Effective Interpolation and Preservation in Guarded Logics

MICHAEL BENEDIKT, University of Oxford
BALDER TEN CATE, LogicBlox and UC Santa Cruz
MICHAEL VANDEN BOOM, University of Oxford

Desirable properties of a logic include decidability, and a model theory that inherits properties of first-order logic, such as interpolation and preservation theorems. It is known that the Guarded Fragment (GF) of first-order logic is decidable and satisfies some preservation properties from first-order model theory; however, it fails to have Craig interpolation. The Guarded Negation Fragment (GNF), a recently defined extension, is known to be decidable and to have Craig interpolation. Here we give the first results on effective interpolation for extensions of GF. We provide an interpolation procedure for GNF whose complexity matches the doubly exponential upper bound for satisfiability of GNF. We show that the same construction gives not only Craig interpolation, but Lyndon interpolation and relativized interpolation, which can be used to provide effective proofs of some preservation theorems. We provide upper bounds on the size of GNF interpolants for both GNF and GF input, and complement this with matching lower bounds.

CCS Concepts: • **Theory of computation** → **Logic**

Additional Key Words and Phrases: Guarded negation fragment, guarded fragment, mosaic method

1. INTRODUCTION

Basic modal logic is known to have many attractive properties, both computationally and in its model theory. Given the limited expressiveness of basic modal logic, logicians have searched for more powerful languages that extend it while preserving its positive features.

Andréka et al. ([1998], Section 4.7) list "metaproperties" of modal logic that one would desire in an extension, including the following:

—The logic should have decidable satisfiability and validity problems.
—The logic should satisfy *interpolation theorems*, like the *Craig Interpolation Property* (CIP): if φ and ψ are in the logic, and φ entails ψ, then there should be a formula θ in the logic such that φ entails θ and θ entails ψ, and θ mentions only relations common to both φ and ψ.

—The logic should satisfy *preservation theorems* such as the Łoś-Tarski theorem: if a sentence in the logic is preserved under superstructures, it should be equivalent to an existential sentence in the logic.
—The logic should have the *finite model property*: a satisfiable sentence should have a finite model.

Interpolation, in particular, is a desirable property since it implies the well-known *Beth Definability Property* (BDP), as well as its extension, the *Projective Beth Definability Property* (PBDP), stating that implicit specifications can be converted to explicit ones. This conversion from implicit to explicit is important in knowledge representation [ten Cate et al. 2011] and databases [Nash et al. 2010]. Some forms of interpolation can be used to prove preservation theorems like Łoś-Tarski [Otto 2000] which state that formulas satisfying certain semantic properties can be converted to specific forms. Interpolation has many other applications, both in simplifying definitions and in verification [McMillan 2004].

Andréka et al. proposed the *Guarded Fragment* (GF) as a candidate for an attractive logic. They showed that it is decidable, that it has the finite model property, and that it satisfies Łoś-Tarski. Unfortunately Hoogland et al. [1999] showed that GF does not satisfy CIP. Later work [Bárány et al. 2013b] showed that PBDP fails for GF.

The *Guarded Negation Fragment* (GNF) is an extension of GF introduced by Bárány et al. [2011]. It also extends the *Unary Negation Fragment* (UNF) introduced by ten Cate and Segoufin [2011]. Both UNF and GNF were proven decidable and shown to have the finite model property [ten Cate and Segoufin 2011; Bárány et al. 2011]. They were also shown to satisfy many of the properties mentioned earlier: in ten Cate and Segoufin [2011], UNF was shown to have CIP and PBDP; in Bárány et al. [2013b], GNF was shown to have CIP, PBDP, and several preservation properties including the analog of Łoś-Tarski. The combination of interpolation and the finite model property for these logics means that interpolation holds when validity is considered over finite models only. This is attractive for applications of these logics in databases, such as rewriting queries using views [Nash et al. 2010], since in databases typically only finite models are considered.

In this work, we provide further evidence that GNF and UNF are attractive extensions of modal logic. We show that these logics satisfy the *Lyndon Interpolation Property* (LIP) and its generalization, the *Relativized Interpolation Property* (RIP). Moreover, we present the first *constructive* procedures for GNF and UNF interpolation, including CIP, LIP, and RIP. Roughly speaking, the LIP says that if φ entails ψ, we can find θ as in CIP, but with a relation symbol occurring positively in θ if and only if it occurs positively in both φ and ψ. The RIP says that the quantification pattern in θ must match the quantification pattern in φ and ψ.

We proceed by extending the *mosaic method* that has been applied in providing decidability for algebras of relations [Németi 1986], modal logics [Mikulás 1998], and temporal logics [Marx et al. 2000]. We present a variant of the method that gives a sound and complete decision procedure for the validity problem for GNF, but also allows interpolants to be read off from a proof of an entailment.

We show how this algorithm gives interpolants that witness the RIP and hence the LIP and CIP for GNF and UNF. Since the methods of Bárány et al. [2013b] and ten Cate and Segoufin [2011] were model theoretic, this gives the first constructive proof of interpolation for GNF and UNF. We deduce bounds on the size of interpolants for GNF, and also on the size of GNF interpolants for GF sentences. In general, our algorithm produces a GNF interpolant (represented as a directed acyclic graph) of size at most doubly exponential in the size of the original validity.

We complement the upper bounds with lower bounds for the size of GNF interpolants for GF and UNF. In particular, we show that the doubly exponential bound on the size of the interpolant produced by our algorithm is optimal.

Finally, we revisit the connection between interpolation and preservation theorems from Otto [2000], and we use the RIP to prove that GNF and GF admit effective preservation theorems, including analogs of the Łoś-Tarski theorem. To the best of our knowledge, these are the first effective preservation theorems, even for GF.

This article is an extended version of Benedikt et al. [2014], providing detailed proofs of the results described there.

Organization. Guarded logics and the mosaic method are described in Section 2. In Section 3, we define mosaics for GF sentences, and show (without proof) how to construct GNF interpolants starting from GF sentences. Section 4 is the main technical section, describing the mosaic method for relativized GNF formulas, and proving effective RIP for GNF using these mosaics. In Section 5, lower bounds for the interpolation results are presented. Finally, in Section 6, we describe how effective RIP can be used to obtain effective preservation theorems for GNF and GF.

2. PRELIMINARIES

2.1. Notation and Conventions

We use $\boldsymbol{x}$, $\boldsymbol{y}$, ... to denote vectors of variables. For a formula ϕ, we write $\phi(\boldsymbol{x})$ to indicate that the free variables in ϕ are contained in $\boldsymbol{x}$, and write free(ϕ) for the actual free variables. The formula ϕ may also use constants con(ϕ). We use α, β, ... to denote atomic formulas, and for such formulas if we write $\alpha(\boldsymbol{x})$, then we assume that free(α) = $\boldsymbol{x}$.

We define occ(ϕ) to be the set describing the *polarity of occurrences of atomic relations* (not including the equality relation). A relation occurs positively (negatively, respectively) in a formula if it is within the scope of an even (odd, respectively) number of negations. For every atomic relation S in ϕ except the equality relation, $(S, +) \in$ occ(ϕ) (($S, -) \in$ occ(ϕ), respectively) if and only if S occurs positively (negatively, respectively) in ϕ.

Size and Width of Formulas. A formula ϕ given as input (to interpolation, preservation, etc.) is assumed to be in the standard tree representation of a formula, and the size of this representation is the number of symbols in ϕ.

For the output, however, we will usually represent the formula and measure its size as a node labeled Directed Acyclic Graph (DAG). The nodes represent formulas, and the edge relation connects a formula to its subformulas. A DAG representation can be exponentially more succinct than its corresponding tree representation. We will also use DAGs with multiple roots to represent a set of formulas.

We often abuse notation by identifying a formula with a representation; given a representation of a formula ϕ, $|\phi|$ represents the size of the representation: the number of symbols if it is a formula, and the number of edges and labels if the representation is a DAG. Similarly, for a set S of formulas represented as a DAG, we write $|S|$ to denote the total number of nodes and edges in the DAG.

The *width* of ϕ, denoted width(ϕ), is the maximum number of free variables of any subformula of ϕ. For GF sentences, the width is bounded by the maximum arity of any relation.

Running Time of Algorithms. For brevity, throughout the article, we give only bounds on the output size, not the running time of the algorithms. However, the proofs will show that the worst-case running time is bounded by a polynomial in the output size.

Classical Query Languages. We will make use of some basic notions of database theory—in particular, the following "classical" query classes: Conjunctive Queries

(CQs), Unions of Conjunctive Queries (UCQs), First-Order logic formulas (FO), existential FO, and positive existential FO. Abiteboul et al. [1995] is a good reference for these languages.

2.2. The Basics of GNF and GF

GNF is built up according to the grammar

$$\phi ::= \top \mid \bot \mid R\,\boldsymbol{t} \mid \exists x.\phi \mid \phi \lor \phi \mid \phi \land \phi \mid \alpha(\boldsymbol{x}) \land \neg\phi(\boldsymbol{x}),$$

where R is either a relation symbol or the equality relation $x = y$, α is an atomic formula (including equality), and $\boldsymbol{t}$ is a tuple over variables and constants. Notice that any use of negation must occur conjoined with an atomic relation that contains all the free variables of the negated formula. Such an atomic relation is a *guard* of the formula. We write guards(ϕ) to denote the set of atoms that appear as guards of a negation in ϕ. We treat $\top$ and $\bot$ as symbols in the vocabulary rather than abbreviations for other formulas; this makes for a cleaner statement of the results later, but is not essential.

The purpose of allowing equalities as guards is to ensure that every formula with at most one free variable can always be guarded. Thus, GNF includes UNF, which is built up as previously, but allowing negation only on formulas with at most one free variable.

GNF should be compared to GF, which can be defined by

$$\phi ::= \top \mid \bot \mid R\,\boldsymbol{t} \mid \exists \boldsymbol{x}.\alpha(\boldsymbol{xy}) \land \phi(\boldsymbol{xy}) \mid \phi \lor \phi \mid \phi \land \phi \mid \neg\phi(\boldsymbol{x}),$$

where R is either a relation symbol or equality, α is an atomic formula (including equality), and $\boldsymbol{t}$ is a tuple over variables and constants. Here it is the quantification that is guarded, rather than negation. As in GNF, we allow equality guards by default.

Normal Form. It is often helpful to consider these formulas in a normal form. Formally, *weak GN-normal form* ϕ are formulas that can be generated by the grammar

$$\phi ::= \bigvee_i \exists \boldsymbol{x}_i . \bigwedge_j \psi_{ij},$$
$$\psi ::= \top \mid \bot \mid R\,\boldsymbol{t} \mid \alpha(\boldsymbol{x}) \land \phi(\boldsymbol{x}) \mid \alpha(\boldsymbol{x}) \land \neg\phi(\boldsymbol{x}),$$

where R is either a relation symbol or the equality relation, $\alpha(\boldsymbol{x})$ is a guard with free(α) = $\boldsymbol{x}$, and $\boldsymbol{t}$ is a vector over variables and constants. The grammar generates UCQ-shaped formulas, where each conjunct is either an atom, an answer-guarded UCQ-shaped formula, or a guarded negation of a UCQ-shaped formula.

The following lemma states that any GNF sentence can be converted to weak GN-normal form with an exponential blowup in overall size. The lemma also tracks a number of other features of the formula that are important later; for instance, that the conversion preserves the polarity of relations and the guards in the formula. The *CQ rank* is also important. Given a formula ϕ in weak GN-normal form, we define rank$_{\mathrm{CQ}}(\phi)$ to be the maximum number of conjuncts ψ_i of any CQ-shaped subformula $\exists \boldsymbol{x}. \bigwedge_i \psi_i$ of ϕ for nonempty $\boldsymbol{x}$. For the purposes of the CQ-rank, formulas $\alpha(\boldsymbol{x}) \land \phi(\boldsymbol{x})$ and $\alpha(\boldsymbol{x}) \land \neg\phi(\boldsymbol{x})$ are treated as single conjuncts in a CQ-shaped subformula.

LEMMA 2.1. *Let ϕ be a GNF formula. We can construct an equivalent ϕ' in* GN-*normal form such that*

—|ϕ'| *is exponential in* |ϕ|;
—width(ϕ') $\leq$ |ϕ| *and* rank$_{\mathrm{CQ}}(\phi') \leq$ |ϕ|;
—occ(ϕ') = occ(ϕ) *and* guards(ϕ') = guards(ϕ).

PROOF. Starting from the innermost formulas, we use the transformations described in Bárány et al. [2011]: $\exists x.(\delta \lor \psi) \rightsquigarrow (\exists x.\delta) \lor (\exists x.\psi)$, $\delta \land (\psi \lor \chi) \rightsquigarrow (\delta \land \psi) \lor (\delta \land \chi)$,

and $(\exists x.\delta) \wedge \psi \rightsquigarrow \exists x'.(\delta[x'/x] \wedge \psi)$, where $\phi[x'/x]$ is the result of substituting some fresh x' for x in ϕ. It is straightforward to check that this results in a formula in weak GN-normal form (in fact, it is in a stronger normal form called GN-normal form in Bárány et al. [2011]). The new formula ϕ' is of size at most $2^{f(m)}$ for $m = |\phi|$ and f some polynomial function, but has CQ-shaped subformulas with at most m conjuncts. Moreover, the number of free variable names needed in any subformula is at most m, so width(ϕ') $\leq m$. These transformation rules do not change the polarity of relations or the guards for negation, so occ(ϕ') = occ(ϕ) and guards(ϕ') = guards(ϕ). □

GF formulas with at most one free variable can be converted to weak GN-normal form with only a polynomial blowup. Note that this result does not hold for arbitrary GF formulas.

LEMMA 2.2. *Let ϕ be a* GF *formula with at most one free variable. We can construct an equivalent ϕ' in weak* GN-*normal form such that*

—$|\phi'|$ *is polynomial in* $|\phi|$;
—width(ϕ') = width(ϕ) *and* rank$_{CQ}$(ϕ') = 1;
—occ(ϕ') = occ(ϕ).

PROOF. The idea is that, for every subformula of ϕ, we find the innermost guard from a guarded quantification, and copy this guard to that subformula. The resulting sentence is in weak GN-normal form.

Formally, fix a signature σ, and let w be the maximum arity of any relation in σ. We prove a slightly stronger statement: for every GF formula $\phi(\boldsymbol{y}) := \alpha(\boldsymbol{y}) \wedge \psi(\boldsymbol{y})$ over signature σ, we can construct $\phi'(\boldsymbol{y}) := \alpha(\boldsymbol{y}) \wedge \psi'(\boldsymbol{y})$ in weak GN-normal form such that $|\psi'| \leq (w|\psi|)^2$, width(ψ') = width(ψ), rank$_{CQ}$(ψ') = 1, and occ(ψ') = occ(ψ). Since any GF formula with at most one free variable can be trivially guarded using an equality guard, this implies the desired result.

We proceed by structural induction on ψ in $\phi := \alpha \wedge \psi$. The base case and the inductive cases for disjunction and negation are trivial.

If ψ is $\psi_1 \wedge \psi_2$, then inductively, construct ϕ'_i for $\phi_i := \alpha \wedge \psi_i$, and set $\psi' := (\alpha \wedge \psi'_1) \wedge (\alpha \wedge \psi'_2)$. Although $(\alpha \wedge \psi'_1) \wedge (\alpha \wedge \psi'_2)$ is a CQ-shaped subformula with two conjuncts, it does not contribute to the CQ rank, since there is no quantification.

If ψ is $\exists \boldsymbol{x}.\beta(\boldsymbol{xy}) \wedge \eta(\boldsymbol{xy})$, then inductively construct η' for $\beta(\boldsymbol{xy}) \wedge \eta(\boldsymbol{xy})$ and set $\psi' := \exists \boldsymbol{x}.\beta(\boldsymbol{xy}) \wedge \eta'(\boldsymbol{xy})$. The outer CQ-shaped subformula in ψ' has rank 1, since β guards all of the variables in η, and any CQ-shaped subformulas in η' have CQ rank 1 by the inductive hypothesis. Hence, rank$_{CQ}$(ψ') = 1.

The other properties can be easily checked. □

Expressivity. Turning to fragments of FO that are common in database theory, it is easy to see that every union of conjunctive queries is expressible in GNF. One can also show that common classes of integrity constraints, such as inclusion dependencies (also known as referential constraints) are expressible in GNF, as well as many of the common dependencies used in data integration and exchange. Going further, the relational translations of many description logic languages used in the semantic web (e.g., $\mathcal{ALC}$ and $\mathcal{ALCHIO}$ [Baader et al. 2003]) are known to be expressible using GF formulas with one free variable, and hence are expressible in GNF.

2.3. Model-Theoretic Properties of GNF

Not only is GNF an expressive fragment of FO, but it was shown to be decidable and to have useful model-theoretic properties, similar to GF.

THEOREM 2.3 ([BÁRÁNY ET AL. 2011]). *Satisfiability for* GNF *formulas is* 2EXPTIME-*complete.* GNF *has the finite model property: a* GNF *formula is satisfiable if and only if it is satisfiable over finite models.* GNF *has the tree-like model property: a* GNF *formula is satisfiable if and only if it is satisfiable over models of bounded tree width.*

We explain the notion of tree-like models further, since this will be used later. A structure $\mathfrak{A}$ has *tree width* $k - 1$ if there is a tree T whose labels are sets $\boldsymbol{a} \subseteq \mathrm{dom}(\mathfrak{A})$ with $|\boldsymbol{a}| \leq k$, satisfying the following conditions:

—for every guarded set $\boldsymbol{a} \subseteq \mathrm{dom}(\mathfrak{A})$, there is some $v \in \mathrm{dom}(T)$ with $\boldsymbol{a} \subseteq T(v)$;
—for every $a \in \mathrm{dom}(\mathfrak{A})$, the set $\{v : a \in T(v)\}$ is a connected subtree of T.

In other words, the structure $\mathfrak{A}$ can be covered with possibly overlapping substructures of size at most k that can be arranged in a tree-like manner. The tree T together with the substructures represented at each node (induced by the elements represented at each node) is called a *tree decomposition* of $\mathfrak{A}$.

2.4. Interpolation

For two formulas φ_L and φ_R, we write $\varphi_L \models \varphi_R$ if every model of φ_L is a model of φ_R, and in this case say that $\varphi_L \models \varphi_R$ is a *validity*. An *interpolant* for a validity is a formula θ such that $\varphi_L \models \theta \models \varphi_R$ and θ mentions only relations present in both φ_L and φ_R. The Craig interpolation theorem [Craig 1957] states that if φ_L and φ_R are in FO, then there is an FO interpolant θ.

Lyndon proved a stronger interpolation property by considering the polarity of relations in the formulas. Recall that $\mathrm{occ}(\phi)$ is the set describing the *polarity of occurrences of atomic relations* (not including the equality relation): for every atomic relation S in ϕ (where S is not the equality relation), $(S, +) \in \mathrm{occ}(\phi)$ ($(S, -) \in \mathrm{occ}(\phi)$, respectively) if and only if S occurs positively (negatively, respectively) in ϕ. Lyndon's interpolation theorem [Lyndon 1959] states that the interpolant θ can be taken so that $\mathrm{occ}(\theta) \subseteq \mathrm{occ}(\varphi_L) \cap \mathrm{occ}(\varphi_R)$.

A more recent relativized interpolation theorem of Otto [2000] states that if φ_L and φ_R both use only quantification *relativized to a set of unary predicates* $\mathbb{U}$, then the same is true of the interpolant. A formula has quantification relativized to $\mathbb{U}$ if every quantification is of the form $\exists \boldsymbol{x}.U_1 x_1 \wedge \cdots \wedge U_k x_k \wedge \psi(\boldsymbol{xy})$ or $\forall \boldsymbol{x}.U_1 x_1 \wedge \cdots \wedge U_k x_k \rightarrow \psi(\boldsymbol{xy})$ where $\boldsymbol{x} = x_1 \ldots x_k$ and $U_i \in \mathbb{U}$ for $i \in \{1, \ldots, k\}$. For instance, if $\mathbb{U} = \{U\}$, then $\exists xyz.Ux \wedge Uy \wedge Uz \wedge Rxy \wedge Ryz$ is $\mathbb{U}$-relativized but $\exists xy.Ux \wedge Rxy$ is not. As shown in Otto [2000], the relativized interpolation theorem gives an alternative means of showing preservation theorems. We explain and expand upon this in Section 6.

Although there are proofs of Craig interpolation for FO that construct interpolants efficiently given proofs that witness the validity (e.g., Fitting [1996]), these do not give bounds on the size of interpolants relative to the size of φ_L and φ_R. In fact, it has been shown that for FO validities, there is no computable bound on the sizes of the interpolants relative to the sizes of the original validities [Friedman 1976].

We say that a logic has the CIP if given φ_L and φ_R in the logic, there is an interpolant that is again in the logic. Similarly, we can talk about the LIP and RIP. Hoogland et al. [1999] showed that the CIP fails for GF. Marx [2007] argued that one of the applications of CIP, the PBDP holds for GF as well as for an extension called the "Packed Fragment." He conjectured that the proof could be made effective using the "mosaic method." The argument in Marx [2007] was flawed, and indeed neither the guarded fragment nor the packed fragment has PBDP [Bárány et al. 2013b]. Nevertheless, here we use the mosaic technique suggested by Marx to show effective CIP, LIP, RIP, and PBDP for GNF (see Theorem 4.10).

2.5. The Mosaic Method

In its simplest form, a mosaic is just a finite set of formulas. When it satisfies certain coherence properties, it can be viewed as a collection of formulas describing part of a model. For instance, in the mosaic system for GF described in Section 4.6 of Andréka et al. [1998], a mosaic can describe the complete subformula type of some guarded set of elements in a model. For modal and guarded logics, it turns out that the existence of a (finite) set of mosaics satisfying certain coherence properties is enough to ensure the existence of a full model. This allows mosaics to be used to decide satisfiability.

We adapt this method to check whether $\varphi_L \wedge \neg\varphi_R$ is unsatisfiable, and if so, extract an interpolant for $\varphi_L \models \varphi_R$ from the mosaics. The purpose of this section is to describe the general framework we are using. We utilize this framework in later sections with different versions of mosaics depending, for instance, on whether the input formulas are in GF versus GNF, or whether we are aiming for Lyndon interpolation versus relativized interpolation. As a result, some of the terms in this section are only defined informally; the formal definitions appear in later sections since they depend on the particular mosaic variant that we are using.

We start by fixing some sentences φ_L and φ_R over signatures σ_L and σ_R, respectively. The first step is to define a finite set Γ of formulas that are relevant to checking satisfiability of $\varphi_L \wedge \neg\varphi_R$. This is usually some extension of the subformula closure of φ_L and φ_R. To be suitable for computing an interpolant for $\varphi_L \models \varphi_R$, we also annotate these formulas with a *provenance* $X \in \{L, R\}$ indicating that the formulas are related to φ_X. This parallels other constructive interpolation methods, for example, Fitting [1996]. Given a provenance X, we write $\tilde{X}$ for the *dual provenance* (i.e., $\tilde{L} = R$ and $\tilde{R} = L$). Throughout the article, we use X, Y, Z to range over provenances L and R.

Given some finite set of *parameters* c, we define $\Gamma(c)$ as the set of formulas from Γ with parameters from c substituted for any free variables. Parameters are just a special type of variable. We write $a, b, \ldots$ to denote vectors of parameters, and given a formula ϕ, we write $\phi(xb)$ to indicate that the free variables in ϕ are contained in x and the parameters in ϕ are contained in b. We write free(ϕ) (par(ϕ), respectively) for the actual free variables parameters, respectively). For atomic α, if we write $\alpha(xb)$, then we assume that free(α) = x and par(α) = b.

A *mosaic* τ over some finite set of parameters c (of size bounded by the width of the original sentences) is then defined to be a subset of formulas from $\Gamma(c)$. Each mosaic is annotated with L or R, and an X-mosaic has additional restrictions on the $\tilde{X}$-formulas appearing in it. This helps prove properties about the interpolants that we later extract from these mosaics.

Mosaics can be *linked* together. For instance, if a formula $Y : \exists x.\psi(xb)$ appears in an X-mosaic τ, then this can be linked to some Y-mosaic τ' using parameters ab and containing the formula $Y : \psi(ab)$ (for some new parameters a). In addition, the linked mosaic τ' must contain the maximum amount of information from τ about formulas using only shared parameters b (subject to any constraints on formulas in τ') and cannot contain any new information about formulas using only shared parameters. We say an existential requirement $Y : \exists x.\psi(xb)$ in τ is *fulfilled* if there is a formula $Y : \psi(ab)$ in τ, or there is some τ' such that τ and τ' are linked via $Y : \exists x.\psi(xb)$.

In order for a mosaic to correspond to a part of a model it must satisfy certain coherence properties. Within a single mosaic, this means that it must be *internally consistent*. For instance, we may want to require that if it contains a conjunction, then it contains both conjuncts. The exact definition will depend on the particular version of mosaics that we are considering. A set of mosaics is coherent if every mosaic is internally consistent and it is *saturated*: every existential requirement is fulfilled by a

mosaic in the set. We then fix some finite set $\mathcal{M}$ of mosaics that suffices to prove the following key *mosaic theorem*:

$\varphi_L \wedge \neg\varphi_R$ is satisfiable if and only if there is a saturated set of internally consistent mosaics from $\mathcal{M}$ containing a mosaic with φ_L and $\neg\varphi_R$.

In one direction, the idea is to extract a saturated set of internally consistent mosaics from a model. In the other (harder) direction, we must show that a saturated set of internally consistent mosaics from $\mathcal{M}$ can be used to construct a model of $\varphi_L \wedge \neg\varphi_R$.

This gives us a procedure for deciding whether or not $\varphi_L \wedge \neg\varphi_R$ is satisfiable. Consider the following *mosaic elimination procedure* over $\mathcal{M}$.

—At stage $i = 1$, every internally inconsistent mosaic is eliminated.
—At stage $i > 1$, a mosaic τ is eliminated if it has an existential requirement that is not fulfilled in τ, and every linked τ' fulfilling it was already eliminated.

This process continues until a fixpoint is reached. The mosaic theorem can then be used to show that $\varphi_L \wedge \neg\varphi_R$ is satisfiable if and only if there is a mosaic containing φ_L and $\neg\varphi_R$ that is never eliminated. For interpolation, we are interested in the case when $\varphi_L \wedge \neg\varphi_R$ is unsatisfiable, since this means $\varphi_L \models \varphi_R$.

It turns out that we can construct a *mosaic interpolant* θ_X^τ for each x-mosaic τ eliminated using the previous procedure. These mosaic interpolants satisfy the property that $\tau_X \models \theta_X^\tau$ and $\theta_X^\tau \models \neg\tau_{\tilde{X}}$, where τ_Y denotes the conjunction of all of the Y-formulas from $\Gamma(\boldsymbol{c})$ in τ and the negation of all of the Y-formulas from $\Gamma(\boldsymbol{c})$ that satisfy the constraints on formulas in τ but are not in τ.

We start by constructing these interpolants for mosaics eliminated because of internal inconsistencies. The interpolants for such mosaics are simple: atomic formulas or the negation of atomic formulas. Proceeding by induction on the stage at which the mosaic is eliminated, we then build up more complicated interpolants for the other mosaics. Intuitively, each mosaic interpolant describes the reason the mosaic was eliminated. Since the sequence of elimination steps for the mosaics provides a proof that $\varphi_L \wedge \neg\varphi_R$ is unsatisfiable, it is not surprising that the desired interpolant for $\varphi_L \models \varphi_R$ can then be constructed from these mosaic interpolants.

It is helpful to compare this to other constructive interpolation approaches. In Fitting [1996], an interpolant is constructed from a proof tree for $\varphi_L \wedge \neg\varphi_R$ in a tableau proof system for FO. The construction starts at the leaves (where very simple interpolants suffice), and builds up to an interpolant that holds for the original validity at the root of the proof tree. Because of the "one-step" nature of the proof rules, there are straightforward interpolation construction rules at each stage.

Using the mosaic method, the definition of the interpolant at the inductive stages is more complicated. Unlike a traditional proof system, however, the mosaics allow precise control over the parameters appearing at each stage. This is advantageous for constructing an interpolant in GNF, since we need to be able to guard certain parameters in order to remain in GNF. Indeed, this precise control over the parameters was one of the motivations for building GNF interpolants from mosaics rather than from a more traditional proof system.

3. GF MOSAIC METHOD

We start by describing mosaics for GF sentences φ_L and φ_R over signatures σ_L and σ_R, respectively. The purpose of this section is to give a more gentle introduction to mosaics, without the complications that come from handling GNF, relativized quantifiers, or equality. As a result, *we omit the proofs in this section, since they are similar to the proofs given for the GNF mosaics described in Section 4.*

We assume the sentences φ_L and $\neg\varphi_R$ are in GF and are in negation normal form (negations are pushed to the inside) and do not use equality. These sentences may use constants, and we let $e := \text{con}(\varphi_L) \cup \text{con}(\varphi_R)$.

