.

We give a high-level description of our approach. To decide joint GF(τ)-consistency of $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ we determine all sets $\Phi \subseteq T(\Xi)$ using at most n variables from $\{x_1, \dots, x_{2n}\}$ that can be satisfied in GF(τ)-bisimilar models in the following sense: there are models $\mathfrak{A}_t$, $t \in \Phi$, realizing t in tuples $\mathbf{a}_t$ in $\text{dom}(\mathfrak{A}_t)$ through assignments v_t such that for any $t_1, t_2 \in \Phi$,

$$\mathfrak{A}_{t_1}, v_{t_1}(\mathbf{x}_{t_1,t_2}) \sim_{\text{GF},\tau} \mathfrak{A}_{t_2}, v_{t_2}(\mathbf{x}_{t_1,t_2}),$$

where $\mathbf{x}_{t_1,t_2}$ are the shared free variables of t_1 and t_2. Such sets Φ will be called *τ-mosaics*. Given the set of all τ-mosaics one can check whether $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ are jointly GF(τ)-consistent by simply checking whether there are types $t_1(\mathbf{x}), t_2(\mathbf{x})$ in a single τ-mosaic Φ such that one can replace the variables $\mathbf{x}_0$ in $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ by variables in $[\mathbf{x}]$ in such a way that $\varphi' \in t_1(\mathbf{x}_1)$, $\psi' \in t_2(\mathbf{x}_2)$ for the resulting formulas φ', ψ'. Thus, in what follows we aim to determine the characteristic properties of τ-mosaics and show that they can be enumerated in triple exponential time in general. If width(Ξ) is fixed, we perform a closer analysis of the set of mosaics and show that double exponential time is sufficient. The characteristic properties of τ-mosaics consist of *internal* properties that can be checked by inspecting a single set Φ of Ξ-types in isolation and one *external* property stating the existence of other τ-mosaics that ensure that τ-mosaics can be attached to each other in such a way that GF(τ)-bisimilar models can be constructed.

To formulate the properties of τ-mosaics, we require some notation. The *restriction* $t(\mathbf{x})_{|X}$ of a Ξ-type $t(\mathbf{x})$ to a set X of variables is the set of $\theta \in t(\mathbf{x})$ with free variables among X. The *restriction* $\Phi_{|X}$ of a set Φ of Ξ-types to X is defined as $\{t(\mathbf{x})_{|X} \mid t(\mathbf{x}) \in \Phi\}$. Types $t(\mathbf{x})$ and $t'(\mathbf{x}')$ *coincide on* X if $t(\mathbf{x})_{|X} = t'(\mathbf{x}')_{|X}$ and sets Φ, Φ' of Ξ-types *coincide on* X if $\Phi_{|X} = \Phi'_{|X}$. A variable x is *free* in a mosaic Φ if Φ contains a type in which x is free.

A formula $Q(\mathbf{x})$ of the form $x = x$ or $\exists \mathbf{y} R(\mathbf{x}\mathbf{y})$ with $R \in \tau$ is called a *τ-guard (for* $\mathbf{x}$*)*. It is called a *strict τ-guard* if it is of the form $x = x$ or $\mathbf{y}$ is empty, respectively. We call a set $\Phi \subseteq T(\Xi)$ a *τ-mosaic* if it satisfies the following conditions:

- Φ is *τ-uniform*: for all τ-guards $Q(\mathbf{z})$ and all $t(\mathbf{x}), s(\mathbf{y}) \in \Phi$ with $[\mathbf{z}] \subseteq [\mathbf{x}] \cap [\mathbf{y}]$, $Q(\mathbf{z}) \in t(\mathbf{x})$ iff $Q(\mathbf{z}) \in s(\mathbf{y})$;
- *closed under restrictions*: if $t(\mathbf{x}) \in \Phi$ and $X \subseteq [\mathbf{x}]$, then $t(\mathbf{x})_{|X} \in \Phi$;
- *GF(τ)-bisimulation saturated*: for all $t(\mathbf{x}) \in \Phi$, all strict τ-guards $Q(\mathbf{y}) \in t(\mathbf{x})$, and all $t'(\mathbf{z}) \in \Phi$ with $[\mathbf{z}] \subseteq [\mathbf{y}]$, there is an $s(\mathbf{y}') \in \Phi$ such that $t'(\mathbf{z}) \subseteq s(\mathbf{y}')$ and $[\mathbf{y}'] = [\mathbf{y}]$.

Intuitively, τ-uniformity reflects that GF(τ)-bisimulations preserve all τ-guards and GF(τ)-bisimulation saturatedness reflects Condition (i) for GF(τ)-bisimulations. Let us illustrate how to read off a mosaic from jointly consistent structures.

Example 3. Let $\mathfrak{A}, \mathfrak{B}$ be the structures from Fig. 1, set $\tau = \{R\}$, and $\Xi = \{\varphi(x), \psi(x)\}$ with φ, ψ as in Example 1. Let Φ be the closure under restrictions of the set containing

$$\text{tp}(\mathfrak{A}, xyz \to ace) \text{ and all types } \text{tp}(\mathfrak{B}, \mathbf{x} \to \mathbf{b})$$

with $\mathbf{x} \in \{xy, yz, zx\}$ and $\mathbf{b} \in \{gb, bd\}$. Thus, for example, Φ contains $\text{tp}(\mathfrak{A}, xy \to ac)$ as well. It can be easily verified that Φ is τ-uniform. To illustrate GF(τ)-bisimulation saturation, consider the types $t(x, y) = \text{tp}(\mathfrak{B}, xy \to bd)$ and $t'(x) = \text{tp}(\mathfrak{A}, x \to a)$, and the strict τ-guard $R(x, y)$ contained in $t(x, y)$. Then GF(τ)-bisimulation saturatedness is witnessed by the type $s(x, y) = \text{tp}(\mathfrak{A}, xy \to ac) \in \Phi$. $\dashv$

In addition to the internal properties above, we have to ensure that τ-mosaics can be linked together. The next two conditions state when this is the case. We say that τ-mosaics Φ_1, Φ_2 are *compatible* if for $\{i, j\} = \{1, 2\}$:

1) for every $t(\mathbf{x}) \in \Phi_i$ there is an $s(\mathbf{y}) \in \Phi_j$ such that $t(\mathbf{x})$ and $s(\mathbf{y})$ coincide on $[\mathbf{x}] \cap [\mathbf{y}]$;
2) if there are $t(\mathbf{x}) \in \Phi_i$ and $s(\mathbf{y}) \in \Phi_j$ and a τ-guard $Q(\mathbf{z}) \in t(\mathbf{x})$ with $[\mathbf{z}] \subseteq [\mathbf{x}] \cap [\mathbf{y}]$, then Φ_i and Φ_j coincide on $[\mathbf{z}]$.

Note that compatibility is a reflexive and symmetric relation. Let $\mathcal{M}$ be a set of τ-mosaics. We call $\Phi \in \mathcal{M}$ *existentially saturated in* $\mathcal{M}$ if for every $t(\mathbf{x}) \in \Phi$ and every formula $\exists \mathbf{y}(R(\mathbf{x}', \mathbf{y}) \wedge \lambda(\mathbf{x}', \mathbf{y})) \in t(\mathbf{x})$ there is a some $\Phi' \in \mathcal{M}$ such that Φ, Φ' are compatible and $R(\mathbf{x}', \mathbf{y}') \wedge \lambda(\mathbf{x}', \mathbf{y}') \in t'(\mathbf{z})$ for some $t'(\mathbf{z}) \in \Phi'$ which coincides with $t(\mathbf{x})$ on $[\mathbf{x}] \cap [\mathbf{z}]$. $\mathcal{M}$ is called *existentially saturated* if every $\Phi \in \mathcal{M}$ is existentially saturated in $\mathcal{M}$.

Example 4. Let $\mathcal{M} = \{\Phi\}$ with Φ as in Example 3. We claim that $\mathcal{M}$ is existentially saturated. Clearly every existentially quantified formula in (any restriction of) $\text{tp}(\mathfrak{A}, xyz \to ace)$ is "realized" in $\text{tp}(\mathfrak{A}, xyz \to ace)$ itself. Consider now, for example, $\exists z' R(z, z') \in t(y,z) := \text{tp}(\mathfrak{B}, yz \to bd)$. Then the type $\text{tp}(\mathfrak{B}, zx \to gb)$ coincides with $t(y,z)$ on $\{z\}$ and contains $R(z,x)$, as required. ⊣

It should be clear that the set of existentially saturated sets of τ-mosaics is closed under unions. Thus, the union of all existentially saturated sets of τ-mosaics is again existentially saturated. This set can be obtained by a purely syntactic elimination procedure, starting with the set of all τ-mosaics with at most n free variables from $\{x_1, \dots, x_{2n}\}$. We fine-tune and analyze this procedure below to obtain our two complexity upper bounds for joint GF(τ)-consistency. To this end, we prove three lemmas about existentially saturated sets of τ-mosaics. The first lemma states that τ-mosaics that are contained in an existentially saturated set behave in the way announced in the high-level overview of the proof.

Lemma 5. *Assume $\mathcal{M}$ is an existentially saturated set of τ-mosaics and let $t_1(\mathbf{x}_1), t_2(\mathbf{x}_2) \in \Psi \in \mathcal{M}$. Then there are pointed models $\mathfrak{A}_1, \mathbf{a}_1$ and $\mathfrak{A}_2, \mathbf{a}_2$ and $v_i : \mathbf{x}_i \mapsto \mathbf{a}_i$ such that*

- *$\mathfrak{A}_i \models t_i(\mathbf{a}_i)$, $i = 1, 2$, and*
- *$\mathfrak{A}_1, v_1([\mathbf{x}_1] \cap [\mathbf{x}_2]) \sim_{GF,\tau} \mathfrak{A}_2, v_2([\mathbf{x}_1] \cap [\mathbf{x}_2])$.*

Proof. Let $\Psi \in \mathcal{M}$. We assume w.l.o.g. that $\mathcal{M}$ is *closed under restrictions* in the sense that for any $\Phi \in \mathcal{M}$ and subset X of the free variables of Φ, $\Phi_{|X} \in \mathcal{M}$. (If it is not closed under restrictions simply add all $\Phi_{|X}$ with $\Phi \in \mathcal{M}$ to $\mathcal{M}$. The resulting set is still existentially saturated.) Define $\widehat{\Psi} := \Psi_{|\emptyset}$, that is, $\widehat{\Psi}$ contains all Ξ-types in Ψ without free variables. By closure under restrictions of $\mathcal{M}$, we have $\widehat{\Psi} \in \mathcal{M}$. Assume $\widehat{\Psi} = \{\hat{t}_1, \dots, \hat{t}_m\}$. We construct structures $\mathfrak{A}_i$, $i = 1, \dots, m$, with $\mathfrak{A}_i$ satisfying $\hat{t}_i$. For the construction, it is useful to employ notation for tree decompositions. A *tree decomposition* of a structure $\mathfrak{A}$ is a triple (T, E, bag) with (T, E) a tree and bag a function that assigns to every $t \in T$ a set $\text{bag}(t) \subseteq \text{dom}(\mathfrak{A})$ such that

1) $\mathfrak{A} = \bigcup_{t \in T} \mathfrak{A}_{|\text{bag}(t)}$;
2) $\{t \in T \mid a \in \text{bag}(t)\}$ is connected in (T, E), for every $a \in \text{dom}(\mathfrak{A})$.

We construct the structures $\mathfrak{A}_i$, $i = 1, \dots, m$ by giving a tree decomposition (T_i, E_i, bag_i) of $\mathfrak{A}_i$. To this end, we define (T_i, E_i, bag_i) and structures $\text{Bag}_i(t)$ with domain $\text{bag}_i(t)$, $t \in T_i$, and then show that (T_i, E_i, bag_i) is a tree decomposition of the union $\mathfrak{A}_i$ of all $\text{Bag}_i(t)$, $t \in T_i$. We start with the definition of (T_i, E_i). Let T_i be the set of all sequences

$$\sigma_n = (t_0(\mathbf{y}_0), \Phi_0), \dots, (t_n(\mathbf{y}_n), \Phi_n)$$

such that $t_0 = \hat{t}_i$, $\Phi_0 = \widehat{\Psi}$, $t_j(\mathbf{y}_j) \in \Phi_j$ for all $j \leq n$, and for all $j < n$:

- Φ_j, Φ_{j+1} are compatible, and
- $t_j(\mathbf{y}_j)$ and $t_{j+1}(\mathbf{y}_{j+1})$ coincide on $[\mathbf{y}_j] \cap [\mathbf{y}_{j+1}]$.

Let E_i be the induced prefix-order on T_i. We call $(t_n(\mathbf{y}_n), \Phi_n)$ the *tail* of σ_n. It remains to define the functions bag_i and Bag_i. We give an inductive definition with the aim to achieve the following: for all $\sigma_n \in T_i$ of the form above the Ξ-type $t_n(\mathbf{y}_n)$ is satisfied in $\mathfrak{A}_i$ under a canonical assignment v_{σ_n} into the set $\text{bag}_i(\sigma_n)$. For the construction, it is important to note that we have $\neg(x = y) \in t$ for any two distinct free variables x, y in any Ξ-type t. Thus we can essentially use (copies of) the variables $\mathbf{y}_n$ to define $\text{bag}_i(\sigma_n)$.

For the inductive definition, start by setting $\text{bag}_i(\sigma_0) = \emptyset$ and $v_{\sigma_0} = \emptyset$ for $\sigma_0 = (\hat{t}_i, \Phi_0)$. In the inductive step, assume that bag_i, $v_{\sigma_{n-1}}$, and Bag_i have been defined on σ_{n-1}, where

$$\sigma_{n-1} = (t_0(\mathbf{y}_0), \Phi_0), \dots, (t_{n-1}(\mathbf{y}_{n-1}), \Phi_{n-1}).$$

Then $\text{bag}_i(\sigma_n)$ contains

- fresh copies y' of the variables $y \in [\mathbf{y}_n] \setminus [\mathbf{y}_{n-1}]$ and
- $v_{\sigma_{n-1}}(y)$ for every $y \in [\mathbf{y}_n] \cap [\mathbf{y}_{n-1}]$,

and $v_{\sigma_n}(y)$ is defined as the copy y' of y for $y \in [\mathbf{y}_n] \setminus [\mathbf{y}_{n-1}]$ and by setting $v_{\sigma_n}(y) := v_{\sigma_{n-1}}(y)$ for $y \in [\mathbf{y}_n] \cap [\mathbf{y}_{n-1}]$. Finally, we define $\text{Bag}_i(\sigma_n)$ by interpreting any relation symbol R in such a way that the atomic formulas in $t_n(\mathbf{y}_n)$ are satisfied under v_{σ_n}, that is, such that $\text{Bag}_i(\sigma_n)$ satisfies $R(v_{\sigma_n}(\mathbf{y}))$ iff $R(\mathbf{y}) \in t_n(\mathbf{y}_n)$.

Let $\mathfrak{A}_i$ be the union of all $\text{Bag}_i(t)$, $t \in T_i$. It is easy to see that (T_i, E_i, bag_i) is a tree decomposition of $\mathfrak{A}_i$. In fact, in the inductive step above, $t_n(\mathbf{y}_n)$ and $t_{n-1}(\mathbf{y}_{n-1})$ coincide on $[\mathbf{y}_n] \cap [\mathbf{y}_{n-1}]$. Thus, the interpretation of any relation symbol R coincides on the intersection of $\text{bag}_i(\sigma_n)$ and $\text{bag}_i(\sigma_{n-1})$. We proceed to show that the GF(τ)-bisimulation mentioned in Lemma 5 indeed exists. To this end, we prove the following auxiliary claim. We call a tuple $\mathbf{a}$ *τ-guarded in* $\mathfrak{A}$ if there exists a τ-guard $Q(\mathbf{x})$ such that $\mathfrak{A} \models Q(\mathbf{a})$. We prove the following in the appendix:

Claim 1. *For all i, j with $1 \leq i, j \leq m$, we have:*

1) *For every $\sigma \in T_i$ with tail$(\sigma) = (t(\mathbf{y}), \Phi)$, we have $\mathfrak{A}_i \models t(v_\sigma(\mathbf{y}))$;*
2) *Let $H_{i,j}$ be the set of all mappings $p_{\sigma,\sigma',\mathbf{z}}$, where*
   - *$\sigma \in T_i, \sigma' \in T_j$, tail$(\sigma) = (t(\mathbf{y}), \Phi)$, and tail$(\sigma') = (t'(\mathbf{y}'), \Phi)$;*
   - *$\mathbf{z}$ is a tuple with $[\mathbf{z}] \subseteq [\mathbf{y}] \cap [\mathbf{y}']$ and $v_\sigma(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_i$ (or, equivalently, $v_{\sigma'}(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_j$);*
   - *$p_{\sigma,\sigma',\mathbf{z}} : v_\sigma(\mathbf{z}) \mapsto v_{\sigma'}(\mathbf{z})$.*

   Then $H_{i,j}$ is a GF(τ)-bisimulation between $\mathfrak{A}_i$ and $\mathfrak{A}_j$.