3.1. Definitions

Closure. In order to handle GF satisfiability testing of $\varphi_L \wedge \neg\varphi_R$, we must consider the subformula closure of φ_L and $\neg\varphi_R$, with a slight extension for atoms. The *closure* $\text{cl}(\varphi)$ of a GF formula φ is the smallest set C of formulas containing φ such that

—if $\psi_1 \wedge \psi_2 \in C$ or $\psi_1 \vee \psi_2 \in C$, then $\psi_1, \psi_2 \in C$;
—if $\exists \boldsymbol{x}.\alpha(\boldsymbol{xy}) \wedge \psi(\boldsymbol{xy}) \in C$, then $\alpha(\boldsymbol{xy}), \psi(\boldsymbol{xy}) \in C$;
—if $\forall \boldsymbol{x}.\alpha(\boldsymbol{xy}) \rightarrow \psi(\boldsymbol{xy}) \in C$, then $\neg\alpha(\boldsymbol{xy}), \psi(\boldsymbol{xy}) \in C$;
—if $\alpha(\boldsymbol{xy}) \in C$, then $\exists \boldsymbol{x}.\alpha(\boldsymbol{xy}) \in C$.

Given parameters $\boldsymbol{c}$, $\text{cl}_L(\boldsymbol{c})$ ($\text{cl}_R(\boldsymbol{c})$, respectively) consists of formulas from $\text{cl}(\varphi_L)$ ($\text{cl}(\neg\varphi_R)$, respectively) with free variables replaced by parameters from $\boldsymbol{c}$ or constants from $\boldsymbol{e}$ and labeled with provenance L (R, respectively).

Observe that the formulas in $\text{cl}_L(\boldsymbol{c}) \cup \text{cl}_R(\boldsymbol{c})$ respect the polarity of relations in the original formulas. This is important for extracting Lyndon-style interpolants later, where the polarity of relations needs to be respected.

Guards. We say α is an L-*atom* (respectively, R-*atom*) if $(\alpha, +) \in \text{occ}(\varphi_L)$ ($(\alpha, -, r) \in \text{occ}(\varphi_R)$, respectively). Usually, any atom is a possible guard in a GF or GNF formula. For the purposes of interpolation, however, we need *atoms in the common signature* (*catoms* for short). We say α is an L-*catom* (R-*catom*, respectively) if $(\alpha, +) \in \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$ ($(\alpha, -) \in \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$, respectively).

Given a collection τ of formulas from $\text{cl}_L(\boldsymbol{c}) \cup \text{cl}_R(\boldsymbol{c})$, we say $\boldsymbol{b} \subseteq \boldsymbol{c}$ *is* x-*guarded in* τ if $\boldsymbol{b}$ is empty, or if there is some x : $\exists \boldsymbol{x}.\alpha(\boldsymbol{xb}) \in \tau$ where α is an x-atom; if α is actually an x-catom, then we say $\boldsymbol{b}$ is x-*cguarded*. We allow $\boldsymbol{x}$ to be empty, in which case the x-guard or x-cguard for $\boldsymbol{b}$ is actually of the form x : $\alpha(\boldsymbol{b})$. We say a formula $\widetilde{x} : \psi$ is x-*safe* in τ if $\text{par}(\psi)$ is x-cguarded in τ.

GF Mosaics. Each GF x-mosaic provides information about an x-guarded set of parameters. Unlike the mosaics in Andréka et al. [1998], we need to have further restrictions on the $\widetilde{x}$-formulas in an x-mosaic. Roughly speaking, for GF mosaics, the $\widetilde{x}$-formulas are the "dangerous" formulas that could potentially introduce a negation when we try to extract GNF interpolants from these mosaics later. As a result, we only allow safe $\widetilde{x}$-formulas in an x-mosaic.

Formally, a *GF* x-*mosaic* τ for $\varphi_L \wedge \neg\varphi_R$ using parameters $\boldsymbol{c}$ is a subset of $\Gamma(\boldsymbol{c}) :=$ $\text{cl}_L(\boldsymbol{c}) \cup \text{cl}_R(\boldsymbol{c})$ such that $\boldsymbol{c}$ is x-guarded in τ, and if $\widetilde{x} : \psi \in \tau$, then $\widetilde{x} : \psi$ is x-safe in τ. We sometimes write $\tau(\boldsymbol{c})$ if we want to emphasize that τ uses parameters $\boldsymbol{c}$.

Let τ be a GF x-mosaic over parameters $\boldsymbol{c}$. Assume Y : $\eta \in \tau$ is an *existential requirement*, that is, η is of the form $\exists \boldsymbol{x}.\psi(\boldsymbol{xb})$. We say η is *fulfilled* in τ if Y : $\psi(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}]$ is in τ for some $\boldsymbol{a} \subseteq \boldsymbol{c} \cup \boldsymbol{e}$. We say τ' is *linked* with τ via Y : η (written $\tau \rightarrow_{Y:\eta} \tau'$) if

(E1) τ' is a GF Y-mosaic over parameters $\boldsymbol{ab}$ with $\boldsymbol{a} \cap \boldsymbol{c} = \emptyset$;
(E2) Y : $\psi(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}] \in \tau'$;
(E3) for all Y : $\psi'(\boldsymbol{b}) \in \tau$, Y : $\psi'(\boldsymbol{b}) \in \tau'$;
(E4) for all $\widetilde{Y} : \psi'(\boldsymbol{b}) \in \tau$ such that $\widetilde{Y} : \psi'(\boldsymbol{b})$ is Y-safe in τ', $\widetilde{Y} : \psi'(\boldsymbol{b}) \in \tau'$;
(E5) for all Z : $\psi'(\boldsymbol{b}) \in \tau'$, Z : $\psi'(\boldsymbol{b}) \in \tau$.

These conditions ensure that the maximum amount of information about the shared parameters $\boldsymbol{b}$ is passed from τ to τ' (subject to the restrictions on $\widetilde{Y}$ formulas in the Y-mosaic τ'), and no new information about $\boldsymbol{b}$ is added in τ'.

Fix a set P of size $2 \cdot \text{width}(\varphi_L \wedge \neg\varphi_R)$, and let $\mathcal{M}$ be the set of GF mosaics over parameters P. The following bound on the number of mosaics is easy to show.

PROPOSITION 3.1. *There is a polynomial function p (independent of φ_L and φ_R) such that if $|\varphi_L| + |\varphi_R| = n$, then there are at most $2^{2^{p(n)}}$ mosaics in $\mathcal{M}$. Moreover, for a fixed bound on the arity of relations, there are at most $2^{p(n)}$ mosaics in $\mathcal{M}$.*

Coherence. We say a GF x-mosaic τ is *internally consistent* if there is no Y : $\bot \in \tau$, there is no atomic $\alpha \in \sigma_L \cup \sigma_R$ such that Y : α, Z : $\neg\alpha \in \tau$, and τ satisfies the following natural *downward closure* properties:

—if Y : $\alpha(\boldsymbol{ab}) \in \tau$, then Y : $\exists \boldsymbol{x}.\alpha(\boldsymbol{xb}) \in \tau$;
—if Y : $\exists \boldsymbol{x}.\alpha(\boldsymbol{xab}) \in \tau$, then Y : $\exists \boldsymbol{xy}.\alpha(\boldsymbol{xyb}) \in \tau$;
—if Y : $\psi_1 \wedge \psi_2 \in \tau$, then Y : ψ_1, Y : $\psi_2 \in \tau$;
—if Y : $\psi_1 \vee \psi_2 \in \tau$, then Y : $\psi_1 \in \tau$ or Y : $\psi_2 \in \tau$;
—if Y : $\forall \boldsymbol{x}.\alpha(\boldsymbol{xb}) \rightarrow \psi(\boldsymbol{xb}) \in \tau$ and there is $\boldsymbol{a} \subseteq \boldsymbol{c} \cup \boldsymbol{e}$ such that Z : $\alpha(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}] \in \tau$ for some Z $\in \{L, R\}$, then Y : $\psi(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}] \in \tau$.

A set of GF mosaics is said to be *saturated* if for every mosaic $\tau(\boldsymbol{c})$ in the set and for every existential requirement Y : $\eta \in \tau$, either Y : η is fulfilled in τ, or τ is linked to some mosaic τ' in the set via Y : η, so η is fulfilled in τ'.

3.2. Soundness and Completeness of GF Mosaic Method

Finding a saturated set of internally consistent mosaics containing the original sentences provides a sound and complete method for GF satisfiability testing.

THEOREM 3.2. $\varphi_L \wedge \neg\varphi_R$ *is satisfiable if and only if there is a saturated set of internally consistent mosaics from $\mathcal{M}$ that contains a mosaic τ such that $L : \varphi_L$, $R : \neg\varphi_R \in \tau$.*

The only if direction is straightforward, since a set of mosaics can easily be read off from a model. The completeness is proven using a standard method for constructing a model from a description of formulas. We omit the argument, since a more general construction will be given in the next section.

This means we can use the mosaic elimination procedure described in the previous section, and Theorem 3.2 ensures that $\varphi_L \wedge \neg\varphi_R$ is satisfiable if and only if there is some mosaic containing $L : \varphi_L$ and $R : \neg\varphi_R$ that is not eliminated using this process.

3.3. Constructive Interpolation for GF

Now assume $\varphi_L \models \varphi_R$, and consider the mosaic elimination procedure for the GF mosaics $\mathcal{M}$ for $\varphi_L \wedge \neg\varphi_R$. We write $\mathcal{N}'$ for the set of mosaics for $\varphi_L \wedge \neg\varphi_R$ that are eliminated using this procedure, and $\mathcal{N}_i$ for the set of mosaics that have been removed by stage i. By Theorem 3.2, we know that every mosaic containing $L : \varphi_L$ and $R : \neg\varphi_R$ must be in $\mathcal{N}'$.

We first construct a *mosaic interpolant* θ_X^τ for each x-mosaic $\tau(\boldsymbol{c})$ in $\mathcal{N}'$. Roughly speaking, θ_X^τ is an interpolant (in the usual sense) for $\tau_X \models \neg\tau_{\widetilde{X}}$, where

$$\tau_X := \bigwedge_{X:\psi \in \tau} \psi \ \wedge \bigwedge_{X:\psi \in \text{cl}_X(\boldsymbol{c}) \setminus \tau} \neg\psi \qquad \text{and} \qquad \tau_{\widetilde{X}} := \bigwedge_{\widetilde{X}:\psi \in \tau} \psi \ \wedge \bigwedge_{\substack{\widetilde{X}:\psi \in \text{cl}_{\widetilde{X}}(\boldsymbol{c}) \setminus \tau \\ \text{s.t. } \widetilde{X} : \psi \text{ is x-safe in } \tau}} \neg\psi \ .$$

That is, τ_Y asserts the Y-formulas in τ, and the negation of the Y-formulas that are not in τ (but satisfy the conditions for being in τ in terms of x-safety if Y = $\widetilde{X}$). Intuitively, θ_X^τ will describe the reason τ was eliminated, that is, it expresses why $\tau_X \wedge \tau_{\widetilde{X}}$ is not satisfiable.

We have the following Lyndon-style mosaic interpolant properties.

LEMMA 3.3. *For each* GF x*-mosaic* $\tau(\boldsymbol{c}) \in \mathcal{N}'$ *(the set of eliminated mosaics), we can construct a DAG representation of a formula* $\theta_{\mathrm{x}}^{\tau}$ *such that*

(Imp) $\tau_{\mathrm{x}} \models \theta_{\mathrm{x}}^{\tau}$ *and* $\theta_{\mathrm{x}}^{\tau} \models \neg \tau_{\tilde{\mathrm{x}}}$;
(Occ) $\mathrm{occ}(\theta_{\mathrm{x}}^{\tau}) \subseteq \mathrm{occ}(\varphi_{\mathrm{L}}) \cap \mathrm{occ}(\varphi_{\mathrm{R}})$ *if* $\mathrm{x} = \mathrm{L}$,
$\mathrm{occ}(\neg\theta_{\mathrm{x}}^{\tau}) \subseteq \mathrm{occ}(\varphi_{\mathrm{L}}) \cap \mathrm{occ}(\varphi_{\mathrm{R}})$ *if* $\mathrm{x} = \mathrm{R}$;
$\mathrm{con}(\theta_{\mathrm{x}}^{\tau}) \subseteq \mathrm{con}(\varphi_{\mathrm{L}}) \cup \mathrm{con}(\varphi_{\mathrm{R}})$;
(Par) $\mathrm{par}(\theta_{\mathrm{x}}^{\tau}) \subseteq \boldsymbol{c}$;
(GN) $\theta_{\mathrm{x}}^{\tau}$ *is in* GNF, *even when parameters are viewed as free variables.*

Moreover, if $n = |\varphi_{\mathrm{L}}| + |\varphi_{\mathrm{R}}|$, *then there is a DAG representation of* $\Theta_i :=$ $\{\theta_{\mathrm{z}}^{\tau} : \tau \text{ is a } \mathrm{z}\text{-mosaic in } \mathcal{N}_i\}$ *such that* $|\Theta_i| \leq i \cdot 2^{2^{p'(n)}}$ $(|\Theta_i| \leq i \cdot 2^{p'(n)}$ *for the fixed arity case), where* p' *is some polynomial function independent of* φ_{L} *and* φ_{R}.

PROOF SKETCH. We proceed by induction on i, the stage at which τ was eliminated. We define the interpolants at each stage, but omit the proof of correctness. During the interpolant construction, we will need to include guards in order to remain in GNF, so for any tuple $\boldsymbol{b}$ of parameters that are x-cguarded in τ, we define a formula $\mathrm{gdd}_{\mathrm{x}}^{\tau}(\boldsymbol{b})$. If $\boldsymbol{b} = \emptyset$, then set $\mathrm{gdd}_{\mathrm{x}}^{\tau}(\boldsymbol{b}) := \top$. Otherwise, there is some formula $\mathrm{x} : \exists \boldsymbol{x}.\beta(\boldsymbol{xb}) \in \tau$ for some x-catom β, and we set $\mathrm{gdd}_{\mathrm{x}}^{\tau}(\boldsymbol{b}) := \exists \boldsymbol{x}.\beta(\boldsymbol{xb})$ (note that $\boldsymbol{x}$ can be empty, in which case this is just an atomic formula). We will also write $\mathrm{gdd}_{\mathrm{x}}^{\tau}(\boldsymbol{b}) \wedge \neg\psi$ to indicate some formula $\exists \boldsymbol{x}.(\beta(\boldsymbol{xb}) \wedge \neg\psi)$.

Base Case. Consider an x-mosaic $\tau(\boldsymbol{c}) \in \mathcal{N}_1$. Then τ has an internal inconsistency, and we consider the following cases.

—Assume one of these conditions holds for $\mathrm{Y} \in \{\mathrm{L}, \mathrm{R}\}$:
   —$\mathrm{Y} : \bot \in \tau$;
   —$\mathrm{Y} : \alpha(\boldsymbol{b}), \mathrm{Y} : \neg\alpha(\boldsymbol{b}) \in \tau$;
   —$\mathrm{Y} : \alpha(\boldsymbol{ab}) \in \tau$, but $\mathrm{Y} : \exists \boldsymbol{x}.\alpha(\boldsymbol{xb}) \notin \tau$;
   —$\mathrm{Y} : \exists \boldsymbol{x}.\alpha(\boldsymbol{xab}) \in \tau$, but $\mathrm{Y} : \exists \boldsymbol{xy}.\alpha(\boldsymbol{xyb}) \notin \tau$;
   —$\mathrm{Y} : \psi_1 \wedge \psi_2 \in \tau$, but $\mathrm{Y} : \psi_1 \notin \tau$ or $\mathrm{Y} : \psi_2 \notin \tau$;
   —$\mathrm{Y} : \psi_1 \vee \psi_2 \in \tau$, but $\mathrm{Y} : \psi_1 \notin \tau$ and $\mathrm{Y} : \psi_2 \notin \tau$;
   —$\mathrm{Y} : \alpha(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}], \mathrm{Y} : \forall \boldsymbol{x}.\alpha(\boldsymbol{xb}) \to \psi(\boldsymbol{xb}) \in \tau$, but $\mathrm{Y} : \psi(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}] \notin \tau$.
   If $\mathrm{Y} = \mathrm{x}$, then $\theta_{\mathrm{x}}^{\tau} := \bot$. If $\mathrm{Y} = \tilde{\mathrm{x}}$, then $\theta_{\mathrm{x}}^{\tau} := \top$.
—If $\mathrm{x} : \alpha(\boldsymbol{b}), \tilde{\mathrm{x}} : \neg\alpha(\boldsymbol{b}) \in \tau$, then $\theta_{\mathrm{x}}^{\tau} := \alpha(\boldsymbol{b})$.
—If $\mathrm{x} : \neg\alpha(\boldsymbol{b}), \tilde{\mathrm{x}} : \alpha(\boldsymbol{b}) \in \tau$, then $\theta_{\mathrm{x}}^{\tau} := \mathrm{gdd}_{\mathrm{x}}^{\tau}(\boldsymbol{b}) \wedge \neg\alpha(\boldsymbol{b})$.
—If $\mathrm{x} : \alpha(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}], \tilde{\mathrm{x}} : \forall \boldsymbol{x}.\alpha(\boldsymbol{xb}) \to \psi(\boldsymbol{xb}) \in \tau$, but $\tilde{\mathrm{x}} : \psi(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}] \notin \tau$, then $\theta_{\mathrm{x}}^{\tau} := \alpha(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}]$.
—If $\tilde{\mathrm{x}} : \alpha(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}], \mathrm{x} : \forall \boldsymbol{x}.\alpha(\boldsymbol{xb}) \to \psi(\boldsymbol{xb}) \in \tau$, but $\mathrm{x} : \psi(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}] \notin \tau$, then $\theta_{\mathrm{x}}^{\tau} := \mathrm{gdd}_{\mathrm{x}}^{\tau}(\boldsymbol{a'b}) \wedge \neg\alpha(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}]$, where $\boldsymbol{a'b} = \mathrm{par}(\alpha(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}])$.

Inductive Case. Now consider an x-mosaic $\tau(\boldsymbol{c}) \in \mathcal{N}_i \setminus \mathcal{N}_{i-1}$. This means there is some $\mathrm{Y} : \eta = \mathrm{Y} : \exists \boldsymbol{x}.\alpha(\boldsymbol{xb}) \wedge \psi(\boldsymbol{xb}) \in \tau$ that is not fulfilled in τ, and if $\tau \to_{\mathrm{Y}:\eta} \tau'$, then $\tau' \in \mathcal{N}_{i-1}$.
   Assume $\mathrm{par}(\psi) = \boldsymbol{b} \subseteq \boldsymbol{c}$, and fix some $\boldsymbol{a} \in P \setminus \boldsymbol{c}$.
   We introduce some additional notation. For $\mathrm{z} \in \{\mathrm{L}, \mathrm{R}\}$, we define $\tau \upharpoonright_{\mathrm{z}}^{\boldsymbol{b}}$ to be the z-mosaic that results from taking the z-formulas in τ that only use parameters from $\boldsymbol{b}$, and then taking the union with the $\tilde{\mathrm{z}}$-formulas in τ whose parameters are contained in $\boldsymbol{b}$ and are z-cguarded in τ. For mosaics υ, υ' and $\mathrm{z} \in \{\mathrm{L}, \mathrm{R}\}$, we write $\upsilon'(\boldsymbol{ab}) \supseteq_{\mathrm{z}} \upsilon$ if υ' uses only parameters from $\boldsymbol{ab}$ and is obtained from υ by adding only z-formulas (so the $\tilde{\mathrm{z}}$-formulas in υ and υ' are identical).

If $Y = X$, then define θ_X^τ to be

$$\bigvee_{\substack{\upsilon'(\boldsymbol{ab}) \supseteq_X \tau \restriction_X^{\boldsymbol{b}} \\ \text{s.t. } \tau \to_{X:\eta} \upsilon'}} \exists \boldsymbol{x}. \left(\bigwedge_{\substack{\tau'(\boldsymbol{ab}) \supseteq_{\widetilde{X}} \upsilon' \\ \text{s.t. } \tau \to_{X:\eta} \tau'}} \theta_X^{\tau'}[\boldsymbol{x}/\boldsymbol{a}] \right).$$

Informally, this interpolant expresses that there is a way to fulfill the existential requirement by adding some x-formulas to $\tau \restriction_X^{\boldsymbol{b}}$, but no matter what $\widetilde{x}$-formulas we add, the resulting linked mosaic τ' has already been eliminated.

If $Y = \widetilde{x}$, then define θ_X^τ to be

$$\mathrm{gdd}_X^\tau(\boldsymbol{b}) \wedge \neg \left(\bigvee_{\substack{\upsilon'(\boldsymbol{ab}) \supseteq_{\widetilde{X}} \tau \restriction_{\widetilde{X}}^{\boldsymbol{b}} \\ \text{s.t. } \tau \to_{\widetilde{X}:\eta} \upsilon'}} \exists \boldsymbol{x}. \left(\bigwedge_{\substack{\tau'(\boldsymbol{ab}) \supseteq_X \upsilon' \\ \text{s.t. } \tau \to_{\widetilde{X}:\eta} \tau'}} \theta_{\widetilde{X}}^{\tau'}[\boldsymbol{x}/\boldsymbol{a}] \right) \right).$$

Notice that in this case we need to ensure the negation is guarded in order to prove (GN). Using the inductive hypothesis, it can be shown that inside the negation, the formula satisfies (GN), and only uses parameters from $\boldsymbol{b}$ (any parameters from $\boldsymbol{a}$ are removed using quantification). Because $\widetilde{x} : \eta$ is a $\widetilde{x}$-formula using $\boldsymbol{b}$ in an x-mosaic, we know that $\mathrm{gdd}_X^\tau(\boldsymbol{b})$ exists. This is enough to conclude that θ_X^τ is in GNF, even when the parameters are viewed as free variables.

In both inductive cases, notice that the parameters $\boldsymbol{ab}$ in the inductively defined interpolants $\theta_Y^{\tau'}$ are guarded by the atom $\alpha(\boldsymbol{ab})$ appearing in the existential requirement $Y : \exists \boldsymbol{x}.\alpha(\boldsymbol{xb}) \wedge \psi(\boldsymbol{xb})$. We cannot necessarily guard the quantification in the interpolant using $\alpha(\boldsymbol{xb})$ since α may not be in the common signature. This is why the resulting interpolant might not be in GF, even when the input sentences are in GF. However, in case $Y = \widetilde{x}$, we can always guard $\boldsymbol{b}$ in the common signature using $\mathrm{gdd}_X^\tau(\boldsymbol{b})$. □

We can prove the following constructive Lyndon interpolation result using the mosaic interpolants from Lemma 3.3.

THEOREM 3.4. *Let φ_L and φ_R be GF sentences without equality over signatures σ_L and σ_R, respectively. If $\varphi_L \models \varphi_R$ and $|\varphi_L| + |\varphi_R| = n$, then we can construct a DAG representation of a GNF interpolant θ such that*

—$\varphi_L \models \theta$ and $\theta \models \varphi_R$;
—$\mathrm{occ}(\theta) \subseteq \mathrm{occ}(\varphi_L) \cap \mathrm{occ}(\varphi_R)$;
—$\mathrm{con}(\theta) \subseteq \mathrm{con}(\varphi_L) \cup \mathrm{con}(\varphi_R)$;
—*the DAG representation of θ is of size at most $2^{2^{p(n)}}$ for some polynomial function p independent of φ_L and φ_R (and this can be improved to a size of at most $2^{p(n)}$ when the bound on the arity of relations is fixed).*

PROOF. Let

$$\theta := \bigvee_{\tau(\emptyset) \supseteq_L \{L:\varphi_L\}} \bigwedge_{\substack{\tau'(\emptyset) \supseteq_R \tau \\ \text{s.t. } \tau' \in \mathcal{N}'}} \theta_L^{\tau'},$$

where $\mathcal{N}'$ is the set of eliminated mosaics and we write $\upsilon'(\emptyset) \supseteq_Z \upsilon$ if υ' is over parameters $\emptyset$ and is obtained from υ by adding only z-formulas.

We prove that $\varphi_L \models \theta \models \varphi_R$. Assume there is a structure $\mathfrak{M}$ such that $\mathfrak{M} \models \varphi_L$. Let $\tau(\emptyset) \supseteq_L \{L : \varphi_L\}$ be the set of sentences from $\mathrm{cl}_L(\emptyset)$ that are true in $\mathfrak{M}$. For any $\tau'(\emptyset) \supseteq_R \tau$ such that $\tau' \in \mathcal{N}'$, $\mathfrak{M} \models \tau'_L$. By Lemma 3.3, this means that $\mathfrak{M} \models \theta_L^{\tau'}$, so $\mathfrak{M} \models \theta$. Now

assume there is a structure $\mathfrak{M}$ such that $\mathfrak{M} \models \theta$. Then there is some $\tau(\emptyset) \supseteq_\mathrm{L} \{\mathrm{L} : \varphi_\mathrm{L}\}$ such that $\mathfrak{M} \models \theta_\mathrm{L}^{\tau'}$ for all $\tau'(\emptyset) \supseteq_\mathrm{R} \tau$ such that $\tau' \in \mathcal{N}'$. Let S' be the set of formulas $\mathrm{R} : \psi$ in $\mathrm{cl}_\mathrm{R}(\emptyset)$ such that $\mathfrak{M} \models \psi$. Consider $\tau' := \tau \cup S' \cup \{\mathrm{R} : \neg\varphi_\mathrm{R}\} \supseteq_\mathrm{R} \tau$. By completeness of the mosaic system (Theorem 3.2), every mosaic containing $\mathrm{L} : \varphi_\mathrm{L}$ and $\mathrm{R} : \neg\varphi_\mathrm{R}$ must be in $\mathcal{N}'$. Hence, $\mathfrak{M} \models \theta_\mathrm{L}^{\tau'}$. By Lemma 3.3 this implies that $\mathfrak{M} \models \neg\tau_\mathrm{R}'$. But every formula $\mathrm{R} : \psi$ in S' was chosen such that $\mathfrak{M} \models \psi$. The only way that $\mathfrak{M} \models \neg\tau_\mathrm{R}'$ is if $\mathfrak{M} \models \varphi_\mathrm{R}$.

The other properties in Theorem 3.4 follow from Lemma 3.3. □

Although GF does not have CIP, it was shown in Hoogland et al. [1999] that GF does admit the following weaker form of interpolation where the interpolant is allowed to freely use guards in addition to relations in the common signature.

THEOREM 3.5 ([HOOGLAND ET AL. 1999]). *For* GF *sentences* $\varphi_\mathrm{L} \models \varphi_\mathrm{R}$ *there is a* GF *sentence* θ *such that* $\varphi_\mathrm{L} \models \theta$ *and* $\theta \models \varphi_\mathrm{R}$ *and* θ *only uses relations in the common signature or relations occurring as a guard for some quantification in* φ_L *or* φ_R.

We remark that our mosaic interpolation method can be modified to produce GF interpolants like this, with the same DAG-size bounds as in Theorem 3.4, by changing the inductive case of Lemma 3.3 to include the guard that is present already in the existential requirement.

D'Agostino and Lenzi recently proved a strengthening of Theorem 3.5 that says GF even has *uniform interpolation* when the interpolant is allowed to freely use guards like this [D'Agostino and Lenzi 2015]. Roughly speaking, this means that given a GF formula φ_L over σ_L and given a subsignature σ of σ_L that contains at least all of the guards of φ_L, there is a GF formula θ over σ that can serve as an interpolant for all GF consequents φ_R of φ_L over this subsignature. The proof of this result requires techniques different from those of this article.

4. GNF MOSAIC METHOD

We now turn to mosaics for GNF. We could define a variant of mosaics for GNF formulas, and use this to prove a Lyndon-style interpolation result similar to Theorem 3.4. Instead, we define mosaics for relativized GNF formulas so that we can prove the stronger relativized interpolation theorem for GNF stated in Theorem 4.10. The consequence of this choice is that the definition of the mosaics is more complicated in order to correctly handle both the unguarded and relativized quantification. For now, we still assume that the formulas do not make use of equality; we address this feature separately at the end of this section. Note that we overload the terminology and notation that was used for GF mosaics.

4.1. Relativized GNF Mosaics

Fix formulas φ_L and φ_R over signatures σ_L and σ_R that are in weak GN-normal form and do not use equality. We assume that $\mathrm{free}(\varphi_\mathrm{L}) = \mathrm{free}(\varphi_\mathrm{R}) = \boldsymbol{z}$ (which may be empty if φ_L and φ_R are sentences), and that these variable names are never used as quantified variables in the formulas. We also view $\boldsymbol{z}$ as parameters, with $\mathrm{par}(\varphi_\mathrm{L}) = \mathrm{par}(\varphi_\mathrm{R}) = \boldsymbol{z}$. Let $\boldsymbol{e} := \mathrm{con}(\varphi_\mathrm{L}) \cup \mathrm{con}(\varphi_\mathrm{R})$.

We also assume φ_L and φ_R are in $\mathbb{U}$-relativized GNF, for $\mathbb{U}$ a distinguished set of unary relations from $\sigma_\mathrm{L} \cup \sigma_\mathrm{R}$. As shorthand, we write $\exists \boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \psi(\boldsymbol{xy})$ for a formula with relativized quantification of the form $\exists \boldsymbol{x}.U_1 x_1 \wedge \cdots \wedge U_k x_k \wedge \psi(\boldsymbol{xy})$ where $\boldsymbol{x} = x_1 \cdots x_k$, $\boldsymbol{U} = U_1 \cdots U_k$, and $U_i \in \mathbb{U}$ for $i \in \{1, \ldots, k\}$.

Closure and Specializations. It turns out that the subformula closure is not sufficient to determine satisfiability of $\varphi_\mathrm{L} \wedge \neg\varphi_\mathrm{R}$ from a set of coherent mosaics. Instead, the mosaics must contain additional information about CQ-shaped subformulas—subformulas of the form $\exists \boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \bigwedge_i \psi_i(\boldsymbol{xy})$. Our solution is to add "specializations"

to the closure. Informally, each specialization describes a way in which the original CQ-shaped formula could be satisfied in some tree-like structure.