To complete the proof of Lemma 5, assume w.l.o.g. that $\hat{t}_i \subseteq t_i(\mathbf{x}_i)$ for $i = 1, 2$. Take $\rho_i = (\hat{t}_i, \widehat{\Psi}) \cdot (t_i(\mathbf{x}_i), \Psi) \in T_i$, for $i = 1, 2$. Consider the tuples $\mathbf{a}_i := v_{\rho_i}(\mathbf{x}_i)$. By Claim 1, $\mathfrak{A}_i \models t_i(\mathbf{a}_i)$. Also by Claim 1, for any tuple $\mathbf{z}$ with $[\mathbf{z}] \subseteq [\mathbf{x}_1] \cap [\mathbf{x}_2]$ and such that $v_{\rho_1}(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_1$ or $\mathfrak{A}_2$, we have $p_{\rho_1,\rho_2,\mathbf{z}} : v_{\rho_1}(\mathbf{z}) \mapsto v_{\rho_2}(\mathbf{z}) \in H_{1,2}$. But then, as any two $p_{\rho_1,\rho_2,\mathbf{z}}$ coincide on the intersection of their domains, we have $\mathfrak{A}_1, v_{\rho_1}([\mathbf{x}_1] \cap [\mathbf{x}_2]) \sim_{\text{GF},\tau} \mathfrak{A}_2, v_{\rho_2}([\mathbf{x}_1] \cap [\mathbf{x}_2])$, as required. ❑

We next show how to read off an existentially saturated set of mosaics from jointly consistent structures, as illustrated in Example 3. We make sure that all mosaics except a single mosaic Ψ use only width(Ξ) many free variables and that also in Ψ only at most two types use more variables.

Lemma 6. *Let $\mathfrak{A}_1, \mathbf{a}_1$ and $\mathfrak{A}_2, \mathbf{a}_2$ be pointed structures with $\mathbf{a}_1$ and $\mathbf{a}_2$ tuples with pairwise distinct elements of length $m \leq fv(\Xi)$ and let τ be a signature. Consider assignments $\mathbf{x}_0 \mapsto \mathbf{a}_i$ with $[\mathbf{x}_0] \subseteq \{x_0, \ldots, x_{2n}\}$. If $\mathfrak{A}_1, \mathbf{a}_1 \sim_{GF,\tau} \mathfrak{A}_2, \mathbf{a}_2$, then there exists an existentially saturated set $\mathcal{M}$ of τ-mosaics and some $\Psi \in \mathcal{M}$ such that*

- *all $\Phi \in \mathcal{M}$ with $\Phi \neq \Psi$ use at most width(Ξ) many free variables;*
- *there exist types $t_1(\mathbf{x}_0), t_2(\mathbf{x}_0) \in \Psi$ such that $t_i(\mathbf{x}_0) = tp(\mathfrak{A}_i, \mathbf{x}_0 \mapsto \mathbf{a}_i)$ for $i = 1, 2$ and all types $t(\mathbf{y}) \in \Psi \setminus \{t_1(\mathbf{x}_0), t_2(\mathbf{x}_0)\}$ use at most width(Ξ) free variables among $[\mathbf{x}_0]$.*

Proof. Assume w.l.o.g. that $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are disjoint. For any tuples $\mathbf{b}_1$ in $\mathfrak{A}_i$ and $\mathbf{b}_2$ in $\mathfrak{A}_j$ with $i, j \in \{1, 2\}$, we use $\text{tp}(\mathbf{x}_1 \mapsto \mathbf{b}_1)$ to denote $\text{tp}(\mathfrak{A}_i, \mathbf{x}_1 \mapsto \mathbf{b}_1)$ and we write $\mathbf{b}_1 \sim_{\text{GF},\tau} \mathbf{b}_2$ if $\mathfrak{A}_i, \mathbf{b}_1 \sim_{\text{GF},\tau} \mathfrak{A}_j, \mathbf{b}_2$. Define $\mathcal{M}$ as follows. Take any tuple $\mathbf{a}$ of distinct elements in $\mathfrak{A}_i$, $i \in \{1, 2\}$. Take a tuple $\mathbf{x}$ from $\{x_1, \ldots, x_{2n}\}$ such that $v : \mathbf{x} \mapsto \mathbf{a}$ is a bijection. Then let $\Phi_{\mathbf{a},\mathbf{x}}$ contain all types $\text{tp}(v' : \mathbf{x}_{|Y} \mapsto \mathbf{b})$ with $Y \subseteq [\mathbf{x}]$ and $\mathbf{b}$ in either $\mathfrak{A}_1$ or $\mathfrak{A}_2$ such that $v(\mathbf{x}_{|Y}) \sim_{\text{GF},\tau} v'(\mathbf{x}_{|Y})$.

Let $\mathcal{M}$ contain all such $\Phi_{\mathbf{a},\mathbf{x}}$ with $\mathbf{a}$ of length at most width(Ξ) and $\mathbf{x}$ from $\{x_1, \ldots, x_{2n}\}$. Moveover, if $m >$ width(Ξ), then add $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ to $\mathcal{M}$, where $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ is obtained from $\Phi_{\mathbf{a}_1,\mathbf{x}_0}$ by removing all t distinct from $t_1(\mathbf{x}_0)$ and $t_2(\mathbf{x}_0)$ using more than width(Ξ) many free variables.

We show in the appendix that $\mathcal{M}$ is as required. ❑

It follows from Lemmas 5 and 6 that the following two conditions are equivalent, where $\mathcal{M}'$ is the maximal existentially saturated set of τ-mosaics using at most width(Ξ) free variables.

1. $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ are jointly GF(τ)-consistent;
2. There exists a τ-mosaic Ψ and Ξ-types $t_1(\mathbf{x}), t_2(\mathbf{x}) \in \Psi$ such that $\mathcal{M} = \{\Psi\} \cup \mathcal{M}'$ is existentially saturated and:
   a) $t_1(\mathbf{x}), t_2(\mathbf{x})$ have fv(Ξ) free variables and one can replace the variables in $[\mathbf{x}_0]$ by variables in $\mathbf{x}$ such that $\varphi' \in t_1(\mathbf{x})$, $\psi' \in t_2(\mathbf{x})$ for the resulting formulas φ', ψ';
   b) all Ξ-types $t(\mathbf{y}) \in \Psi \setminus \{t_1(\mathbf{x}), t_2(\mathbf{x})\}$ use at most width(Ξ) free variables among $[\mathbf{x}]$;

Hence, it suffices to provide an algorithm deciding Condition 2.

Lemma 7. *On input $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$, Condition 2 can be decided in time triple exponential in the size of $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ in general, and double exponential in the size of $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ if width(Ξ) is bounded by a constant.*

Proof. First determine $\mathcal{M}'$ by exhaustively removing τ-mosaics that are not existentially saturated from the list of all τ-mosaics with at most width(Ξ) free variables. It can be verified that the fixpoint is existentially saturated. Next we proceed as follows: for every pair $t_1(\mathbf{x}), t_2(\mathbf{x})$ of Ξ-types that satisfies Condition 2(a) enumerate all τ-mosaics Ψ satisfying Condition 2(b), that is, $t_1(\mathbf{x}), t_2(\mathbf{x}) \in \Psi$ and all types in Ψ except $t_1(\mathbf{x}), t_2(\mathbf{x})$ use at most width(Ξ) free variables among $[\mathbf{x}]$. Accept if at least one $\{\Psi\} \cup \mathcal{M}'$ is existentially saturated. Reject otherwise.

Correctness of the algorithm is straightforward, so it remains to analyze its run time. For this purpose, let r be the number of subformulas (of formulas) in Ξ and $\ell \geq 0$. Observe that a subformula with ℓ free variables has at most $(2n)^\ell$ instantiations with variables from $x_1, \ldots, x_{2n}$. Since for every such instantiated formula either the formula itself or its negation is contained in any type, there are at most $2^{r(2n)^\ell}$ many types with ℓ free variables. Thus, there are only double exponentially many choices for $t_1(\mathbf{x}), t_2(\mathbf{x})$ and Ψ. Moreover, the set of all τ-mosaics with at most width(Ξ) free variables is of size triple exponential in the size of $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ in general, and double exponential in the size of $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ if width(Ξ) is bounded by a constant. The upper bounds now follow from the observation that checking whether some Φ is existentially saturated in some set $\mathcal{M}_0$ of mosaics can be done in time polynomial in the size of $\mathcal{M}_0$. ❑

From the equivalence of Conditions 1 and 2, and Lemma 7 we finally obtain that joint GF(τ)-consistency is in 3ExpTime in general, and in 2ExpTime if the arity of relation symbols is bounded by a constant.

B. Lower Bounds

We reduce the word problem for exponentially and double exponentially space bounded alternating Turing machines, respectively. An *alternating Turing machine (ATM)* is a tuple $M = (Q, \Theta, \Gamma, q_0, \Delta)$ where $Q = Q_\exists \uplus Q_\forall$ is the set of states that consists of *existential states* in $Q_\exists$ and *universal states* in $Q_\forall$. Further, Θ is the input alphabet and Γ is the tape alphabet that contains a *blank symbol* $\square \notin \Theta$, $q_0 \in Q_\exists$ is the *starting state*, and the *transition relation* Δ is of the form $\Delta \subseteq Q \times \Gamma \times Q \times \Gamma \times \{L, R\}$. The set $\Delta(q, a) := \{(q', a', M) \mid (q, a, q', a', M) \in \Delta\}$ must contain exactly two or zero elements for every $q \in Q$ and $a \in \Gamma$. Moreover, the state q' must be from $Q_\forall$ if $q \in Q_\exists$ and from $Q_\exists$ otherwise, that is, existential and universal states alternate. We use a slightly non-standard acceptance condition (note that there are no accepting states): The ATM accepts an input w if it runs forever on all branches and rejects otherwise. Starting from the standard ATM model, this can be achieved by assuming that exponentially (resp., double exponentially) space bounded ATMs terminate on every input and then modifying them to enter an infinite loop from the accepting state.

More formally, a *configuration* of an ATM is a word wqw' with $w, w' \in \Gamma^*$ and $q \in Q$. We say that wqw' is *existential* if q is, and likewise for *universal*. *Successor configurations* are defined in the usual way. Note that every configuration has either zero or two successor configurations. A *computation* *tree* of an ATM M on input w is an infinite tree whose nodes are labeled with configurations of M such that

- the root is labeled with the initial configuration $q_0 w$;
- if a node is labeled with an existential configuration wqw', then it has a single successor and this successor is labeled with a successor configuration of wqw';
- if a node is labeled with a universal configuration wqw', then it has two successors and these successors are labeled with the two successor configurations of wqw'.

An ATM M *accepts* an input w if there is a computation tree of M on w. It is well-known that the word problem for 2^n-space bounded and 2^{2^n}-space bounded ATMs is 2EXPTIME-hard and 3EXPTIME-hard, respectively [54].

1) Bounded Arity: For didactic reasons, we start with showing 2EXPTIME-hardness for the bounded arity case. Let M be a 2^n-space bounded ATM and w an input. The idea of the reduction is as follows. We set

$$\tau = \{R, S, X, Z, B_\forall, B_\exists^1, B_\exists^2\} \cup \{A_\sigma \mid \sigma \in \Gamma \cup (Q \times \Gamma)\},$$

where R, S are binary relation symbols, and the remaining symbols are unary. We aim to construct φ such that M accepts w iff $\varphi \wedge A(x)$ and $\varphi \wedge \neg A(x)$ are jointly GF(τ)-consistent. The sentence φ is a conjunction of several GF-sentences, which are, except for one, also FO^2-sentences. The first conjunct, φ_0 below, is this exception and enforces that every element satisfying A is involved in a three-element R-loop (similar to Example 1):

$$\varphi_0 = \forall x\big(A(x) \rightarrow \exists yz(G(x,y,z) \wedge X(x) \wedge \neg X(y) \wedge \neg X(z) \wedge R(x,y) \wedge R(y,z) \wedge R(z,x)\big)$$

Now, if $\varphi \wedge A(x)$ and $\varphi \wedge \neg A(x)$ are jointly GF(τ)-consistent, there exist models $\mathfrak{A}$ and $\mathfrak{B}$ of φ and elements a, b such that $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$, and $\mathfrak{A}, a \sim_{\text{GF},\tau} \mathfrak{B}, b$. If the latter holds, then from $a \in A^{\mathfrak{A}}$ and φ_0 it follows that b has an infinite outgoing path ρ along R on which every third element satisfies X and is guarded τ-bisimilar to a. Let us call these elements the X-*elements*. As guarded bisimilarity is an equivalence relation, all X-elements are actually guarded τ-bisimilar. The other conjuncts of φ will enforce that along the X-elements on ρ, a counter counts modulo 2^n using relation symbols not in τ. Moreover, in every X-element of ρ starts an infinite tree along symbol S that is supposed to mimick the computation tree of M. Along this tree, two counters are maintained:

- one counter starting at 0 and counting modulo 2^n to divide the tree in subpaths of length 2^n; each such path of length 2^n represents a configuration;
- another counter starting at the value of the counter along ρ and also counting modulo 2^n.

To link successive configurations we use the fact that all X-elements on ρ are guarded τ-bisimilar and thus each X-element is the starting point of trees along S with identical τ-decorations. As on the mth such tree the second counter starts at all nodes at distances $k \cdot 2^n - m$, for all $k \geq 1$, we are in the position to coordinate all positions at all successive configurations.

In detail, let $w = a_0, \dots, a_{n-1}$ be an input to M of length n. We will be using unary symbols A_i, U_i, V_i, $1 \leq i \leq n$ to represent the aforementioned binary counters; we will refer to them with A-counter, U-counter, and V-counter, respectively.

The sentences below enforce that the A-counter along the R-path ρ is incremented (precisely) at every X-element. In order to avoid that the counter is stipulated at a (which would lead to a contradiction), we use an additional symbol $I \notin \tau$ that is satisfied along the entire path and acts as an additional guard:[1]

$$\forall xy\big(R(x,y) \rightarrow (\neg A(x) \wedge X(x) \rightarrow I(x)\big)$$
$$\forall xy\big(R(x,y) \rightarrow (I(x) \leftrightarrow I(y))\big)$$
$$\forall xy\big(R(x,y) \wedge I(x) \wedge \neg X(y) \rightarrow \text{Eq}(x,y))\big)$$
$$\forall xy\big(R(x,y) \wedge I(x) \wedge X(y) \rightarrow \text{Succ}(x,y))\big)$$

Here, atoms $\text{Eq}(x,y)$ and $\text{Succ}(x,y)$ are abbreviations for formulas that express that the A-counter value at x equals (resp., is the predecessor of) the A-counter value at y, that is:

$$\text{Eq}(x,y) = \textstyle\bigwedge_i A_i(x) \leftrightarrow A_i(y))$$
$$\text{Succ}(x,y) = \textstyle\bigvee_i \Big(A_i(y) \wedge \neg A_i(x) \wedge \bigwedge_{j<i}(\neg A_j(y) \wedge A_j(x)) \wedge \bigwedge_{j>i}(A_j(y) \leftrightarrow A_j(x))\Big)$$

Now, we start a tree along S from all X-elements on the infinite R-path. Along the path, we maintain the U- and V-counter, which are initialized to 0 and the value of the A-counter, respectively:

$$\forall x \exists y S(x,y)$$
$$\forall xy\big(S(x,y) \rightarrow (I(x) \leftrightarrow I(y))\big)$$
$$\forall x\big(I(x) \wedge X(x) \rightarrow \text{Min}_U(x)\big)$$
$$\forall x\big(I(x) \wedge X(x) \rightarrow \textstyle\bigwedge_i(V_i(x) \leftrightarrow A_i(x))\big)$$

Here, $\text{Min}_U(x)$ is an abbreviation for the formula that expresses that the U-counter is 0 at x; we use similar abbreviations such as $\text{Max}_V(x)$ below. The U and V-counters are incremented along S analogously to how the A-counter is incremented along R, but on every S-step; we omit details. Configurations of M are represented between two consecutive elements having U-counter value 0. We next enforce the structure of the computation tree, assuming that $q_0 \in Q_\forall$.

$$\forall x\big(I(x) \wedge X(x) \rightarrow B_\forall(x)\big)$$
$$\forall xy\big(S(x,y) \wedge I(x) \wedge \neg\text{Max}_U(x) \rightarrow (B_\forall(x) \leftrightarrow B_\forall(y))\big)$$
$$\forall xy\big(S(x,y) \wedge I(x) \wedge \neg\text{Max}_U(x) \rightarrow \textstyle\bigwedge_{j=1}^{2}(B_\exists^j(x) \leftrightarrow B_\exists^j(y))\big)$$
$$\forall xy\big(S(x,y) \wedge I(x) \wedge \text{Max}_U(x) \rightarrow (B_\forall(x) \leftrightarrow \neg B_\forall(y))\big)$$
$$\forall x\big(I(x) \wedge \text{Max}_U(x) \rightarrow \exists y(S(x,y) \wedge Z(y)) \wedge \exists y(S(x,y) \wedge \neg Z(y))\big)$$
$$\forall x\big(I(x) \wedge \neg B_\forall(x) \rightarrow (B_\exists^1(x) \leftrightarrow \neg B_\exists^2(x))\big)$$

These sentences enforce that all nodes which represent a configuration satisfy exactly one of $B_\forall, B_\exists^1, B_\exists^2$, indicating the

[1] Throughout, we assume that $\wedge$ has higher precedence than $\rightarrow$. Moreover, some formulas are not syntactically guarded but can easily be rewritten.

kind of configuration and, if existential, also a choice of the transition function, indicated in the superscript of $B^j_\exists$. The symbol $Z \in \tau$ enforces the branching.