Consider a weak GN-normal form formula $\phi(\boldsymbol{y})$ of the form $\exists \boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \bigwedge_i \psi_i(\boldsymbol{x}\boldsymbol{y})$. A *(relativized) specialization* of ϕ is a formula ϕ' obtained from ϕ by the following operations:

—select a subset $\boldsymbol{z}$ of $\boldsymbol{x}$ (call variables from $\boldsymbol{yz}$ the *inside variables* and variables from $\boldsymbol{x} \setminus \boldsymbol{z}$ the *outside variables*);
—select a partition $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ of the outside variables, with the property that for every ψ_j, either ψ_j has no outside variables or all of its outside variables are contained in some partition element $\boldsymbol{x}_i$;
—let $\boldsymbol{U}_0$ be the conjunction of atoms from $\boldsymbol{U}(\boldsymbol{x})$ using only inside variables, and let $\boldsymbol{U}_j$ be the conjunction of atoms from $\boldsymbol{U}(\boldsymbol{x})$ using outside variables in $\boldsymbol{x}_j$;
—let χ_0 be the conjunction of the ψ_i using only inside variables, and let χ_j be the conjunction of the ψ_i using outside variables and satisfying $\text{free}(\psi_i) \subseteq \boldsymbol{x}_j\boldsymbol{yz}$;
—set $\phi'(\boldsymbol{yz})$ to be

$$\boldsymbol{U}_0(\boldsymbol{z}) \wedge \chi_0(\boldsymbol{yz}) \wedge \bigwedge_{j \in \{1,\ldots,k\}} \exists \boldsymbol{x}_j.\boldsymbol{U}_j(\boldsymbol{x}_j) \wedge \chi_j(\boldsymbol{x}_j\boldsymbol{yz}).$$

We say a specialization $\boldsymbol{U}_0(\boldsymbol{z}) \wedge \chi_0(\boldsymbol{yz}) \wedge \bigwedge_j \exists \boldsymbol{x}_j.\boldsymbol{U}_j(\boldsymbol{x}_j) \wedge \chi_j(\boldsymbol{x}_j\boldsymbol{yz})$ is *nontrivial* if either there are no outside variables (thus the specialization is only $\boldsymbol{U}_0 \wedge \chi_0$), or χ_0 is nonempty, or the partition of the outside variables is nontrivial.

The following lemma states the relationship between the original CQ-shaped subformula ϕ and specializations of ϕ: every tree-like structure satisfying ϕ induces a nontrivial specialization that is also satisfied in the structure, and every specialization of ϕ logically implies ϕ.

LEMMA 4.1. *Let $\phi(\boldsymbol{y})$ be a CQ-shaped formula $\exists \boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \bigwedge_i \psi_i(\boldsymbol{x}\boldsymbol{y})$.*

—*For all structures $\mathfrak{M}$ and tree decompositions T of $\mathfrak{M}$, if there exists a node v with $\boldsymbol{b} \subseteq T(v)$ and $\mathfrak{M}, \boldsymbol{b} \models \phi(\boldsymbol{y})$, then there is a nontrivial specialization $\phi'(\boldsymbol{yz})$ of ϕ and a node w with a tuple $\boldsymbol{c}$ such that $\boldsymbol{bc} \in T(w)$ and $\mathfrak{M}, \boldsymbol{bc} \models \phi'(\boldsymbol{yz})$.*
—*For all structures $\mathfrak{M}$ and specializations $\phi'(\boldsymbol{yz})$ of ϕ, if $\mathfrak{M}, \boldsymbol{bc} \models \phi'(\boldsymbol{yz})$, then $\mathfrak{M}, \boldsymbol{b} \models \phi(\boldsymbol{y})$.*

PROOF. For elements $\boldsymbol{d}$ and a node w with a neighboring node w' in a tree decomposition T we say $\boldsymbol{d}$ *is contained in the direction of w'* if (i) w' is the parent of w and $\boldsymbol{d}$ appears in the tree resulting from removing from T the subtree rooted at w, or (ii) w' is a child of w and $\boldsymbol{d}$ appears in the subtree rooted at w'.

We prove the first part of the lemma. There must be some tuple $\boldsymbol{a} = a_1 \ldots a_m$ of elements (corresponding to $\boldsymbol{x} = x_1 \ldots x_m$) such that $\mathfrak{M}, \boldsymbol{ab} \models \boldsymbol{U}(\boldsymbol{x}) \wedge \bigwedge_i \psi_i(\boldsymbol{x}\boldsymbol{y})$. If there is a node w in the tree decomposition T with $\boldsymbol{ab} \subseteq T(w)$, then it is easy (take the specialization where all variables $\boldsymbol{x}$ are inside variables). Otherwise, there is a node w in T with $\boldsymbol{b} \subseteq T(w)$ and such that $\boldsymbol{ab}$ is not contained in the direction of any neighbor of w (if not, then starting at the node v containing $\boldsymbol{b}$, we could eventually reach a node w' with $\boldsymbol{ab} \subseteq T(w')$, which we are assuming is not possible). Let $\boldsymbol{z}$ be the tuple of variables from $\boldsymbol{x}$ corresponding to elements in $\boldsymbol{c} := T(w) \cap \boldsymbol{a}$ (i.e., $x_i \in \boldsymbol{z}$ if and only if $a_i \in T(w)$). Take $\boldsymbol{yz}$ to be the inside variables, corresponding to elements $\boldsymbol{bc}$ (the elements inside $T(w)$). Let O be the nonempty set of elements from $\boldsymbol{a}$ that are not in $T(w)$ (i.e., the elements that correspond to outside variables). Let $O_{w'}$ be the set of elements from O that are contained in the direction of a neighbor w' of w. Because $o \in O$ do not appear in $T(w)$ and $|O|$ is finite, $\{O_{w'} : w' \text{ is a neighbor of } w \text{ and } O_{w'} \neq \emptyset\}$ is a partition of O into a finite number of partitions. This induces a partition of the

outside variables based on what variables belong together because the witnesses are contained in the same direction from w. Taking the resulting nontrivial specialization ϕ', we have $\mathfrak{M}, \boldsymbol{bc} \models \phi'(\boldsymbol{yz})$. $\square$

We are now ready to define the closure. Each formula is labeled with a *provenance* $\mathrm{X} \in \{\mathrm{L}, \mathrm{R}\}$ and a *polarity* $p \in \{+, -\}$. The *closure* $\mathrm{cl}(\mathrm{X}^p : \varphi)$ is the smallest set C of formulas containing $\mathrm{X}^p : \varphi$ and such that

—if $\mathrm{X}^+ : \alpha(\boldsymbol{y}) \wedge \neg\psi(\boldsymbol{y}) \in C$, then $\mathrm{X}^+ : \alpha(\boldsymbol{y}), \mathrm{X}^- : \psi(\boldsymbol{y}) \in C$;
—if $\mathrm{X}^- : \alpha(\boldsymbol{y}) \wedge \neg\psi(\boldsymbol{y}) \in C$, then $\mathrm{X}^- : \alpha(\boldsymbol{y}), \mathrm{X}^+ : \psi(\boldsymbol{y}) \in C$;
—if $\mathrm{X}^p : \bigwedge_i \psi_i \in C$ or $\mathrm{X}^p : \bigvee_i \psi_i \in C$, then $\mathrm{X}^p : \psi_i \in C$ for all i;
—if $\mathrm{X}^+ : \exists \boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \psi(\boldsymbol{xy}) \in C$, then $\mathrm{X}^+ : \boldsymbol{U}(\boldsymbol{x}) \wedge \psi(\boldsymbol{xy}) \in C$;
—if $\mathrm{X}^- : \exists \boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \psi(\boldsymbol{xy}) \in C$ and $\exists \boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \psi(\boldsymbol{xy})$ has a specialization

$$\boldsymbol{U}_0(\boldsymbol{z}) \wedge \chi_0(\boldsymbol{yz}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{yz}) \quad \text{with} \quad \boldsymbol{U}_0(\boldsymbol{z}) = U_1 z_1 \wedge \cdots \wedge U_j z_j,$$

then $\mathrm{X}^- : U_i z_i \in C$ for all $i \in \{1, \dots, j\}$ and $\mathrm{X}^- : \chi_0(\boldsymbol{yz}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{yz}) \in C$;
—if $\mathrm{X}^+ : \alpha(\boldsymbol{xy}) \in C$ for $\boldsymbol{x} = \{x_1, \dots, x_k\}$ and $\mathrm{X}^+ : U_1 z_1, \dots, \mathrm{X}^+ : U_k z_k \in C$ for $U_i \in \mathbb{U}$, then $\mathrm{X}^+ : \exists \boldsymbol{x}.U_1 x_1 \wedge \cdots \wedge U_k x_k \wedge \alpha(\boldsymbol{xy}) \in C$.

Given parameters $\boldsymbol{c}$, we let $\mathrm{cl}_{\mathrm{L}}(\boldsymbol{c})$ ($\mathrm{cl}_{\mathrm{R}}(\boldsymbol{c})$, respectively) consist of formulas $\mathrm{X}^p : \psi$ from $\mathrm{cl}(\mathrm{L}^+ : \varphi_{\mathrm{L}})$ ($\mathrm{cl}(\mathrm{R}^- : \varphi_{\mathrm{R}})$, respectively) with the free variables replaced by parameters from $\boldsymbol{c}$ or constants from $\boldsymbol{e}$, and such that the resulting formulas have parameters contained in $\boldsymbol{c}$. Recall that we are viewing any variables $\boldsymbol{z}$ that were free in φ_{X} as parameters, so these will never be replaced by other parameters. For instance, $\mathrm{L}^+ : \psi(\boldsymbol{yz'})[\boldsymbol{b}/\boldsymbol{y}] \in \mathrm{cl}_{\mathrm{L}}(\boldsymbol{c})$ if $\mathrm{L}^+ : \psi(\boldsymbol{yz'}) \in \mathrm{cl}(\mathrm{L}^+ : \varphi_{\mathrm{L}})$, $\boldsymbol{z'} \subseteq \boldsymbol{z}$ are parameters corresponding to original free variables and $\boldsymbol{z'} \subseteq \boldsymbol{c}$, and $\boldsymbol{b} \subseteq \boldsymbol{c} \cup \boldsymbol{e}$.

Each formula in $\mathrm{cl}_{\mathrm{L}}(\boldsymbol{c}) \cup \mathrm{cl}_{\mathrm{R}}(\boldsymbol{c})$ is labeled with a provenance L or R and a polarity $+$ or $-$ to indicate whether the formula is built from a subformula of φ_{L} or $\neg\varphi_{\mathrm{R}}$, and whether this subformula occurred positively or negatively, that is, whether it occurred within the scope of an even or odd number of negations. Given a formula ψ and a polarity p, we write $p\psi$ to mean ψ ($\neg\psi$, respectively) if p is $+$ ($-$, respectively).

It is straightforward to check that the formulas in $\mathrm{cl}_{\mathrm{L}}(\boldsymbol{c}) \cup \mathrm{cl}_{\mathrm{R}}(\boldsymbol{c})$ respect the polarity of relations, guards, and constants in the original formulas.

PROPOSITION 4.2. *For all $\tau \subseteq \mathrm{cl}_{\mathrm{L}}(\boldsymbol{c}) \cup \mathrm{cl}_{\mathrm{R}}(\boldsymbol{c})$,*

—*if $\mathrm{L}^p : \psi \in \tau$, then $\mathrm{occ}(p\psi) \subseteq \mathrm{occ}(\varphi_{\mathrm{L}})$, $\mathrm{guards}(\psi) \subseteq \mathrm{guards}(\varphi_{\mathrm{L}})$, and $\mathrm{con}(\psi) \subseteq \mathrm{con}(\varphi_{\mathrm{L}}) \cup \mathrm{con}(\varphi_{\mathrm{R}})$;*
—*if $\mathrm{R}^p : \psi \in \tau$, then $\mathrm{occ}(\neg p\psi) \subseteq \mathrm{occ}(\varphi_{\mathrm{R}})$, $\mathrm{guards}(\psi) \subseteq \mathrm{guards}(\varphi_{\mathrm{R}})$, and $\mathrm{con}(\psi) \subseteq \mathrm{con}(\varphi_{\mathrm{L}}) \cup \mathrm{con}(\varphi_{\mathrm{R}})$.*

Guards and Relativizers. We introduce terminology to refer to relations that are used as guards or relativizers, and to indicate which of these relations are in the common signature. We say α is an L-*atom* (R-*atom*, respectively) if $(\alpha, +) \in \mathrm{occ}(\varphi_{\mathrm{L}})$ ($(\alpha, -) \in \mathrm{occ}(\varphi_{\mathrm{R}})$, respectively). Likewise, we say α is an L-*catom* (R-*catom*, respectively) if $(\alpha, +) \in \mathrm{occ}(\varphi_{\mathrm{L}}) \cap \mathrm{occ}(\varphi_{\mathrm{R}})$ ($(\alpha, -) \in \mathrm{occ}(\varphi_{\mathrm{L}}) \cap \mathrm{occ}(\varphi_{\mathrm{R}})$, respectively). We say α is an X-*guard* (X-*cguard*, respectively) if α is an X-atom (X-catom, respectively) and $\alpha \in \mathrm{guards}(\varphi_{\mathrm{L}}) \cup \mathrm{guards}(\varphi_{\mathrm{R}})$. We say U is an X-*relativizer* (X-*crelativizer*, respectively) if U is an X-atom (X-catom, respectively) and $U \in \mathbb{U}$. The idea is that X-catoms and X-crelativizers are predicates in the common signature that can appear in the interpolant—positively if $\mathrm{X} = \mathrm{L}$ and negatively if $\mathrm{X} = \mathrm{R}$.

Now let τ be a collection of formulas from $\mathrm{cl}_{\mathrm{L}}(\boldsymbol{c}) \cup \mathrm{cl}_{\mathrm{R}}(\boldsymbol{c})$ and let $\boldsymbol{d} \subseteq \boldsymbol{c}$ be a distinguished subset of these parameters. We say $b \in \boldsymbol{c}$ is X-*crelativized in τ* if there is some $\mathrm{X}^+ : Ub \in \tau$

where $U \in \mathbb{U}$ and U is an x-catom. We say $\boldsymbol{b}$ is x-crelativized if every $b \in \boldsymbol{b}$ is x-crelativized. We say $\boldsymbol{b} \subseteq \boldsymbol{c}$ *is* x-*cguarded in* τ if $|\boldsymbol{b}| \leq 1$, or if there is some $\textsc{x}^+ : \exists \boldsymbol{x}.U(\boldsymbol{x}) \wedge \alpha(\boldsymbol{x}\boldsymbol{b'}) \in \tau$ where α and $\boldsymbol{U}$ are x-catoms, $\alpha \in \text{guards}(\varphi_{\textsc{l}}) \cup \text{guards}(\varphi_{\textsc{r}})$, $\boldsymbol{U} \subseteq \mathbb{U}$, $\boldsymbol{b'} \supseteq \boldsymbol{b}$, and $\boldsymbol{b'} \cap \boldsymbol{d}$ is x-crelativized. We remark that the definition of x-cguardedness here differs from the previous section due to technical requirements when proving soundness and completeness for relativized GNF mosaics. In particular, unary relations are always considered to be x-cguarded.

We say a formula $\widetilde{\textsc{x}}^p : \psi$ is x-*safe in* τ if

—$\widetilde{\textsc{x}}^p = \widetilde{\textsc{x}}^+$ and $\text{par}(\psi) \cap \boldsymbol{d}$ is x-crelativized in τ and $\text{par}(\psi)$ is x-cguarded in τ, or
—$\widetilde{\textsc{x}}^p = \widetilde{\textsc{x}}^-$ and $\text{par}(\psi) \cap \boldsymbol{d}$ is x-crelativized in τ.

We emphasize that these definitions (x-cguarded, x-crelativized, and x-safe) depend on the distinguished set of parameters $\boldsymbol{d}$ in τ, so when we talk about being, for example, x-cguarded in τ, we will assume that τ has some distinguished set of parameters $\boldsymbol{d}$.

Relativized GNF *Mosaics.* Let $\boldsymbol{c}$ be a tuple of parameters with $|\boldsymbol{c}| \leq \text{width}(\varphi_{\textsc{l}} \wedge \neg\varphi_{\textsc{r}})$. A *(relativized)* GNF x-*mosaic* τ for $\varphi_{\textsc{l}} \wedge \neg\varphi_{\textsc{r}}$ using parameters $\boldsymbol{c}$ and relativizing $\boldsymbol{d} \subseteq \boldsymbol{c}$ is a subset of $\Gamma(\boldsymbol{c}) := \text{cl}_{\textsc{l}}(\boldsymbol{c}) \cup \text{cl}_{\textsc{r}}(\boldsymbol{c})$ with distinguished parameters $\boldsymbol{d}$ and satisfying the requirement that if $\widetilde{\textsc{x}}^p : \psi \in \tau$, then $\widetilde{\textsc{x}}^p : \psi$ is x-safe in τ with respect to $\boldsymbol{d}$. We sometimes write $\tau(\boldsymbol{c})$ to emphasize that τ uses parameters $\boldsymbol{c}$.

Let τ be a GNF x-mosaic over parameters $\boldsymbol{c}$ and relativizing $\boldsymbol{d}$. Assume $\textsc{y}^+ : \eta \in \tau$ for η an *existential requirement* of the form $\exists \boldsymbol{x}.\psi(\boldsymbol{x}\boldsymbol{b})$. We say $\textsc{y}^+ : \eta$ is *fulfilled* in τ if $\textsc{y}^+ : \psi(\boldsymbol{x}\boldsymbol{b})[\boldsymbol{a}/\boldsymbol{x}] \in \tau$ for some $\boldsymbol{a} \subseteq \boldsymbol{c} \cup \boldsymbol{e}$. We say τ is *linked* to τ' via $\textsc{y}^+ : \eta$ (written $\tau \to_{\textsc{y}^+:\eta} \tau'$) if

(E1) τ' is a GNF y-mosaic over parameters $\boldsymbol{a}\boldsymbol{b}$ relativizing $\boldsymbol{a} \cup (\boldsymbol{b} \cap \boldsymbol{d})$ ($\boldsymbol{a}$, respectively) if $\textsc{y} = \textsc{x}$ ($\textsc{y} = \widetilde{\textsc{x}}$, respectively), and with $\boldsymbol{a} \cap \boldsymbol{c} = \emptyset$;
(E2) $\textsc{y}^+ : \psi(\boldsymbol{x}\boldsymbol{b})[\boldsymbol{a}/\boldsymbol{x}] \in \tau'$;
(E3) for all $\textsc{y}^p : \psi'(\boldsymbol{b}) \in \tau$, $\textsc{y}^p : \psi'(\boldsymbol{b}) \in \tau'$;
(E4) for all $\widetilde{\textsc{y}}^p : \psi'(\boldsymbol{b}) \in \tau$ such that $\widetilde{\textsc{y}}^p : \psi'(\boldsymbol{b})$ is y-safe in τ', $\widetilde{\textsc{y}}^p : \psi'(\boldsymbol{b}) \in \tau'$;
(E5) for all $\textsc{z}^p : \psi'(\boldsymbol{b}) \in \tau'$, $\textsc{z}^p : \psi'(\boldsymbol{b}) \in \tau$.

These conditions ensure that the maximum amount of information about shared parameters is passed through a link, and no new information about shared parameters is added. We will explain later the intuition behind the relativized parameters in τ' specified by condition (E1).

Fix some set P of parameters of size $2 \cdot \text{width}(\varphi_{\textsc{l}} \wedge \neg\varphi_{\textsc{r}})$ that includes the parameters corresponding to the original free variables in the input formulas $\text{par}(\varphi_{\textsc{l}}) = \text{par}(\varphi_{\textsc{r}})$. Let $\mathcal{M}$ be the set of relativized GNF mosaics over parameters P.

Bound On Number of Relativized GNF *Mosaics.* There is a doubly exponential upper bound for the number of relativized GNF mosaics for $\varphi_{\textsc{l}} \wedge \neg\varphi_{\textsc{r}}$ in $\mathcal{M}$. The following proposition states this bound more precisely.

PROPOSITION 4.3. *There is a polynomial function p such that if $\varphi := \varphi_{\textsc{l}} \wedge \neg\varphi_{\textsc{r}}$ for $\varphi_{\textsc{l}}$ and $\varphi_{\textsc{r}}$ in weak GN-normal form without equality, and $n = |\varphi_{\textsc{l}}| + |\varphi_{\textsc{r}}|$, $w = \text{width}(\varphi)$, and $h = \text{rank}_{\text{CQ}}(\varphi)$, then the number of relativized GNF mosaics in $\mathcal{M}$ for φ is at most $2^{p(n)^{p(hw)}}$. These bounds apply even when the input formulas $\varphi_{\textsc{l}}$ and $\varphi_{\textsc{r}}$ are given as DAG representations.*

PROOF. Fix $\varphi := \varphi_{\textsc{l}} \wedge \neg\varphi_{\textsc{r}}$ with $n = |\varphi_{\textsc{l}}| + |\varphi_{\textsc{r}}|$, $w = \text{width}(\varphi)$, and $h = \text{rank}_{\text{CQ}}(\varphi)$. Let $l := |\text{con}(\varphi)|$. Note that $w, h, l \leq n$.

We first calculate the size of $\text{cl}(\textsc{y}^p : \varphi_{\textsc{y}})$, where $p = +$ if $\textsc{y} = \textsc{l}$ and $p = -$ if $\textsc{y} = \textsc{r}$. There are at most $2n$ polarity-labeled subformulas of $\varphi_{\textsc{y}}$. However, $\text{cl}(\textsc{y}^p : \varphi_{\textsc{y}})$ contains

additional formulas according to the definition of GNF closure. For each positive atomic formula α in $\mathrm{Y}^p : \varphi_\mathrm{Y}$, there are at most w free variables, and hence at most 2^w ways to choose variables for quantification. There can be at most n predicates in $\mathbb{U}$, and at most n^w ways to relativize the quantified variables using these unary predicates. Hence, for each positive atomic formula α in $\mathrm{Y}^p : \varphi_\mathrm{Y}$, there are at most $2^w n^w$ additional formulas that may be added. For each negative CQ-shaped formula ϕ in $\mathrm{Y}^p : \varphi_\mathrm{Y}$, there are at most 2^h CQ-shaped subformulas ϕ' obtained by choosing some subset of the conjuncts in ϕ. In each ϕ', there are at most 2^w choices of the inside variables, and w^w ways to partition the outside variables resulting in specializations ϕ'' of ϕ'. Notice that these specializations ϕ'' only have CQ-shaped subformulas resulting from taking some subset of the conjuncts in the original CQ-shaped formula ϕ, so the specializations of these subformulas have already been accounted for. Hence, for each negative CQ-shaped subformula in $\mathrm{Y}^p : \varphi_\mathrm{Y}$, there are at most $2^h 2^w w^w$ specializations that may be added. This means that the size of $\mathrm{cl}(\mathrm{Y}^p : \varphi_\mathrm{Y})$ is at most $2n(2^w n^w + 2^h 2^w w^w)$.

Now recall that P is a fixed set of parameters of size at most $2w$. Each formula in $\mathrm{cl}(\mathrm{Y}^p : \varphi_\mathrm{Y})$ has at most w variables that can be mapped to parameters from P or constants from $\boldsymbol{e}$, so the size of $\mathrm{cl}_\mathrm{Y}(P)$ is $M := 2n(2^w n^w + 2^h 2^w w^w)(2w+l)^w$. Every mosaic in $\mathcal{M}$ contains a subset of the formulas in $\mathrm{cl}_\mathrm{L}(P) \cup \mathrm{cl}_\mathrm{R}(P)$, is labeled L or R, and has some subset of its at most w parameters distinguished as the relativized parameters. Hence, $|\mathcal{M}| \leq 2 \cdot 2^w \cdot 2^{2M} \leq 2^{p(n)p^{(hw)}}$ for some polynomial function p that is independent of φ_L and φ_R.

We remark that these bounds apply even when n is the size of a DAG representation of φ, since the size of the closure set depends on the number of distinct subformulas. □

Coherence. A GNF x-mosaic τ is *internally consistent* if there is no $\mathrm{Y} \in \{\mathrm{L}, \mathrm{R}\}$ such that $\mathrm{Y}^+ : \bot \in \tau$ or $\mathrm{Y}^- : \top \in \tau$, there is no atomic formula α and no $\mathrm{Y}, \mathrm{Z} \in \{\mathrm{L}, \mathrm{R}\}$ such that $\mathrm{Y}^+ : \alpha \in \tau$ and $\mathrm{Z}^- : \alpha \in \tau$, and τ is *downward closed*:

(i) if $\mathrm{Y}^+ : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$, then $\mathrm{Y}^+ : \alpha(\boldsymbol{b}) \in \tau$ and $\mathrm{Y}^- : \psi(\boldsymbol{b}) \in \tau$;

(ii) if $\mathrm{Y}^- : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$ and $\mathrm{Z}^+ : \alpha(\boldsymbol{b}) \in \tau$, then $\mathrm{Y}^+ : \psi(\boldsymbol{b}) \in \tau$;

(iii) if $\mathrm{Y}^+ : \bigwedge_i \psi_i \in \tau$ ($\mathrm{Y}^- : \bigvee_i \psi_i \in \tau$, respectively), $\mathrm{Y}^+ : \psi_i \in \tau$ ($\mathrm{Y}^- : \psi_i \in \tau$, respectively) for all i;

(iv) if $\mathrm{Y}^+ : \bigvee_i \psi_i \in \tau$ ($\mathrm{Y}^- : \bigwedge_i \psi_i \in \tau$, respectively), $\mathrm{Y}^+ : \psi_i \in \tau$ ($\mathrm{Y}^- : \psi_i \in \tau$, respectively) for some i;

(v) if $\mathrm{Y}^- : \exists\boldsymbol{x}.\boldsymbol{U}(\boldsymbol{x}) \wedge \psi(\boldsymbol{x}\boldsymbol{b}) \in \tau$, and there is some specialization

$$\boldsymbol{U}_0(\boldsymbol{c}') \wedge \chi_0(\boldsymbol{b}\boldsymbol{c}') \wedge \bigwedge_k \exists\boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{b}\boldsymbol{c}') \quad \text{with} \quad \boldsymbol{U}_0(\boldsymbol{c}') = U_1 c_1 \wedge \cdots \wedge U_j c_j$$

and $\mathrm{Z}_i^+ : U_i c_i \in \tau$ for all $i \in \{1, \ldots, j\}$, then

$$\mathrm{Y}^- : \chi_0(\boldsymbol{b}\boldsymbol{c}') \wedge \bigwedge_k \exists\boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{b}\boldsymbol{c}') \in \tau.$$

The last rule in this definition is a generalization of the downward closure rule for universal statements in the GF mosaics. It helps ensure that when we try to construct a model from a set of mosaics, a CQ-shaped subformula asserted to be false in one mosaic, does not become true in the constructed structure.

A set of relativized GNF mosaics is said to be *saturated* if for every mosaic $\tau(\boldsymbol{c})$ in the set and for every existential requirement $\mathrm{Y}^+ : \eta \in \tau$ either $\mathrm{Y}^+ : \eta$ is fulfilled in τ, or τ is linked to some mosaic τ' in the set via $\mathrm{Y}^+ : \eta$, so $\mathrm{Y}^+ : \eta$ is fulfilled in τ'.

Before giving the proof of soundness and completeness, we first give some intuition behind what the distinguished, relativized parameters represent in a saturated set of internally consistent mosaics. Roughly speaking, the distinguished relativized parameters $\boldsymbol{d}$ in an x-mosaic represent parameters that were introduced due to some existential requirement in an x-formula (and hence were originally x-relativized). This helps explain why for linked mosaics τ and τ' over $\mathrm{Y}^+ : \exists\boldsymbol{x}.\psi(\boldsymbol{x}\boldsymbol{b})$, the "new" parameters

$\boldsymbol{a}$ used in τ' always become part of the relativized parameters in the Y-mosaic τ'. If we have a series of linked mosaics all due to X existential requirements, we want to remember all of these X-relativized parameters. As soon as we link to a $\widetilde{x}$-mosaic, however, then we can forget about these X-relativized parameters, and concentrate on the new $\widetilde{x}$-relativized parameters. Thus, the set of distinguished relativized parameters $\boldsymbol{b} \cap \boldsymbol{d}$ is carried through a link between X-mosaics, but not carried through a link between an X-mosaic and a $\widetilde{x}$-mosaic. This forgetfulness is acceptable because in the interpolation construction for the Y-mosaic τ' we will only quantify out (and hence need relativizers for) the most recently introduced Y-parameters. An alternative approach would be to annotate every parameter with a provenance and carry information about all of the parameters through a link. We could then place restrictions on L-formulas that use R-parameters (and vice versa).

4.2. Soundness and Completeness of GNF Mosaic Method

We now prove soundness and completeness of the GNF mosaic method.

THEOREM 4.4. $\varphi_L \wedge \neg\varphi_R$ *is satisfiable if and only if there is a saturated set of internally consistent* GNF *mosaics from* $\mathcal{M}$ *that contains a mosaic* τ *such that* $L^+ : \varphi_L, R^- : \varphi_R \in \tau$.

In order to prove this theorem, we must relate our coherency requirements for GNF mosaics to "realizability" in a structure. We start with some additional definitions. Fix parameters $\boldsymbol{c}$ and some $\boldsymbol{d} \subseteq \boldsymbol{c}$. Given a structure $\mathfrak{M}$ over $\sigma_L \cup \sigma_R \cup \boldsymbol{c}$ such that $|\boldsymbol{c}| \leq \mathrm{width}(\varphi_L \wedge \neg\varphi_R)$, let $\upsilon_Y^{\mathfrak{M}}$ be the unique Y-mosaic over $\boldsymbol{c}$ relativizing $\boldsymbol{d}$ such that

$$\upsilon_Y^{\mathfrak{M}} := \left\{ Y^p : \psi \mid Y^p : \psi \in \mathrm{cl}_Y(\boldsymbol{c}) \text{ and } \mathfrak{M} \models p\psi \right\},$$

where $\mathfrak{M} \models p\psi$ means $\mathfrak{M} \models \psi$ ($\mathfrak{M} \models \neg\psi$, respectively) if $p = +$ ($p = -$, respectively). The X-mosaic

$$\tau := \upsilon_X^{\mathfrak{M}} \cup \left\{ \widetilde{x}^p : \psi \mid \widetilde{x}^p : \psi \in \upsilon_{\widetilde{x}}^{\mathfrak{M}} \text{ and } \widetilde{x}^p : \psi \text{ is X-safe in } \upsilon_X^{\mathfrak{M}} \right\}$$

is called the X-*mosaic* τ *of* $\boldsymbol{c}$ *relativizing* $\boldsymbol{d}$ *in* $\mathfrak{M}$. We say an X-mosaic τ over $\boldsymbol{c}$ relativizing $\boldsymbol{d}$ is *realizable* if there is some structure $\mathfrak{M}$ such that τ is the X-mosaic of $\boldsymbol{c}$ relativizing $\boldsymbol{d}$ in $\mathfrak{M}$. We now prove two propositions about realizability that are useful for proving soundness for the GNF mosaic method.

PROPOSITION 4.5. *If* τ *is realizable, then* τ *is internally consistent.*

PROOF. Let $\mathfrak{M}$ be a structure realizing the X-mosaic $\tau(\boldsymbol{c})$ relativizing $\boldsymbol{d}$. By definition, $z^p : \psi \in \tau$ implies $\mathfrak{M} \models p\psi$. This means it is not possible for there to be some $Y^+ : \bot$, $Y^- : \top$ or $Y^+ : \alpha$ and $z^- : \alpha$ in τ. We must check that the downward closure properties are satisfied in τ.

If $Y = X$, then it is straightforward to check that the downward closure properties hold. For instance, consider (i). If $Y^+ : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$, then $\mathfrak{M} \models \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b})$. Since $\mathfrak{M} \models \alpha(\boldsymbol{b})$ and $\mathfrak{M} \models \neg\psi(\boldsymbol{b})$ and $Y^+ : \alpha(\boldsymbol{b}) \in \mathrm{cl}_Y(\boldsymbol{c})$ and $Y^- : \psi(\boldsymbol{b}) \in \mathrm{cl}_Y(\boldsymbol{c})$, this is enough to ensure that $Y^+ : \alpha(\boldsymbol{b}), Y^- : \psi(\boldsymbol{b}) \in \tau$.

If $Y = \widetilde{x}$, then we use a similar argument, but in order to conclude that the required formulas appear in τ, we must also show that these formulas are X-safe in $\upsilon_X^{\mathfrak{M}}$.

For instance, consider (ii). Assume $Y^- : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$ and $z^+ : \alpha(\boldsymbol{b}) \in \tau$. Then we know that $\mathfrak{M} \models \alpha(\boldsymbol{b})$ and $\mathfrak{M} \models \alpha(\boldsymbol{b}) \rightarrow \psi(\boldsymbol{b})$, hence $\mathfrak{M} \models \psi(\boldsymbol{b})$. This means $Y^+ : \psi(\boldsymbol{b}) \in \upsilon_Y^{\mathfrak{M}}$. It remains to show that this formula is X-safe in $\upsilon_X^{\mathfrak{M}}$. If $Y = z$, then $z^+ : \alpha(\boldsymbol{b}) \in \tau$ implies that $z^+ : \alpha(\boldsymbol{b})$ is X-safe in $\upsilon_X^{\mathfrak{M}}$, so $Y^+ : \psi(\boldsymbol{b})$ is also X-safe in $\upsilon_X^{\mathfrak{M}}$. If $Y = \widetilde{z}$, then α appears positively in an X-formula (since $z = x$), and negatively in a $\widetilde{x}$-formula (since $Y = \widetilde{x}$). Making use of Proposition 4.2, this means that α is an X-cguard. Moreover,

Y^- : $\alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$ implies that $\boldsymbol{b} \cap \boldsymbol{d}$ is x-crelativized in τ. This means $\boldsymbol{b}$ is x-cguarded by X^+ : $\alpha(\boldsymbol{b})$ in τ, so Y^+ : $\psi(\boldsymbol{b})$ is x-safe in $\upsilon_X^{\mathfrak{M}}$.

Now consider (v), still for $Y = \widetilde{x}$. Assume Y^- : $\exists \boldsymbol{x}.U(\boldsymbol{x}) \wedge \psi(\boldsymbol{xb}) \in \tau$ and there is some specialization Y^- : $\boldsymbol{U}_0(\boldsymbol{c}') \wedge \chi_0(\boldsymbol{bc}') \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k \boldsymbol{bc}')$ such that $\boldsymbol{U}_0(\boldsymbol{c}') = U_1 c_1 \wedge \cdots \wedge U_j c_j$ and Z_i^+ : $U_i c_i \in \tau$ for all $i \in \{1, \ldots, j\}$. We first claim that Y^- : $\chi_0(\boldsymbol{bc}') \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_j(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k \boldsymbol{bc}')$ is x-safe in $\upsilon_X^{\mathfrak{M}}$. Since the formula is negative, we only need to ensure that all parameters in $\boldsymbol{bc}' \cap \boldsymbol{d}$ are x-crelativized in τ. We already know that the parameters in $\boldsymbol{b} \cap \boldsymbol{d}$ are x-crelativized since Y^- : $\exists \boldsymbol{x}.\psi(\boldsymbol{xb}) \in \tau$. Now consider parameters $c_i \in \boldsymbol{c}'$. If $Z_i = \widetilde{x}$ and $c_i \in \boldsymbol{d}$, then Z_i^+ : $U_i c_i \in \tau$ implies that c_i is x-crelativized in τ. If $Z_i = X$ and $c_i \in \boldsymbol{d}$, then Z_i^+ : $U_i c_i \in \tau$ is the x-crelativizer for c_i. Assume for the sake of contradiction that Y^- : $\chi_0(\boldsymbol{bc}') \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_j(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k \boldsymbol{bc}') \notin \tau$. Since this formula is in $\mathrm{cl}_Y(\boldsymbol{c})$ and is x-safe in $\upsilon_X^{\mathfrak{M}}$, the only way it is not in τ is if $\mathfrak{M} \models \chi_0(\boldsymbol{bc}') \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_j(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k \boldsymbol{bc}')$. But this would mean $\mathfrak{M} \models \boldsymbol{U}_0(\boldsymbol{c}') \wedge \chi_0(\boldsymbol{bc}') \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_j(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k \boldsymbol{bc}')$, since Z_i^+ : $U_i c_i \in \tau$. By Lemma 4.1, every specialization of a formula implies the original formula, contradicting $\mathfrak{M} \models \neg\exists \boldsymbol{x}.U(\boldsymbol{x}) \wedge \psi(\boldsymbol{xb})$. □

PROPOSITION 4.6. *If υ is realizable, then there is a saturated set $\mathcal{S} \subseteq \mathcal{M}$ of realizable mosaics that includes υ.*

PROOF. We inductively construct sets $\mathcal{S}_i$, ensuring at each stage i that $\mathcal{S}_i$ contains υ and other realizable mosaics, and that any existential requirement in $\mathcal{S}_{i-1}$ is fulfilled in $\mathcal{S}_i$. This is enough to ensure that the set $\mathcal{S} := \bigcup \mathcal{S}_i$ is a saturated set of realizable mosaics that includes υ.

At stage $i = 1$, we set $\mathcal{S}_1 := \{\upsilon\}$, which is realizable by assumption.

At stage $i > 1$, if every existential requirement in $\mathcal{S}_{i-1}$ is fulfilled, then set $\mathcal{S}_i := \mathcal{S}_{i-1}$.

Otherwise, consider some x-mosaic $\tau(\boldsymbol{c})$ relativizing $\boldsymbol{d}$ in $\mathcal{S}_{i-1}$ with Y^+ : $\eta \in \tau$ for η of the form $\exists \boldsymbol{x}.\psi(\boldsymbol{xb})$ and $\mathrm{par}(\eta) = \boldsymbol{b}$ for which there is no $\boldsymbol{a} \subseteq \boldsymbol{c} \cup \boldsymbol{e}$ with Y^+ : $\psi(\boldsymbol{xb})[\boldsymbol{a}/\boldsymbol{x}] \in \tau$, and for all τ' such that $\tau \to_{Y^+:\eta} \tau'$, we have $\tau' \notin \mathcal{S}_{i-1}$.

Let $\mathfrak{M}$ be a structure realizing τ. Take parameters $\boldsymbol{a} \subseteq P$ corresponding to $\boldsymbol{x}$ such that $\boldsymbol{a} \cap \boldsymbol{c} = \emptyset$. This is possible since $|\boldsymbol{a}| \leq \mathrm{width}(\varphi_L \wedge \neg\varphi_R)$ and $|P| = 2 \cdot \mathrm{width}(\varphi_L \wedge \neg\varphi_R)$. Since $\mathfrak{M} \models \tau$, there are elements $\boldsymbol{a}' \subseteq M$, and an expansion $\mathfrak{M}'$ of $\mathfrak{M}$ with the interpretation $\boldsymbol{a}^{\mathfrak{M}'} := \boldsymbol{a}'$ such that $\mathfrak{M}' \models \psi(\boldsymbol{ab})$. Let τ' be the Y-mosaic of $\boldsymbol{ab}$ relativizing $\boldsymbol{a} \cup (\boldsymbol{b} \cap \boldsymbol{d})$ ($\boldsymbol{a}$, respectively) in $\mathfrak{M}'$ if $Y = X$ ($Y = \widetilde{x}$, respectively). This means τ' is realizable.

We claim that $\tau \to_{Y^+:\eta} \tau'$. (E1) and (E2) are satisfied by construction: by the choice of $\boldsymbol{a}^{\mathfrak{M}'}$, we have $\mathfrak{M}' \models \psi(\boldsymbol{ab})$, so Y^+ : $\psi(\boldsymbol{ab}) \in \tau'$. (E3) and (E4) follow by the definition of realizable mosaics, and the fact that $\mathfrak{M}$ and $\mathfrak{M}'$ are identical with respect to formulas that use only parameters from $\boldsymbol{b}$. It remains to check (E5). Assume there is some Z^p : $\psi'(\boldsymbol{b}) \in \tau'$. We have $\mathfrak{M}' \models p\psi'$ since τ' is realized by $\mathfrak{M}'$. Recall that τ is an x-mosaic, τ' is a Y-mosaic, and ψ' is a Z^p-formula. The result is immediate if $Z^p = X^p$. Otherwise, if $Z^p = \widetilde{x}^p$ it suffices to show that Z^p : $\psi'(\boldsymbol{b})$ is x-safe in τ.

—Assume $Z^p = \widetilde{x}^p$ and $X = Y$. Then $\mathrm{par}(\psi') \cap \boldsymbol{d}$ must be x-crelativized in τ', and hence x-crelativized in τ. Moreover, if $p = +$, then since $\widetilde{x}^+$: $\psi' \in \tau'$ and τ' is an x-mosaic, $\mathrm{par}(\psi')$ must be x-cguarded in τ' by some X^+ : $\exists \boldsymbol{w}.U(\boldsymbol{w}) \wedge \beta(\boldsymbol{wa'b'})$ where $\boldsymbol{a}' \subseteq \boldsymbol{a}$ and $\mathrm{par}(\psi') \subseteq \boldsymbol{b}' \subseteq \boldsymbol{b}$ and such that parameters in $\boldsymbol{a}'\boldsymbol{b}' \cap (\boldsymbol{a} \cup (\boldsymbol{b} \cap \boldsymbol{d}))$ are x-crelativized in τ'. But this means that $\mathfrak{M} \models \exists \boldsymbol{wx}.U(\boldsymbol{w}) \wedge U'(\boldsymbol{x}) \wedge \beta(\boldsymbol{wxb'})$ for some $U'(\boldsymbol{x})$ conjunction of x-crelativizers. Hence, X^+ : $\exists \boldsymbol{wx}.U(\boldsymbol{w}) \wedge U'(\boldsymbol{x}) \wedge \beta(\boldsymbol{wxb'}) \in \tau$. Since the parameters in $\boldsymbol{b}' \cap \boldsymbol{d}$ are also x-crelativized in τ, $\mathrm{par}(\psi')$ is x-cguarded in τ. Overall, this means that $\widetilde{x}^p$: ψ' is x-safe in τ.

—Assume $Z^p = \widetilde{x}^p$ and $X \neq Y$. Then $\mathrm{par}(\psi')$ must be x-cguarded in τ and $\mathrm{par}(\psi') \cap \boldsymbol{d}$ must be x-crelativized in τ, since Y^+ : $\eta \in \tau$ with $\mathrm{par}(\eta) = \boldsymbol{b}$ and $X \neq Y$. This is enough to ensure that $\widetilde{x}^p$: ψ' is x-safe in τ.

We have constructed a realizable $\textsc{y}$-mosaic τ' such that $\tau \to_{\textsc{y}^+ : \eta} \tau'$, so we add τ' to $\mathcal{S}_i$. Repeating this procedure for all such τ (and all such $\textsc{y}^+ : \eta \in \tau$) in $\mathcal{S}_{i-1}$, we will end up with a set $\mathcal{S}_i$ of realizable mosaics such that any existential requirement in $\mathcal{S}_{i-1}$ is fulfilled in $\mathcal{S}_i$. $\square$

We now work towards proving a proposition useful for completeness, a sort of converse of Proposition 4.6. In order to do this, we first prove a lemma about the transfer of formulas between linked mosaics.

LEMMA 4.7. *Let* $\tau_0 \to \cdots \to \tau_k$ *be linked mosaics such that every* τ_i *includes parameters* $\boldsymbol{b}$.

—*Backward transfer property: If* $\textsc{z}^p : \psi(\boldsymbol{b}) \in \tau_k$, *then* $\textsc{x}^p : \psi(\boldsymbol{b}) \in \tau_0$.
—*Forward relativizer transfer property: If* $\textsc{z}^+ : Ub \in \tau_0$ *for* $b \in \boldsymbol{b}$, *then* $\textsc{z}^+ : Ub \in \tau_k$.
—*Forward negative transfer property: If* $\textsc{z}^- : \psi(\boldsymbol{b}) \in \tau_0$, *then* $\textsc{z}^- : \psi(\boldsymbol{b}) \in \tau_k$.

PROOF. We make use of the linking conditions (E1)–(E5).

Backward Transfer Property. The backward transfer property follows easily from property (E5) of linked mosaics.

Forward Relativizer Transfer Property. We proceed by induction on k. The base case for $k = 0$ is immediate. Otherwise, assume $k > 0$. Consider $\tau_0 \to \tau_1$ where τ_0 is an $\textsc{x}$-mosaic relativizing $\boldsymbol{d}$ and τ_1 is a $\textsc{y}$-mosaic.

—If $\textsc{z} = \textsc{y}$, then (E3) implies that $\textsc{z}^+ : Ub \in \tau_1$.
—If $\textsc{z} = \widetilde{\textsc{y}}$ and $\textsc{y} = \textsc{x}$, then $\widetilde{\textsc{y}}^+ : Ub \in \tau_0$ implies that $b \cap \boldsymbol{d}$ is $\textsc{x}$-crelativized in τ_0. If $b \notin \boldsymbol{d}$, then $\widetilde{\textsc{y}}^+ : Ub$ trivially satisfies the requirements for $\textsc{y}$-safety in τ_1. Otherwise, if $b \in \boldsymbol{d}$, then b is $\textsc{x}$-crelativized by some $\textsc{x}^+ : U'b \in \tau_0$. Since $\textsc{x}^+ : U'b$ is trivially $\textsc{y}$-safe in τ_1, (E3) ensures that $\textsc{x}^+ : U'b \in \tau_1$. Hence, b is $\textsc{y}$-crelativized in τ_1. Moreover, b is trivially $\textsc{y}$-cguarded (since there is only one parameter). This means that $\widetilde{\textsc{y}}^+ : Ub$ is $\textsc{y}$-safe in τ_1. Hence, in either case, we can apply (E4) to ensure that $\textsc{z}^+ : Ub \in \tau_1$.
—If $\textsc{z} = \widetilde{\textsc{y}}$ and $\textsc{y} = \widetilde{\textsc{x}}$, then it is trivially $\textsc{y}$-safe: it is $\textsc{y}$-cguarded because there is only one parameter, and it does not need to be $\textsc{y}$-crelativized (since no shared parameters are relativized when moving from an $\textsc{x}$ to $\widetilde{\textsc{x}}$ mosaic). Hence, we can apply (E4) to get $\textsc{z}^+ : Ub \in \tau_1$.

Since we have guaranteed that $\textsc{z}^+ : Ub \in \tau_1$, we can apply the inductive hypothesis to $\tau_1 \to \cdots \to \tau_k$ to get the result.

Forward Negative Transfer Property. We again proceed by induction on k. The base case when $k = 0$ is trivial, so assume $k > 0$. Consider $\tau_0 \to \tau_1$ where τ_0 is an $\textsc{x}$-mosaic relativizing $\boldsymbol{d}$ and τ_1 is a $\textsc{y}$-mosaic. In order to apply the inductive hypothesis, we must show that $\textsc{z}^- : \psi(\boldsymbol{b}) \in \tau_1$.

—If $\textsc{z} = \textsc{y}$, then $\textsc{z}^- : \psi(\boldsymbol{b}) \in \tau_1$ by (E3).
—If $\textsc{z} = \widetilde{\textsc{y}}$ and $\textsc{y} = \textsc{x}$, then $\widetilde{\textsc{y}}^- : \psi(\boldsymbol{b}) \in \tau_0$ implies that $\boldsymbol{b} \cap \boldsymbol{d}$ is $\textsc{x}$-crelativized in τ_0.

By the forward relativizer transfer property, we know that $\boldsymbol{b} \cap \boldsymbol{d}$ is also $\textsc{x}$-crelativized in τ_1 (i.e., we know that all of the $\textsc{x}$-crelativizers in τ_0 are in τ_1). This is enough to ensure that $\widetilde{\textsc{y}}^- : \psi(\boldsymbol{b})$ is $\textsc{y}$-safe in τ_1, so (E4) guarantees that $\textsc{z}^- : \psi(\boldsymbol{b}) \in \tau_1$.

—If $\textsc{z} = \widetilde{\textsc{y}}$ and $\textsc{y} = \widetilde{\textsc{x}}$, then it is trivially $\textsc{y}$-safe: no parameters need to be $\textsc{y}$-crelativized (since no shared parameters are crelativized when moving from an $\textsc{x}$ to $\widetilde{\textsc{x}}$ mosaic). Hence, we can apply (E4) to get $\textsc{z}^- : \psi(\boldsymbol{b}) \in \tau_1$.

Since $\textsc{z}^+ : Ub \in \tau_1$, we can apply the inductive hypothesis to $\tau_1 \to \cdots \to \tau_k$ to get the result. $\square$

We are now ready to show that whenever there is a set of saturated mosaics, we can construct a model—in fact, a tree-like model—satisfying the formulas in these mosaics.

PROPOSITION 4.8. *Let $\mathcal{S} \subseteq \mathcal{M}$ be a saturated set of internally consistent mosaics. Then for all $\tau(\boldsymbol{c}) \in \mathcal{S}$, there is a tree-like model $\mathfrak{M}$ over $\sigma_L \cup \sigma_R \cup \boldsymbol{c}$ such that for all $\mathrm{x}^p : \psi \in \tau$, $\mathfrak{M} \models p\psi$.*

PROOF. Fix $\tau \in \mathcal{S}$. Note that τ is internally consistent by assumption. Our goal is to construct a model $\mathfrak{M}$ witnessing the satisfiability of the formulas in τ. We will use the mosaics in $\mathcal{S}$ as building blocks.

We build inductively a tree decomposition T of $\mathfrak{M}$, together with a map f from nodes v in the tree decomposition to a mosaic in $\mathcal{S}$. We also think of f as a map from the elements in a node v to the parameters in the corresponding mosaic.

At stage 0, let T_0 consist of a single node v_0 containing a copy of the parameters in τ and constants $\boldsymbol{e}$, and set $f(v_0) = \tau$.

At stage i, consider the set

$$S = \{(v, \mathrm{Y}^+ : \eta) : v \text{ is a leaf in } T_{i-1} \text{ and } \mathrm{Y}^+ : \eta(f(\boldsymbol{b})) \in f(v) \text{ is not fulfilled in } f(v)$$
$$\text{for } \eta \text{ of the form } \exists \boldsymbol{x}.\psi'(\boldsymbol{x}f(\boldsymbol{b}))\}.$$

For each $(v, \mathrm{Y}^+ : \eta) \in S$, we know that there is some $\tau' \in \mathcal{S}$ such that $f(v) \rightarrow_{\mathrm{Y}^+ : \eta} \tau'$ (since $\mathcal{S}$ is saturated). We construct T_i by extending T_{i-1} according to the following procedure: for each $(v, \mathrm{Y}^+ : \eta) \in S$, we construct a child v' of v with $T(v') = \boldsymbol{a} \cup \boldsymbol{b} \cup \boldsymbol{e}$ for new elements $\boldsymbol{a}$, and define $f(v') = \tau'$ such that $\mathrm{Y}^+ : \psi'(f(\boldsymbol{a}\boldsymbol{b})) \in \tau'$.

Let T be the limit of this process, and let $\mathfrak{M}$ be the structure with elements $M = \bigcup_{v \in T} T(v)$, and such that $\mathfrak{M} \models \alpha(\boldsymbol{a})$ if and only if there exists v such that $\boldsymbol{a}$ appears in $T(v)$ and $\mathrm{x}^+ : \alpha(f(\boldsymbol{a})) \in f(v)$.

We must now show that for all v in T and for all $\boldsymbol{a}$ in $T(v)$, if $\mathrm{z}^p : \psi(f(\boldsymbol{a})) \in f(v)$, then $\mathfrak{M} \models p\psi(\boldsymbol{a})$. The proof is by induction on the structure of ψ.

—Assume $\mathrm{z}^+ : \psi$ is an atom. Then $\mathfrak{M} \models \psi$ by construction.
—Assume $\mathrm{z}^- : \psi$ is an atom $\alpha(f(\boldsymbol{a}))$, and suppose for the sake of contradiction that $\mathfrak{M} \models \alpha(\boldsymbol{a})$. Then there is some w such that $\mathrm{x}^+ : \alpha(f(\boldsymbol{a})) \in f(w)$ and $f(\boldsymbol{a})$ appears in $f(w')$ for all w' on the path between v and w. Using the backwards transfer property (Lemma 4.7), there must be some node containing both $\mathrm{z}^- : \alpha(f(\boldsymbol{a}))$ and $\mathrm{x}^+ : \alpha(f(\boldsymbol{a}))$, which contradicts internal consistency of mosaics in $\mathcal{S}$.
—Assume $\mathrm{z}^p : \psi$ is of the form $\bigwedge_i \psi_i$ or $\bigvee_i \psi_i$. The result follows using internal consistency and the inductive hypothesis.
—Assume $\mathrm{z}^+ : \psi$ is $\alpha(f(\boldsymbol{b})) \wedge \neg\psi'(f(\boldsymbol{b}))$. By internal consistency, $\mathrm{z}^+ : \alpha(f(\boldsymbol{b})) \in f(v)$ and $\mathrm{z}^- : \psi'(f(\boldsymbol{b})) \in f(v)$. The result follows using the inductive hypothesis.
—Assume $\mathrm{z}^- : \psi$ is $\alpha(f(\boldsymbol{b})) \wedge \neg\psi'(f(\boldsymbol{b}))$, and suppose for the sake of contradiction that $\mathfrak{M} \models \alpha(\boldsymbol{b}) \wedge \neg\psi'(\boldsymbol{b})$. Then there is some w such that $\mathrm{x}^+ : \alpha(f(\boldsymbol{b})) \in f(w)$ and $f(\boldsymbol{b})$ appears in $f(w')$ for all w' on the path between v and w. Using the backwards transfer property (Lemma 4.7), there must be some node containing both $\mathrm{z}^- : \alpha(f(\boldsymbol{b})) \wedge \neg\psi'(f(\boldsymbol{b}))$ and $\mathrm{x}^+ : \alpha(f(\boldsymbol{b}))$. By internal consistency, this implies $\mathrm{z}^+ : \psi'(f(\boldsymbol{b}))$ is also in this node, so $\mathfrak{M} \models \psi'(\boldsymbol{b})$ by the inductive hypothesis, which contradicts $\mathfrak{M} \models \alpha(\boldsymbol{b}) \wedge \neg\psi'(\boldsymbol{b})$.
—Assume $\mathrm{z}^+ : \psi$ is of the form $\mathrm{z}^+ : \exists \boldsymbol{x}.\psi'(\boldsymbol{x}f(\boldsymbol{b}))$. Then either $\mathrm{z}^+ : \psi'(f(\boldsymbol{a}\boldsymbol{b})) \in f(v)$ for some $\boldsymbol{a} \subseteq T(v)$, or there is some child w of v with $f(v) \rightarrow_{\mathrm{z}^+ : \psi} f(w)$ and $\mathrm{z}^+ : \psi'(f(\boldsymbol{a}\boldsymbol{b})) \in f(w)$ for $\boldsymbol{a} \subseteq T(w)$. In either case, the inductive hypothesis implies $\mathfrak{M} \models \psi'(\boldsymbol{a}\boldsymbol{b})$, so $\mathfrak{M} \models \psi$.
—Assume $\mathrm{z}^- : \psi$ is of the form $\mathrm{z}^- : \exists \boldsymbol{x}.\psi'(\boldsymbol{x}f(\boldsymbol{b}))$, and assume for the sake of contradiction that $\mathfrak{M} \models \exists \boldsymbol{x}.\psi'(\boldsymbol{x}\boldsymbol{b})$. By Lemma 4.1, there is a node w with $\boldsymbol{b}\boldsymbol{c} \subseteq T(w)$ and a

nontrivial specialization $\phi'(\boldsymbol{yz})$ of ψ such that

$$\mathfrak{M}, \boldsymbol{bc} \models (\boldsymbol{U}_0(\boldsymbol{z}) \wedge \chi_0(\boldsymbol{yz})) \wedge \bigwedge_{j \in \{1,\ldots,k\}} \exists \boldsymbol{x}_j . \boldsymbol{U}_j(\boldsymbol{x}_j) \wedge \chi_j(\boldsymbol{x}_j \boldsymbol{yz}).$$

By the backward/forward transfer properties (Lemma 4.7), $z^- : \psi \in f(w)$. Let $\boldsymbol{U}_0(\boldsymbol{z}) = U_1 z_1 \wedge \cdots \wedge U_l z_l$. Since $\mathfrak{M} \models \boldsymbol{U}_0(\boldsymbol{c})$, for each $i \in \{1, \ldots, l\}$, there is a node w_i such that $z_i^+ : U_i f(c_i) \in f(w_i)$, and the parameter $f(c_i)$ is in every mosaic on the path between w_i and w. By the backward transfer property and forward relativizer transfer properties (Lemma 4.7), $z_i^+ : U_i f(c_i) \in f(w)$. By internal consistency, $z^- : \phi'(f(\boldsymbol{bc})) \in f(w)$. Likewise, by internal consistency, we have that $z^- : \boldsymbol{U}_0(f(\boldsymbol{c})) \wedge \chi_0(f(\boldsymbol{bc})) \in f(w)$ or $z^- : \exists \boldsymbol{x}_j . \boldsymbol{U}_j(\boldsymbol{x}_j) \wedge \chi_j(\boldsymbol{x}_j f(\boldsymbol{bc})) \in f(w)$ for some j. Since it is a nontrivial specialization, we can apply the inductive hypothesis to this formula in $f(w)$, and get a contradiction.

We remark that this is the only case where we need both forward and backward transfer properties, since we need to be able to move $z^- : \psi$ (and the relativizing formulas) to the node w related to the the nontrivial specialization.

Thus, we have succeeded in constructing a model $\mathfrak{M}$ that satisfies the formulas appearing in the mosaics from $\mathcal{S}$ that were used during the construction—in particular, satisfying the formulas from τ. $\square$

We are now ready to prove Theorem 4.4, the soundness and completeness of the GNF mosaic method. For soundness we must show the following:

If $\varphi_L \wedge \neg \varphi_R$ is satisfiable, then there is a saturated set of internally consistent mosaics from $\mathcal{M}$ that contains some τ such that $L^+ : \varphi_L, R^- : \varphi_R \in \tau$.