We next set the initial configuration, for input $w = a_0, \ldots, a_{n-1}$. Below, we use $\forall y(S^i(x,y) \to \psi(y))$ to abbreviate the GF-formula that enforces ψ at all elements y that are reachable in i steps via S from x.

$$\forall x(I(x) \wedge X(x) \to A_{q_0,a}(x))$$
$$\forall x\big(I(x) \wedge X(x) \to \forall y(S^k(x,y) \to A_{a_k}(y))\big), \quad 0 < k < n$$
$$\forall x\big(I(x) \wedge X(x) \to \forall y(S^n(x,y) \to \text{Blank}(y))\big)$$
$$\forall x(\text{Blank}(x) \to A_\Box(x))$$
$$\forall x(\text{Blank}(x) \wedge \neg\text{Max}_U(x) \to \forall y(S(x,y) \to \text{Blank}(y)))$$

We next coordinate consecutive configurations, focusing on cells that are not at the border of a configuration; these corner cases can be dealt with accordingly. To this end, we associate with M functions f_i, $i \in \{1,2\}$ that map the content of three consecutive cells of a configuration to the content of the middle cell in the i-th successor configuration (assuming an arbitrary order on the $\Delta(q,a)$). Moreover, for each triple $(\sigma_1,\sigma_2,\sigma_3) \in (\Gamma \cup (Q \times \Gamma))^3$, we fix a GF-formula $\psi_{\sigma_1,\sigma_2,\sigma_3}(x)$ that is satisfied at an element a of the computation tree iff a is labeled with A_{σ_2}, a has an S-predecessor labeled with A_{σ_1}, and a has an S-successor labeled with A_{σ_3}. Now, in each configuration, we synchronize elements with V-counter 0, by including for every $\vec{\sigma} = (\sigma_1,\sigma_2,\sigma_3)$ and $i \in \{1,2\}$ the following sentences:

$$\forall x\big(I(x) \wedge \text{Min}_V(x) \wedge \neg\text{Min}_U(x) \wedge \neg\text{Max}_U(x) \wedge B_\forall(x) \to$$
$$(\psi_{\vec{\sigma}}(x) \to A^1_{f_1(\vec{\sigma})}(x) \wedge A^2_{f_2(\vec{\sigma})}(x))\big)$$
$$\forall x\big(I(x) \wedge \text{Min}_V(x) \wedge \neg\text{Min}_U(x) \wedge \neg\text{Max}_U(x) \wedge B^i_\exists(x) \to$$
$$(\psi_{\vec{\sigma}}(x) \to A^i_{f_i(\vec{\sigma})}(x))\big)$$

The unary symbols A^i_σ are used as markers (not in τ) and are propagated along S for 2^n steps, exploiting the V-counter. The superscript $i \in \{1,2\}$ determines the successor configuration that the symbol is referring to. After crossing the end of a configuration, the symbol σ is propagated using further unary symbols A'_σ (the superscript is not needed anymore because the branching happens at the end of the configuration, based on Z):

$$\forall x\big(\neg\text{Max}_U(x) \wedge A^i_\sigma(x) \to \forall y(S(x,y) \to A^i_\sigma(y))\big)$$
$$\forall x\big(\text{Max}_U(x) \wedge B_\forall(x) \wedge A^1_\sigma(x) \to$$
$$\forall y(S(x,y) \to (Z(y) \to A'_\sigma(y)))\big)$$
$$\forall x\big(\text{Max}_U(x) \wedge B_\forall(x) \wedge A^2_\sigma(x) \to$$
$$\forall y(S(x,y) \to (\neg Z(y) \to A'_\sigma(y)))\big)$$
$$\forall x\big(\text{Max}_U(x) \wedge B^i_\exists(x) \wedge A^i_\sigma(x) \to \forall y(S(x,y) \to A'_\sigma(y))\big)$$
$$\forall x\big(\neg\text{Max}_V(x) \wedge A'_\sigma(x) \to \forall y(S(x,y) \to A'_\sigma(x))\big)$$
$$\forall x\big(\text{Max}_V(x) \wedge A'_\sigma(x) \to \forall y(S(x,y) \to A_\sigma(x))\big)$$

For those (q,a) with $\Delta(q,a) = \emptyset$, we add the sentence

$$\forall x\, \neg A_{q,a}(x)$$

to ensure that such halting states are never reached. Correctness of the reduction is established in the appendix.

Lemma 8. *M accepts the input w iff there exists models $\mathfrak{A}, \mathfrak{B}$ of φ and elements $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$ such that $\mathfrak{A}, a \sim_{GF,\tau} \mathfrak{B}, b$.*

2) General Case: We reduce the word problem of 2^{2^n}-space bounded ATMs using the very same idea as in the previous section. However, we need double exponential counters instead of the single exponential counters for A, U, V above. These counters are encoded in a way similar to the 2EXPTIME-hardness proof for satisfiability in the guarded fragment [7]. The mentioned encoding is based on *pairs of elements*, so we "lift" the above reduction to pairs of elements and consequently double the arity of all involved symbols. More precisely, we set

$$\tau = \{R, S, X, Z, B_\forall, B^1_\exists, B^2_\exists\} \cup \{A_\sigma \mid \sigma \in \Gamma \cup (Q \times \Gamma)\},$$

where R, S are 4-ary relation symbols, and the remaining symbols are binary. The sentence φ is a conjunction of several sentences. The first conjunct, φ_0 below, enforces that every pair of elements satisfying A is involved in a three-element R-loop as follows:[2]

$$\varphi_0 = \forall xx'\big(Axx' \to \exists yy'zz'(Gxx'yy'zz' \wedge Xxx' \wedge$$
$$\neg Xyy' \wedge \neg Xzz' \wedge Rxx'yy' \wedge Ryy'zz' \wedge Rzz'xx')\big)$$

As above, we aim to construct φ such that M accepts w iff there exist models $\mathfrak{A}$ and $\mathfrak{B}$ of φ and pairs $\mathbf{a}, \mathbf{b}$ such that $\mathbf{a} \in A^{\mathfrak{A}}$, $\mathbf{b} \notin A^{\mathfrak{B}}$, and $\mathfrak{A}, \mathbf{a} \sim_{\text{GF},\tau} \mathfrak{B}, \mathbf{b}$. If the latter holds then from $\mathbf{a} \in A^{\mathfrak{A}}$ and φ_0 it follows that $\mathbf{b}$ has an infinite outgoing "path" ρ along R on which every third pair of element satisfies X and is guarded τ-bisimilar to $\mathbf{a}$. Let us call these pairs the *X-pairs*. Observe that all X-pairs are guarded τ-bisimilar.

The main difference to the reduction above is the realization of the counters, so we will concentrate on this and leave the (straightforward) remainder of the proof to the reader. For realizing the A-counter, we use an n-ary relation symbol D and associate a counter to every pair of elements (a, a') as follows. We assume the order $a < a'$ which induces an order $<$ on tuples $\mathbf{a} \in \{a, a'\}^n$. Thus, every tuple $\mathbf{a} \in \{a, a'\}^n$ corresponds to a number $r(\mathbf{a}) < 2^n$, the rank of $\mathbf{a}$ according to $<$. Now the sequence of truth values on all these tuples in D can be viewed as the binary representation of a number $< 2^{2^n}$.

The A-counter along the R-path ρ is enforced by the following sentences:

$$\forall xx'yy'\big(Rxx'yy' \to (\neg Axx' \wedge Xxx' \to Ixx')\big)$$
$$\forall xx'yy'\big(Rxx'yy' \to (Ixx' \leftrightarrow Iyy')\big)$$
$$\forall xx'yy'\big(Rxx'yy' \wedge Ixx' \to (\neg Xyy' \to \text{Eq}(xx'yy'))\big)$$
$$\forall xx'yy'\big(Rxx'yy' \wedge Ixx' \to (Xyy' \to \text{Succ}(xx'yy'))\big)$$

Again, the I acts as an additional guard that disables the counting at $\mathbf{a}$. It remains to define the formulas $\text{Eq}(xx'yy')$ and $\text{Succ}(xx'yy')$. We show in the appendix that we can

[2]We omit commas and/or parantheses when no confusion can arise.

axiomatize a $(4n+4)$-ary predicate E such that, for pairs a, a' and b, b' where b, b' represents a successor node of a, a', and for $\mathbf{a}, \mathbf{a}' \in \{a, a'\}^n$ and $\mathbf{b}, \mathbf{b}' \in \{b, b'\}^n$, we have

$$E(\mathbf{a}\mathbf{a}'aa'\mathbf{b}\mathbf{b}'bb') \quad \text{iff} \quad r(\mathbf{a}) = r(\mathbf{b}) \text{ and } r(\mathbf{a}') = r(\mathbf{b}').$$

Then the formulas Eq and Succ can be defined as follows:

$$\begin{aligned}
\text{Eq}(xx'yy') &= \forall \mathbf{x}\mathbf{y}\mathbf{x}'\mathbf{y}' \big(E\mathbf{x}\mathbf{x}'xx'\mathbf{y}\mathbf{y}'yy' \rightarrow (D\mathbf{x} \leftrightarrow D\mathbf{y})\big) \\
\text{Succ}(xx'yy') &= \exists \mathbf{x}\mathbf{y} \big(E\mathbf{x}\mathbf{x}xx'\mathbf{y}\mathbf{y}yy' \wedge \neg D\mathbf{x} \wedge D\mathbf{y} \\
&\quad \wedge \forall \mathbf{x}'\mathbf{y}' \big(E\mathbf{x}\mathbf{x}'xx'\mathbf{y}\mathbf{y}'yy' \rightarrow \\
&\qquad (\text{less}(\mathbf{x}'\mathbf{x}xx') \rightarrow D\mathbf{x}' \wedge \neg D\mathbf{y}') \wedge \\
&\qquad (\text{less}(\mathbf{x}\mathbf{x}'xx') \rightarrow (D\mathbf{x}' \leftrightarrow D\mathbf{y}'))\big)\big)
\end{aligned}$$

where, for $\mathbf{x} = x_0 \dots x_{n-1}$ and $\mathbf{x}' = x'_0 \dots x'_{n-1}$, we have

$$\text{less}(\mathbf{x}\mathbf{x}'xx') = \bigvee_{i<n} \left(x'_i = x' \wedge x_i = x \wedge \bigwedge_{j>i} x_j = x'_j\right).$$

Thus, $\text{less}(\mathbf{x}\mathbf{x}'xx')$ compares the positions of $\mathbf{x}$ and $\mathbf{x}'$ according to the order $x < x'$. Moreover, $\text{Eq}(xx'yy')$ is true iff the counters stipulated by x, x' and y, y' have precisely the same bits set. Finally, $\text{Succ}(xx'yy')$ asserts the existence of a position k such that *(i)* in the counter stipulated by x, x' bit k is set to 0 while in the counter stipulated by y, y' bit k is set to 1, *(ii)* on all positions k' less than k, the bits in the former counter are 1 while the bits in the latter are 0, and *(iii)* on all positions k' greater than k the counters agree on their bits.

Having the adapted counters available, the proof then proceeds along the lines of the proof given for the bounded arity case, always replacing single elements/variables with pairs of elements/variables as exemplified above.

VI. Deciding Joint $\text{FO}^2(\tau)$-consistency

We prove Theorem 1 (ii). We proceed similarly to the proof for GF by proving a N2ExpTime upper bound for joint $\text{FO}^2(\tau)$-consistency and then applying Lemma 2 to obtain a coN2ExpTime upper bound for FO^2-interpolant existence. For the complexity lower bound we consider joint $\text{FO}^2(\tau)$-consistency for an input of the form given in Lemma 4.

A. Upper Bound

We show the N2ExpTime upper bound for joint $\text{FO}^2(\tau)$-consistency by proving that if two FO^2-formulas are jointly $\text{FO}^2(\tau)$-consistent, then there exist $\text{FO}^2(\tau)$-bisimilar models satisfying the formulas of at most double exponential size:

Theorem 2. *If $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ are jointly $FO^2(\tau)$-consistent, then there are pointed models $\mathfrak{B}_1, \mathbf{b}_1$ and $\mathfrak{B}_2, \mathbf{b}_2$ of at most double exponential size such that $\mathfrak{B}_1 \models \varphi(\mathbf{b}_1)$, $\mathfrak{B}_2 \models \psi(\mathbf{b}_2)$ and $\mathfrak{B}_1, \mathbf{b}_1 \sim_{FO^2,\tau} \mathfrak{B}_2, \mathbf{b}_2$.*

The remainder of this section is devoted to the proof. We first simplify the input formulas. Generalizing [19], we show in the appendix that one can assume w.l.o.g. that the input formulas only use relation symbols of arity at most two. Then one can easily extend the normal form for FO^2 sentences provided in [19] to the following normal form for formulas: for any FO^2-formula $\chi(\mathbf{x})$ only using relation symbols of arity at most two one can construct in polynomial time an FO^2-formula $\chi'(\mathbf{x})$ of the form

$$R_0(\mathbf{x}) \wedge \forall x \forall y \alpha \wedge \bigwedge_{i=1}^{m} \forall x \exists y \beta_i(x, y),$$

where R_0 is a relation symbol and α and β_i are quantifier-free such that all relations symbols in $\chi'(\mathbf{x})$ have arity at most two and

1) $\chi' \models \chi$;
2) every model of χ can be expanded to a model of χ'.

In what follows we can thus assume that the input formulas $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$ are of this form. Let $\Xi = \{\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)\}$. We use $\text{cl}(\Xi)$ to denote the closure under single negation of the set of all subformulas of φ and ψ with at most the variable x free and all formulas of the form $R(x)$ and $R(x, x)$ with R a unary or, respectively, binary relation symbol in φ, ψ. The *1-type $t_{\mathfrak{A}}(a)$ realized in a pointed structure* $\mathfrak{A}, a$ is defined as

$$t_{\mathfrak{A}}(a) := \{\chi(x) \mid \mathfrak{A} \models \chi(a), \chi \in \text{cl}(\Xi)\}$$

A *1-type* t is any subset of $\text{cl}(\Xi)$ such that there exists a pointed structure $\mathfrak{A}, a$ with $t = t_{\mathfrak{A}}(a)$. A *link-type* l contains $x \neq y$ and for any binary relation symbol R in Ξ either $R(x, y)$ or $\neg R(x, y)$ and $R(y, x)$ or $\neg R(y, x)$. The *link-type $l_{\mathfrak{A}}(a, b)$ realized in a pointed structure* $\mathfrak{A}, a, b$ contains $R(x, y)$ iff $\mathfrak{A} \models R(a, b)$ and it contains $R(y, x)$ iff $\mathfrak{A} \models R(b, a)$.

For a pair (l, s) with l a link-type and s a 1-type we say that nodes d, d' *satisfy* (l, s) in $\mathfrak{A}$ if $l = l_{\mathfrak{A}}(d, d')$ and $s = t_{\mathfrak{A}}(d')$.

Now assume that φ and ψ are jointly $\text{FO}^2(\tau)$-consistent. Then we find pointed models $\mathfrak{A}_1, \mathbf{a}_1$ and $\mathfrak{A}_2, \mathbf{a}_2$ satisfying φ and ψ, respectively, such that $\mathfrak{A}_1, \mathbf{a}_1 \sim_{\text{FO}^2,\tau} \mathfrak{A}_2, \mathbf{a}_2$. We extract from $\mathfrak{A}_1$ and $\mathfrak{A}_2$ new pointed models $\mathfrak{B}_1, \mathbf{b}_1$ and $\mathfrak{B}_2, \mathbf{b}_2$ which still witness joint $\text{FO}^2(\tau)$-consistency of φ and ψ but which are of at most double exponential size in φ and ψ. In what follows we assume that $\text{dom}(\mathfrak{A}_1) \cap \text{dom}(\mathfrak{A}_2) = \emptyset$. We write $d \sim_{\text{FO}^2,\tau} e$ if there are $i, j \in \{1, 2\}$ with $d \in \text{dom}(\mathfrak{A}_i)$, $e \in \text{dom}(\mathfrak{A}_j)$, and $\mathfrak{A}_i, d \sim_{\text{FO}^2,\tau} \mathfrak{A}_j, e$.

A *mosaic* m is a pair (Φ_1, Φ_2) with Φ_1, Φ_2 sets of 1-types. The *mosaic $m(d) = (\Phi_1, \Phi_2)$ generated by* $d \in \text{dom}(\mathfrak{A}_1) \cup \text{dom}(\mathfrak{A}_2)$ is defined by setting

$$\Phi_j = \{t_{\mathfrak{A}_j}(e) \mid e \in \text{dom}(\mathfrak{A}_j), d \sim_{\text{FO}^2,\tau} e\},$$

for $j = 1, 2$. The set $\mathcal{M}$ of all mosaics generated in $\mathfrak{A}_1, \mathfrak{A}_2$ is then defined as

$$\mathcal{M} = \{m(d) \mid d \in \text{dom}(\mathfrak{A}_1) \cup \text{dom}(\mathfrak{A}_2)\}.$$

Observe that since $\text{FO}^2(\tau)$-bisimulations are global, $\mathcal{M} = \{m(d) \mid d \in \text{dom}(\mathfrak{A}_i)\}$, for $i = 1, 2$. The set $\mathcal{K} \subseteq \mathcal{M}$ of *king mosaics* is defined as the set of all $m(d) \in \mathcal{M}$ such that for all e with $m(d) = m(e)$ we have $d \sim_{\text{FO}^2,\tau} e$. Let $\mathcal{C} = \mathcal{M} \setminus \mathcal{K}$ be the set of *pawn mosaics*. If $m(d) = (\Phi_1, \Phi_2)$ is a king mosaic, then call any $t \in \Phi_i$ such that there exists exactly one e with $t = t_{\mathfrak{A}_i}(e)$ and $d \sim_{\text{FO}^2,\tau} e$ an *i-king* in (Φ_1, Φ_2). Any $t \in \Phi_i$ that is not an i-king in (Φ_1, Φ_2) is called an *i-pawn* in (Φ_1, Φ_2). All $t \in \Phi_i$ with (Φ_1, Φ_2) a pawn mosaic are called *i-pawns* in (Φ_1, Φ_2). Note that we generalize a few notions introduced in the single exponential size model property proof for FO^2 presented in [19]. In that proof, 1-types that are realized exactly once in a model played a special roles and were called kings. Here we generalize that notion to king mosaics and kings within king mosaics.

We are now in the position to define the domains of $\mathfrak{B}_1, \mathfrak{B}_2$ as follows. Let s be the size of the input $\varphi(\mathbf{x}_0), \psi(\mathbf{x}_0)$. Then the number of mosaics is bounded by $m_{\varphi,\psi} = 2^{2^{s+1}}$. Let $k_1 = 2^{4s} \times m_{\varphi,\psi}$ and let $k_2 = 2^{3s} \times k_1^2$.

Take k_1 many copies $(t,1),(t,2),\ldots,(t,k_1)$ of every 1-type t and take k_2 many copies $(m,1),(m,2),\ldots,(m,k_2)$ of every pawn mosaic m. Then the domain $\text{dom}(\mathfrak{B}_i)$ of $\mathfrak{B}_i$ contains, for $i = 1,2$:

1) *new i-kings* (t,m), for $m \in \mathcal{K}$ and t an i-king in m;
2) *semi i-pawns* $((t,1),m),\ldots,((t,k_1),m)$ for $m \in \mathcal{K}$ and t an i-pawn in m;
3) *full i-pawns* $((t,1),(m,j)),\ldots,((t,k_1),(m,j))$, for m a pawn mosaic, t an i-pawn in m, and $1 \leq j \leq k_2$.

Observe that $\text{dom}(\mathfrak{B}_1) \cup \text{dom}(\mathfrak{B}_2)$ is of double exponential size in φ, ψ. To simplify notation we

- denote copies of types t by t' and copies of pawn mosaics m by m';
- often regard a king mosaic m as a copy m' of itself and an i-king t in a king mosaic m as a copy t' of itself.

We aim to construct $\mathfrak{B}_1$ and $\mathfrak{B}_2$ such that the following two conditions hold (where, as announced, t' and m' also range over i-kings and king mosaics, respectively):

1) Any pair (t',m') realizes the 1-type of which t' is a copy. More precisely, for $i = 1,2$, if $(t',m') \in \text{dom}(\mathfrak{B}_i)$ and t' is a copy of 1-type t and $\beta(x) \in \text{cl}(\Xi)$, then
$$\mathfrak{B}_i \models \beta(t',m') \quad \Leftrightarrow \quad \beta(x) \in t.$$
2) For any copy m' of a mosaic, all (t',m') are $\text{FO}^2(\tau)$-bisimilar. More precisely, for all t_1', t_2', m' such that $(t_1',m') \in \text{dom}(\mathfrak{B}_i)$ and $(t_2',m') \in \text{dom}(\mathfrak{B}_j)$ for some $i,j \in \{1,2\}$: $\mathfrak{B}_i,(t_1',m') \sim_{\text{FO}^2,\tau} \mathfrak{B}_j,(t_2',m')$.

We first define the interpretation of relation symbols on singleton subsets of $\text{dom}(\mathfrak{B}_i)$ in the obvious way by setting $(t',m') \in R^{\mathfrak{B}_i}$ iff $R(x) \in t$, for R unary, and $((t',m'),(t',m')) \in R^{\mathfrak{B}_i}$ iff $R(x,x) \in t$, for R binary. It thus remains to define the link-types $l_{\mathfrak{B}_i}((t_1',m_1'),(t_2',m_2'))$ between distinct nodes (t_1',m_1') and (t_2',m_2') in $\mathfrak{B}_i$, $i = 1,2$. To this end, we will carefully associate

- with every copy m' of a mosaic a *generator* $g \in \text{dom}(\mathfrak{A}_1) \cup \text{dom}(\mathfrak{A}_2)$ such that m' is a copy of $m = m(g)$;
- with every node $(t',m') \in \text{dom}(\mathfrak{B}_i)$ a *witness* $d \in \text{dom}(\mathfrak{A}_i)$ for (t',m') such that $d \sim_{\text{FO}^2,\tau} g$ for the generator g of m' and t' is a copy of $t_{\mathfrak{A}_i}(d)$.

If (t_1',m_1') and (t_2',m_2') contain a new i-king, then we will define $l_{\mathfrak{B}_i}((t_1',m_1'),(t_2',m_2'))$ as $l_{\mathfrak{A}_i}(d_1,d_2)$ for the selected witnesses d_1 and d_2 for (t_1',m_1') and (t_2',m_2'), respectively. For i-pawns, $l_{\mathfrak{B}_i}((t_1',m_1'),(t_2',m_2'))$ will be defined using 'global' constraints and will not in general be the induced link type from $\mathfrak{A}_i$. We now give the detailed construction.

For king mosaics m we simply select as its generator any g with $m = m(g)$ and for new i-kings (t,m) we take the unique $d \sim_{\text{FO}^2,\tau} g$ with $t = t_{\mathfrak{A}_i}(d)$ as its witness. The definition of link-types between new i-kings is then as announced: if (t_1,m_1) and (t_2,m_2) are new distinct i-kings, then set $l_{\mathfrak{B}_i}((t_1,m_1),(t_2,m_2)) := l_{\mathfrak{A}_i}(d_1,d_2)$ for the witnesses d_1,d_2 for (t_1,m_1) and (t_2,m_2), respectively.

Link-types between new i-kings and semi i-pawns. Assume d is the witness for an i-king (t,m) and (d,d') satisfies (l,s) for a link-type l and 1-type s, where d' is a node realizing an i-pawn and $m(d')$ is a king-mosaic. Then we aim to ensure that the link-type realized by $((t,m),(s',m(d'))$ equals l, for some copy s' of s. To obtain these link-types we carefully choose the witnesses d' for semi-i-pawns and then take, as announced, the link-type between the selected witnesses as given by $\mathfrak{A}_i$:

(P1) Let (t,m) be a new i-king. Let (s,n) be a pair such that n is a king mosaic and s an i-pawn in n. Let l be a link-type. If d is the witness of (t,m) and (d,d') satisfies (l,s) for some d' with $n = m(d')$, then pick a copy s' of s, pick such a d' as the witness for (s',n), and set $l_{\mathfrak{B}_i}((t_{\mathfrak{A}_i}(e),m(e)),(s',n)) := l_{\mathfrak{A}_i}(e,d')$, for all witnesses e for new i-kings. There are sufficiently many fresh copies s' of 1-types s as $k_1 \geq m_1 m_2$, where m_1 is the number of new i-kings and m_2 is the number of link-types.

For any pair (s',n) with n a king mosaic and s' a copy of an i-pawn s in n not selected according to (P1), pick any d' with $n = m(d')$ such that s' is a copy of $t_{\mathfrak{A}_i}(d')$ as the witness for (s',n) and let $l_{\mathfrak{B}_i}((t_{\mathfrak{A}_i}(e),m(e)),(s',n)) := l_{\mathfrak{A}_i}(e,d')$, for all witnesses e for new i-kings.

Link-types between semi i-pawns. For any king mosaics m_1,m_2 and i-pawns $t_1 \in m_1$ and $t_2 \in m_2$ define
$$\begin{aligned} L_i((t_1,m_1),(t_2,m_2)) = \{ & l_{\mathfrak{A}_i}(d_1,d_2) \mid d_1 \neq d_2, \\ & t_1 = t_{\mathfrak{A}_i}(d_1), m_1 = m(d_1), \\ & t_2 = t_{\mathfrak{A}_i}(d_2), m_2 = m(d_2)\}. \end{aligned}$$

(Note that $(t_1,m_1) = (t_2,m_2)$ is possible.) Using the fact that the number of copies of any 1-type exceeds 4×2^{2s}, it is straightforward to define the link-types $l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2))$, where t_1' and t_2' are copies of t_1 and t_2, in such a way that the following holds:

(P2) If t_1' is a copy of t_1 and l a link-type, then there exists a copy t_2' of t_2 such that $l = l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2))$ iff $l \in L_i((t_1,m_1),(t_2,m_2))$.

Selecting generators for pawn mosaics. To define link-types for pairs of nodes that include full i-pawns, we first fix the generators of copies of pawn mosaics as follows:

(M) Let (t',n) be either a new i-king or semi i-pawn. Let l be a link-type and s a 1-type. If d is the witness for (t',n) and (d,d') satisfies (l,s) for some d' such that $d' \not\sim_{\text{FO}^2,\tau} g$ for any g generating a king mosaic, then take such a d' and a copy m' of the pawn mosaic m generated by d' and select as generator of m' any g' with

$m(g') = m$ and $d' \sim_{\mathrm{FO}^2,\tau} g'$. There are sufficiently many copies of pawn mosaics as $k_2 \geq 2m_1m_2$, where m_1 is the number of new i-kings and semi i-pawns and m_2 is the number of link-types.

For any copy m' of a pawn mosaic m for which no generator has yet been selected in (M) choose an arbitrary g with $m = m(g)$ as a generator.

Link-types between new i-kings and full i-pawns. These link-types are now defined similarly to the link-types between new i-kings and semi i-pawns.

(P3) Let (t, m) be a new i-king. Let (s, n) be a pair such that n' is a copy of the pawn mosaic n and s is an i-pawn in n. Let l be a link-type. If d is the witness of (t, m) and g the generator of n' and (d, d') satisfies (l, s) for some d' with $d' \sim_{\mathrm{FO}^2,\tau} g$, then pick a copy s' of s, pick such a d' as the witness for (s', n'), and set $l_{\mathfrak{B}_i}((t_{\mathfrak{A}_i}(e), m(e)), (s', n')) := l_{\mathfrak{A}_i}(e, d')$, for all witnesses e for new i-kings.

As in (P1), there are sufficiently many copies for this to work and for any full i-pawn (s', n') not yet selected, pick any d' with $d' \sim_{\mathrm{FO}^2,\tau} g$ for the generator g of n' and s' a copy of $t_{\mathfrak{A}_i}(d')$ as the witness for (s', n') and let $l_{\mathfrak{B}_i}((t_{\mathfrak{A}_i}(e), m(e)), (s', n)) := l_{\mathfrak{A}_i}(e, d')$, for all witnesses e for new i-kings.

Link-types between semi i-pawns and full i-pawns. For any king mosaic m, i-pawn $t \in m$, copy n' of a pawn mosaic n, and any i-pawn $s \in n$, let g be the generator of n' and set

$$L_i((t, m), (s, n')) = \{l_{\mathfrak{A}_i}(d_1, d_2) \mid t = t_{\mathfrak{A}_i}(d_1), m = m(d_1), s = t_{\mathfrak{A}_i}(d_2), d_2 \sim_{\mathrm{FO}^2,\tau} g\}$$

Similarly to (P2), it is now straightforward to define link-types $l_{\mathfrak{B}_i}((t', m), (s', n'))$ in such a way that the following holds:

(P4) (a) If t' is a copy of t and l a link-type, then there exists a copy s' of s such that $l = l_{\mathfrak{B}_i}((t', m), (s', n'))$ iff $l \in L_i((t, m), (s, n'))$.
(b) If s' is a copy of s and l a link-type, then there exists a copy t' of t such that $l = l_{\mathfrak{B}_i}((t', m), (s', n'))$ iff $l \in L_i((t, m), (s, n'))$.

Link-types between full i-pawns. For any pawn mosaics m_1, m_2 and i-pawns $t_1 \in m_1$ and $t_2 \in m_2$, define $L_i((t_1, m_1), (t_2, m_2))$ in exactly the same way as in the definition of link-types between semi i-pawns. Then one can define the link types $l_{\mathfrak{B}_i}((t_1', m_1'), (t_2', m_2'))$ in such a way that the following holds:

(P5) If m_1' and m_2' are copies of m_1 and m_2, t_1' is a copy of t_1, and l is a link-type, then there exists a copy t_2' of t_2 such that $l = l_{\mathfrak{B}_i}((t_1', m_1'), (t_2', m_2'))$ iff $l \in L_i((t_1, m_1), (t_2, m_2))$.

This finishes the definition of $\mathfrak{B}_1$ and $\mathfrak{B}_2$. It is not difficult to show that Conditions 1 and 2 above hold. Assume w.l.o.g. that $\mathbf{a}_1 = (a_{11}, a_{12})$ and $\mathbf{a}_2 = (a_{21}, a_{22})$ with $a_{11} \neq a_{12}$. Then, $a_{21} \neq a_{22}$ as otherwise $\mathfrak{A}_1, \mathbf{a}_1 \not\sim_{\mathrm{FO}^2,\tau} \mathfrak{A}_2, \mathbf{a}_2$. It is straightforward to ensure in the construction of $\mathfrak{B}_1$ and $\mathfrak{B}_2$ above that $a_{11}, a_{12}, a_{21}, a_{22}$ are witnesses for domain elements $(t_{11}, m_1), (t_{12}, m_2)$ of $\mathfrak{B}_1$ and $(t_{21}, m_1), (t_{22}, m_2)$ of $\mathfrak{B}_2$ and that $l_{\mathfrak{B}_i}((t_{i1}, m_1), (t_{i2}, m_2)) = l_{\mathfrak{A}_i}(a_{i1}, a_{i2})$, for $i = 1, 2$. Then we have that $\mathfrak{B}_1 \models \varphi((t_{11}, m_1), (t_{12}, m_2))$ and $\mathfrak{B}_2 \models \psi((t_{21}, m_1), (t_{22}, m_2))$, by Condition 1. By Condition 2, $\mathfrak{B}_1, (t_{11}, m_1), (t_{12}, m_2) \sim_{\mathrm{FO}^2,\tau} \mathfrak{B}_2, (t_{21}, m_1), (t_{22}, m_2)$.

B. Lower Bound

The lower bound proof is essentially a modification of the lower bound for (the bounded arity case for) GF. In fact, it is also an reduction from the word problem of exponentially space bounded ATMs which uses the same signature τ. Again, we aim to construct an FO^2-sentence φ' such that such an ATM M accepts input w iff $\varphi' \wedge A(x)$ and $\varphi' \wedge \neg A(x)$ are jointly FO$^2(\tau)$-consistent. The sentence φ' is obtained from the sentence φ constructed for GF by replacing the first conjunct φ_0 with φ_0' (recall that all other conjuncts are already in FO^2). Recall that φ_0 enforced a cycle of length three using a ternary relation, which is impossible in FO^2. Instead, we proceed similar to Example 2. Indeed, φ_0' enforces that every element satisfying A is involved in such a cycle:

$$\varphi_0' = \forall x\,(Y(x) \to X(x) \wedge \varphi_3(x) \wedge \forall y\,(Y(y) \to x = y)) \wedge \forall x\,(A(x) \to Y(x))$$

where φ_3 is as in Example 2, that is, it enforces the existence of a path of length three to an element satisfying Y, which is enforced to be a singleton. Now, if $\varphi' \wedge A(x)$ and $\varphi' \wedge \neg A(x)$ are jointly FO$^2(\tau)$-consistent, there exist models $\mathfrak{A}$ and $\mathfrak{B}$ of φ' and elements a, b such that $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$, and $\mathfrak{A}, a \sim_{\mathrm{FO}^2,\tau} \mathfrak{B}, b$. If the latter holds, then from $a \in A^{\mathfrak{A}}$ and φ_0' it follows that b has an infinite outgoing path ρ along R on which every third element satisfies X. As FO$^2(\tau)$-bisimilarity is an equivalence relation, all these elements satisfying X are actually FO$^2(\tau)$-bisimilar. Now, the synchronization of the successor configurations works in the very same way as for GF; we prove correctness in the appendix.

Lemma 9. *M accepts the input w iff there exists models $\mathfrak{A}, \mathfrak{B}$ of φ' and elements $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$ such that $\mathfrak{A}, a \sim_{FO^2,\tau} \mathfrak{B}, b$.*

To prove the second part of Theorem 1 (ii) we replace in φ' every occurrence of any formula of the form $E(x)$ and $E(y)$ for a unary symbol $E \in \mathrm{sig}(\varphi') \setminus (\tau \cup \{A\})$ by the formula

$$\chi_E(x) = \exists y(R_E(x, y) \wedge \exists x(N(y, x) \wedge \exists y(N(x, y) \wedge A(y))))$$

and the formula $\chi_E(y)$ obtained from $\chi_E(x)$ by swapping x and y, respectively. Here R_E, $E \in \mathrm{sig}(\varphi') \setminus (\tau \cup \{A\})$, and N are fresh binary relation symbols. An analogue of Lemma 9 is proved in the appendix for the resulting formula φ'' and the signature τ' containing all relation symbols in φ'' except A.