Let $\mathrm{par}(\varphi_L) = \mathrm{par}(\varphi_R) = \boldsymbol{z}$. Since $\varphi_L \wedge \neg \varphi_R$ is satisfiable, there is a model $\mathfrak{M}$ over signature $\sigma_L \cup \sigma_R \cup \boldsymbol{z}$ such that $\mathfrak{M} \models \varphi_L \wedge \neg \varphi_R$. Let τ be the L-mosaic of $\boldsymbol{z}$ relativizing $\emptyset$ in $\mathfrak{M}$, that is, the set of formulas from $\mathrm{cl}_L(\boldsymbol{z})$ that are true in $\mathfrak{M}$, and the set of formulas from $\mathrm{cl}_R(\boldsymbol{z})$ that are true in $\mathfrak{M}$ and satisfy the L-safety conditions on R-formulas in an L-mosaic. Since $\varphi_L \wedge \neg \varphi_R$ holds in $\mathfrak{M}$, τ contains $L^+ : \varphi_L$. Since $R^- : \varphi_R$ is trivially L-safe (since it is a negative formula, and there are no relativized parameters), τ also contains $R^- : \varphi_R$. Since τ is a realizable mosaic, Proposition 4.6 implies that there is a saturated set $\mathcal{S}$ of realizable mosaics including τ. Moreover, by Proposition 4.5, $\mathcal{S}$ must contain only internally consistent mosaics, as desired.

For completeness of the mosaic method, we must show the following:

If there is a saturated set of internally consistent mosaics from $\mathcal{M}$ containing τ such that $L^+ : \varphi_L, R^- : \varphi_R \in \tau$, then $\varphi_L \wedge \neg \varphi_R$ is satisfiable.

This follows immediately from Proposition 4.8.

This concludes the proof of Theorem 4.4. Together with the mosaic elimination algorithm described in Section 2.5, this yields an alternative proof of the 2ExpTime-bound on satisfiability testing for GNF.

4.3. Mosaic Interpolants for Relativized GNF

Our goal in this section is to prove a mosaic interpolation lemma (like Lemma 3.3) for the relativized GNF mosaics. We continue to restrict our attention to input formulas φ_L and φ_R that do not use equality, but remark that even with equality-free input, the mosaic interpolants that we construct may use trivial equality guards for negated formulas with at most one free variable.

Fix $\mathbb{U}$-relativized GNF formulas φ_{L} and φ_{R} without equality as in the previous section. That is, φ_{L} and φ_{R} are $\mathbb{U}$-relativized formulas in weak GN-normal form and without equality over signatures σ_{L} and σ_{R}. Moreover, $\mathrm{par}(\varphi_{\mathrm{L}}) = \mathrm{par}(\varphi_{\mathrm{R}})$. Let $\mathcal{M}$ be the set of relativized GNF mosaics for $\varphi_{\mathrm{L}} \wedge \neg \varphi_{\mathrm{R}}$.

Consider the mosaic elimination procedure over $\mathcal{M}$ described in Section 2.5. We write $\mathcal{N}'$ for the set of mosaics for $\varphi_{\mathrm{L}} \wedge \neg \varphi_{\mathrm{R}}$ that are eliminated using this procedure, and $\mathcal{N}_i$ for the set of mosaics that have been removed by stage i.

For an x-mosaic $\tau(\boldsymbol{c})$, let

$$\tau_{\mathrm{X}} := \bigwedge_{\mathrm{X}^p : \psi \in \tau} p\psi \;\wedge\; \bigwedge_{\mathrm{X}^p : \psi \in \mathrm{cl}_{\mathrm{X}}(\boldsymbol{c}) \setminus \tau} \neg p\psi \quad \text{and} \quad \tau_{\widetilde{\mathrm{X}}} := \bigwedge_{\widetilde{\mathrm{X}}^p : \psi \in \tau} p\psi \;\wedge\; \bigwedge_{\substack{\widetilde{\mathrm{X}}^p : \psi \in \mathrm{cl}_{\widetilde{\mathrm{X}}}(\boldsymbol{c}) \setminus \tau \\ \text{s.t. } \widetilde{\mathrm{X}}^p : \psi \text{ is x-safe in } \tau}} \neg p\psi \;.$$

That is, τ_{Y} asserts the Y-formulas in τ, and the negation of the Y-formulas that are not in τ (but satisfy the conditions for being in τ in terms of x-safety if $\mathrm{Y} = \widetilde{\mathrm{X}}$).

We now define mosaic interpolants for every x-mosaic $\tau(\boldsymbol{c}) \in \mathcal{N}'$.

LEMMA 4.9. *For each GNF x-mosaic $\tau(\boldsymbol{c}) \in \mathcal{N}_i$ relativizing $\boldsymbol{d}$, we can construct a DAG representation of a formula $\theta_{\mathrm{X}}^{\tau}$ such that*

(Imp) $\tau_{\mathrm{X}} \models \theta_{\mathrm{X}}^{\tau}$ *and* $\theta_{\mathrm{X}}^{\tau} \models \neg \tau_{\widetilde{\mathrm{X}}}$*;*
(Occ) $\mathrm{occ}(\theta_{\mathrm{X}}^{\tau}) \subseteq \mathrm{occ}(\varphi_{\mathrm{L}}) \cap \mathrm{occ}(\varphi_{\mathrm{R}})$ *if* $\mathrm{X} = \mathrm{L}$*;*
     $\mathrm{occ}(\neg \theta_{\mathrm{X}}^{\tau}) \subseteq \mathrm{occ}(\varphi_{\mathrm{L}}) \cap \mathrm{occ}(\varphi_{\mathrm{R}})$ *if* $\mathrm{X} = \mathrm{R}$*;*
     $\mathrm{con}(\theta_{\mathrm{X}}^{\tau}) \subseteq \mathrm{con}(\varphi_{\mathrm{L}}) \cup \mathrm{con}(\varphi_{\mathrm{R}})$*;*
(Par) $\mathrm{par}(\theta_{\mathrm{X}}^{\tau}) \subseteq \boldsymbol{c}$ *and* $\mathrm{par}(\theta_{\mathrm{X}}^{\tau}) \cap \boldsymbol{d}$ *is x-crelativized in* τ*;*
(Rel) $\theta_{\mathrm{X}}^{\tau}$ *is in $\mathbb{U}$-relativized GNF, even when parameters are viewed as free variables.*

Moreover, if $n = |\varphi_{\mathrm{L}}| + |\varphi_{\mathrm{R}}|$, $w = \mathrm{width}(\varphi_{\mathrm{L}} \wedge \neg \varphi_{\mathrm{R}})$, and $h = \mathrm{rank}_{\mathrm{CQ}}(\varphi_{\mathrm{L}} \wedge \neg \varphi_{\mathrm{R}})$, then there is a DAG representation of $\Theta_i := \{\theta_{\mathrm{Z}}^{\tau} : \tau \text{ is a z-mosaic in } \mathcal{N}_i\}$ with $|\Theta_i| \leq i \cdot 2^{q(n)^{q(hw)}}$ where q is a polynomial function independent of φ_{L} and φ_{R}.

PROOF. We fix some notation. For any tuple $\boldsymbol{b}$ of parameters that is x-cguarded in τ, we define a formula $\mathrm{gdd}_{\mathrm{X}}^{\tau}(\boldsymbol{b})$. If $\boldsymbol{b} = \emptyset$, then we set $\mathrm{gdd}_{\mathrm{X}}^{\tau}(\boldsymbol{b}) := \top$. If $\boldsymbol{b} = b$ (a single parameter), then we set $\mathrm{gdd}_{\mathrm{X}}^{\tau}(\boldsymbol{b}) := (b = b)$ (a trivial equality guard). Otherwise, there is some $\mathrm{X}^+ : \exists \boldsymbol{x}.(\boldsymbol{U}(\boldsymbol{x}) \wedge \beta(\boldsymbol{x}\boldsymbol{b}')) \in \tau$ where β and $\boldsymbol{U}$ are x-catoms, $\boldsymbol{U} \subseteq \mathbb{U}$, $\boldsymbol{b}' \supseteq \boldsymbol{b}$, and $\boldsymbol{b}' \cap \boldsymbol{d}$ is x-crelativized, and we set $\mathrm{gdd}_{\mathrm{X}}^{\tau}(\boldsymbol{b}) := \exists \boldsymbol{x}.(\boldsymbol{U}(\boldsymbol{x}) \wedge \beta(\boldsymbol{x}\boldsymbol{b}'))$. In this last case, when we write $\mathrm{gdd}_{\mathrm{X}}^{\tau}(\boldsymbol{b}) \wedge \neg \psi$ we mean the formula $\exists \boldsymbol{x}.(\boldsymbol{U}(\boldsymbol{x}) \wedge \beta(\boldsymbol{x}\boldsymbol{b}') \wedge \neg \psi)$. Observe that in all of these cases, when $\mathrm{gdd}_{\mathrm{X}}^{\tau}(\boldsymbol{b})$ is defined, we have $\tau_{\mathrm{X}} \models \mathrm{gdd}_{\mathrm{X}}^{\tau}(\boldsymbol{b})$.

The proof proceeds by induction on i, the stage at which τ was eliminated. We intersperse the definitions of the interpolants and proofs of correctness.

Base case. Assume there is an x-mosaic $\tau(\boldsymbol{c}) \in \mathcal{N}_1$. Then τ has an internal inconsistency. Assume one of the following conditions holds for $\mathrm{Y} \in \{\mathrm{L}, \mathrm{R}\}$:

—$\mathrm{Y}^+ : \bot \in \tau$ or $\mathrm{Y}^- : \top \in \tau$;
—$\mathrm{Y}^+ : \alpha(\boldsymbol{b}), \mathrm{Y}^- : \alpha(\boldsymbol{b}) \in \tau$;
—$\mathrm{Y}^+ : \bigwedge_i \psi_i \in \tau$ ($\mathrm{Y}^- : \bigvee_i \psi_i \in \tau$, respectively), but $\mathrm{Y}^+ : \psi_i \notin \tau$ ($\mathrm{Y}^- : \psi_i \notin \tau$, respectively) for some i;
—$\mathrm{Y}^+ : \bigvee_i \psi_i \in \tau$ ($\mathrm{Y}^- : \bigwedge_i \psi_i \in \tau$, respectively), but $\mathrm{Y}^+ : \psi_i \notin \tau$ ($\mathrm{Y}^- : \psi_i \notin \tau$, respctively) for all i;
—$\mathrm{Y}^+ : \alpha(\boldsymbol{b}) \wedge \neg \psi(\boldsymbol{b}) \in \tau$, but $\mathrm{Y}^+ : \psi(\boldsymbol{b}) \notin \tau$ or $\mathrm{Y}^- : \psi(\boldsymbol{b}) \notin \tau$;
—$\mathrm{Y}^+ : \alpha(\boldsymbol{b}) \in \tau$ and $\mathrm{Y}^- : \alpha(\boldsymbol{b}) \wedge \neg \psi(\boldsymbol{b}) \in \tau$, but $\mathrm{Y}^+ : \psi(\boldsymbol{b}) \notin \tau$.

These internal inconsistencies produce very simple interpolants: if $\mathrm{X} = \mathrm{Y}$, then $\theta_{\mathrm{X}}^{\tau} := \bot$; if $\mathrm{X} \neq \mathrm{Y}$, then $\theta_{\mathrm{X}}^{\tau} := \top$.

We now consider the more interesting cases.

Assume $x^+ : \alpha(\boldsymbol{b}), \widetilde{x}^- : \alpha(\boldsymbol{b}) \in \tau$. Then we set $\theta_x^\tau := \alpha(\boldsymbol{b})$. This clearly satisfies (Imp), (Par), and (Rel).

Assume $x^- : \alpha(\boldsymbol{b}), \widetilde{x}^+ : \alpha(\boldsymbol{b}) \in \tau$ then $\theta_x^\tau := \mathrm{gdd}_x^\tau(\boldsymbol{b}) \wedge \neg\alpha(\boldsymbol{b})$. Although $\neg\alpha(\boldsymbol{b})$ would suffice for (Imp), we need to guard this formula with $\mathrm{gdd}_x^\tau(\boldsymbol{b})$ in order to ensure that we remain in GNF. Because $\boldsymbol{b}$ occurs in a $\widetilde{x}^+$-formula in the x-mosaic τ, $\mathrm{gdd}_x^\tau(\boldsymbol{b})$ exists and is entailed by τ_x. Hence, (Imp) holds, and θ_x^τ is in GNF as required for (Rel). Property (Par) is immediate.

Assume $x^+ : \alpha(\boldsymbol{b})$ and $\widetilde{x}^- : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$, but $\widetilde{x}^+ : \psi(\boldsymbol{b}) \notin \tau$. Then we set $\theta_x^\tau := \alpha(\boldsymbol{b})$. (Par) and (Rel) clearly hold. For (Imp), it is clear that $\tau_x \models \theta_x^\tau$. To prove the other part of (Imp), we argue that $\tau_{\widetilde{x}} \models \neg\theta_x^\tau$. Suppose for the sake of contradiction that there is some model $\mathfrak{M}$ of $\tau_{\widetilde{x}} \wedge \alpha(\boldsymbol{b})$. Since $\widetilde{x}^- : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$, this means that $\mathfrak{M} \models \neg(\alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b})) \wedge \alpha(\boldsymbol{b})$, so $\mathfrak{M} \models \psi(\boldsymbol{b})$. By assumption, $\widetilde{x}^+ : \psi(\boldsymbol{b}) \notin \tau$. Notice $x^+ : \alpha(\boldsymbol{b}) \in \tau$ for α appearing positively in an x-formula and negatively in an $\widetilde{x}$-formula. Since occurrences are preserved in moving to mosaics (Proposition 4.2), this means that α is an x-cguard. Likewise, $\boldsymbol{b} \cap \boldsymbol{d}$ is x-crelativized in τ, since $\widetilde{x}^- : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$. Hence, $\boldsymbol{b}$ is x-safe in τ. We can conclude that $\neg\psi(\boldsymbol{b})$ is a conjunct in $\tau_{\widetilde{x}}$, so $\mathfrak{M} \models \neg\psi(\boldsymbol{b})$, a contradiction.

Assuming $\widetilde{x}^+ : \alpha(\boldsymbol{b})$ and $x^- : \alpha(\boldsymbol{b}) \wedge \neg\psi(\boldsymbol{b}) \in \tau$, but $x^+ : \psi(\boldsymbol{b}) \notin \tau$, a similar argument shows that if we set $\theta_x^\tau := \mathrm{gdd}_x^\tau(\boldsymbol{b}) \wedge \neg\alpha(\boldsymbol{b})$ the desired properties are satisfied.

Assume $\widetilde{x}^- : \phi \in \tau$ for ϕ of the form $\exists \boldsymbol{x}.\psi(\boldsymbol{xb}) \in \tau$, but there is some specialization missing from τ. If there is an internal inconsistency in τ like this—and no other type of internal inconsistency—then without loss of generality, we can assume that the specialization uses at most one additional variable c. That is, we can assume there is some specialization $Uc \wedge \chi_0(\boldsymbol{bc}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{bc})$ of ϕ such that $z^+ : Uc \in \tau$ but $\widetilde{x}^- : \chi_0(\boldsymbol{bc}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{bc}) \notin \tau$. Based on the formulas in τ, it is straightforward to check that $\boldsymbol{bc} \cap \boldsymbol{d}$ is x-crelativized. This implies that $\widetilde{x}^- : \chi_0(\boldsymbol{bc}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{bc}) \notin \tau$ is x-safe in τ. Set θ_x^τ to be Uc if $z = x$ and $\top$ otherwise. (Par) is satisfied since if c appears in θ_x^τ, then c is x-crelativized in τ. (Rel) is trivially satisfied. For (Imp), it is clear that $\tau_x \models \theta_x^\tau$. Now suppose for the sake of contradiction that there is some model $\mathfrak{M}$ such that $\mathfrak{M} \models \theta_x^\tau \wedge \tau_{\widetilde{x}}$. Then $\mathfrak{M} \models Uc$ and $\mathfrak{M} \models \chi_0(\boldsymbol{bc}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{bc})$ using the fact that $\widetilde{x}^- : \chi_0(\boldsymbol{bc}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{bc}) \notin \tau$ is x-safe in τ. This means this specialization of ϕ holds in $\mathfrak{M}$. But every specialization of ϕ implies ϕ (Lemma 4.1), a contradiction of $\mathfrak{M} \models \neg\phi$.

Assume $x^- : \phi \in \tau$ for ϕ of the form $\exists \boldsymbol{x}.\psi(\boldsymbol{xb}) \in \tau$. As in the previous case, we can assume there is some specialization $x^- : Uc \wedge \chi_0(\boldsymbol{bc}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{bc})$ of ϕ such that $z^+ : Uc \in \tau$ but $x^- : \chi_0(\boldsymbol{bc}) \wedge \bigwedge_k \exists \boldsymbol{x}_k.\boldsymbol{U}_k(\boldsymbol{x}_k) \wedge \chi_k(\boldsymbol{x}_k\boldsymbol{bc}) \notin \tau$. Set θ_x^τ to $(c = c) \wedge \neg Uc$ if $z = \widetilde{x}$, and $\bot$ otherwise. The proof of correctness is similar to the previous case.

In all of these cases, (Occ) follows from Proposition 4.2.

Inductive Case. Now assume there is an x-mosaic $\tau(\boldsymbol{c}) \in \mathcal{N}_i \setminus \mathcal{N}_{i-1}$ relativizing $\boldsymbol{d} \subseteq \boldsymbol{c}$. Then there is some $y^+ : \eta = y^+ : \exists \boldsymbol{x}.\psi(\boldsymbol{xb}) \in \tau$, and if $\tau \rightarrow_{y^+:\eta} \tau'$, then $\tau' \in \mathcal{N}_{i-1}$.

Let $\boldsymbol{r}$ be the parameters that are relativized in any such τ' (that is, $\boldsymbol{r}$ is $\boldsymbol{a} \cup (\boldsymbol{b} \cap \boldsymbol{d})$ if $y = x$, and $\boldsymbol{r}$ is $\boldsymbol{a}$ if $y = \widetilde{x}$).

We introduce some additional notation. For $z \in \{L, R\}$, we define $\tau \restriction_z^{\boldsymbol{b},\boldsymbol{r}}$ to be the z-mosaic over parameters $\boldsymbol{b}$ and relativizing $\boldsymbol{r}$ that consists of the union of (i) the z-formulas in τ that only use parameters from $\boldsymbol{b}$, and (ii) the $\widetilde{z}$-formulas in τ whose parameters are contained in $\boldsymbol{b}$ and are z-safe in the formulas added in step (i). For mosaics υ, υ' and $z \in \{L, R\}$, we write $\upsilon'(\boldsymbol{ab}) \supseteq_z \upsilon$ if υ' uses only parameters from $\boldsymbol{ab}$ and is obtained from υ by adding only z-formulas (so the $\widetilde{z}$-formulas in υ and υ' are identical). For a y-mosaic υ', we write $\mathrm{rel}_y^{\upsilon'}(\boldsymbol{a})$ for the conjunction of all $U\boldsymbol{a}$ such that $U \in \mathbb{U}$ is a y-catom and $y^+ : U\boldsymbol{a} \in \upsilon'$ (so $y^+ : U\boldsymbol{a}$ is a y-crelativizer in υ').

Assume $\textsc{y} = \textsc{x}$. Then let

$$\bigvee_{\substack{\upsilon'(\boldsymbol{ab}) \supseteq_{\textsc{x}} \tau \restriction_{\textsc{x}}^{\boldsymbol{b},\boldsymbol{r}} \\ \text{s.t. } \tau \to_{\textsc{x}^+:\eta} \upsilon'}} \exists \boldsymbol{x}'. \left(\mathrm{rel}_{\textsc{x}}^{\upsilon'}(\boldsymbol{a})[\boldsymbol{x}/\boldsymbol{a}] \wedge \bigwedge_{\substack{\tau'(\boldsymbol{ab}) \supseteq_{\widetilde{\textsc{x}}} \upsilon' \\ \text{s.t. } \tau \to_{\textsc{x}^+:\eta} \tau'}} \theta_{\textsc{x}}^{\tau'}[\boldsymbol{x}/\boldsymbol{a}] \right),$$

where $\boldsymbol{x}' \subseteq \boldsymbol{x}$ is the subset of $\boldsymbol{x}$ appearing in $\mathrm{rel}_{\textsc{x}}^{\upsilon'}(\boldsymbol{a})[\boldsymbol{x}/\boldsymbol{a}]$.

We now begin the proof of correctness, starting with (Imp).

Assume there is a model $\mathfrak{M}$ for $\tau_{\textsc{x}}$. Since $\textsc{x}^+ : \eta \in \tau$, this means that $\mathfrak{M} \models \exists \boldsymbol{x}.\psi(\boldsymbol{xb})$, so there are elements $\boldsymbol{a}' \in M$ and an expansion $\mathfrak{M}'$ of $\mathfrak{M}$ with the interpretation $\boldsymbol{a}^{\mathfrak{M}'} := \boldsymbol{a}'$ such that $\mathfrak{M}' \models \psi(\boldsymbol{ab})$. Take υ' to be the union of $\tau \restriction_{\textsc{x}}^{\boldsymbol{b},\boldsymbol{r}}$ and the set of formulas $\textsc{x}^p : \psi'$ in $\mathrm{cl}_{\textsc{x}}(\boldsymbol{ab})$ such that $\mathfrak{M}' \models p\psi'$. It is straightforward to check that $\tau \to_{\textsc{x}^+:\eta} \upsilon'$. By construction, $\mathfrak{M}' \models \upsilon'_{\textsc{x}}$ and $\mathfrak{M}' \models \mathrm{rel}_{\textsc{x}}^{\upsilon'}(\boldsymbol{a})$. Moreover, for any choice of $\tau'(\boldsymbol{ab}) \supseteq_{\widetilde{\textsc{x}}} \upsilon'$, $\mathfrak{M}' \models \tau'_{\textsc{x}}$ (since the x-formulas are identical in υ' and τ'). By the inductive hypothesis, this means that $\mathfrak{M}' \models \theta_{\textsc{x}}^{\tau'}$. Hence, $\mathfrak{M}' \models \mathrm{rel}_{\textsc{x}}^{\upsilon'}(\boldsymbol{a}) \wedge \bigwedge_{\tau'(\boldsymbol{ab}) \supseteq_{\widetilde{\textsc{x}}} \upsilon' \text{ s.t. } \tau \to_{\textsc{x}^+:\eta} \tau'} \theta_{\textsc{x}}^{\tau'}$ and $\mathfrak{M} \models \theta_{\textsc{x}}^\tau$.

Now assume that there is a model $\mathfrak{M}$ of $\theta_{\textsc{x}}^\tau$. Then there is some $\upsilon'(\boldsymbol{ab}) \supseteq_{\textsc{x}} \tau \restriction_{\textsc{x}}^{\boldsymbol{b},\boldsymbol{r}}$ with $\tau \to_{\textsc{x}^+:\eta} \upsilon'$, elements $\boldsymbol{a}' \in M$, and an expansion $\mathfrak{M}'$ of $\mathfrak{M}$ with the interpretation $\boldsymbol{a}^{\mathfrak{M}'} := \boldsymbol{a}'$, such that $\mathfrak{M}' \models \mathrm{rel}_{\textsc{x}}^{\upsilon'}(\boldsymbol{a}) \wedge \bigwedge_{\tau'(\boldsymbol{ab}) \supseteq_{\widetilde{\textsc{x}}} \upsilon' \text{ s.t. } \tau \to_{\textsc{x}^+:\eta} \tau'} \theta_{\textsc{x}}^{\tau'}$. Note that for all $\tau'(\boldsymbol{ab}) \supseteq_{\widetilde{\textsc{x}}} \upsilon'$ with $\tau \to_{\textsc{x}^+:\eta} \tau'$, $\mathfrak{M}' \models \theta_{\textsc{x}}^{\tau'}$ and consequently, the inductive hypothesis implies $\mathfrak{M}' \models \neg\tau'_{\widetilde{\textsc{x}}}$. In particular, consider $\tau' := \upsilon' \cup S' \supseteq_{\widetilde{\textsc{x}}} \upsilon'$ where S' is the set of formulas $\widetilde{\textsc{x}}^p : \psi'$ in $\mathrm{cl}_{\widetilde{\textsc{x}}}(\boldsymbol{ab})$ such that $\mathfrak{M}' \models p\psi'$, $\mathrm{par}(\psi') \cap \boldsymbol{a} \neq \emptyset$, and $\widetilde{\textsc{x}}^p : \psi'$ is x-safe in υ'. It is clear that $\tau'(\boldsymbol{ab}) \supseteq_{\widetilde{\textsc{x}}} \upsilon'$ and $\tau \to_{\textsc{x}^+:\eta} \tau'$, so $\mathfrak{M}' \models \neg\tau'_{\widetilde{\textsc{x}}}$ as observed previously. Since $\mathfrak{M}' \models \neg\tau'_{\widetilde{\textsc{x}}}$, there is some conjunct χ in $\tau'_{\widetilde{\textsc{x}}}$ such that $\mathfrak{M}' \models \neg\chi$. Consider some conjunct $\chi(\boldsymbol{ab})$ in $\tau'_{\widetilde{\textsc{x}}}$ that actually uses some parameters from $\boldsymbol{a}$. Then by choice of S', $\mathfrak{M}' \models \chi(\boldsymbol{ab})$, so this formula cannot witness the fact that $\mathfrak{M}' \models \neg\tau'_{\widetilde{\textsc{x}}}$. This means there must be some conjunct $\chi(\boldsymbol{b})$ in $\tau'_{\widetilde{\textsc{x}}}$ that only uses parameters from $\boldsymbol{b}$ such that $\mathfrak{M}' \models \neg\chi(\boldsymbol{b})$. By definition of $\tau \to_{\textsc{x}^+:\eta} \tau'$, $\chi(\boldsymbol{b})$ must also be a conjunct in $\tau_{\widetilde{\textsc{x}}}$. Hence, $\mathfrak{M} \models \neg\tau_{\widetilde{\textsc{x}}}$.

For (Occ), the desired occurrences of constants and polarities follow from the inductive hypothesis and the fact that occurrences are preserved in mosaics (Proposition 4.2).

The last thing to prove is (Par) and (Rel). The inductively defined interpolants $\theta_{\textsc{x}}^{\tau'}$ only use parameters in τ'. Moreover, any parameters in $\boldsymbol{a} \cup (\boldsymbol{b} \cap \boldsymbol{d})$ are x-crelativized in τ'. By construction, τ' and υ' have the same x-formulas, so these x-crelativizers are also in υ'. In particular, the x-crelativizers for parameters in $\boldsymbol{a}$ appear as conjuncts in $\mathrm{rel}_{\textsc{x}}^{\upsilon'}(\boldsymbol{a})$. This means that any parameters from $\boldsymbol{a}$ in $\theta_{\textsc{x}}^{\tau'}$ are removed using existential quantification, and this quantification is $\mathbb{U}$-relativized by $\mathrm{rel}_{\textsc{x}}^{\upsilon'}(\boldsymbol{a})$. The remaining parameters in $\theta_{\textsc{x}}^\tau$ are from $\boldsymbol{b}$, so $\mathrm{par}(\theta_{\textsc{x}}^\tau) \subseteq \boldsymbol{c}$. Moreover, any x-relativizers for $\boldsymbol{b} \cap \boldsymbol{d}$ in υ' must also be in τ (since $\tau \to_{\textsc{x}^+:\eta} \upsilon'$ and parameters $\boldsymbol{b}$ occur in both τ and υ'). Hence, $\mathrm{par}(\theta_{\textsc{x}}^\tau) \cap \boldsymbol{d}$ is x-crelativized in τ. Overall, this means that (Par) and (Rel) hold.

Assume $\textsc{y} = \widetilde{\textsc{x}}$. Then let

$$\mathrm{gdd}_{\textsc{x}}^\tau(\boldsymbol{b}) \wedge \neg \left(\bigvee_{\substack{\upsilon'(\boldsymbol{ab}) \supseteq_{\widetilde{\textsc{x}}} \tau \restriction_{\widetilde{\textsc{x}}}^{\boldsymbol{b},\boldsymbol{r}} \\ \text{s.t. } \tau \to_{\widetilde{\textsc{x}}^+:\eta} \upsilon'}} \exists \boldsymbol{x}'. \left(\mathrm{rel}_{\widetilde{\textsc{x}}}^{\upsilon'}(\boldsymbol{a})[\boldsymbol{x}/\boldsymbol{a}] \wedge \bigwedge_{\substack{\tau'(\boldsymbol{ab}) \supseteq_{\textsc{x}} \upsilon' \\ \text{s.t. } \tau \to_{\widetilde{\textsc{x}}^+:\eta} \tau'}} \theta_{\widetilde{\textsc{x}}}^{\tau'}[\boldsymbol{x}/\boldsymbol{a}] \right) \right),$$

where $\boldsymbol{x}' \subseteq \boldsymbol{x}$ is the subset of $\boldsymbol{x}$ appearing in $\mathrm{rel}_{\widetilde{\textsc{x}}}^{\upsilon'}(\boldsymbol{a})[\boldsymbol{x}/\boldsymbol{a}]$.

We prove (Imp) by showing $\tau_{\widetilde{\textsc{x}}} \models \neg\theta_{\textsc{x}}^\tau$ and $\neg\theta_{\textsc{x}}^\tau \models \neg\tau_{\textsc{x}}$.