VII. Conclusion

We have shown tight complexity bounds for interpolant and explicit definition existence in GF and coN2ExpTime/2ExpTime upper and, respectively, lower bounds for FO^2. Many questions remain to be explored. First we conjecture that these problems are actually coN2ExpTime-complete in FO^2. Then it would be of interest to determine the size of interpolants/explicit definitions in GF and FO^2 if they exist. Note that recently the size and computation of interpolants in GNF has been studied in depth [45]. In contrast to GF, GNF enjoys CIP and PBDP and it is not difficult to show using the complexity lower bound proof given above that in GF minimal interpolants/explicit definitions are, in the worst case, at least by one exponential larger than in GNF.

There are many logics without the CIP and PBDP for which the complexity of interpolant and explicit definition existence remain to be explored, examples include the extension of FO^2 with counting, FO^2 without equality, the extension of GF with constants, and the Horn fragment of GF introduced in [55].

References

[1] J. Van Benthem, "The many faces of interpolation," *Synthese*, pp. 451–460, 2008.

[2] K. L. McMillan, "Interpolation and sat-based model checking," in *Proc. of CAV*. Springer, 2003, pp. 1–13.

[3] A. Cimatti, A. Griggio, and R. Sebastiani, "Efficient interpolant generation in satisfiability modulo theories," in *Proc. of TACAS*. Springer, 2008, pp. 397–412.

[4] B. ten Cate, E. Franconi, and İ. Seylan, "Beth definability in expressive description logics," *J. Artif. Intell. Res.*, vol. 48, pp. 347–414, 2013.

[5] M. Benedikt, J. Leblay, B. ten Cate, and E. Tsamoura, *Generating Plans from Proofs: The Interpolation-based Approach to Query Reformulation*, ser. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2016.

[6] H. Andréka, I. Németi, and J. van Benthem, "Modal languages and bounded fragments of predicate logic," *J. Philosophical Logic*, vol. 27, no. 3, pp. 217–274, 1998.

[7] E. Grädel, "On the restraining power of guards," *J. Symb. Log.*, vol. 64, no. 4, pp. 1719–1742, 1999.

[8] E. Grädel and I. Walukiewicz, "Guarded fixed point logic," in *Proc. of LICS*. IEEE Computer Society, 1999, pp. 45–54.

[9] E. Grädel, C. Hirsch, and M. Otto, "Back and forth between guarded and modal logics," in *Proc. of LICS*. IEEE Computer Society, 2000, pp. 217–228.

[10] M. Otto, "Highly acyclic groups, hypergraph covers, and the guarded fragment," *J. ACM*, vol. 59, no. 1, pp. 5:1–5:40, 2012.

[11] V. Bárány, B. ten Cate, and L. Segoufin, "Guarded negation," *J. ACM*, vol. 62, no. 3, pp. 22:1–22:26, 2015.

[12] V. Bárány, G. Gottlob, and M. Otto, "Querying the guarded fragment," *Log. Methods in Comput. Sci.*, vol. 10, no. 2, 2014.

[13] A. Hernich, C. Lutz, F. Papacchini, and F. Wolter, "Dichotomies in ontology-mediated querying with the guarded fragment," *ACM Trans. Comput. Log.*, vol. 21, no. 3, pp. 1–47, 2020.

[14] E. Hoogland, M. Marx, and M. Otto, "Beth definability for the guarded fragment," in *Proc. of LPAR*. Springer, 1999, pp. 273–285.

[15] V. Bárány, M. Benedikt, and B. ten Cate, "Rewriting guarded negation queries," in *Proc. of MFCS*. Springer, 2013, pp. 98–110.

[16] L. Henkin, "Logical systems containing only a finite number of symbols," Department of Mathematics, University of Montreal, Report, 1967.

[17] M. Otto, *Bounded Variable Logics and Counting: A Study in Finite Models*, ser. Lecture Notes in Logic. Cambridge University Press, 2017, vol. 9.

[18] M. Grohe, "Finite variable logics in descriptive complexity theory," *Bulletin of Symbolic Logic*, vol. 4, no. 4, p. 345–398, 1998.

[19] E. Grädel, P. G. Kolaitis, and M. Y. Vardi, "On the decision problem for two-variable first-order logic," *Bull. Symb. Log.*, vol. 3, no. 1, pp. 53–69, 1997.

[20] M. Y. Vardi, "Why is modal logic so robustly decidable?" in *Proc. of DIMACS Workshop on Descriptive Complexity and Finite Models*, vol. 31. DIMACS/AMS, 1996, pp. 149–183.

[21] E. Grädel, "Why are modal logics so robustly decidable?" in *Current Trends in Theoretical Computer Science, Entering the 21st Centurey*. World Scientific, 2001, pp. 393–408.

[22] E. Kieronski, I. Pratt-Hartmann, and L. Tendera, "Two-variable logics with counting and semantic constraints," *ACM SIGLOG News*, vol. 5, no. 3, pp. 22–43, 2018.

[23] S. D. Comer, "Classes without the amalgamation property." *Pacific J. Math.*, vol. 28, no. 2, pp. 309–318, 1969.

[24] D. Pigozzi, "Amalgamation, congruence-extension, and interpolation properties in algebras," *Algebra Univers.*, no. 1, pp. 269–349, 1971.

[25] H. Andreka, S.D.Comer, and I. Nemeti, "Epimorphisms in cylindric algebras," Tech. Rep., 1983.

[26] H.Andreka, S.D.Comer, J.X.Madarasz, I. Nemeti, and T. Sayed-Ahmed, "Epimorphisms in cylindric algebras and definability in finite variable logic," *Algebra Univers.*, vol. 61, pp. 261–282, 2009.

[27] J. Lehmann and P. Hitzler, "Concept learning in description logics using refinement operators," *Mach. Learn.*, vol. 78, pp. 203–250, 2010.

[28] D. M. L. Martins, "Reverse engineering database queries from examples: State-of-the-art, challenges, and research opportunities," *Inf. Syst.*, 2019.

[29] E. Krahmer and K. van Deemter, "Computational generation of referring expressions: A survey," *Comput. Linguist.*, vol. 38, no. 1, pp. 173–218, 2012.

[30] J. C. Jung, C. Lutz, H. Pulicini, and F. Wolter, "Logical separability of incomplete data under ontologies," in *Proc. of KR*, 2020.

[31] J. C. Jung, C. Lutz, H. Pulcini, and F. Wolter, "Separating positive and negative data examples by concepts and formulas: The case of restricted signatures," in *Proc. of DL*, 2020.

[32] M. Marx, "Queries determined by views: Pack your views," in *Proc. of PODS*. ACM, 2007, p. 23–30.

[33] D. Toman and G. E. Weddell, *Fundamentals of Physical Design and Query Compilation*, ser. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2011.

[34] E. Hoogland and M. Marx, "Interpolation and definability in guarded fragments," *Stud. Log.*, vol. 70, no. 3, pp. 373–409, 2002.

[35] V. Goranko and M. Otto, "Model theory of modal logic," in *Handbook of Modal Logic*. Elsevier, 2007, pp. 249–329.

[36] E. Grädel and M. Otto, "The freedoms of (guarded) bisimulation," in *Johan van Benthem on Logic and Information Dynamics*. Springer International Publishing, 2014, pp. 3–31.

[37] J. Barwise, "On moschovakis closure ordinals," *J. Symb. Log.*, vol. 42, no. 2, pp. 292–296, 1977.

[38] N. Immerman, "Upper and lower bounds for first order expressibility," *J. Comput. Syst. Sci.*, vol. 25, no. 1, pp. 76–98, 1982.

[39] T. Place and M. Zeitoun, "Separating regular languages with first-order logic," *Log. Methods Comput. Sci.*, vol. 12, no. 1, 2016.

[40] K. Henkell, "Pointlike sets: the finest aperiodic cover of a finite semigroup," *J. Pure Appl. Algebra*, vol. 55, no. 1-2, pp. 85–126, 1988.

[41] K. Henkell, J. Rhodes, , and B. Steinberg, "Aperiodic pointlikes and beyond," *Internat. J. Algebra Comput.*, vol. 20, no. 2, pp. 287–305, 2010.

[42] A. Artale, J. C. Jung, A. Mazzullo, A. Ozaki, and F. Wolter, "Living without beth and craig: Explicit definitions and interpolants in description logics with nominals and role hierarchies," in *Proc. of AAAI*, 2021.

[43] E. Hoogland and M. Marx, "Interpolation in guarded fragments," University of Amsterdam, Tech. Rep., 1999.

[44] V. Bárány, M. Benedikt, and B. ten Cate, "Some model theory of guarded negation," *J. Symb. Log.*, vol. 83, no. 4, pp. 1307–1344, 2018.

[45] M. Benedikt, B. ten Cate, and M. Vanden Boom, "Effective interpolation and preservation in guarded logics," *ACM Trans. Comput. Log.*, vol. 17, no. 2, pp. 8:1–8:46, 2016.

[46] G. D'Agostino and G. Lenzi, "Bisimulation quantifiers and uniform interpolation for guarded first order logic," *Theor. Comput. Sci.*, vol. 563, pp. 75–85, 2015.

[47] J. C. Jung, C. Lutz, M. Martel, T. Schneider, and F. Wolter, "Conservative extensions in guarded and two-variable fragments," in *Proc. of ICALP*. Schloss Dagstuhl – LZI, 2017, pp. 108:1–108:14.

[48] M. Marx and C. Areces, "Failure of interpolation in combined modal logics," *Notre Dame J. Formal Log.*, vol. 39, no. 2, pp. 253–273, 1998.

[49] J. Barwise and J. van Benthem, "Interpolation, preservation, and pebble games," *J. Symb. Log.*, vol. 64, no. 2, pp. 881–903, 1999.

[50] H. Andreka and I. Nemeti, "Two-variable logic has weak, but not strong, beth definability," Tech. Rep., 2020.

[51] L. Maksimova and D. Gabbay, *Interpolation and Definability in Modal and Intuitionistic Logics*. Clarendon Press, 2005.

[52] C. Areces, P. Blackburn, and M. Marx, "Hybrid logics: Characterization, interpolation and complexity," *J. Symb. Log.*, vol. 66, no. 3, pp. 977–1010, 2001.

[53] B. ten Cate, "Interpolation for extended modal languages," *J. Symb. Log.*, vol. 70, no. 1, pp. 223–234, 2005.

[54] A. K. Chandra, D. C. Kozen, and L. J. Stockmeyer, "Alternation," *J. ACM*, vol. 28, pp. 114–133, 1981.

[55] J. C. Jung, F. Papacchini, F. Wolter, and M. Zakharyaschev, "Model comparison games for horn description logics," in *Proc. of LICS*. IEEE, 2019, pp. 1–14.

Appendix

Proofs for Section IV

Lemma 2. *Let $\mathcal{L}$ be either FO^2 or GF. Let $\varphi(\mathbf{x}), \psi(\mathbf{x})$ be $\mathcal{L}$-formulas and let $\tau = sig(\varphi) \cap sig(\psi)$. Then the following conditions are equivalent:*

1) *there does not exist an $\mathcal{L}$-interpolant for $\varphi(\mathbf{x}), \psi(\mathbf{x})$;*
2) *$\varphi(\mathbf{x}), \neg\psi(\mathbf{x})$ are jointly $\mathcal{L}(\tau)$-consistent.*

Proof. ($\Leftarrow$) Assume there is an $\mathcal{L}$-interpolant $\theta(\mathbf{x})$ and let $\mathfrak{A}, \mathfrak{B}$ be structures and $\mathbf{a}, \mathbf{b}$ be tuples such that $\mathfrak{A} \models \varphi(\mathbf{a})$ and $\mathfrak{B} \models \neg\psi(\mathbf{b})$. Suppose further that $\mathfrak{A}, \mathbf{a} \sim_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b}$. Since $\varphi(\mathbf{x}) \models \theta(\mathbf{x})$, we have $\mathfrak{A} \models \theta(\mathbf{a})$. By Lemma 1, we obtain $\mathfrak{B} \models \theta(\mathbf{b})$. Finally, as $\theta(\mathbf{x}) \models \psi(\mathbf{x})$, we obtain $\mathfrak{B} \models \psi(\mathbf{b})$, a contradiction.

($\Rightarrow$) Suppose that for all structures $\mathfrak{A}, \mathfrak{B}$ and tuples $\mathbf{a}, \mathbf{b}$ such that $\mathfrak{A} \models \varphi(\mathbf{a})$ and $\mathfrak{B} \models \neg\psi(\mathbf{b})$ we have $\mathfrak{A}, \mathbf{a} \not\sim_{\mathcal{L},\tau} \mathfrak{B}, \mathbf{b}$. Let Φ be defined by taking

$$\Phi = \{\varphi'(\mathbf{x}) \in \mathcal{L}(\tau) \mid \varphi(\mathbf{x}) \models \varphi'(\mathbf{x})\}.$$

Clearly, $\varphi(\mathbf{x}) \models \Phi$. We claim that also $\Phi \models \psi(\mathbf{x})$. To see this, let $\mathfrak{B}, \mathbf{b}$ such that $\mathfrak{B} \models \Phi(\mathbf{b})$. Let $\mathfrak{B}'$ be an ω-saturated elementary extension of $\mathfrak{B}$ and let $\mathfrak{A}, \mathbf{a}$ be an ω-saturated pointed structure realizing $\{\chi(\mathbf{x}) \in \mathcal{L}(\tau) \mid \mathfrak{B} \models \chi(\mathbf{b})\} \cup \{\varphi\}$ in $\mathbf{a}$ ($\mathfrak{A}, \mathbf{a}$ exists by compactness and the definition of Φ). By definition of Φ and Lemma 1, we have $\mathfrak{A}, \mathbf{a} \sim_{\mathcal{L},\tau} \mathfrak{B}', \mathbf{b}$. By the initial assumption, we cannot have $\mathfrak{B}' \models \neg\psi(\mathbf{b})$ and thus $\mathfrak{B} \models \psi(\mathbf{b})$. By compactness, there is a finite subset Φ' of Φ such that $\Phi' \models \psi(\mathbf{x})$. The conjunction of the formulas in Φ' is the required interpolant. ❑

Lemma 3. *Let $\mathcal{L}$ be either FO^2 or GF. There is a polynomial time reduction of explicit $\mathcal{L}$-definability to $\mathcal{L}$-interpolant existence.*

Proof. Assume φ, $\theta(\mathbf{x})$, and τ are given. Then $\theta(\mathbf{x})$ is explicitly definable under φ iff there exists an $\mathcal{L}$-interpolant for $\varphi \wedge \theta(\mathbf{x}), \varphi' \rightarrow \theta'(\mathbf{x})$, where φ' and θ' are obtained from φ and θ, respectively, by renaming all non-τ symbols R to fresh R' of the same arity. ❑

Proofs for Section V

Claim 1. *For all i, j with $1 \leq i, j \leq m$, we have:*

1) *For every $\sigma \in T_i$ with tail$(\sigma) = (t(\mathbf{y}), \Phi)$, we have $\mathfrak{A}_i \models t(v_\sigma(\mathbf{y}))$;*
2) *Let $H_{i,j}$ be the set of all mappings $p_{\sigma,\sigma',\mathbf{z}}$, where*
   - *$\sigma \in T_i, \sigma' \in T_j$, tail$(\sigma) = (t(\mathbf{y}), \Phi)$, and tail$(\sigma') = (t'(\mathbf{y}'), \Phi)$;*
   - *$\mathbf{z}$ is a tuple with $[\mathbf{z}] \subseteq [\mathbf{y}] \cap [\mathbf{y}']$ and $v_\sigma(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_i$ (or, equivalently, $v_{\sigma'}(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_j$);*
   - *$p_{\sigma,\sigma',\mathbf{z}} : v_\sigma(\mathbf{z}) \mapsto v_{\sigma'}(\mathbf{z})$.*

   Then $H_{i,j}$ is a GF(τ)-bisimulation between $\mathfrak{A}_i$ and $\mathfrak{A}_j$.

Proof. For Point 1, we prove by induction that, for all $\sigma \in T_i$ with tail$(\sigma) = (t(\mathbf{y}), \Phi)$ and all formulas $\varphi(\mathbf{z})$ with $[\mathbf{z}] \subseteq [\mathbf{y}]$, we have:

$$\varphi(\mathbf{z}) \in t(\mathbf{y}) \quad \text{iff} \quad \mathfrak{A}_i \models \varphi(v_\sigma(\mathbf{z}))$$

The induction base is given by the definition of bag$_i(\sigma)$. If φ is of the shape $\neg\varphi'$, $\varphi' \wedge \varphi''$, or $\varphi' \vee \varphi''$, the statement is immediate from the hypothesis. Consider now $\varphi(\mathbf{z}) = \exists \mathbf{x}(R(\mathbf{z}, \mathbf{x}) \wedge \lambda(\mathbf{z}, \mathbf{x}))$.

($\Rightarrow$) Since $\mathcal{M}$ is existentially saturated, there is a $\Phi' \in \mathcal{M}$ such that Φ, Φ' are compatible and $R(\mathbf{z}, \mathbf{x}') \wedge \lambda(\mathbf{z}, \mathbf{x}') \in t'(\mathbf{y}')$ for some $t'(\mathbf{y}') \in \Phi'$ such that $t(\mathbf{y})$ and $t'(\mathbf{y}')$ coincide on $[\mathbf{y}] \cap [\mathbf{y}']$. By definition of T_i and compatibility of Φ, Φ', we have $\sigma' = \sigma \cdot (t'(\mathbf{y}'), \Phi') \in T_i$. Moreover, by induction, we obtain that $\mathfrak{A}_i$ satisfies $R(\mathbf{z}, \mathbf{x}') \wedge \lambda(\mathbf{z}, \mathbf{x}')$ under $v_{\sigma'}$. By definition of bag$_i(\sigma)$ and bag$_i(\sigma')$, we get $\mathfrak{A}_i \models \varphi(v_\sigma(\mathbf{z}))$.

($\Leftarrow$) Conversely, assume $\mathfrak{A}_i \models \varphi(v_\sigma(\mathbf{z}))$. By construction, there is some $\sigma' \in T_i$ such that $v_\sigma(\mathbf{z}) = v_{\sigma'}(\mathbf{z})$ and $\mathfrak{A}_i$ satisfies $R(\mathbf{z}, \mathbf{x}') \wedge \lambda(\mathbf{z}, \mathbf{x}')$ under $v_{\sigma'}$, for some $\mathbf{x}'$. By induction hypothesis, $R(\mathbf{z}, \mathbf{x}') \wedge \lambda(\mathbf{z}, \mathbf{x}') \in t'(\mathbf{y}')$, where tail$(\sigma') = (t'(\mathbf{y}'), \Phi')$. Thus, $\exists \mathbf{x}(R(\mathbf{z}, \mathbf{x}) \wedge \lambda(\mathbf{z}, \mathbf{x})) = \varphi(\mathbf{z}) \in t'(\mathbf{y}')$. As $v_\sigma(\mathbf{z}) = v_{\sigma'}(\mathbf{z})$, the construction of T_i implies that $t'(\mathbf{y}')$ and $t(\mathbf{y})$ coincide on all subformulas over $\mathbf{z}$, hence $\varphi(\mathbf{z}) \in t(\mathbf{y})$.

For Point 2, observe first that the $p_{\sigma,\sigma',\mathbf{z}}$ are partial τ-isomorphisms between τ-guarded tuples since all $\Phi \in \mathcal{M}$ are τ-uniform. (In addition, the observation that $v_\sigma(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_i$ iff $v_{\sigma'}(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_j$ follows from the condition that Φ is τ-uniform.) By symmetry, it suffices to prove Condition (i) for GF(τ)-bisimulations.

Let $p \in H_{i,j}$. Then we have $\sigma \in T_i, \sigma' \in T_j$ with tail$(\sigma) = (t(\mathbf{y}), \Phi)$ and tail$(\sigma') = (t'(\mathbf{y}'), \Phi)$ and we have a tuple $\mathbf{z}$ such that $[\mathbf{z}] \subseteq [\mathbf{y}] \cap [\mathbf{y}']$ and $v_\sigma(\mathbf{z})$ is τ-guarded in $\mathfrak{A}_i$ and $p = p_{\sigma,\sigma',\mathbf{z}}$. Thus, there is a τ-guard $Q(\mathbf{z})$ with $\mathfrak{A}_i \models Q(v_\sigma(\mathbf{z}))$ and $\mathfrak{A}_j \models Q(v_{\sigma'}(\mathbf{z}))$. Consider any tuple $\mathbf{b}$ with $\mathfrak{A}_i \models R(\mathbf{b})$ for some $R \in \tau$. We have to show that there exists a mapping $p_{\rho,\rho',\mathbf{z}'} \in H_{i,j}$ with domain $[\mathbf{b}]$ which coincides with $p_{\sigma,\sigma',\mathbf{z}}$ on $[v_\sigma(\mathbf{z})] \cap [\mathbf{b}]$. We distinguish on whether or not that intersection is empty.

Case 1. $[v_\sigma(\mathbf{z})] \cap [\mathbf{b}] = \emptyset$. The existence of such a mapping follows from GF(τ)-bisimulation saturatedness: to see this, observe that, as we have a tree decomposition, there exists $\rho_0 \in T_i$ such that $[\mathbf{b}] \subseteq \text{dom}(\text{bag}(\rho_0))$. Let tail$(\rho_0) = (s(\mathbf{x}_0), \Omega)$. Then there exists a tuple $\mathbf{y}_0$ with $[\mathbf{y}_0] \subseteq [\mathbf{x}_0]$ such that $v_{\rho_0}(\mathbf{y}_0) = \mathbf{b}$. We have $R(\mathbf{y}_0) \in s(\mathbf{x}_0)$. As $\hat{t}_j \in \Omega$, by GF(τ)-bisimulation saturatedness of Ω, there exists $s'(\mathbf{y}_0') \in \Omega$ such that $\hat{t}_j \subseteq s'(\mathbf{y}_0')$ and $[\mathbf{y}_0'] = [\mathbf{y}_0]$. But then $R(\mathbf{y}_0) \in s'(\mathbf{y}_0')$. Also $\rho = (\hat{t}_j, \widehat{\Psi}) \cdot (s'(\mathbf{y}_0'), \Omega) \in T_j$. Thus $p_{\rho_0,\rho,\mathbf{y}_0'}$ is as required.

Case 2. $[v_\sigma(\mathbf{z})] \cap [\mathbf{b}] \neq \emptyset$. As we have a tree decomposition, there exists $\rho_0 \in T_i$ such that $[\mathbf{b}] \subseteq \text{dom}(\text{bag}(\rho_0))$. Let tail$(\rho_0) = (s(\mathbf{x}_0), \Omega)$. Then there exists a tuple $\mathbf{z}'$ with $[\mathbf{z}'] \subseteq [\mathbf{x}_0]$ such that $v_{\rho_0}(\mathbf{z}') = \mathbf{b}$. We distinguish the following cases:

(a) $\rho_0 = \sigma$;

(b) $\rho_0 \neq \sigma$.

Assume first that (a) holds. Then $(s(\mathbf{x}_0), \Omega) = (t(\mathbf{y}), \Phi)$ and $\mathbf{b} = v_\sigma(\mathbf{z}')$. We use GF($\tau$)-bisimulation saturatedness of Φ. Consider the restriction $\mathbf{z}''$ of $\mathbf{z}'$ to $[\mathbf{z}] \cap [\mathbf{z}']$ and the restriction $t'(\mathbf{y}')_{|[\mathbf{z}'']}$ of $t'(\mathbf{y}')$ to $[\mathbf{z}'']$. Then there exists $s'(\mathbf{z}_0') \in \Phi$ such that $t'(\mathbf{y}')_{|[\mathbf{z}'']} \subseteq s'(\mathbf{z}_0') \in \Phi$ and $[\mathbf{z}_0'] = [\mathbf{z}']$. Let $\sigma'' = \sigma' \cdot (s'(\mathbf{z}_0'), \Phi) \in T_j$. Then $p_{\sigma,\sigma'',\mathbf{z}_0'}$ is as required, as Φ is τ-uniform.

Assume now that Point (b) holds. Consider the restriction $\mathbf{z}''$ of $\mathbf{z}'$ to $[\mathbf{z}] \cap [\mathbf{z}']$ and the restriction $t'(\mathbf{y}')_{|[\mathbf{z}'']}$ of $t'(\mathbf{y}')$ to $[\mathbf{z}'']$. Consider the restriction $\Phi_{|[\mathbf{z}'']}$ of Φ to $[\mathbf{z}'']$. By closure under restrictions, $\Phi_{|[\mathbf{z}'']} \in \mathcal{M}$. Observe that $\Phi, \Phi_{|[\mathbf{z}'']}$ and $\Phi_{|[\mathbf{z}'']}, \Omega$ are compatible: indeed, in the tree decomposition all bags on the path from σ to ρ_0 have a tail $(\cdot, \Omega')$ satisfying $\Phi_{|[\mathbf{z}'']} \subseteq \Omega'$. Thus $t'(\mathbf{y}')_{|[\mathbf{z}'']} \in \Omega$. Using the fact that Ω is GF(τ)-bisimulation saturated, one can now show that there exists $s'(\mathbf{z}_0') \in \Omega$ such that $t'(\mathbf{y}')_{|[\mathbf{z}'']} \subseteq s'(\mathbf{z}_0')$ and $[\mathbf{z}_0'] = [\mathbf{z}']$. We then have

$$\rho = \sigma' \cdot (t'(\mathbf{y}')_{|[\mathbf{z}'']}, \Phi_{|[\mathbf{z}'']}) \cdot (s'(\mathbf{z}_0'), \Omega) \in T_j$$

and $p_{\rho_0,\rho,\mathbf{z}_0'}$ is as required. ❑

Lemma 6 *Let $\mathfrak{A}_1, \mathbf{a}_1$ and $\mathfrak{A}_2, \mathbf{a}_2$ be pointed structures with $\mathbf{a}_1$ and $\mathbf{a}_2$ tuples with pairwise distinct elements of length $m \leq fv(\Xi)$ and let τ be a signature. Consider assignments $\mathbf{x}_0 \mapsto \mathbf{a}_i$ with $[\mathbf{x}_0] \subseteq \{x_0, \ldots, x_{2n}\}$. If $\mathfrak{A}_1, \mathbf{a}_1 \sim_{GF,\tau} \mathfrak{A}_2, \mathbf{a}_2$, then there exists an existentially saturated set $\mathcal{M}$ of τ-mosaics and some $\Psi \in \mathcal{M}$ such that*

- *all $\Phi \in \mathcal{M}$ with $\Phi \neq \Psi$ use at most width(Ξ) many free variables;*
- *there exist types $t_1(\mathbf{x}_0), t_2(\mathbf{x}_0) \in \Psi$ such that $t_i(\mathbf{x}_0) = tp(\mathfrak{A}_i, \mathbf{x}_0 \mapsto \mathbf{a}_i)$ for $i = 1, 2$ and all types $t(\mathbf{y}) \in \Psi \setminus \{t_1(\mathbf{x}_0), t_2(\mathbf{x}_0)\}$ use at most width(Ξ) free variables among $[\mathbf{x}_0]$.*

Proof. Assume w.l.o.g. that $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are disjoint. For any tuples $\mathbf{b}_1$ in $\mathfrak{A}_i$ and $\mathbf{b}_2$ in $\mathfrak{A}_j$ with $i, j \in \{1, 2\}$, we use tp$(\mathbf{x}_1 \mapsto \mathbf{b}_1)$ to denote tp$(\mathfrak{A}_i, \mathbf{x}_1 \mapsto \mathbf{b}_1)$ and we write $\mathbf{b}_1 \sim_{\mathrm{GF},\tau} \mathbf{b}_2$ if $\mathfrak{A}_i, \mathbf{b}_1 \sim_{\mathrm{GF},\tau} \mathfrak{A}_j, \mathbf{b}_2$. Define $\mathcal{M}$ as follows. Take any tuple $\mathbf{a}$ of distinct elements in $\mathfrak{A}_i$, $i \in \{1, 2\}$. Take a tuple $\mathbf{x}$ from $\{x_1, \ldots, x_{2n}\}$ such that $v : \mathbf{x} \mapsto \mathbf{a}$ is a bijection. Then let $\Phi_{\mathbf{a},\mathbf{x}}$ contain all types tp$(v' : \mathbf{x}_{|Y} \mapsto \mathbf{b})$ with $Y \subseteq [\mathbf{x}]$ and $\mathbf{b}$ in either $\mathfrak{A}_1$ or $\mathfrak{A}_2$ such that $v(\mathbf{x}_{|Y}) \sim_{\mathrm{GF},\tau} v'(\mathbf{x}_{|Y})$.

Let $\mathcal{M}$ contain all such $\Phi_{\mathbf{a},\mathbf{x}}$ with $\mathbf{a}$ of length at most width(Ξ) and $\mathbf{x}$ from $\{x_1, \ldots, x_{2n}\}$. Moveover, if $m >$ width(Ξ), then add $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ to $\mathcal{M}$, where $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ is obtained from $\Phi_{\mathbf{a}_1,\mathbf{x}_0}$ by removing all t distinct from $t_1(\mathbf{x}_0)$ and $t_2(\mathbf{x}_0)$ using more than width(Ξ) many free variables.

We show that $\mathcal{M}$ is as required. By definition, tp$(\mathfrak{A}_1, \mathbf{x}_0 \mapsto \mathbf{a}_1)$, tp$(\mathfrak{A}_2, \mathbf{x}_0 \mapsto \mathbf{a}_2) \in \Phi_{\mathbf{a}_1,\mathbf{x}_0} \in \mathcal{M}$.

For the next steps we first assume that instead of $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ we have $\Phi_{\mathbf{a}_1,\mathbf{x}_0}$ in $\mathcal{M}$. Then observe that if we have any $\Phi \in \mathcal{M}$ and $t(\mathbf{x}'), s(\mathbf{x}'') \in \Phi$, then we can assume that $\Phi = \Phi_{\mathbf{a},\mathbf{x}}$, we have a bijection v from $\mathbf{a}$ to $\mathbf{x}$, $\mathbf{x}' = \mathbf{x}_{|Y'}$ and $\mathbf{x}'' = \mathbf{x}_{|Y''}$ for appropriate sets of variables $Y', Y'' \subseteq [\mathbf{x}]$, and there are $v' : \mathbf{x}_{|Y'} \mapsto \mathfrak{A}_i$ and $v'' : \mathbf{x}_{Y''} \mapsto \mathfrak{A}_j$ such that $v'(\mathbf{x}_{|Y'}) \sim_{\mathrm{GF},\tau} v(\mathbf{x}_{|Y'})$ and $v''(\mathbf{x}_{|Y''}) \sim_{\mathrm{GF},\tau} v(\mathbf{x}_{|Y''})$. Then $v'(\mathbf{x}_{|Y' \cap Y''}) \sim_{\mathrm{GF},\tau} v''(\mathbf{x}_{|Y' \cap Y''})$. We show that each $\Phi_{\mathbf{a},\mathbf{x}}$ is τ-uniform and GF(τ)-bisimulation saturated.

1) Every $\Phi_{\mathbf{a},\mathbf{x}} \in \mathcal{M}$ is τ-uniform: let $t(\mathbf{x}'), s(\mathbf{x}'') \in \Phi_{\mathbf{a},\mathbf{x}}$ be as above and assume that $Q(\mathbf{z})$ is a τ-guard with $[\mathbf{z}] \subseteq [\mathbf{x}'] \cap [\mathbf{x}'']$. Then $[\mathbf{z}] \subseteq Y' \cap Y''$ and so $Q(\vec{z}) \in t(\mathbf{x}')$ iff $Q(\mathbf{z}) \in s(\mathbf{x}'')$ since $v'(\mathbf{x}_{|Y' \cap Y''}) \sim_{\mathrm{GF},\tau} v''(\mathbf{x}_{|Y' \cap Y''})$, as required.
2) To show GF(τ)-bisimulation saturatedness let $\Phi_{\mathbf{a},\mathbf{x}} \in \mathcal{M}$ and $t(\mathbf{x}'), s(\mathbf{x}'') \in \Phi_{\mathbf{a},\mathbf{x}}$ be as above and let $R(\mathbf{y}) \in t(\mathbf{x}')$ with $[\mathbf{x}''] \subseteq [\mathbf{y}]$ be a strict τ-guard. We have $Y'' \subseteq [\mathbf{y}] \subseteq Y'$ and $v'(\mathbf{x}_{|Y''}) \sim_{\mathrm{GF},\tau} v''(\mathbf{x}_{|Y''})$. Let H be the GF(τ)-bisimulation witnessing this. By the definition of GF(τ)-bisimulations, there exists $p \in H$ with domain $v'(\mathbf{x}_{|[\mathbf{y}]})$ such that $p \circ v'_{|Y''} = v''$. Now we expand v'' to the domain $[\mathbf{y}]$ by setting $\hat{v} := p \circ v'_{|\mathbf{x}_{|[\mathbf{y}]}}$. Let $\mathbf{b}'$ be the image of $\mathbf{x}_{|[\mathbf{y}]}$ under $\hat{v}$. Then the type tp$(\hat{v} : \mathbf{x}_{|[\mathbf{y}]} \mapsto \mathbf{b}')$ is as required.