Assume there is a model $\mathfrak{M}$ for $\tau_{\widetilde{\textsc{x}}}$. Since $\widetilde{\textsc{x}}^+ : \eta \in \tau$, this means that $\mathfrak{M} \models \exists \boldsymbol{x}.\psi(\boldsymbol{xb})$, so there are elements $\boldsymbol{a}' \in M$ and an expansion $\mathfrak{M}'$ of $\mathfrak{M}$ with the interpretation

$\boldsymbol{a}^{\mathfrak{M}'} := \boldsymbol{a}$ such that $\mathfrak{M}' \models \psi(\boldsymbol{ab})$. Take υ' to be the union of $\tau \restriction_{\widetilde{\mathbf{x}}}^{\boldsymbol{b};\boldsymbol{r}}$ and the set of formulas $\widetilde{\mathbf{x}}^p : \psi'$ in $\mathrm{cl}_{\widetilde{\mathbf{x}}}(\boldsymbol{ab})$ such that $\mathfrak{M}' \models p\psi'$. It is straightforward to check that $\tau \to_{\widetilde{\mathbf{x}}^+:\eta} \upsilon'$. By construction, $\mathfrak{M}' \models \upsilon'_{\widetilde{\mathbf{x}}}$ and $\mathfrak{M}' \models \mathrm{rel}_{\widetilde{\mathbf{x}}}^{\upsilon'}(\boldsymbol{a})$. Moreover, for any choice of $\tau'(\boldsymbol{ab}) \supseteq_{\mathbf{x}} \upsilon'$, $\mathfrak{M}' \models \tau'_{\widetilde{\mathbf{x}}}$ (since the $\widetilde{\mathbf{x}}$-formulas are identical in υ' and τ'). By the inductive hypothesis, this means that $\mathfrak{M}' \models \theta_{\widetilde{\mathbf{x}}}^{\tau'}$. Hence, $\mathfrak{M}' \models \mathrm{rel}_{\widetilde{\mathbf{x}}}^{\upsilon'}(\boldsymbol{a}) \wedge \bigwedge_{\tau'(\boldsymbol{ab}) \supseteq_{\mathbf{x}} \upsilon' \text{ s.t. } \tau \to_{\widetilde{\mathbf{x}}^+:\eta} \tau'} \theta_{\widetilde{\mathbf{x}}}^{\tau'}$ so $\mathfrak{M} \models \neg\theta_{\mathbf{x}}^{\tau}$.

Now assume that there is a model $\mathfrak{M}$ of $\neg\theta_{\mathbf{x}}^{\tau}$.

If $\mathfrak{M} \models \neg\mathrm{gdd}_{\mathbf{x}}^{\tau}(\boldsymbol{b})$, then it must be the case that $|\boldsymbol{b}| > 1$ (since $\neg\top$ and $\neg(b=b)$ cannot hold in $\mathfrak{M}$). Thus, $\mathfrak{M} \models \neg\tau_{\mathbf{x}}$ since $\mathbf{x}^+ : \mathrm{gdd}_{\mathbf{x}}^{\tau}(\boldsymbol{b}) \in \tau$ (guaranteed since $\widetilde{\mathbf{x}}^+ : \eta(\boldsymbol{b}) \in \tau$).

Otherwise, there is some $\upsilon'(\boldsymbol{ab}) \supseteq_{\widetilde{\mathbf{x}}} \tau \restriction_{\widetilde{\mathbf{x}}}^{\boldsymbol{b};\boldsymbol{r}}$ with $\tau \to_{\widetilde{\mathbf{x}}^+:\eta} \upsilon'$, elements $\boldsymbol{a}' \in M$, and an expansion $\mathfrak{M}'$ of $\mathfrak{M}$ with the interpretation $\boldsymbol{a}^{\mathfrak{M}'} := \boldsymbol{a}'$, such that $\mathfrak{M}' \models \mathrm{rel}_{\widetilde{\mathbf{x}}}^{\upsilon'}(\boldsymbol{a}) \wedge \bigwedge_{\tau'(\boldsymbol{ab}) \supseteq_{\mathbf{x}} \upsilon' \text{ s.t. } \tau \to_{\widetilde{\mathbf{x}}^+:\eta} \tau'} \theta_{\widetilde{\mathbf{x}}}^{\tau'}$.

Note that for all $\tau'(\boldsymbol{ab}) \supseteq_{\mathbf{x}} \upsilon'$ with $\tau \to_{\widetilde{\mathbf{x}}^+:\eta} \tau'$, $\mathfrak{M}' \models \theta_{\widetilde{\mathbf{x}}}^{\tau'}$ and consequently, the inductive hypothesis implies $\mathfrak{M}' \models \neg\tau'_{\mathbf{x}}$. In particular, consider $\tau' := \upsilon' \cup S' \supseteq_{\mathbf{x}} \upsilon'$ where S' is the set of formulas $\mathbf{x}^p : \psi'$ in $\mathrm{cl}_{\mathbf{x}}(\boldsymbol{ab})$ such that $\mathfrak{M}' \models p\psi'$, $\mathrm{par}(\psi') \cap \boldsymbol{a} \neq \emptyset$, and $\mathbf{x}^p : \psi'$ is $\widetilde{\mathbf{x}}$-safe in υ'. It is clear that $\tau'(\boldsymbol{ab}) \supseteq_{\mathbf{x}} \upsilon'$ and $\tau \to_{\widetilde{\mathbf{x}}^+:\eta} \tau'$, so $\mathfrak{M}' \models \neg\tau'_{\mathbf{x}}$ as observed previously.

Since $\mathfrak{M}' \models \neg\tau'_{\mathbf{x}}$, there is some conjunct χ in $\tau'_{\mathbf{x}}$ such that $\mathfrak{M}' \models \neg\chi$. Consider some conjunct $\chi(\boldsymbol{ab})$ in $\tau'_{\mathbf{x}}$ that actually uses some parameters from $\boldsymbol{a}$. Then by choice of S', $\mathfrak{M}' \models \chi(\boldsymbol{ab})$, so this formula cannot witness the fact that $\mathfrak{M}' \models \neg\tau'_{\mathbf{x}}$. This means there must be some conjunct $\chi(\boldsymbol{b})$ in $\tau'_{\mathbf{x}}$ that only uses parameters from $\boldsymbol{b}$ such that $\mathfrak{M}' \models \neg\chi(\boldsymbol{b})$. By definition of $\tau \to_{\widetilde{\mathbf{x}}^+:\eta} \tau'$, $\chi(\boldsymbol{b})$ must also be a conjunct in $\tau_{\mathbf{x}}$. Hence, $\mathfrak{M} \models \neg\tau_{\mathbf{x}}$.

For (Occ), the desired occurrences of constants and polarities follow from the inductive hypothesis and the fact that occurrences are preserved in mosaics (Proposition 4.2). As an example, we give the proof for (Occ) if $\mathbf{x} = \mathbf{R}$ and $\mathrm{gdd}_{\mathbf{x}}^{\tau}(\boldsymbol{b})$ corresponds to an atomic formula. Let $(\alpha, p) \in \mathrm{occ}(\neg\theta_{\mathbf{x}}^{\tau})$ (we abuse notation slightly and write this to mean that the relation S used in α satisfies $(S, p) \in \mathrm{occ}(\neg\theta_{\mathbf{x}}^{\tau})$). If (α, p) is $(\mathrm{gdd}_{\mathbf{x}}^{\tau}(\boldsymbol{b}), -)$ (since $\mathrm{gdd}_{\mathbf{x}}^{\tau}(\boldsymbol{b})$ occurs negatively in $\neg\theta_{\mathbf{x}}^{\tau}$), then by definition of an $\mathbf{x}$-catom for $\mathbf{x} = \mathbf{R}$, $(\mathrm{gdd}_{\mathbf{x}}^{\tau}(\boldsymbol{b}), -) \in \mathrm{occ}(\varphi_{\mathbf{L}}) \cap \mathrm{occ}(\varphi_{\mathbf{R}})$. If (α, p) is $(U, +)$ for a conjunct Ua in $\mathrm{rel}_{\widetilde{\mathbf{x}}}^{\upsilon'}(\boldsymbol{a})$, then by definition of a $\widetilde{\mathbf{x}}$-crelativizer for $\mathbf{x} = \mathbf{R}$, we must have $(U, +) \in \mathrm{occ}(\varphi_{\mathbf{L}}) \cap \mathrm{occ}(\varphi_{\mathbf{R}})$. Otherwise, (α, p) comes from $\mathrm{occ}(\theta_{\widetilde{\mathbf{x}}}^{\tau'})$ for some τ' such that $\tau \to_{\widetilde{\mathbf{x}}^+:\eta} \tau'$. By the inductive hypothesis, this means $(\alpha, p) \in \mathrm{occ}(\varphi_{\mathbf{L}}) \cap \mathrm{occ}(\varphi_{\mathbf{R}})$. Overall, this means that $\mathrm{occ}(\neg\theta_{\mathbf{x}}^{\tau}) \subseteq \mathrm{occ}(\varphi_{\mathbf{L}}) \cap \mathrm{occ}(\varphi_{\mathbf{R}})$ so (Occ) holds for $\theta_{\mathbf{x}}^{\tau}$.

The last thing to prove is (Par) and (Rel). The inductively defined interpolants $\theta_{\widetilde{\mathbf{x}}}^{\tau'}$ only use parameters in τ'. Moreover, any parameters in $\theta_{\widetilde{\mathbf{x}}}^{\tau'}$ from $\boldsymbol{a}$ are $\widetilde{\mathbf{x}}$-crelativized in τ'. By construction, τ' and υ' have the same $\widetilde{\mathbf{x}}$-formulas, so these $\widetilde{\mathbf{x}}$-crelativizers are also in υ'. In particular, the $\widetilde{\mathbf{x}}$-crelativizers for parameters in $\boldsymbol{a}$ appear as conjuncts in $\mathrm{rel}_{\widetilde{\mathbf{x}}}^{\upsilon'}(\boldsymbol{a})$. This means that any parameters from $\boldsymbol{a}$ in $\theta_{\mathbf{x}}^{\tau'}$ are removed using existential quantification, and this quantification is $\mathbb{U}$-relativized by $\mathrm{rel}_{\widetilde{\mathbf{x}}}^{\upsilon'}(\boldsymbol{a})$. The remaining parameters in the negated subformula of $\theta_{\mathbf{x}}^{\tau}$ are from $\boldsymbol{b}$. Because $\widetilde{\mathbf{x}}^+ : \eta$ is a $\widetilde{\mathbf{x}}^+$-formula using $\boldsymbol{b}$ in an $\mathbf{x}$-mosaic, we know that any parameters in $\boldsymbol{b} \cap \boldsymbol{d}$ are $\mathbf{x}$-crelativized in τ and $\boldsymbol{b}$ is $\mathbf{x}$-cguarded in τ. This means that $\mathrm{gdd}_{\mathbf{x}}^{\tau}(\boldsymbol{b})$ is defined and can guard the parameters $\boldsymbol{b}$ in the negated subformula. Overall, this means (Par) and (Rel) hold.

Size of Interpolants. Finally, we bound the size of the shared DAG representation Θ_i of all $\theta_{\mathbf{x}}^{\tau}$ coming from $\tau \in \mathcal{N}_i$, such that $|\Theta_i| \leq i \cdot 2^{q(n)^{q(hw)}}$ for some polynomial function q independent of $\varphi_{\mathbf{L}}$ and $\varphi_{\mathbf{R}}$. Let $n = |\varphi_{\mathbf{L}}| + |\varphi_{\mathbf{R}}|$, $w = \mathrm{width}(\varphi_{\mathbf{L}} \wedge \neg\varphi_{\mathbf{R}})$, and $h = \mathrm{rank}_{\mathrm{CQ}}(\varphi_{\mathbf{L}} \wedge \neg\varphi_{\mathbf{R}})$. Let $f(n, h, w) = 2^{p(n)^{p(hw)}}$ be the bound on the total number of mosaics for $\varphi_{\mathbf{L}}$ and $\varphi_{\mathbf{R}}$ in $\mathcal{M}$ given by Proposition 4.3.

Let Θ_0 be the empty graph. We claim that at each stage $i > 0$ in the interpolant construction, the number of nodes and edges added to Θ_{i-1} in order to represent the new formulas added in stage i is polynomial in $f(n, h, w)$. The desired bound for the size of the DAG representations follows from this claim.

First consider $i = 1$. To build Θ_1, we start with the graph consisting of nodes labeled by atomic formulas (without provenance labels) from $\mathrm{cl}_{\mathrm{L}}(P) \cup \mathrm{cl}_{\mathrm{R}}(P)$, as well as nodes labeled with $\top$ and $\bot$. The number of nodes in this graph is at most $f(n, h, w)$. Constructing a formula $\theta_{\mathrm{x}}^{\tau}$ for $\tau \in \mathcal{N}_1$ results in a formula with at most $w + 2$ nonatomic subformulas (for a formula of the form $\exists \boldsymbol{x}.\beta(\boldsymbol{xb}) \wedge \neg\alpha(\boldsymbol{b})$), so $w + 2$ new nodes may be needed in the DAG representation in addition to the nodes for atomic formulas. Since $|\mathcal{N}_1| \leq f(n, h, w)$, this means the total size of Θ_1 is quadratic in $f(n, h, w)$.

Now consider $i > 1$. For each $\tau \in \mathcal{N}_i$, the formula $\theta_{\mathrm{x}}^{\tau}$ results in formulas with (at most) the following new subformulas:

—w subformulas from existentially quantified variables in gdd;
—three subformulas from the guarded negation; and
—$f(n, h, w)$ subformulas from the disjunction over υ', each with w existential subformulas, w subformulas from rel, and $f(n, h, w)$ subformulas from the conjunction over τ'.

Since $|\mathcal{N}_i| \leq f(n, h, w)$, this means that the number of new nodes and edges added to the DAG representation of Θ_{i-1} in order to construct the DAG representation of Θ_i is polynomial (cubic) in $f(n, h, w)$.

This concludes the proof of Lemma 4.9. □

4.4. Constructive Interpolation using Mosaics

We are now ready to prove the main result of this work, effective interpolation for the guarded negation fragment:

THEOREM 4.10. *Let ψ_{L} and ψ_{R} be GNF (equality-free UNF, respectively) formulas over signatures σ_{L} and σ_{R}, respectively. If $\psi_{\mathrm{L}} \models \psi_{\mathrm{R}}$ and $|\psi_{\mathrm{L}}| + |\psi_{\mathrm{R}}| = n$, then we can construct a DAG representation of a GNF (UNF, respectively) interpolant θ such that*

—$\psi_{\mathrm{L}} \models \theta$ *and* $\theta \models \psi_{\mathrm{R}}$;
—$\mathrm{occ}(\theta) \subseteq \mathrm{occ}(\psi_{\mathrm{L}}) \cap \mathrm{occ}(\psi_{\mathrm{R}})$;
—$\mathrm{free}(\theta) \subseteq \mathrm{free}(\psi_{\mathrm{L}}) \cap \mathrm{free}(\psi_{\mathrm{R}})$;
—$\mathrm{con}(\theta) \subseteq \mathrm{con}(\psi_{\mathrm{L}}) \cup \mathrm{con}(\psi_{\mathrm{R}})$;
—*if ψ_{L} and ψ_{R} are $\mathbb{U}$-relativized, for $\mathbb{U}$ a distinguished set of unary relations from $\sigma_{\mathrm{L}} \cup \sigma_{\mathrm{R}}$, then θ is $\mathbb{U}$-relativized;*
—*the DAG representation of θ is of size at most $2^{2^{p(n)}}$ for some polynomial function p independent of ψ_{L} and ψ_{R} (and this can be improved to a size of at most $2^{p(n)}$ when ψ_{L} and ψ_{R} are GF sentences without equality and the bound on the arity of relations is fixed).*

We emphasize that the running time of the algorithm is polynomial in the output size, so in general, computing the GNF interpolant can be done in 2ExpTime (and in ExpTime when starting with GF sentences with relations of bounded arity).

Using this interpolation method, we can also construct an *FO* interpolant that can be represented by a doubly exponential sized formula—for example, by applying the preceding theorem, and then converting efficiently from a DAG representation to a formula [Avigad 2003]. As another corollary of Theorem 4.10, we get doubly exponential bounds on the DAG size of explicit GNF definitions coming from Projective Beth Definability. The argument is standard (see Hoogland [2000]).

We now prove Theorem 4.10. with the help of the mosaic interpolation lemma. We consider various types of input, but continue to assume that the input formulas ψ_{L} and ψ_{R} are equality-free and free(ψ_{L}) = free(ψ_{R}) = $\boldsymbol{z}$. We will discuss how to handle equality and the case of free(ψ_{L}) $\neq$ free(ψ_{R}) in Sections 4.5 and 4.6.

Relativized GNF *Input in Weak* GN-*Normal Form.* Assume ψ_{L} and ψ_{R} are $\mathbb{U}$-relativized GNF formulas without equality, and are already in weak GN-normal form.

Let $\varphi_{\text{L}} := \psi_{\text{L}}$ and $\varphi_{\text{R}} := \psi_{\text{R}}$.

Use the mosaic elimination procedure over the set $\mathcal{M}$ of relativized GNF mosaics for $\varphi_{\text{L}} \wedge \neg\varphi_{\text{R}}$. Then construct mosaic interpolants for the eliminated mosaics $\mathcal{N}' \subseteq \mathcal{M}$ using Lemma 4.9, and let

$$\theta := \bigvee_{\tau(\boldsymbol{z}) \supseteq_{\text{L}} \{\text{L}^+ : \varphi_{\text{L}}\}} \bigwedge_{\substack{\tau'(\boldsymbol{z}) \supseteq_{\text{R}} \tau \\ \text{s.t. } \tau' \in \mathcal{N}'}} \theta_{\text{L}}^{\tau'},$$

where $\mathcal{N}'$ is the set of eliminated mosaics and we write $\upsilon'(\boldsymbol{z}) \supseteq_{\text{Z}} \upsilon$ if υ' is over parameters $\boldsymbol{z}$ and relativizing $\emptyset$, and is obtained from υ be adding only z-formulas.

We claim θ is a GNF interpolant for $\varphi_{\text{L}} \models \varphi_{\text{R}}$ satisfying the properties in Theorem 4.10. We first prove $\varphi_{\text{L}} \models \theta \models \varphi_{\text{R}}$.

Assume there is a structure $\mathfrak{M}$ such that $\mathfrak{M} \models \varphi_{\text{L}}$. Let $\tau(\boldsymbol{z}) \supseteq_{\text{L}} \{\text{L}^+ : \varphi_{\text{L}}\}$ be the set of formulas from $\text{cl}_{\text{L}}(\boldsymbol{z})$ that are true in $\mathfrak{M}$. For any $\tau'(\boldsymbol{z}) \supseteq_{\text{R}} \tau$, $\mathfrak{M} \models \tau_{\text{L}}'$. By Lemma 4.9, this means that $\mathfrak{M} \models \theta_{\text{L}}^{\tau'}$. Overall, $\mathfrak{M} \models \theta$.

Now assume there is a structure $\mathfrak{M}$ such that $\mathfrak{M} \models \theta$. Then there is some $\tau(\boldsymbol{z}) \supseteq_{\text{L}} \{\text{L}^+ : \varphi_{\text{L}}\}$ such that $\mathfrak{M} \models \bigwedge_{\substack{\tau'(\boldsymbol{z}) \supseteq_{\text{R}} \tau \\ \text{s.t. } \tau' \in \mathcal{N}'}} \theta_{\text{L}}^{\tau'}$. Let S' be the set of formulas $\text{R}^p : \psi$ in $\text{cl}_{\text{R}}(\boldsymbol{z})$ such that $\mathfrak{M} \models p\psi$, and $\text{R}^p : \psi$ is L-safe in τ. Consider the L-mosaic over $\boldsymbol{z}$ (relativizing $\emptyset$) such that $\tau' := \tau \cup S' \cup \{\text{R}^- : \varphi_{\text{R}}\}$. By completeness of the mosaic system (Proposition 4.8, which is part of Theorem 4.4), every mosaic containing $\text{L}^+ : \varphi_{\text{L}}$ and $\text{R}^- : \varphi_{\text{R}}$ must be in $\mathcal{N}'$. Hence, $\tau'(\boldsymbol{z}) \supseteq_{\text{R}} \tau$ and $\tau' \in \mathcal{N}'$, so $\mathfrak{M} \models \theta_{\text{L}}^{\tau'}$. By Lemma 4.9, this implies that $\mathfrak{M} \models \neg\tau_{\text{R}}'$. But every formula $\text{R}^p : \psi$ in S' was chosen such that $\mathfrak{M} \models p\psi$. The only way that $\mathfrak{M} \models \neg\tau_{\text{R}}'$ is if $\mathfrak{M} \models \varphi_{\text{R}}$.

Since each $\theta_{\text{L}}^{\tau'}$ is in $\mathbb{U}$-relativized GNF by Lemma 4.9, this formula is in $\mathbb{U}$-relativized GNF. Note that it is not necessarily in GN-normal form or weak GN-normal form.

The other required properties follow from Lemma 4.9.

Let if $n = |\varphi_{\text{L}}| + |\varphi_{\text{R}}|$, $w = \text{width}(\varphi_{\text{L}} \wedge \neg\varphi_{\text{R}})$, and $h = \text{rank}_{\text{CQ}}(\varphi_{\text{L}} \wedge \neg\varphi_{\text{R}})$. Since the combined size of the DAG representation of all $\theta_{\text{L}}^{\tau'}$ used in θ is at most $2^{q(n)^{q(hw)}}$ (for polynomial function q given by Lemma 4.9 that are independent of the input formulas), and there are at most $(2^{q(n)^{q(hw)}})^2$ new nodes that need to be added to this DAG representation to construct θ, the overall size of the DAG representation of θ is at most $2^{r(n)^{r(hw)}}$ for some polynomial function r independent of the input formulas. Since $w, h \leq n$, this means that overall, the size of this DAG representation for θ is at most doubly exponential in n.

Arbitrary Relativized GNF *Input (not necessarily in weak* GN-*normal form).* Assume ψ_{L} and ψ_{R} are in $\mathbb{U}$-relativized GNF, but are not in weak GN-normal form. Using Lemma 2.1, we convert these formulas to equivalent φ_{L} and φ_{R} in weak GN-normal form. This conversion can be done in a way that preserves $\mathbb{U}$-relativization, so we can then proceed to find the interpolant for φ_{L} and φ_{R} as described in the previous case. Because the conversion to weak GN-normal form keeps the width and CQ rank bounded by n and the overall size exponential in n, the size of the DAG representation of the interpolant is still at most doubly exponential in n.

GNF *Input with Ordinary (Unrelativized) Quantifiers.* Assume ψ_{L} and ψ_{R} are in GNF and have ordinary (unrelativized) quantifiers. We introduce a new unary relation U that is not in σ_{L} or σ_{R}. Let ψ_{X}^U be the result of relativizing all quantifiers in ψ_{X} to U. Consider the validity $\bigwedge_{z \in \boldsymbol{z}} Uz \wedge \bigwedge_{e \in \boldsymbol{e}} Ue \wedge \psi_{\mathrm{L}}^U \models \psi_{\mathrm{R}}^U$ where $\boldsymbol{z} := \mathrm{par}(\psi_{\mathrm{L}}) = \mathrm{par}(\psi_{\mathrm{R}})$ and $\boldsymbol{e} := \mathrm{con}(\psi_{\mathrm{L}}) \cup \mathrm{con}(\psi_{\mathrm{R}})$. Then apply the interpolation result for this relativized GNF input, which yields a $\mathbb{U}$-relativized interpolant θ^U for $\mathbb{U} = \{U\}$. Replacing all occurrences of Ux in θ^U with $\top$ results in a GNF interpolant θ for the original validity.

UNF *Input.* Assume ψ_{L} and ψ_{R} are in UNF (and for simplicity that they are already in weak GN-normal form). Observe that in the interpolation construction in Lemma 4.9, the only guards for negation that are used for an interpolant θ_{X}^τ are equality guards of the form $x = x$ or x-cguards in τ. By definition, x-cguards must have appeared as guards in one of the original formulas ψ_{L} and ψ_{R}. Since the only guards in the original formulas were unary guards, this means that θ_{X}^τ uses only unary negation.
For UNF input that is not in weak GN-normal form, the conversion into weak GN-normal form using Lemma 2.1 keeps the formula in UNF and does not change the guards, so the same argument applies.

GF *Input.* Assume ψ_{L} and ψ_{R} are GF sentences. We can convert these formulas to weak GN-normal form sentences φ_{L} and φ_{R} as described in Lemma 2.2, and then argue as in the previous cases. Because the conversion to weak GN-normal form keeps the width the same, CQ rank 1, and size polynomial in n, this means that the size of the DAG representation for the interpolant is doubly exponential in n in general, but singly exponential in n when the width is considered fixed.

4.5. Equality
In order to extend our interpolation results to handle input formulas with equality, we first show that any GNF formula with equality can be converted to a form with a very limited use of equality. Let us say that a GNF formula ϕ is *equality normalized* if

(1) every occurrence of a relational atomic formula $R\boldsymbol{t}$ in ϕ appears in conjunction with

$$\text{all-distinct}(\boldsymbol{t}) := \bigwedge_{t \in \boldsymbol{t}} (t = t) \wedge \bigwedge_{t,t' \in \boldsymbol{t}, t \neq t'} \neg(t = t'),$$

(2) whenever equalities are used as guards for negations, then these equality guards are of the form $x = x$, and
(3) there are no occurrences of equality in ϕ other than those described in (i) and (ii).

Let $\phi(\boldsymbol{x})$ be a GNF formula containing a set $\boldsymbol{e}$ of constants, and let $\equiv$ be any equivalence relation over $\boldsymbol{x} \cup \boldsymbol{e}$. We denote by $\xi_\equiv$ the FO formula $\bigwedge_{s \equiv t} (s = t) \wedge \bigwedge_{s \not\equiv t} \neg(s = t)$, that is, the conjunction of all equalities and inequalities corresponding to $\equiv$. Note that this formula does not necessarily belong to GNF as it may contain unguarded inequalities. We denote by $\xi_\equiv^{\mathrm{gn}}$ the GNF formula $\bigwedge_{s \equiv t} (s = t) \wedge \bigwedge_{s \not\equiv t, s \in \boldsymbol{e}, t \in \boldsymbol{x} \cup \boldsymbol{e}} \neg(s = t)$, where we only include inequalities that are always expressible in GNF (i.e., inequalities where at least one of the terms is a constant).

LEMMA 4.11. *Let $\phi(\boldsymbol{x})$ be a GNF formula in weak GN-normal form, and let $\equiv$ be an equivalence relation over the free variables and constants in ϕ. We can construct a DAG representation of an equality-normalized $\phi'_\equiv(\boldsymbol{x})$ in weak GN-normal form such that*

—$\models \forall \boldsymbol{x}(\xi_{\equiv}(\boldsymbol{x}) \to (\phi(\boldsymbol{x}) \leftrightarrow \phi'_{\equiv}(\boldsymbol{x})))$;

—$|\phi'_{\equiv}| \leq p(|\phi|^w)$, where w is the width of ϕ and p is a polynomial function independent of ϕ;

—$\mathrm{width}(\phi'_{\equiv}) \leq \mathrm{width}(\phi)$ and $\mathrm{rank}_{\mathrm{CQ}}(\phi'_{\equiv}) \leq \mathrm{rank}_{\mathrm{CQ}}(\phi)$.

Furthermore, $\mathrm{occ}(\phi'_{\equiv}) \subseteq \mathrm{occ}(\phi)$, *and, if ϕ is $\mathbb{U}$-relativized, then $\phi'_{\equiv}$ is as well.*

PROOF. We may assume that every equality guard in ϕ is of the form $x = x$ (if an equality guard is of the form $x = y$, we can freely replace all occurrences of y in the subformula by x). This takes care of condition (2) in the definition of equality-normalized formulas.

Let $\boldsymbol{x}$ be the free variables and let $\boldsymbol{e}$ be the set of constants in ϕ. For each $\equiv$-equivalence class over $\boldsymbol{x} \cup \boldsymbol{e}$, we fix an arbitrary representative—a constant whenever possible. We replace all occurrences of each variable and constant in $\boldsymbol{x} \cup \boldsymbol{e}$ by the representative of its equivalence class. Next, we replace subformulas of the form $s = t$, with $s, t \in \boldsymbol{x} \cup \boldsymbol{e}$, by $\top$ if $s = t$ and by $\bot$ otherwise; finally, we conjoin every relational atom $R\,\boldsymbol{t}$ with $\boldsymbol{t} \subseteq \boldsymbol{x} \cup \boldsymbol{e}$ with all-distinct($\boldsymbol{t}$).

At this point, the conditions (1) and (2) in the definition of equality-normalized formulas are satisfied for equalities that have free variables or constants on both sides. It remains only to take care of the equalities that involve a quantified variable.

Consider any subformula of the form $\psi(\boldsymbol{z}) = \exists \boldsymbol{y}.\chi(\boldsymbol{y}, \boldsymbol{z})$. We will essentially do a case distinction, for each quantified variable $y_i \in \boldsymbol{y}$, of the possible values that y_i may take. More precisely, we replace $\psi(\boldsymbol{z})$ by the disjunction, for each map $f : \boldsymbol{y} \to (\boldsymbol{y} \cup \boldsymbol{z} \cup \boldsymbol{e})$, of the formula ψ_f obtained from ψ by (i) replacing each $y_i \in \boldsymbol{y}$ by $f(y_i)$, (ii) replacing $y_i = y_j$ by $\top$ if $f(y_i) = f(y_j)$ and by $\bot$ otherwise, (iii) replacing $y_i = t$ or or $t = y_i$ for $t \in \boldsymbol{z} \cup \boldsymbol{e}$ by $\top$ if $f(y_i) = t$ and by $\bot$ otherwise (and dropping the quantifiers corresponding to quantified variables that no longer occur in the formula). Finally, we conjoin every atom $R\,\boldsymbol{t}$ with $\bigwedge_{t \in \boldsymbol{t}}(t = t) \wedge \bigwedge_{t,t' \in \boldsymbol{t}, t \neq t'} \neg(t = t')$. This clearly preserves the semantics of the formula over structures satisfying $\xi_{\equiv}$. It does not change the CQ rank because $\boldsymbol{t}$ is guarded. The size blowup involved in this procedure is at most exponential in the width of ϕ.