Finally we show that every $\Phi \in \mathcal{M}$ is existentially saturated in $\mathcal{M}$. Assume $\Phi_{\mathbf{a},\mathbf{x}}$ is given. Assume $\exists \mathbf{y}(R(\mathbf{x}', \mathbf{y}) \wedge \lambda(\mathbf{x}', \mathbf{y})) \in t(\mathbf{x}_{|Y}) = $ tp$(v' : \mathbf{x}_{|Y} \mapsto \mathbf{b})$ with $Y \subseteq [\mathbf{x}]$ and $\mathbf{b}$ w.l.o.g. in $\mathfrak{A}_1$. Then $\mathfrak{A}_1 \models_{v'} \exists \mathbf{y}(R(\mathbf{x}', \mathbf{y}) \wedge \lambda(\mathbf{x}', \mathbf{y}))$. Then we find an assignment v'' for the variables in $[\mathbf{x}'\mathbf{y}]$ which coincides with v' on $[\mathbf{x}']$ such that $\mathfrak{A}_1 \models_{v''} R(\mathbf{x}', \mathbf{y}) \wedge \lambda(\mathbf{x}', \mathbf{y})$. Take a tuple $\mathbf{c}$ of distinct elements with $[\mathbf{c}] = [v''(\mathbf{x}'\mathbf{y})]$ and a tuple $\mathbf{y}'$ of variables in $\{x_1, \ldots, x_{2n}\}$ such that $[\mathbf{x}'] = [\mathbf{x}] \cap [\mathbf{y}']$ and we have a bijection $\rho : \mathbf{y}' \mapsto \mathbf{c}$ which coincides with v' on $[\mathbf{x}']$. Then $\rho(\mathbf{y}'_{|[\mathbf{x}']}) \sim_{\mathrm{GF},\tau} v(\mathbf{x}_{|[\mathbf{x}']})$ and so $\Phi_{\mathbf{a},\mathbf{x}}$ and $\Phi_{\mathbf{c},\mathbf{y}'}$ are compatible and $\Phi_{\mathbf{c},\mathbf{y}'}$ is as required.

For the proof with $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ instead of $\Phi_{\mathbf{a}_1,\mathbf{x}_0}$ in $\mathcal{M}$ observe that $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ is τ-uniform and GF(τ)-bisimulation saturated as $\widehat{\Phi}_{\mathbf{a}_1,\mathbf{x}_0}$ behaves in exactly the same way as $\Phi_{\mathbf{a}_1,\mathbf{x}_0}$ regarding τ-guarded $Q(\mathbf{y})$. For the same reason all elements of $\mathcal{M}$ are still existentially saturated in $\mathcal{M}$. ❑

2ExpTime *Lower Bound*

Lemma 8. *M accepts the input w iff there exists models $\mathfrak{A}, \mathfrak{B}$ of φ and elements $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$ such that $\mathfrak{A}, a \sim_{GF,\tau} \mathfrak{B}, b$.*

Proof. ($\Rightarrow$) If M accepts w, there is a computation tree of M on w. We construct a single model $\mathfrak{A}$ of φ as follows. Let $\mathfrak{A}^*$ be the infinite tree-shaped structure that represents the computation tree of M on w as described above, that is, configurations are represented by sequences of 2^n elements linked by S. Moreover, all elements of a configuration are labeled with $B_\forall$, $B^1_\exists$, or $B^2_\exists$ depending on whether the configuration is universal or existential, and in the latter case the superscript indicates which choice has been made for the existential state. Finally, the first element of the first successor configuration of a universal configuration is labeled with Z. In particular, $\mathfrak{A}^*$ only interprets the symbols in τ non-empty. Now, we obtain structures $\mathfrak{A}_k$, $k < 2^n$ from $\mathfrak{A}^*$ by interpreting non-τ-symbols as follows:

- the entire domain of $\mathfrak{A}_k$ satisfies I;
- the U-counter starts at 0 at the root and counts modulo 2^n along each S-path;

- the V-counter starts at k at the root and counts modulo 2^n along each S-path;
- the auxiliary concept names of the shape A^i_σ and A'_σ are interpreted in a minimal way so as to satisfy the sentences listed above. Note that the sentences are Horn, thus there is no choice.

Now obtain $\mathfrak{A}$ from $\mathfrak{A}^*$ and the $\mathfrak{A}_k$ as follows. First, create a both side infinite R-path

$$\dots b_{-2}Rb_{-1}Rb_0Rb_1Rb_2\dots$$

and realize the corresponding A-counter along the path and label every b_{3k}, $k \in \mathbb{Z}$, with X. Then, add all $\mathfrak{A}^*_k$ to every node b_{3k}, $k \in \mathbb{Z}$, on the path by identifying the roots of the $\mathfrak{A}_k$ with the respective node on the path. Moreover, add to $\mathfrak{A}$ three elements a_0, a_1, a_2 such that $(a_0, a_1, a_2) \in G^{\mathfrak{A}}$, $(a_0, a_1), (a_1, a_2), (a_2, a_0) \in R^{\mathfrak{A}}$, $a_0 \in X^{\mathfrak{A}}$, and $a_0 \in A^{\mathfrak{A}}$. Finally, add a copy of $\mathfrak{A}^*$ to $\mathfrak{A}$ by identifying the root of $\mathfrak{A}^*$ with a_0. We claim that $\mathfrak{A}$ is as required. In particular, $\mathfrak{A}, a_0$ is a model of $\varphi \wedge A(x)$, $\mathfrak{A}, b_0$ is a model of $\varphi \wedge \neg A(x)$, and the set S of all mappings

- $(a_i, a_{i+1}) \mapsto (b_{i+3k}, b_{i+3k+1})$ with $k \in \mathbb{Z}$, $i \in \{0, 1, 2\}$, and $a_3 := a_0$,
- $(e, f) \mapsto (e', f')$ with $(e, f) \in S^{\mathfrak{B}}$ and e', f' copies of e, f in some $\mathfrak{A}_k$, and
- all restrictions of the above,

is a GF(τ)-bisimulation on $\mathfrak{A}$ with $a_0 \mapsto b_0 \in S$.

($\Leftarrow$) Let $\mathfrak{A}, \mathfrak{B}$ be a models of φ such that $\mathfrak{A}, a \sim_{\text{GF},\tau} \mathfrak{B}, b$ for some elements a, b with $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$. As it was argued above, due to the three-element R-loop enforced at a via φ_0, from b there has to be an outgoing infinite R-path on which all S-trees are guarded τ-bisimilar. (There is also an incoming infinite R-path with this property, but it is not relevant for the proof.) All those S-trees are additionally labeled with some auxiliary relation symbols not in τ, depending on the distance from b. However, it can be shown using the arguments that accompanied the construction of φ that all S-trees contain a computation tree of M on input w. Hence, M accepts w. ❑

3EXPTIME *Lower Bound*

We show how to axiomatize the predicate E as announced in the main part, that is, for pairs a, a' and b, b', where b, b' represents a successor node of a, a', and for all $\mathbf{a}, \mathbf{a}' \in \{a, a'\}^n$ and $\mathbf{b}, \mathbf{b}' \in \{b, b'\}^n$, we have

$$E(\mathbf{a}\mathbf{a}'aa'\mathbf{b}\mathbf{b}'bb') \quad \text{iff} \quad r(\mathbf{a}) = r(\mathbf{b}) \text{ and } r(\mathbf{a}') = r(\mathbf{b}'). \quad (1)$$

We abbreviate the tuples $\mathbf{x}\mathbf{x}'$ and $\mathbf{y}\mathbf{y}'$ with $\mathbf{u}$ and $\mathbf{v}$, respectively; thus $\mathbf{u} = u_0 \dots u_{2n-1}$ and $\mathbf{v} = v_0 \dots v_{2n-1}$ are tuples of length $2n$. Moreover, let Σ be the set of all substitutions $[u_i/x, v_i/y]$ and $[u_i/x', v_i/y']$, for all $i < 2n$. Now, add the following sentences:

$$\forall xx'yy'\big(R(xx'yy') \to E(x^{2n}xx'y^{2n}yy')\big)$$
$$\forall \mathbf{u}xx'\mathbf{v}yy'\big(E(\mathbf{u}xx'\mathbf{v}yy') \to \bigwedge_{\sigma \in \Sigma} E(\sigma(\mathbf{u})xx'\sigma(\mathbf{v})yy')\big)$$
$$\forall \mathbf{u}xx'\mathbf{v}yy'\big(E(\mathbf{u}xx'\mathbf{v}yy') \to \bigwedge_{i<2n} (u_i = x \wedge v_i = y) \vee (u_i = x' \wedge v_i = y')\big)$$

These sentences axiomatize E as required, since the last sentence enforces "only if" of Property (1) while the first and second sentence together enforce "if".

We finish noting that $\text{Min}_U(xx')$ can be expressed by the formula

$$\text{Min}_U(xx') = \forall \mathbf{x}\big(D_U(\mathbf{x}) \to \bigvee_{i<n} (x_i \neq x \wedge x_i \neq x')\big).$$

PROOFS FOR SECTION VI

Proofs for the Upper Bound

Lemma 10. *Joint FO$^2(\tau)$-consistency can be reduced in polynomial time to joint FO$^2(\tau)$-consistency for formulas using relation symbols of arity at most two.*

Proof. We show that the construction given in the finite model property proof in [19] also works for joint FO$^2(\tau)$-consistency.

Consider FO$^2(\tau)$-formulas φ and ψ. We may assume that $\text{sig}(\varphi) \cap \text{sig}(\psi) = \tau$. For any relation symbol R of arity at least three that occurs in φ or ψ we do the following: for any atomic formula $R(v_1, \dots, v_n)$ that occurs in φ or ψ introduce a fresh relation symbol $R^{v_1,\dots,v_n}$ of arity two if both x and y occur in $v_1, \dots, v_n$ and of arity one otherwise.

If both x and y occur in $v_1, \dots, v_n$, then replace in φ and ψ every occurrence of $R(v_1, \dots, v_n)$ in φ, ψ by $R^{v_1,\dots,v_n}(x, y)$. If only x occurs in $v_1, \dots, v_n$, then replace $R(v_1, \dots, v_n)$ by $R^{v_1,\dots,v_n}(x)$ and if only y occurs in $v_1, \dots, v_n$ then replace $R(v_1, \dots, v_n)$ by $R^{v_1,\dots,v_n}(y)$. Let φ' and ψ' be the resulting formulas.

It remains to capture the logical relationships between different formulas $R(v_1, \dots, v_n)$ and $R(v'_1, \dots, v'_n)$ using implications between the fresh atomic formulas. For example, if $R(v_1, \dots, v_n)$ and $R(v'_1, \dots, v'_n)$ are both subformulas of φ or ψ and $R(v'_1, \dots, v'_n)$ is obtained from $R(v_1, \dots, v_n)$ by replacing x by y and y by x, then we take the implication

$$\forall x \forall y (R^{v_1,\dots,v_n}(x, y) \leftrightarrow R^{v'_1,\dots,v'_n}(y, x))$$

We also take for any R in φ or ψ and any two distinct $R(v_1, \dots, v_n)$ and $R(v'_1, \dots, v'_n)$ occurring in φ or ψ the implication:

$$\forall x (R^{v_1,\dots,v_n}(x, x) \leftrightarrow R^{v'_1,\dots,v'_n}(x, x))$$

Let χ_R be the conjunction of all these implications between the fresh atomic formulas. Now let

$$\varphi^\dagger = \varphi' \wedge \bigwedge_{R \text{ occurs in } \varphi} \chi_R, \quad \psi^\dagger = \psi' \wedge \bigwedge_{R \text{ occurs in } \psi} \chi_R$$

and let τ' contain all relation symbols of arity at most two in τ and all fresh $R^{v_1,\dots,v_n}$ for $R \in \tau$.

We show that φ and ψ are jointly FO$^2(\tau)$-consistent iff $\varphi^\dagger$ and $\psi^\dagger$ are jointly FO$^2(\tau')$-consistent.

Assume $\mathfrak{A}, \mathbf{a} \sim_{\mathrm{FO}^2,\tau} \mathfrak{B}, \mathbf{b}$, $\mathfrak{A} \models \varphi(\mathbf{a})$, and $\mathfrak{B} \models \psi(\mathbf{b})$. Define the structure $\mathfrak{A}'$ in the same way as $\mathfrak{A}$ except that for relation symbols R of arity ≥ 3:

- $(a,b) \in (R^{v_1,\dots,v_n})^{\mathfrak{A}'}$ if $\mathfrak{A} \models_v R(v_1,\dots,v_n)$ for $v(x) = a$ and $v(y) = b$, if x and y occur in $v_1,\dots,v_n$ and $R^{v_1,\dots,v_n}$ occurs in φ or ψ.
- $a \in (R^{v_1,\dots,v_n})^{\mathfrak{A}'}$ if $\mathfrak{A} \models_v R(v_1,\dots,v_n)$ for $v(x) = a$, if only x occurs in $v_1,\dots,v_n$ and $R^{v_1,\dots,v_n}$ occurs in φ or ψ.
- $a \in (R^{v_1,\dots,v_n})^{\mathfrak{A}'}$ if $\mathfrak{A} \models_v R(v_1,\dots,v_n)$ for $v(y) = a$, if only y occurs in $v_1,\dots,v_n$ and $R^{v_1,\dots,v_n}$ occurs in φ or ψ.

$\mathfrak{B}'$ is defined in the same way using $\mathfrak{B}$. It is readily checked that $\mathfrak{A}', \mathbf{a} \sim_{\mathrm{FO}^2,\tau'} \mathfrak{B}', \mathbf{b}$, $\mathfrak{A}' \models \varphi^\dagger(\mathbf{a})$, and $\mathfrak{B}' \models \psi^\dagger(\mathbf{b})$.

Conversely, assume $\mathfrak{A}, \mathbf{a} \sim_{\mathrm{FO}^2,\tau'} \mathfrak{B}, \mathbf{b}$, $\mathfrak{A} \models \varphi^\dagger(\mathbf{a})$, and $\mathfrak{B} \models \psi^\dagger(\mathbf{b})$. We define the structure $\mathfrak{A}'$ in the same way as $\mathfrak{A}$ except that for relation symbols R of arity ≥ 3:

- $(v(v_1),\dots,v(v_n)) \in R^{\mathfrak{A}'}$ if $R(v_1,\dots,v_n)$ occurs in φ or ψ such that for the assignment v it holds that $\mathfrak{A} \models_v R^{v_1,\dots,v_n}(x,y)$ (or, if only x or only y occur in $v_1,\dots,v_n$, $\mathfrak{A} \models_v R^{v_1,\dots,v_n}(x)$ or $\mathfrak{A} \models_v R^{v_1,\dots,v_n}(y)$ respectively).
- no other tuples are in $R^{\mathfrak{A}'}$.

$\mathfrak{B}'$ is defined in the same way using $\mathfrak{B}$. Using the conjuncts χ_R it can be shown that $\mathfrak{A}', \mathbf{a} \sim_{\mathrm{FO}^2,\tau} \mathfrak{B}', \mathbf{b}$ and $\mathfrak{A}' \models \varphi(\mathbf{a})$ and $\mathfrak{B}' \models \psi(\mathbf{b})$. ❑

We first show that one can achieve Condition (P2) for links between semi i-pawns. Recall that for any king mosaics m_1, m_2 and i-pawns $t_1 \in m_1$ and $t_2 \in m_2$

$$\begin{aligned} L_i((t_1,m_1),(t_2,m_2)) \quad = \quad & \{l_{\mathfrak{A}_i}(d_1,d_2) \mid d_1 \neq d_2, \\ & t_1 = t_{\mathfrak{A}_i}(d_1), m_1 = m(d_1), \\ & t_2 = t_{\mathfrak{A}_i}(d_2), m_2 = m(d_2)\}. \end{aligned}$$

We aim to define link-types $l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2))$, where t_1' and t_2' are copies of t_1 and t_2, in such a way that the following holds:

(P2) If t_1' is a copy of t_1 and l a link-type, then there exists a copy t_2' of t_2 such that $l = l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2))$ iff $l \in L_i((t_1,m_1),(t_2,m_2))$.

In the construction, we use the fact that there are $\geq 4 \times 2^{2s}$ many copies of any i-pawn and that the number of link types does not exceed 2^{2s}. Assume first that $(t_1,m_1) \neq (t_2,m_2)$. Then partition, for $j = 1,2$, the set $\{(t_j,1),\dots,(t_j,k_1)\}$ of copies of t_j into two sets M_1^j, M_2^j such that $|M_1^j|, |M_2^j| \geq 2^{2s}$. Now we define the link types between any pair (t_1',m_1) and (t_2',m_2) as follows

- for every $t_1' \in M_1^1$ do the following: take for any link type $l \in L_i((t_1,m_1),(t_2,m_2))$ some (t_2',m_2) with $t_2' \in M_1^2$ and set
$$l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2)) := l$$
There are sufficiently many (t_2',m_2) with $t_2' \in M_1^2$ since $|M_1^2| \geq 2^{2s}$.
- for every $t_1' \in M_2^1$ do the following: take for any link type $l \in L_i((t_1,m_1),(t_2,m_2))$ some (t_2',m_2) with $t_2' \in M_2^2$ and set
$$l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2)) := l$$
There are sufficiently many (t_2',m_2) with $t_2' \in M_2^2$ since $|M_2^2| \geq 2^{2s}$.
- for every $t_2' \in M_1^2$ do the following: take for any link type $l \in L_i((t_1,m_1),(t_2,m_2))$ some (t_1',m_1) with $t_1' \in M_2^1$ and set
$$l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2)) := l$$
There are sufficiently many (t_1',m_1) with $t_1' \in M_2^1$ since $|M_2^1| \geq 2^{2s}$.
- for every $t_2' \in M_2^2$ do the following: take for any link type $l \in L_i((t_1,m_1),(t_2,m_2))$ some (t_1',m_1) with $t_1' \in M_1^1$ and set
$$l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2)) := l$$
There are sufficiently many (t_1',m_1) with $t_1' \in M_1^1$ since $|M_1^1| \geq 2^{2s}$.