We obtain the desired equality-normalized $\phi'_{\equiv}$ in weak GN-normal form by performing this rewriting in a bottom-up fashion starting with the innermost quantifiers. By using a DAG representation, the overall size of this equality-normalized form is at most exponential in w; we remark that using a tree representation, it would be exponential in both w and the maximal nesting depth of quantifiers in the formula. Furthermore, even though we introduce disjunctions, the resulting formula is easily seen to be still in weak GN-normal form. The other properties can easily be checked. □

THEOREM 4.12. *Let φ_{L} and φ_{R} be GNF formulas in weak GN-normal form with equality such that $\varphi_{\mathrm{L}} \models \varphi_{\mathrm{R}}$.*

—*For all equivalence relations $\equiv$ over the free variables and constants in φ_{L} and φ_{R},*

$$(\varphi_{\mathrm{L}})'_{\equiv} \models (\varphi_{\mathrm{R}})'_{\equiv}.$$

—*Assume $\theta_{\equiv}$ denotes an interpolant for $(\varphi_{\mathrm{L}})'_{\equiv} \models (\varphi_{\mathrm{R}})'_{\equiv}$ for each equivalence relation $\equiv$ over the free variables and constants in φ_{L} and φ_{R}. Then*

$$\theta := \bigvee_{\equiv} \xi_{\equiv}^{\mathrm{gn}} \wedge \theta_{\equiv}$$

is an interpolant for $\varphi_{\mathrm{L}} \models \varphi_{\mathrm{R}}$. Moreover, if the $\theta_{\equiv}$ are all in GNF and have a DAG representation of size doubly exponential (singly exponential, respectively) in the size of φ_{L} and φ_{R}, then θ is in GNF and has a DAG representation of doubly exponential (singly exponential, respectively) size.

PROOF. Let $\boldsymbol{x}$ be the free variables and $\boldsymbol{e}$ the constants occurring in φ_L and φ_R. Applying Lemma 4.11, we obtain that $\varphi_L \models \varphi_R$ holds if and only if, for each equivalence relation $\equiv$ on $\boldsymbol{x} \cup \boldsymbol{e}$,

$$\xi_\equiv \wedge (\varphi_L)'_\equiv \models (\varphi_R)'_\equiv.$$

We must show that this entailment is valid if and only if $(\varphi_L)'_\equiv \models (\varphi_R)'_\equiv$ is valid. One direction is trivial. For the other direction, we proceed by contraposition: let $\mathfrak{M} \models (\varphi_L)'_\equiv \wedge \neg(\varphi_R)'_\equiv$. Recall that $(\varphi_L)'_\equiv$ and $(\varphi_R)'_\equiv$ contain only the representative of each $\equiv$-equivalence class. We can freely expand $\mathfrak{M}$ so that all other variables and constants denote the same element that is denoted by the representative of their equivalence class. The only potential source of difficulty is that $\mathfrak{M}$ may interpret two constants or free variables, s and t, as the same element a, even though they are representatives of different $\equiv$-equivalence classes. However, in this case, we can simply "pull s and t apart," that is, we can replace a by two copies, a_1 and a_2, such that a_1 and a_2 participate in all the same facts that a did. This modification of the structure does not affect the truth of $(\varphi_L)'_\equiv$ and the falsity of $(\varphi_R)'_\equiv$, due to the fact that these formulas are equality normalized, as can be shown by a straightforward formula induction.

Next, we argue that $\theta := \bigvee_\equiv \xi_\equiv^{gn} \wedge \theta_\equiv$ is an interpolant for $\varphi_L \models \varphi_R$. Suppose there is a model $\mathfrak{M}$ of φ_L. Let $\equiv_\mathfrak{M}$ be the equivalence relation over $\boldsymbol{x} \cup \boldsymbol{e}$ induced by the interpretation of these variables and constants in $\mathfrak{M}$. Then $\mathfrak{M} \models \xi_{\equiv_\mathfrak{M}} \wedge (\varphi_L)'_{\equiv_\mathfrak{M}}$ by Lemma 4.11, so $\mathfrak{M} \models \xi_{\equiv_\mathfrak{M}}^{gn} \wedge \theta_{\equiv_\mathfrak{M}}$ and therefore $\mathfrak{M} \models \theta$. Now consider a model $\mathfrak{M}$ for θ, and assume $\mathfrak{M} \models \xi_\equiv^{gn} \wedge \theta_\equiv$. This means $\mathfrak{M} \models \xi_\equiv^{gn} \wedge (\varphi_R)'_\equiv$, by the definition of the interpolants $\theta_\equiv$. We must prove that $\xi_\equiv^{gn} \wedge (\varphi_R)'_\equiv \models \varphi_R$. We can assume that φ_R and $(\varphi_R)'_\equiv$ are UCQ-shaped subformulas. Consider the CQ-shaped formula $\exists \boldsymbol{y}.\chi(\boldsymbol{x}, \boldsymbol{y})$ in φ_R corresponding to some $\exists \boldsymbol{y}'.\chi'(\boldsymbol{x}', \boldsymbol{y}')$ in $(\varphi_R)'_\equiv$ that is satisfied in $\mathfrak{M}$. By construction of $(\varphi_R)'_\equiv$ we must have $\boldsymbol{x}' \subseteq \boldsymbol{x}$ and $\boldsymbol{y}' \subseteq \boldsymbol{y}$. We claim that χ must be satisfied in $\mathfrak{M}$, and hence φ_R is satisfied in $\mathfrak{M}$. Consider a conjunct $\alpha \wedge \psi$ in χ (a guarded subformula). In χ', this conjunct becomes $\alpha' \wedge \text{all-distinct}(\text{free}(\alpha')) \wedge \psi'_{\equiv'}$ for some $\equiv'$ over $\text{free}(\alpha) \cup \boldsymbol{e}$ that agrees with $\equiv$ on $\text{free}(\alpha) \cup \boldsymbol{e}$. By assumption, $\mathfrak{M}, \boldsymbol{a} \models \alpha' \wedge \text{all-distinct}(\text{free}(\alpha')) \wedge \psi'_{\equiv'}$. This induces some corresponding potential witness $\boldsymbol{a}$ for $\alpha \wedge \psi$ (based on the relationship between the free variables in χ and χ'). We know that $\mathfrak{M}, \boldsymbol{a} \models \alpha$; this is true even if α is an equality guard of the form $s = t$, since s and t must be in the same equivalence class in $\equiv'$ (otherwise in χ' this would become $\bot$). Observe that $\mathfrak{M}, \boldsymbol{a} \models \xi_\equiv^{gn} \wedge \text{all-distinct}(\text{free}(\alpha'))$ implies that $\mathfrak{M}, \boldsymbol{a} \models \xi_{\equiv'}$, since $\text{all-distinct}(\text{free}(\alpha'))$ asserts all of the inequalities that are missing from $\xi_{\equiv'}^{gn}$. Hence, Lemma 4.11 ensures that $\mathfrak{M}, \boldsymbol{a} \models \psi$ if and only if $\mathfrak{M}, \boldsymbol{a} \models \psi'_\equiv$. The same argument works when the conjunct is of the form $\alpha \wedge \neg\psi$. Since this holds for every conjunct of χ, we know that $\mathfrak{M} \models \varphi_R$. Hence, θ is an interpolant for $\varphi_L \models \varphi_R$.

The number of equivalence relations $\equiv$ over $\boldsymbol{x} \cup \boldsymbol{e}$ is at most exponential in $|\boldsymbol{x} \cup \boldsymbol{e}| \leq |\varphi_L| + |\varphi_R|$. Likewise, the size of $\xi_\equiv^{gn}$ is at most polynomial in $|\boldsymbol{x} \cup \boldsymbol{e}| \leq |\varphi_L| + |\varphi_R|$. Hence, the overall size of the DAG representation for θ is still doubly exponential (singly exponential, respectively) in $|\varphi_L| + |\varphi_R|$ if each $\theta_\equiv$ has a DAG representation of at most doubly exponential (singly exponential, respectively) size. □

We have thus reduced the problem at hand to obtaining interpolants for entailments of the form $(\varphi_L)'_\equiv \models (\varphi_R)'_\equiv$, where both formulas involved are equality normalized. We can do this by adapting the mosaic method from before, basically treating equality as an ordinary relation symbol, but restricting to mosaics with limited use of equality. We say that an x-mosaic is *equality trivial* if (i) $Y^+ : t = t' \in \tau$ only if t and t' are the same term (parameter or constant), and (ii) $Y^- : t = t'$ only if t and t' are distinct terms (parameters or constants). To work with these equality-trivial mosaics, we change the linking definition slightly to allow constants to be used to witness existential requirements, and update the downward closure rules to handle subformulas all-distinct($\boldsymbol{t}$)

so nontrivial equalities are not introduced. A close inspection of the soundness and completeness proofs show that, for equality-normalized input formulas ψ_L and ψ_R, we can restrict attention to equality-trivial mosaics like this. Furthermore, in the completeness proof, the model that is constructed is such that distinct terms are realized by distinct elements of the model. Consequently, whenever an equality-trivial mosaic contains an equality or inequality, then this equality or inequality is guaranteed to be realized by every corresponding pair of elements of the model. We can conclude that, for equality-normalized input sentences, the mosaic method constitutes a sound and complete decision procedure.

THEOREM 4.13. *Let ψ_L and ψ_R be equality normalized* GNF *formulas in weak* GN-*normal form, and let $\mathcal{M}$ be the set of equality-normalized* GNF *mosaics for $\psi_L \wedge \neg \psi_R$ over some set of parameters of size twice the width of $\psi_L \wedge \neg \psi_R$. Then, $\psi_L \models \psi_R$ holds if and only if there is a saturated set of internally consistent equality-trivial* GNF *mosaics from $\mathcal{M}$ that contains a mosaic τ such that $L^+ : \psi_L, R^- : \psi_R \in \tau$.*

A GNF *interpolant θ can be constructed from the mosaic interpolants as before, using only trivial equality guards.*

The interpolation results in Theorem 4.10 for (relativized or unrelativized) GNF input formulas with equality follow from Lemma 2.1, Lemma 4.11, Theorem 4.12, and Theorem 4.13. For GF sentences with equality and a fixed bound on the arity of relations, the single exponential bound comes from observing that the conversion to weak GN-normal form and then to equality normal form using Lemmas 2.2 and 4.11 only results in a polynomial blowup in the size of the formula when the width is fixed. For UNF input with equality, we can use a similar proof strategy, but in this case, we must use a different equality-normal form, replacing all-distinct($\boldsymbol{t}$) with distinct$_{UNF}(\boldsymbol{t})$ which only asserts inequalities expressible in UNF.

Using this reduction for equality described in this section, we can make no claims about the polarity of occurrences of equality in the interpolant—and this is why no such claims are made in our interpolation result, Theorem 4.10. The conversion to equality-normalized GN-normal form may change the polarity of occurrences of equality and equality may be introduced as unary guards in both polarities in the interpolant. This implies that the polarity of occurrences of equality in the interpolant are unrelated to the polarity of equality in the original formulas (in contrast to the other relations in the common signature, which can only appear in the interpolant in polarity p if they occurred in both φ_L and φ_R in polarity p). Indeed, even if the original formulas did not use equality, equality may appear in both polarities in the resulting interpolant due to the introduction of trivial equality guards for unary negation. However, the only nontrivial equalities and inequalities come from the $\xi_{\equiv}^{gn}$ that is added to the mosaic interpolants by Theorem 4.12.

We remark that this approach to handling equality is different from a standard approach used in Motohashi [1984]. In that work, the input formulas are augmented with formulas axiomatizing equality. A standard interpolation algorithm can then be used on this augmented formula, treating equality like any other relation. We cannot apply the same technique here because we cannot express the equality axioms in GNF.

4.6. Free Variables

We conclude by describing how to extend the interpolation results to handle $\psi_L(\boldsymbol{xz})$ and $\psi_R(\boldsymbol{yz})$ with different free variables, while still ensuring that the interpolant only uses the common free variables $\boldsymbol{z}$. We suppose that $\psi_L(\boldsymbol{xz})$ and $\psi_R(\boldsymbol{yz})$ are in $\mathbb{U}$-relativized weak GN-normal form.

First, observe that it is easy to handle extra free variables in the antecedent. If $\psi_L(\boldsymbol{xz}) \models \psi_R(\boldsymbol{yz})$ is valid, then $\exists \boldsymbol{x}. U' x_1 \wedge \cdots \wedge U' x_k \wedge \psi_L(\boldsymbol{xz}) \models \psi_R(\boldsymbol{yz})$ is valid, for a fresh

relativizer U'. Moreover, an interpolant for this new validity can be used directly as an interpolant for the original validity since the new relativizer U' is not in the common signature.

Hence, it suffices to consider the case when $\psi_L(\mathbf{z}) \models \psi_R(\mathbf{yz})$, say for $\mathbf{y} = \{y_1, \ldots, y_k\}$. To handle this case, we use a restricted set of mosaics. We say a mosaic τ is $\mathbf{y}$-restricted if the following additional conditions hold: (i) for all $x^+ : \psi \in \tau$ such that $par(\psi) \cap \mathbf{y} \neq \emptyset$, $par(\psi) \cap \mathbf{z} = \emptyset$ and there is some $y_i \in \mathbf{y}$ such that $par(\psi) \cap \mathbf{y} = \{y_i\}$, and (ii) for all $L^p : \psi \in \tau$ such that $par(\psi) \cap \mathbf{y} \neq \emptyset$, $par(\psi) \cap \mathbf{y}$ is R-crelativized in τ.

The proof for this case proceeds in two main steps. We first argue that $\mathbf{y}$-restricted mosaics yield a sound and complete system for testing validity of formulas like this (Theorem 4.14). We then argue that the mosaic interpolants built from these $\mathbf{y}$-restricted mosaics have stronger properties that enable us to construct an interpolant that only uses $\mathbf{z}$ (Lemma 4.15).

The first step is proving soundness and completeness for $\mathbf{y}$-restricted mosaics.

THEOREM 4.14. *Let $\mathcal{M}$ be the set of $\mathbf{y}$-restricted GNF mosaics for $\psi_L(\mathbf{z}) \wedge \neg\psi_R(\mathbf{yz})$ over some set of parameters of size twice the width of $\psi_L \wedge \neg\psi_R$ and containing $\mathbf{yz}$. Then $\psi_L \models \psi_R$ holds if and only if there is a saturated set of internally consistent $\mathbf{y}$-restricted GNF mosaics from $\mathcal{M}$ that contains a mosaic τ such that $L^+ : \psi_L, R^- : \psi_R \in \tau$.*

We omit the proof, but the intuition for why this theorem holds comes from examining the mosaic properties and downward closure rules that govern how $\mathbf{y}$ can appear in the mosaics. When we are looking at mosaics for $\psi_L(\mathbf{z})$ and $\psi_R(\mathbf{yz})$, we know that we will need to consider mosaics with $\mathbf{y}$ present at least in $R^- : \psi_R(\mathbf{yz})$ and (if we are using equality) $R^+ : y_i = y_i$ for all $y_i \in \mathbf{y}$. The downward closure rules in mosaics may force variables from $\mathbf{y}$ to enter additional formulas. However, supposing we have a set of formulas that is $\mathbf{y}$-restricted, applying the downward closure rules can only add a positive formula containing variables from $\mathbf{y}$ that is guarded by a positive formula—which must contain at most one of the variables from $\mathbf{y}$ by assumption. Likewise, the downward closure rules that would force $\mathbf{y}$ to enter an L-formula via some $L^- : \eta$ require that this happens in the presence of some relativizers—either the desired R-crelativizers themselves, or L-relativizers, which would only be allowed in the mosaic if there were already corresponding R-crelativizers for these variables from $\mathbf{y}$.

The advantage of using $\mathbf{y}$-restricted mosaics is that the resulting mosaic interpolants also have a stronger property on the use of variables from $\mathbf{y}$. Taking advantage of this property, we can use these mosaic interpolants to build an interpolant—without $\mathbf{y}$—for the original validity. This is the second step mentioned previously.

LEMMA 4.15. *For any $\mathbf{y}$-restricted L-mosaic $\tau(\mathbf{yz})$, if the mosaic interpolant θ_L^τ uses variables from $\mathbf{y}$, then it uses at most one variable from $\mathbf{y}$ (and this variable is R-crelativized in τ) and no variables from $\mathbf{z}$. An interpolant for $\psi_L(\mathbf{z}) \models \psi_R(\mathbf{yz})$ using only variables $\mathbf{z}$ and satisfying the properties from Theorem 4.10 can be constructed from these mosaic interpolants.*

PROOF. The first part of this lemma follows by observing that in the mosaic interpolation construction given in Section 4.3, parameters can only appear together in a mosaic interpolant if they appear together in a positive formula.

We now proceed to use mosaic interpolants like this in order to construct an interpolant for $\psi_L(\mathbf{z}) \models \psi_R(\mathbf{yz})$. Let $\mathbf{y} = \{y_1, \ldots, y_k\}$. We adapt some notation from earlier. For $z \in \{L, R\}$, we define $\tau \restriction_z^{\mathbf{b},r}$ to be the $\mathbf{y}$-restricted z-mosaic over parameters $\mathbf{b}$ and relativizing r that consists of the union of (i) the z-formulas in τ that only use parameters from $\mathbf{b}$, and (ii) the $\tilde{z}$-formulas in τ whose parameters are contained in $\mathbf{b}$ and are z-safe in the formulas remaining after step (i). For mosaics υ, υ' and $z \in \{L, R\}$, we write $\upsilon'(\mathbf{ab}) \supseteq_z \upsilon$

if υ' uses only parameters from $\boldsymbol{ab}$ and is obtained from υ by adding only $\boldsymbol{y}$-restricted z-formulas from the closure (so the $\tilde{\mathrm{z}}$-formulas in υ and υ' are identical). For a Y-mosaic υ', we write $\mathrm{rel}_{\mathrm{Y}}^{\upsilon'}(\boldsymbol{a})$ for the conjunction of all Ua such that $U \in \mathbb{U}$ is a Y-catom and $\mathrm{Y}^+ : Ua \in \upsilon'$ (so $\mathrm{Y}^+ : Ua$ is a Y-crelativizer in υ').

We first claim that

$$\theta(\boldsymbol{yz}) := \bigwedge_{\substack{\tau(\boldsymbol{yz}) \supseteq_{\mathrm{R}} \{\mathrm{R}^- : \psi_{\mathrm{R}}\} \\ \text{s.t. } \tau \in \mathcal{M}}} \left(\neg\,\mathrm{rel}_{\mathrm{R}}^{\tau}(\boldsymbol{y}) \ \vee \bigvee_{\substack{\tau'(\boldsymbol{yz}) \supseteq_{\mathrm{L}} \tau \\ \text{s.t. } \tau'' := \tau' |_{\mathrm{L}}^{\boldsymbol{yz};\emptyset} \in \mathcal{N}'}} \theta_{\mathrm{L}}^{\tau''}(\boldsymbol{yz}) \right)$$

is an interpolant for $\psi_{\mathrm{L}}(\boldsymbol{z}) \models \psi_{\mathrm{R}}(\boldsymbol{yz})$, where $\mathcal{M}$ is the set of $\boldsymbol{y}$-restricted mosaics, and $\mathcal{N}'$ is the set of eliminated mosaics. Note that this interpolant may still use $\boldsymbol{y}$. Although $\neg\,\mathrm{rel}_{\mathrm{R}}^{\tau}(\boldsymbol{y})$ is not in GNF, it can easily be rewritten in GNF by pushing the negation inside and adding trivial equality guards.

Suppose $\mathfrak{M} \models \psi_{\mathrm{L}}(\boldsymbol{z})$. Let υ be the set of L-formulas $\mathrm{L}^p : \psi \in \mathrm{cl}_{\mathrm{L}}(\boldsymbol{yz})$ such that $\mathfrak{M} \models p\psi$. For all $\tau(\boldsymbol{yz}) \supseteq_{\mathrm{R}} \{\mathrm{R}^- : \psi_{\mathrm{R}}\}$ that are $\boldsymbol{y}$-restricted mosaics, consider $\tau'' := (\tau \cup \upsilon) |_{\mathrm{L}}^{\boldsymbol{yz};\emptyset}$. By construction, υ contains $\mathrm{L}^+ : \psi_{\mathrm{L}}$, and τ contains $\mathrm{R}^- : \psi_{\mathrm{R}}$. Neither of these formulas are eliminated by the operation $(\tau \cup \upsilon) |_{\mathrm{L}}^{\boldsymbol{yz};\emptyset}$, since they satisfy the conditions for being in a $\boldsymbol{y}$-restricted L-mosaic. Hence, by completeness of the mosaic method, $\tau'' \in \mathcal{N}'$ and $\mathfrak{M} \models \tau_{\mathrm{L}}''$, hence $\mathfrak{M} \models \theta_{\mathrm{L}}^{\tau''}$ and $\mathfrak{M} \models \theta$.

Now suppose that $\mathfrak{M} \models \theta$. Let S be the set of R-formulas $\mathrm{R}^p : \psi \in \mathrm{cl}_{\mathrm{R}}(\boldsymbol{yz})$ such that $\mathfrak{M} \models p\psi$. Let $\tau := S \cup \{\mathrm{R}^- : \psi_{\mathrm{R}}\}$. By construction, $\mathfrak{M} \models \mathrm{rel}_{\mathrm{R}}^{\tau}(\boldsymbol{y})$, so there is some $\tau' \supseteq_{\mathrm{L}} \tau$ and $\tau'' := \tau' |_{\mathrm{L}}^{\boldsymbol{yz};\emptyset} \in \mathcal{N}'$ such that $\mathfrak{M} \models \theta_{\mathrm{L}}^{\tau''}$. Hence, $\mathfrak{M} \models \neg\tau_{\mathrm{R}}''$. Note that τ'' must contain $\mathrm{R}^- : \psi_{\mathrm{R}}$ since it trivially satisfies the conditions for being in a $\boldsymbol{y}$-restricted mosaic, regardless of the L-formulas in τ'. Since $\mathfrak{M} \models \neg\tau_{\mathrm{R}}''$, there is some conjunct in τ_{R}'' that is not satisfied in $\mathfrak{M}$. This conjunct cannot be $p\psi$ for some $\mathrm{R}^p : \psi \in S$, since $\mathfrak{M} \models p\psi$ for such formulas. This conjunct cannot be $\neg p\psi$ for some $\mathrm{R}^p : \psi \notin \tau''$ since this would imply that $\mathrm{R}^p : \psi$ is L-safe in τ'', and hence $\mathrm{R}^p : \psi \notin S$ and $\mathfrak{M} \models \neg p\psi$. Therefore, the only possibility is that the conjunct $\neg\psi_{\mathrm{R}}$ in τ_{R}'' coming from $\mathrm{R}^- : \psi_{\mathrm{R}} \in \tau''$ is not satisfied, so $\mathfrak{M} \models \psi_{\mathrm{R}}$. Except for the fact that this interpolant makes use of $\boldsymbol{y}$, the other properties stated in Theorem 4.10 follow from the properties of the mosaic interpolants.

Our remaining task is to transform θ into an interpolant that does not use $\boldsymbol{y}$.

For $\tau(\boldsymbol{yz}) \supseteq_{\mathrm{R}} \{\mathrm{R}^- : \psi_{\mathrm{R}}\}$ such that $\tau \in \mathcal{M}$, let Θ^{τ} be the set of mosaic interpolants $\theta_{\mathrm{L}}^{\tau''}$ such that $\tau'' := \tau' |_{\mathrm{L}}^{\boldsymbol{yz};\emptyset} \in \mathcal{N}'$ for some $\tau'(\boldsymbol{yz}) \supseteq_{\mathrm{L}} \tau$. Let Θ_0^{τ} be the subset of Θ^{τ} using no variables from $\boldsymbol{y}$ (i.e., only variables from $\boldsymbol{z}$), and for $i \in \{1, \ldots, k\}$, let Θ_i^{τ} be the subset of Θ^{τ} using only variable y_i. By the properties of the mosaic interpolants, $\Theta_0^{\tau}, \ldots, \Theta_k^{\tau}$ partition Θ^{τ}, so θ is logically equivalent to

$$\theta'(\boldsymbol{yz}) := \bigwedge_{\substack{\tau(\boldsymbol{yz}) \supseteq_{\mathrm{R}} \{\mathrm{R}^- : \psi_{\mathrm{R}}\} \\ \text{s.t. } \tau \in \mathcal{M}}} \left(\bigvee_{\theta_0 \in \Theta_0^{\tau}} \theta_0(\boldsymbol{z}) \vee \bigvee_{i \in \{1, \ldots, k\}} \left(\neg\,\mathrm{rel}_{\mathrm{R}}^{\tau}(y_i) \vee \bigvee_{\theta_i \in \Theta_i^{\tau}} \theta_i(y_i) \right) \right),$$

which is polynomial in the size of θ.

Now observe that if $\theta'(\boldsymbol{yz})$ is an interpolant for $\psi_{\mathrm{L}}(\boldsymbol{z}) \models \psi_{\mathrm{R}}(\boldsymbol{yz})$, then $\forall \boldsymbol{y}.\theta'(\boldsymbol{yz})$ is clearly an interpolant as well. The problem is that this is not in general in GNF. However, we can express this in GNF, by pushing the universal quantification inside the conjunction, and applying appropriate relativized *unary* universal quantification over y_i to the disjuncts using y_i (and no other variables). In other words, $\forall \boldsymbol{y}.\theta'(\boldsymbol{yz})$ is

logically equivalent to

$$\theta''(\boldsymbol{z}) := \bigwedge_{\substack{\tau(\boldsymbol{yz}) \supseteq_{\mathrm{R}} \{\mathrm{R}^-:\psi_{\mathrm{R}}\} \\ \text{s.t. } \tau \in \mathcal{M}}} \left(\bigvee_{\theta_0 \in \Theta_0^\tau} \theta_0(\boldsymbol{z}) \vee \bigvee_{i \in \{1,\ldots,k\}} \neg \exists y_i . \left(\mathrm{rel}_{\mathrm{R}}^\tau(y_i) \wedge \neg \bigvee_{\theta_i \in \Theta_i^\tau} \theta_i(y_i) \right) \right).$$

This interpolant $\theta''(\boldsymbol{z})$ for $\psi_{\mathrm{L}}(\boldsymbol{z}) \models \psi_{\mathrm{R}}(\boldsymbol{yz})$ satisfies the desired properties: namely, it is in relativized GNF, uses only the variables $\boldsymbol{z}$ common to the antecedent and consequent, and has a DAG representation doubly exponential in the size of the validity. □

Putting Theorem 4.14 and Lemma 4.15 together, this means that we can use our mosaic interpolation system on $\boldsymbol{y}$-restricted mosaics to construct an interpolant that only uses free variables that are common to ψ_{L} and ψ_{R}. Moreover, using the reductions to relativized interpolation described in Section 4.4, we can ensure that this property of the free variables holds for the other classes of input formulas considered there. Therefore, Theorem 4.10 holds even for input formulas with different free variables.

5. LOWER BOUNDS

We have matching lower bounds for the interpolation results stated in Theorem 4.10. In particular:

THEOREM 5.1. *There is a polynomial function p and a family of* UNF *sentences* $\chi_n \models \chi_n'$ *without equality such that* $|\chi_n| + |\chi_n'| \leq p(n)$, *and there is no* GNF *interpolant* θ_n *for* $\chi_n \models \chi_n'$ *with size at most* $2^{2^{n-1}}$, *even when* θ_n *is represented via a DAG.*

We prove this theorem and then mention some related lower bound results.

5.1. Family of UNF Formulas

The family of sentences χ_n and χ_n' witnessing the lower bound in Theorem 5.1 describe certain *indexed AND/OR trees*, finite trees where the leaves are labeled with 0 or 1, and other nodes are labeled with AND or OR. Moreover, each node is indexed by a tree that describes the depth of the node (up to depth 2^{2^n}).

Formally, an indexed AND/OR tree is a structure over signature ρ_n consisting of the following predicates:

(1) Unary predicates IndValOne, $\mathrm{IndDepth}_0, \ldots, \mathrm{IndDepth}_n$, and a ternary connective IndChild describe the index of one node. Informally, $\mathrm{IndDepth}_0, \ldots, \mathrm{IndDepth}_n$ and IndChild describe the structure of a binary tree of depth n (with $\mathrm{IndDepth}_0$ labeling the root), while IndValOne distinguishes the leaves of this tree with value one. Each such tree can be assigned a value, when viewing the leaves as 2^n bits of a binary number.

(2) Unary predicates InputOne and IsOr along with binary predicate AOChild describe the AND/OR structure, with IsOr representing an OR node and InputOne distinguishing the inputs on the leaves that have value one.

The signatures of χ_n and χ_n' extend ρ_n using two copies of a predicate AOValOne for the intermediate value of the circuit computation at each node. We write σ_n and σ_n' for the signatures with AOValOne and AOValOne$'$, respectively. The common signature $\sigma_n \cap \sigma_n'$ is ρ_n.

We will be interested in structures where the nodes in the AND/OR tree are roots of index trees describing the depth. Roughly speaking, the formula χ_n will assert

there is an indexed AND/OR tree of depth 2^{2^n}, and the value (using unary relation AOValOne) of any AND/OR tree in the structure is 1.