For semi i-pawns $(t_1',m_1),(t_2',m_2)$ that have not yet been connected by any of the four steps above, choose an arbitrary link type l from $L_i((t_1,m_1),(t_2,m_2))$ and set $l_{\mathfrak{B}_i}((t_1',m_1),(t_2',m_2)) := l$. It is readily checked that (P2) is satisfied.

Now assume that $(t_1,m_1) = (t_2,m_2)$. Then partition the set $\{(t_1,1),\dots,(t_1,k_1)\}$ of copies of t_1 into four sets M_1^j, M_2^j such that $|M_1^j|, |M_2^j| \geq 2^{2s}$, $j = 1,2$, and define $l_{\mathfrak{B}_i}(((t_1,k),m_1),((t_1,k'),m_1))$ in exactly the same way as above for $(t_1,k),(t_1,k') \in M_r^1 \times M_{r'}^2$, $r,r' \in \{1,2\}$. For any $((t_1,k),m_1),((t_1,k'),m_1))$ with $k \neq k'$ for which $l_{\mathfrak{B}_i}(((t_1,k),m_1),((t_1,k'),m_1))$ has not yet been defined choose an arbitrary link type l from $L_i((t_1,m_1),(t_2,m_2))$ and set $l_{\mathfrak{B}_i}(((t_1,k),m_1),((t_1,k'),m_2)) := l$. Then (P2) is satisfied.

We now show that Conditions (1) and (2) are satisfied, starting with Condition (1).

Lemma 11. *Let t' be a copy of t and m' a copy of m. For $i = 1,2$, all $(t',m') \in dom(\mathfrak{B}_i)$, the witness d of (t',m') in $\mathfrak{A}_i$, and all $\gamma(x) \in cl(\Xi)$:*

$$\mathfrak{B}_i \models \gamma(t',m') \quad \Leftrightarrow \quad \gamma(x) \in t_{\mathfrak{A}_i}(d) \quad \Leftrightarrow \quad \gamma(x) \in t.$$

Proof. The equivalence '$\gamma(x) \in t_{\mathfrak{A}_i}(d)$ iff $\gamma(x) \in t$' follows from the definition of witnesses d of (t',m'). We thus show the first equivalence. For $\gamma(x)$ of the form $R(x)$ or $R(x,x)$ the equivalence holds by definition. It thus suffices to show the first equivalence for existentially quantified $\gamma(x) = \exists y \beta(x,y)$ with $\beta(x,y)$ quantifier-free.

($\Rightarrow$) It suffices to observe that the following holds for all $(t_1',m_1') \in \mathrm{dom}(\mathfrak{B}_i)$: if $l = l_{\mathfrak{B}_i}((t_1',m_1'),(t_2',m_2'))$ for some $(t_2',m_2') \in \mathrm{dom}(\mathfrak{B}_i)$, then there exist d_1, d_2 with $m_j = m(d_j)$ and $t_j = t_{\mathfrak{A}_j}(d_j)$ for $j = 1,2$ such that $l = l_{\mathfrak{A}_i}(d_1,d_2)$.

($\Leftarrow$) We show the following

Claim 1. Let d_1 be the witness for (t'_1, m'_1). If $l = l_{\mathfrak{A}_i}(d_1, d_2)$ for some $d_2 \in \text{dom}(\mathfrak{A}_i)$, then there exists (t'_2, m'_2) such that $m_2 = m(d_2)$, $t_2 = t_{\mathfrak{A}_i}(d_2)$, and $l = l_{\mathfrak{B}_i}((t'_1, m'_1), (t'_2, m'_2))$.

For the proof of Claim 1 let d_1 be the witness for (t'_1, m'_1) and $l = l_{\mathfrak{A}_i}(d_1, d_2)$ for some $d_2 \in \text{dom}(\mathfrak{A}_i)$.

Case 1. $t_{\mathfrak{A}_i}(d_1)$ is an i-king in king mosaic $m = m(d_1)$.

If $t_{\mathfrak{A}_i}(d_2)$ is an i-king in king mosaic $m(d_2)$, then $(t_{\mathfrak{A}_i}(d_2), m(d_2))$ is as required.

If $t_{\mathfrak{A}_i}(d_2)$ is an i-pawn in king mosaic $m(d_2)$, then by (P1) there exists $(t'_2, m(d_2))$ such that $t_2 = t_{\mathfrak{A}_i}(d_2)$ and $l = l_{\mathfrak{B}_i}((t_{\mathfrak{A}_i}(d_1), m(d_1)), (t'_2, m(d_2))$. Then $(t'_2, m(d_2))$ is as required.

If $t_{\mathfrak{A}_i}(d_2)$ is an i-pawn in pawn mosaic $m(d_2)$, then by (M) and (P3) there exists a full i-pawn (t'_2, m') such that $t_2 = t_{\mathfrak{A}_i}(d_2)$ and $l = l_{\mathfrak{B}_i}((t_{\mathfrak{A}_i}(d_1), m(d_1)), (t'_2, m'))$, as required.

Case 2. (t'_1, m'_1) is a semi i-pawn. The claim follows from (P1) if $t_{\mathfrak{A}_i}(d_2)$ is an i-king in king mosaic $m(d_2)$. For $t_{\mathfrak{A}_i}(d_2)$ an i-pawn in king mosaic $m(d_2)$, the claim follows from (P2). For $t_{\mathfrak{A}_i}(d_2)$ an i-pawn in a pawn mosaic $m(d_2)$, the claim follows from (P4).

Case 3. (t'_1, m'_1) is a full i-pawn. The claim follows from (P3) if $t_{\mathfrak{A}_i}(d_2)$ is an i-king in king mosaic $m(d_2)$. For $t_{\mathfrak{A}_i}(d_2)$ an i-pawn in king mosaic $m(d_2)$, the claim follows from (P4). For $t_{\mathfrak{A}_i}(d_2)$ an i-pawn in a pawn mosaic $m(d_2)$, the claim follows from (P5). ❑

We now prove Condition (2). The *restriction* $l_{|\tau}$ of a link-type l to a signature τ is the set of all $R(x, y)$ and $R(y, x)$ in l with $R \in \tau$. Any such restriction is called a *τ-link*.

Note first that for all $(t'_1, m'), (t'_2, m') \in \text{dom}(\mathfrak{B}_1) \cup \text{dom}(\mathfrak{B}_2)$ there exists a generator g of m' and witnesses d_1 for (t'_1, m') and d_2 for (t'_2, m'). Thus $d_1 \sim_{\text{FO}^2,\tau} d_2$ and $t_i = d_{\mathfrak{A}_j}(d_i)$ for appropriate $j \in \{1, 2\}$. Thus $\chi(x) \in t_1$ iff $\chi(x) \in t_2$ for any formula $\chi(x)$ of the form $R(x)$ or $R(x, x)$ with $R \in \tau$. We obtain from Lemma 11 that (t'_1, m') and (t'_2, m') satisfy the same atomic formulas $R(x)$ and $R(x, x)$ with $R \in \tau$. To fully check Conditions (i) and (ii) for $\text{FO}^2(\tau)$-bisimulations, we introduce some notation. For $(t'_1, m'_1) \in \text{dom}(\mathfrak{B}_i)$, any copy m'_2 of a mosaic, and any τ-link h, we write

$$(t'_1, m'_1) \rightarrow_h m'_2$$

if there exists $(t'_2, m'_2) \in \text{dom}(\mathfrak{B}_i)$ such that

$$h = l_{\mathfrak{B}_i}((t'_1, m'_1), (t'_2, m'_2))_{|\tau}$$

Then it suffices to show the following

Lemma 12. *For $(t'_1, m'_1), (t'_2, m'_1) \in$ dom$(\mathfrak{B}_1) \cup$ dom$(\mathfrak{B}_2)$, any copy m'_2 of a mosaic, and any τ-link h:*

$$(t'_1, m'_1) \rightarrow_h m'_2 \quad \Leftrightarrow \quad (t'_2, m'_1) \rightarrow_h m'_2$$

Proof. We make a case distinction and first assume that m_1 or m_2 is a king mosaic. By construction, then $(t'_1, m'_1) \rightarrow_h m'_2$ implies that for the generators g_1, g_2 of m'_1, m'_2 there are $d_1 \sim_{\text{FO}^2,\tau} g_1$ and $d_2 \sim_{\text{FO}^2,\tau} g_2$ such that $h = l_{\mathfrak{A}_i}(d_1, d_2)_{|\tau}$, for appropriate i. But then, by the definition of mosaics and $\text{FO}^2(\tau)$-bisimulations, for all $d_1 \sim_{\text{FO}^2,\tau} g_1$ there exists $d_2 \sim_{\text{FO}^2,\tau} g_2$ such that $h = l_{\mathfrak{A}_i}(d_1, d_2)_{|\tau}$, for appropriate i. But then $(t'_2, m'_1) \rightarrow_h m'_2$ for all (t'_2, m'_1).

Now assume that m_1 and m_2 are both pawn mosaics. By construction, then $(t'_1, m'_1) \rightarrow_h m'_2$ iff there are d_1, d_2 with $m_1 = m(d_1)$ and $m_2 = m(d_2)$ such that $t_1 = t_{\mathfrak{A}_i}(d_1)$ and $h = l_{\mathfrak{A}_i}(d_1, d_2)_{|\tau}$, for appropriate i. But by the definition of mosaics and $\text{FO}^2(\tau)$-bisimulations, the latter is the case iff there are d_1, d_2 with $m_1 = m(d_1)$ and $m_2 = m(d_2)$ such that $h = l_{\mathfrak{A}_i}(d_1, d_2)_{|\tau}$, for appropriate i. This condition does not depend on t'_1 and so $(t'_2, m'_1) \rightarrow_h m'_2$ follows. ❑

Lemma 9. *M accepts the input w iff there exists models $\mathfrak{A}, \mathfrak{B}$ of φ' and elements $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$ such that $\mathfrak{A}, a \sim_{FO^2,\tau} \mathfrak{B}, b$.*

Proof. The proof is essentially the same as the proof for Lemma 8; we give it here for the sake of completeness.

($\Rightarrow$) If M accepts w, there is a computation tree of M on w. We construct a single model $\mathfrak{A}$ of φ' as follows. Let $\mathfrak{A}^*$ be the infinite tree-shaped structure that represents the computation tree of M on w as described above, that is, configurations are represented by sequences of 2^n elements linked by S. Moreover, all elements of a configuration are labeled with $B_\forall$, $B^1_\exists$, or $B^2_\exists$ depending on whether the configuration is universal or existential, and in the latter case the superscript indicates which choice has been made for the existential state. Finally, the first element of the first successor configuration of a universal configuration is labeled with Z. In particular, $\mathfrak{A}^*$ only interprets the symbols in τ non-empty. Now, we obtain structures $\mathfrak{A}_k$, $k < 2^n$ from $\mathfrak{A}^*$ by interpreting non-τ-symbols as follows:

- the entire domain of $\mathfrak{A}_k$ satisfies I;
- the U-counter starts at 0 at the root and counts modulo 2^n along each S-path;
- the V-counter starts at k at the root and counts modulo 2^n along each S-path;
- the auxiliary concept names of the shape A^i_σ and A'_σ are interpreted in a minimal way so as to satisfy the sentences listed above. Note that the sentences are Horn, thus there is no choice.

Now obtain $\mathfrak{A}$ from $\mathfrak{A}^*$ and the $\mathfrak{A}_k$ as follows. First, create a both side infinite R-path

$$\ldots b_{-2} R b_{-1} R b_0 R b_1 R b_2 \ldots$$

and realize the corresponding A-counter along the path and label every b_{3k}, $k \in \mathbb{Z}$, with X. Then, add all $\mathfrak{A}^*_k$ to every node b_{3k}, $k \in \mathbb{Z}$, on the path by identifying the roots of the $\mathfrak{A}_k$ with the respective node on the path. Moreover, add to $\mathfrak{A}$ three elements a_0, a_1, a_2 such that $(a_0, a_1), (a_1, a_2), (a_2, a_0) \in R^{\mathfrak{A}}$ $a_0 \in X^{\mathfrak{A}}$, $a_0 \in Y^{\mathfrak{A}}$, and $a_0 \in A^{\mathfrak{A}}$. Finally, add a copy of $\mathfrak{A}^*$ to $\mathfrak{A}$ by identifying the root of $\mathfrak{A}^*$ with a_0. We claim that $\mathfrak{A}$ is as required. In particular, $\mathfrak{A}, a_0$ is a model of $\varphi \wedge A(x)$, $\mathfrak{A}, b_0$ is a model of $\varphi \wedge \neg A(x)$, and the set S of all pairs

- (a_i, b_{i+3k}) with $k \in \mathbb{Z}$, $i \in \{0,1,2\}$, and
- (e, e') with e' copy of e in some $\mathfrak{A}_k$,

is an $\mathrm{FO}^2(\tau)$-bisimulation on $\mathfrak{A}$ with $(a_0, b_0) \in S$.

($\Leftarrow$) Let $\mathfrak{A}, \mathfrak{B}$ be a models of φ such that $\mathfrak{A}, a \sim_{\mathrm{FO}^2,\tau} \mathfrak{B}, b$ for some elements a, b with $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$. As it was argued above, due to the three-element R-loop enforced at a via φ_0', from b there has to be an outgoing infinite R-path on which every third element $\mathrm{FO}^2(\tau)$-bisimilar, and thus the S-trees starting at these elements are also $\mathrm{FO}^2(\tau)$-bisimilar. (There is also an incoming infinite R-path with this property, but it is not relevant for the proof.) All those S-trees are additionally labeled with some auxiliary relation symbols not in τ, depending on the distance from b. However, it can be shown using the arguments that accompanied the construction of φ' that all S-trees contain a computation tree of M on input w. Hence, M accepts w. ❏

Recall the definition of φ'' and τ' from the paper. Then it suffices to prove the following.

Lemma 13. *M accepts the input w iff there exists models $\mathfrak{A}, \mathfrak{B}$ of φ'' and elements $a \in A^{\mathfrak{A}}$, $b \notin A^{\mathfrak{B}}$ such that $\mathfrak{A}, a \sim_{FO^2,\tau'} \mathfrak{B}, b$.*

Proof. ($\Leftarrow$) is an immediate consequence of Lemma 9.

($\Rightarrow$) We expand the model $\mathfrak{A}$ constructed in the proof of Lemma 9 and obtain a model $\mathfrak{A}'$ of φ'' such that $a_0 \in A^{\mathfrak{A}'}$, $b_0 \notin A^{\mathfrak{A}'}$ and $\mathfrak{A}', a_0 \sim_{\mathrm{FO}^2,\tau'} \mathfrak{A}', b_0$.

In detail, to define $\mathfrak{A}'$ we keep $\mathfrak{A}$ but attach to every X-element d of $\mathfrak{A}$ an N^{-1} path of length 2 from d to a fresh node (d, E), for every $E \in \mathrm{sig}(\varphi') \setminus (\tau \cup \{A\})$. Now we proceed as follows for every $E \in \mathrm{sig}(\varphi') \setminus (\tau \cup \{A\})$: add $(d, (a_0, E))$ to $R_E^{\mathfrak{A}'}$ iff $d \in E^{\mathfrak{A}}$, for all $d \in \mathrm{dom}(\mathfrak{A})$. This ensures that $\mathfrak{A}' \models \chi_E(d)$ iff $d \in E^{\mathfrak{A}}$ for all such d. Let $\Delta_0, \Delta_1, \ldots$ be the maximal subsets of $\mathrm{dom}(\mathfrak{A})$ such that all elements of Δ_i are $\mathrm{FO}^2(\tau)$-bisimilar in $\mathfrak{A}$. We add additional pairs to $R_E^{\mathfrak{A}'}$ in such a way that all elements of any Δ_i are also $\mathrm{FO}^2(\tau')$-bisimilar in $\mathfrak{A}'$:

- if $\Delta_i \supseteq E^{\mathfrak{A}}$ then also add $(d, (d', E))$ to $R_E^{\mathfrak{A}'}$ for all X-elements d' and $d \in \Delta_i$;
- if $\Delta_i \cap E^{\mathfrak{A}} = \emptyset$, then we do not add any $(d, (d', E))$ to $R_E^{\mathfrak{A}'}$, for X-elements d' and $d \in \Delta_i$;
- otherwise we make sure that for every X-element d' there exist both $d \in \Delta_i$ with $(d, (d', E)) \in R_E^{\mathfrak{A}'}$ and $e \in \Delta_i$ with $(e, (d', E)) \notin R_E^{\mathfrak{A}}$ and we make sure for every $d \in \Delta_i$ there exist both an X-element d' with $(d, (d', E)) \in R_E^{\mathfrak{A}'}$ and an X-element e' with $(d, (e', E)) \notin R_E^{\mathfrak{A}'}$. This is easily achieved without adding any additional pairs of the form $(d, (a_0, E))$ to $R_E^{\mathfrak{A}'}$.

This finishes the definition of $\mathfrak{A}'$. It is easy to see that $\mathfrak{A}'$ is as required. ❏