Likewise, χ'_n will assert

the value (using unary relation $\text{AOValOne}'$) of any AND/OR tree in the structure is 1.

We first define the following auxiliary UNF formulas (technically, these should all be indexed by n). We write out some of these auxiliary formulas to give a flavor of how these notions can be expressed in UNF. We will use the convention that built-in predicates like IndChild will be written $\text{IndChild}\,xx_lx_r$, whereas our derived auxiliary formulas like IndChild_d that have free variables x and x_d will be written $\text{IndChild}_d(x, x_d)$. Note that some unary negations in these UNF formulas are not explicitly guarded by an atomic relation, but it is straightforward to modify the formulas to explicitly guard the negations with relations from ρ_n.

—$\text{IndChild}_d(x, x_d)$: x_d is the d-child of x for $d \in \{\mathbf{l}, \mathbf{r}\}$.
—$\text{Link}(x_n, x'_n)$: x_n and x'_n are leaves in corresponding positions in the same index tree.
—$\text{LinkNext}(x_n, y_n)$: x_n and y_n are leaves in corresponding positions in neighboring index trees.

$$\exists x_0 \cdots x_{n-1} y_0 \cdots y_{n-1}.$$

$$\Bigg(\text{AOChild}\,x_0 y_0 \wedge \text{IndDepth}_0\,x_0 \wedge \text{IndDepth}_0\,y_0 \quad \wedge$$

$$\bigwedge_{i=0}^{n-1} \Big((\text{IndChild}_\mathbf{l}(x_i, x_{i+1}) \wedge \text{IndChild}_\mathbf{l}(y_i, y_{i+1})) \vee (\text{IndChild}_\mathbf{r}(x_i, x_{i+1})$$

$$\wedge\, \text{IndChild}_\mathbf{r}(y_i, y_{i+1})) \Big) \Bigg).$$

—$\text{Before}(x_n, x'_n)$: x_n and x'_n are leaves in the same index tree, with x_n to the left of x'_n.
—IndexedAOTrees: Every node satisfies exactly one IndDepth_i, AOChild only connects IndDepth_0-nodes, and every IndDepth_0-node is the root of an index tree.
—$\text{IndMin}(x_0)$: The index tree rooted at x_0 is the minimum (all 0's).
—$\text{IndMax}(x_0)$: The index tree rooted at x_0 is the maximum (all 1's).
—$\text{AORoot}(x_0)$: x_0 is the root of an index tree with no incoming AOChild-edges.
—$\text{AOLeaf}(x_0)$: x_0 is the root of an index tree with no outgoing AOChild-edges.
—IndexIsDepth: Every root of an index tree with no incoming AOChild-edges has minimal index, and for every index tree with nonmaximal value l, every AOChild-neighboring tree has index $l + 1$.

$$\neg\exists x.(\text{AORoot}(x) \wedge \neg\text{IndMin}(x)) \quad \wedge$$

$$\neg\exists x.\text{IndDepth}_n\,x \wedge \neg\text{IndValOne}\,x \wedge \neg\exists x'.(\text{Before}(x, x') \wedge \neg\text{IndValOne}\,x') \quad \wedge$$

$$\begin{pmatrix} \exists y.(\text{LinkNext}(x, y) \wedge \neg\text{IndValOne}\,y) & \vee \\ \exists x'y'.(\text{LinkNext}(x', y') \wedge \text{Before}(x, x') \wedge (\neg\text{IndValOne}\,x' \vee \text{IndValOne}\,y')) & \vee \\ \exists x'y'.(\text{LinkNext}(x', y') \wedge \text{Before}(x', x) \wedge ((\text{IndValOne}\,x' \wedge \neg\text{IndValOne}\,y') & \vee \\ (\neg\text{IndValOne}\,x' \wedge \text{IndValOne}\,y'))) & \end{pmatrix}.$$

—ConsistentAOValues: For all leaves in an AND/OR tree, AOValOne matches InputOne, and for all other nodes in an AND/OR tree, AOValOne matches the

evaluation of the AND/OR tree rooted at that node.

$\neg\exists x.(\text{AOLeaf}(x) \land ((\text{InputOne}\,x \land \neg\text{AOValOne}\,x) \lor (\neg\text{InputOne}\,x \land \text{AOValOne}\,x))) \land$

$\neg\exists x.(\text{IndDepth}_0\,x \land \text{IsOr}\,x \land \text{AOValOne}\,x \land \neg\exists y.(\text{AOChild}\,xy \land \text{AOValOne}\,y)) \land$

$\neg\exists x.(\text{IndDepth}_0\,x \land \neg\text{IsOr}\,x \land \text{AOValOne}\,x \land \exists y.(\text{AOChild}\,xy \land \neg\text{AOValOne}\,y)) \land$

$\neg\exists x.(\text{IndDepth}_0\,x \land \text{IsOr}\,x \land \neg\text{AOValOne}\,x \land \exists y.(\text{AOChild}\,xy \land \text{AOValOne}\,y)) \land$

$\neg\exists x.(\text{IndDepth}_0\,x \land \neg\text{IsOr}\,x \land \neg\text{AOValOne}\,x \land \neg\exists y.(\text{AOChild}\,xy \land \neg\text{AOValOne}\,y)).$

Using these auxiliary formulas, we can now define χ_n, χ_n' as UNF sentences over signatures σ_n and σ_n', such that

$\chi_n := \text{IndexedAOTrees} \land \text{IndexIsDepth} \land \text{ConsistentAOValues} \land$
$\qquad \exists x.(\text{AORoot}(x)) \land$
$\qquad \neg\exists x.(\text{IndDepth}_0\,x \land \neg\text{IndMax}(x) \land \neg\exists y.(\text{IndDepth}_0\,y \land \text{AOChild}\,xy)) \land$
$\qquad \neg\exists x.(\text{AORoot}(x) \land \neg\text{AOValOne}\,x),$

$\chi_n' := \neg[\text{ConsistentAOValues}'] \lor \neg\exists x.(\text{AORoot}(x) \land \neg\text{AOValOne}'\,x),$

where ConsistentAOValues$'$ is the result of replacing AOValOne with AOValOne$'$ in ConsistentAOValues.

Note that in UNF we cannot enforce that a structure has an actual tree-shaped part corresponding to an indexed AND/OR tree; we can only pick out parts of the structure that can be viewed as indexed AND/OR trees (more precisely, parts of the structure that are "GN bisimilar" to such trees). Thus, χ_n expresses that (up to GN bisimulation) there is an indexed AND/OR tree in the structure and every indexed AND/OR tree has depth 2^{2^n}. Moreover, AOValOne describes the internal value computations of every AND/OR tree, and AOValOne holds at the root of every such tree. Likewise, χ_n' expresses that AOValOne$'$ describes the internal value computations of every AND/OR tree in the structure, and AOValOne holds at the root of every such tree.

We have $\chi_n \models \chi_n'$, and it is straightforward to check $|\chi_n| + |\chi_n'|$ is polynomial in n.

5.2. GN Bisimulation Game

In order to help prove that there is no "small" GNF interpolant for $\chi_n \models \chi_n'$, we introduce a variant of GN bisimulation [Bárány et al. 2011].

A position in the *k-width GN bisimulation game* between $\mathfrak{A}$ and $\mathfrak{B}$ (relative to some signature σ without constants) is a *partial rigid homomorphism* f from $\mathfrak{A}$ to $\mathfrak{B}$ or vice versa with $|\text{dom}(f)| \le k$. This means that f is a partial homomorphism and for any guarded tuple $\boldsymbol{c} \in \text{dom}(f)$, $f\!\restriction_{\boldsymbol{c}}$ (the restriction of f to $\boldsymbol{c}$) is a partial isomorphism. We say the *active structure* is the structure containing $\text{dom}(f)$.

Starting in position f, one round of the game consists of the following:

Restrict Spoiler can restrict to some subset $\boldsymbol{c} \subseteq \text{dom}(f)$, and then the game proceeds from position $f\!\restriction_{\boldsymbol{c}}$; or

Switch if $\text{dom}(f)$ is a guarded tuple $\boldsymbol{c}$, then Spoiler can choose to switch structures, and the game proceeds from position $(f\!\restriction_{\boldsymbol{c}})^{-1}$; or

Add if $|\text{dom}(f)| < k$, Spoiler can select an element c such that c is in the active structure but $c \notin \text{dom}(f)$, Duplicator can choose d in the inactive structure such that $f[c \mapsto d]$ is a partial rigid homomorphism, and then the game proceeds from position $f[c \mapsto d]$.

The *m-round game* consists of a sequence of at most m rounds like this. Duplicator wins if she never gets stuck (i.e., in a round where Spoiler adds an element c, Duplicator is always able to choose some d such that $f[c \mapsto d]$ is a partial rigid homomorphism).

We write $\mathfrak{A}, \boldsymbol{a} \rightarrow_{\mathrm{gn}[\sigma]}^{k,m} \mathfrak{B}, \boldsymbol{b}$ if Duplicator has a winning strategy in the k-width m-round GN-bisimulation game relative to signature σ starting from a partial rigid homomorphism $g : \boldsymbol{a} \mapsto \boldsymbol{b}$. A winning strategy for Duplicator implies agreement between $\mathfrak{A}$ and $\mathfrak{B}$ on certain GNF formulas.

PROPOSITION 5.2. *Assume* $\mathfrak{A}, \boldsymbol{a} \rightarrow_{\mathrm{gn}[\sigma]}^{k,m} \mathfrak{B}, \boldsymbol{b}$, *that is, Duplicator has a winning strategy in the k-width, m-round GN-bisimulation game between $\mathfrak{A}$ and $\mathfrak{B}$ (relative to signature σ without constants) starting from position $\boldsymbol{a} \mapsto \boldsymbol{b}$. If $\varphi(\boldsymbol{x})$ is a DAG representation of a GNF formula over σ such that $|\varphi| \leq m$,* width$(\varphi) \leq k$, *and* free$(\varphi) = \boldsymbol{x}$, *then* $\mathfrak{A} \models \varphi(\boldsymbol{a})$ *implies* $\mathfrak{B} \models \varphi(\boldsymbol{b})$.

PROOF. Assume that $\mathfrak{A}, \boldsymbol{a} \rightarrow_{\mathrm{gn}[\sigma]}^{k,m} \mathfrak{B}, \boldsymbol{b}$, and $g : \boldsymbol{a} \mapsto \boldsymbol{b}$ is a witnessing partial rigid homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$.

We proceed by induction on m. The base case is trivial.

Let $m > 0$ and consider some DAG-represented GNF formula $\varphi(\boldsymbol{x})$ such that $|\varphi| \leq m$, width$(\varphi) \leq k$, free$(\varphi) = \boldsymbol{x}$, and $\mathfrak{A} \models \varphi(\boldsymbol{a})$. We must show that $\mathfrak{B} \models \varphi(\boldsymbol{b})$.

Assume φ is of the form $\alpha(\boldsymbol{x})$. The result follows since g is a partial homomorphism.

Assume φ is of the form $\psi_1 \wedge \psi_2$. For each $i \in \{1, 2\}$, we can consider the partial rigid homomorphism after Spoiler restricts to $\boldsymbol{a}_i \subseteq \boldsymbol{a}$ corresponding to free$(\psi_i) \subseteq \boldsymbol{x}$. Since this is a valid move for Spoiler, Duplicator has a winning strategy in the $m-1$ round game from $g' := g\!\restriction_{\boldsymbol{a}_i}$. Since $|\psi_i| \leq m-1$ and width$(\psi_i) \leq k$, the inductive hypothesis implies that $\mathfrak{B} \models \psi_i(g'(\boldsymbol{a}_i))$ for $i \in \{1, 2\}$, so $\mathfrak{B} \models \varphi(\boldsymbol{b})$. The proof is similar for φ of the form $\psi_1 \vee \psi_2$.

Assume φ is of the form $\alpha(\boldsymbol{x}) \wedge \neg\psi(\boldsymbol{x}')$ where $\boldsymbol{x}' \subseteq \boldsymbol{x}$. Since $\mathfrak{A} \models \varphi(\boldsymbol{a})$, $\mathfrak{A} \models \alpha(\boldsymbol{a})$ and hence $\boldsymbol{a}$ is guarded in $\mathfrak{A}$. We also know that $\mathfrak{B} \models \alpha(\boldsymbol{b})$ since g is a partial homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$. Let $\boldsymbol{a}'$ be the restriction of $\boldsymbol{a}$ to $\boldsymbol{x}'$. Since $\boldsymbol{a}$ is guarded, $\boldsymbol{a}'$ is guarded in $\mathfrak{A}$. Consider the play where Spoiler restricts to the guarded tuple $\boldsymbol{a}'$, and then switches structures. Since these are valid moves for Spoiler in the game, Duplicator has a winning strategy in the $m-2$ round game starting from $g' := (g\!\restriction_{\boldsymbol{a}'})^{-1}$. Let $\boldsymbol{b}' = g(\boldsymbol{a}')$. Assume for the sake of contradiction that $\mathfrak{B} \models \psi(\boldsymbol{b}')$. Since $|\psi| \leq m-2$, the inductive hypothesis ensures that $\mathfrak{B} \models \psi(\boldsymbol{b}')$ implies $\mathfrak{A} \models \psi(g'(\boldsymbol{b}'))$, so we must have $\mathfrak{A} \models \psi(\boldsymbol{a}')$. But this contradicts the assumption that $\mathfrak{A} \models \varphi(\boldsymbol{a})$. Therefore, we must have $\mathfrak{B} \models \neg\psi(\boldsymbol{b}')$, so $\mathfrak{B} \models \varphi(\boldsymbol{b})$.

Assume φ is the form $\exists y.\psi(\boldsymbol{x}y)$. Since $\mathfrak{A} \models \varphi(\boldsymbol{a})$, there must be some c such that $\mathfrak{A} \models \psi(\boldsymbol{a}c)$. If $c \in \boldsymbol{a}$, then let $g' := g$. Otherwise, let g' be Duplicator's choice of partial rigid homomorphism after Spoiler adds c. This is a valid move in the game, since width$(\varphi) \leq k$ implies $|\boldsymbol{x}| < k$ and hence $|\boldsymbol{a}| < k$. In both cases, we must have $\mathfrak{A}, \boldsymbol{a}c \rightarrow_{\mathrm{gn}[\sigma]}^{k,m-1} \mathfrak{B}, \boldsymbol{b}d$ for $d := g'(c)$. Since $|\psi| \leq m-1$, the inductive hypothesis implies that $\mathfrak{B} \models \psi(\boldsymbol{b}d)$, so $\mathfrak{B} \models \varphi(\boldsymbol{b})$ as desired. □

5.3. Family of Counterexamples

The specific family of counterexamples that we use to prove Theorem 5.1 are an inductively defined family of indexed AND/OR trees $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ of increasing depth. The value of $\mathfrak{A}_i^n$ ($\mathfrak{B}_i^n$, respectively) is 1 (0, respectively), but the trees have some identical subtrees that make them hard to distinguish with "small" GNF sentences.

Roughly speaking, we show that for $i = 2^{2^n-1}$, Duplicator has a winning strategy in the i-round i-width GN-bisimulation game relative to signature ρ_n between the indexed AND/OR trees $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ (starting from the empty homomorphism). By Proposition 5.2, this is enough to ensure that $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ cannot be distinguished by GNF sentences with a DAG representation of size at most i. Supposing that there is a "small" (size at most 2^{2^n-1}) DAG representation of an interpolant θ_n for $\chi_n \models \chi_n'$ implies that $\mathfrak{A}_i^n \models \theta_n$

since $\mathfrak{A}_i^n \models \chi_n$ by construction. But $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ cannot be distinguished by such "small" formulas, so this means that $\mathfrak{B}_i^n \models \theta_n$, and hence $\mathfrak{B}_i^n \models \chi'_n$. This is impossible since $\mathfrak{B}_i^n$ is an indexed AND/OR tree with value 0, not value 1 as asserted by χ'_n.

We now describe in more detail this family of counterexamples $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ for all $i \leq 2^{2^n-1}$. For notational simplicity in the following description, we fix n and write $\mathfrak{A}_i$ and $\mathfrak{B}_i$ instead of $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$.

We write a to denote a single a-labeled node, and $a(T_1, T_2)$ to denote a tree where the root is labeled a and the children are the roots of trees T_1 and T_2. The basic structure of these AND/OR trees is defined recursively, where $\mathfrak{A}_0 := 1$ and $\mathfrak{B}_0 := 0$, and for $i > 0$

$$\mathfrak{A}_i := \text{AND}(\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{A}_{i-1}), (\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{A}_{i-1})),$$
$$\mathfrak{B}_i := \text{AND}(\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{A}_{i-1}), (\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{B}_{i-1})).$$

Evaluating $\mathfrak{A}_i$ ($\mathfrak{B}_i$, respectively) gives value 1 (0, respectively).

A separate index tree (binary tree of depth n) is then attached to each node in the trees described previously, where the index correctly describes the depth of this node in the AND/OR tree.

We view these indexed AND/OR trees as structures over the signature $\rho_n = \sigma_n \cap \sigma'_n$. In particular, this means that there are no internal computations using AOValOne or AOValOne$'$. Although the children of nodes in the AND/OR tree are unordered, we will refer to the left (direction **l**) and right (direction **r**) child based on the order of the children as shown previously. This is just to aid in the following description.

Let $h : \boldsymbol{b} \mapsto \boldsymbol{a}$ be a position (partial rigid homomorphism) in the k-width GN-bisimulation game between some $\mathfrak{B}_i$ and $\mathfrak{A}_i$. We can assume that h is restricted to elements in the AND/OR tree (rather than the elements in the index trees). This is without loss of generality since h can always be chosen to preserve the depth of positions, and the local structure of the index trees is identical in $\mathfrak{A}_i$ and $\mathfrak{B}_i$.

As mentioned earlier, we want to show that Duplicator has a winning strategy in certain GN-bisimulation games between $\mathfrak{B}_i$ and $\mathfrak{A}_i$. In order to prove this inductively, we need to introduce the notion of "safe" positions h, which represent good starting positions for Duplicator.

We first introduce some additional notation. We write $\mathfrak{B}_i(b)$ for the restriction of $\mathfrak{B}_i$ to the subtree rooted at b. For $i > 0$, let C_i (D_i, respectively) denote the set of nodes in $\mathfrak{A}_i$ ($\mathfrak{B}_i$, respectively) at depth 2 (these are the roots of the $\mathfrak{A}_{i-1}$ and $\mathfrak{B}_{i-1}$ subtrees of $\mathfrak{A}_i$ and $\mathfrak{B}_i$). We write $\boldsymbol{b}\restriction_b$ for the restriction of $\boldsymbol{b}$ to those elements appearing in $\mathfrak{B}_i(b)$.

We now define inductively on i what it means to be (i, k, m)-*safe*. A strategy in the m-round game is (i, k, m)-safe if it uses only (i, k, m')-safe positions where m' is the number of remaining moves in the game. A position $h : \boldsymbol{b} \mapsto \boldsymbol{a}$ is (i, k, m)-safe if it satisfies the following properties:

(S1) for all $b \in \boldsymbol{b}$, depth$(b) = $ depth$(h(b))$;
(S2) if $i > 0$, then there exists a mapping $f : D_i \to C_i$ such that for all $b, b' \in D_i$,
  —if b, b' are siblings, then $f(b)$, $f(b')$ are siblings;
  —if $d \in \boldsymbol{b}$ is the parent of b, then $h(d)$ is the parent of $f(b)$;
  —if $d \in \boldsymbol{b}$ appears in $\mathfrak{B}_i(b)$, then $h(d)$ appears in $\mathfrak{A}_i(f(b))$;
  —$\mathfrak{B}_i(b), \boldsymbol{b}\restriction_b \to_{\text{gn}[\rho_n]}^{k,m} \mathfrak{A}_i(f(b)), h(\boldsymbol{b}\restriction_b)$ via a strategy for Duplicator that is $(i-1, k, m)$-safe.

Likewise, we say a position $h : \boldsymbol{a} \mapsto \boldsymbol{b}$ is (i, k, m)-safe if the conditions (S1) and (S2) hold with $\mathfrak{B}_i$, $\boldsymbol{b}$, D_i exchanged with $\mathfrak{A}_i$, $\boldsymbol{a}$, C_i, respectively.

Observe that the empty partial rigid homomorphism h from $\mathfrak{B}_i$ to $\mathfrak{A}_i$ is (i, k, m)-safe: (S1) vacuously holds, and (S2) holds since (if $i > 0$) we can choose f such that sibling relationships are preserved, and $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtrees in $\mathfrak{B}_i$ are

mapped to $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtrees in $\mathfrak{A}_i$; similarly for the empty partial rigid homomorphism from $\mathfrak{A}_i$ to $\mathfrak{B}_i$.

More importantly, we can show that Duplicator has a winning (i, k, m)-safe strategy when starting from an (i, k, m)-safe position.

LEMMA 5.3. *For all k, $i \le 2^{2^n-1}$, and $m \le i$,*

—*if $\boldsymbol{b} \mapsto \boldsymbol{a}$ is (i, k, m)-safe, then $\mathfrak{B}_i, \boldsymbol{b} \to^{k,m}_{\mathrm{gn}[\rho_n]} \mathfrak{A}_i, \boldsymbol{a}$ via an (i, k, m)-safe strategy for Duplicator;*

—*if $\boldsymbol{a} \mapsto \boldsymbol{b}$ is (i, k, m)-safe, then $\mathfrak{A}_i, \boldsymbol{a} \to^{k,m}_{\mathrm{gn}[\rho_n]} \mathfrak{B}_i, \boldsymbol{b}$ via an (i, k, m)-safe strategy for Duplicator.*

PROOF. We fix k, and proceed by induction on i and $m \le i$.

The base cases are covered (for $i = 0$ and $m = 0$, or $i > 0$ and $m = 0$), since Duplicator wins by default when $m = 0$.

Assume $i > 0$ and $m > 0$ and the starting position h is (i, k, m)-safe. We assume that the active structure is $\mathfrak{B}_i$, the current position is $h : \boldsymbol{b} \to \boldsymbol{a}$, and (S2) is witnessed by a map $f : D_i \to C_i$. The proof is similar when the active structure is $\mathfrak{A}_i$. For each possible move of Spoiler, we describe Duplicator's move, leading to an $(i, k, m-1)$-safe position h'. By the inductive hypothesis (on m), this is enough to ensure that we can extend this to a full winning strategy for Duplicator.

Restrict. If Spoiler restricts to elements $\boldsymbol{b}' \subseteq \boldsymbol{b}$, then the new position $h' := h\!\restriction_{\boldsymbol{b}'}$ is still (i, k, m)-safe, and hence $(i, k, m-1)$-safe.

Switch. If Spoiler switches structures to the new position $h' := \boldsymbol{a} \mapsto \boldsymbol{b}$, then $\boldsymbol{b}$ must be a single position or neighboring positions in the AND/OR tree (since $\boldsymbol{b}$ must be guarded). We must show h' is $(i, k, m-1)$-safe. (S1) trivially holds since no new elements have been added. We must define $f' : C_i \to D_i$ satisfying (S2) (for i, k, and $m-1$).

Assume $\boldsymbol{b}$ corresponds to a single element b.

If b is at depth at least 2, then let $d \in D_i$ be the node at depth 2 such that $\mathfrak{B}_i(d)$ contains b. Let $a := h(b)$ and $c := f(d) \in C_i$. We define f' such that $f'(c) := d$. Then $\mathfrak{A}_i(c)$ contains a and $\mathfrak{B}_i(d)$ contains b, and we can use the $(i-1, k, m)$-safe strategy from $\mathfrak{B}_i(d), b \to^{k,m}_{\mathrm{gn}[\rho_n]} \mathfrak{A}_i(f(d)), h(b)$ guaranteed by (S2) for h, to see that $\mathfrak{A}_i(c), a \to^{k,m-1}_{\mathrm{gn}[\rho_n]} \mathfrak{B}_i(f'(c)), h'(a)$ as desired. The sibling c' of c can be mapped by f' to the sibling d' of d (in fact, it would also work to map to d). If the sibling of c corresponds to an $\mathfrak{A}_{i-1}$ subtree, then $\mathfrak{A}_i(c')$ may correspond to an $\mathfrak{A}_{i-1}$ subtree and $\mathfrak{B}_i(d')$ to a $\mathfrak{B}_{i-1}$ subtree. Because $\mathrm{dom}(h')$ does not include any elements from these mismatched subtrees, the inductive hypothesis (on i) ensures that $\mathfrak{A}_{i-1}, \emptyset \to^{k,m-1}_{\mathrm{gn}[\rho_n]} \mathfrak{B}_{i-1}, \emptyset$ using an $(i-1, k, m-1)$-safe strategy, so $\mathfrak{A}_i(c'), \emptyset \to^{k,m-1}_{\mathrm{gn}[\rho_n]} \mathfrak{B}_i(d'), \emptyset$ using an $(i-1, k, m-1)$-safe strategy as desired. The remaining elements $c' \in C_i$ can be mapped to elements in positions **ll** and **lr** in $\mathfrak{B}_i$ such that if c' is the root of an $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtree in $\mathfrak{A}_i$, then $f'(c')$ is the root of an $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtree in $\mathfrak{B}_i$. In these subgames, Duplicator clearly has an $(i-1, k, m-1)$-safe winning strategy, since the arenas are isomorphic.

If b is at depth 1, then $h(b)$ is also at depth 1 by (S1). Let c_1 and c_2 be the children of $h(b)$ in $\mathfrak{A}_i$. Set $f'(c_1)$ and $f'(c_2)$ to be the corresponding children of b in $\mathfrak{B}_i$. As in the previous case, this may result in mapping a $\mathfrak{A}_{i-1}$ subtree in $\mathfrak{A}_i$ to a $\mathfrak{B}_{i-1}$ subtree in $\mathfrak{B}_i$, but the inductive hypothesis (on i) ensures that (S2) still holds for these positions. The remaining elements $c' \in C_i$ can be mapped to elements in positions **ll** and **lr** in $\mathfrak{B}_i$ such that if c' is the root of an $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtree in $\mathfrak{A}_i$, then $f'(c')$ is the root of an $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtree in $\mathfrak{B}_i$. In these subgames, Duplicator clearly has an $(i-1, k, m-1)$-safe winning strategy, since the arenas are isomorphic.

If b is the root in $\mathfrak{B}_i$, then $h(b)$ must be the root in $\mathfrak{A}_i$. Let f' be the map $\mathbf{ll} \mapsto \mathbf{ll}$, $\mathbf{lr} \mapsto \mathbf{lr}$, $\mathbf{rl} \mapsto \mathbf{ll}$, $\mathbf{rr} \mapsto \mathbf{lr}$. Notice that we are "cheating" for nodes in the right subtree of $\mathfrak{A}_i$ by mapping them to the left subtree in $\mathfrak{B}_i$. This clearly satisfies (S2), since we are mapping $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtrees in $\mathfrak{A}_i$ to $\mathfrak{A}_{i-1}$ ($\mathfrak{B}_{i-1}$, respectively) subtrees in $\mathfrak{B}_i$.

Now assume $\boldsymbol{b}$ is a pair of neighboring elements. If $\boldsymbol{b}$ are neighboring positions both of depth at least 2, then the argument is similar to the preceding depth 2 case. If $\boldsymbol{b}$ spans the root and depth 1, then the argument is similar to the preceding depth 1 case. If $\boldsymbol{b}$ spans depth 1 and depth 2, then the argument is similar to the preceding depth 2 case.

Add. If Spoiler adds an element b in $\mathfrak{B}_i$, then we must select a in $\mathfrak{A}_i$ such that $h' := h[b \mapsto a]$ is $(i, k, m-1)$-safe.

If b is at depth at least 2, then let b_2 be the node at depth 2 such that b is in $\mathfrak{B}_i(b_2)$. By (S2), $\mathfrak{B}_i(b_2), \boldsymbol{b}|_{b_2} \to^{k,m}_{\mathrm{gn}[\rho_n]} \mathfrak{A}_i(f(b_2)), h(\boldsymbol{b}|_{b_2})$ using an $(i-1, k, m)$-safe strategy. Select a based on Duplicator's move according to this strategy when Spoiler plays b in $\mathfrak{B}_i(b_2)$ (the strategy will give a position in $\mathfrak{A}_i(f(b_2))$, so a should be the element in the corresponding position in $\mathfrak{A}_i$). We know that $h[b \mapsto a]$ is $(i-1, k, m-1)$-safe restricted to $\mathfrak{B}_i(b_2)$ and $\mathfrak{A}_i(f(b_2))$. This ensures that (S1) holds. We set $f' := f$. We must check that $h' := h[b \mapsto a]$ is a partial rigid homomorphism from $\mathfrak{B}_i$ to $\mathfrak{A}_i$ (not just restricted to the $(i-1)$-subtrees). The delicate situation is if $b = b_2$, and the parent d of b is in $\boldsymbol{b}$, since this means AOChild db holds (recall that AOChild is the successor relation in the AND/OR tree). In that case, (S2) for h ensures that the parent of $f(b)$ is $h(d)$. But in order for h' to be $(i-1, k, m-1)$-safe restricted to $\mathfrak{B}_i(b_2)$ and $\mathfrak{A}_i(f(b_2))$, a must be $f(b)$. Since the parent of a is $h(d) = h'(d)$, this ensures that AOChild $h'(d)h'(b)$ holds, so h' is still a partial rigid homomorphism. The conditions of (S2) for $(i, k, m-1)$-safety follow since h was (i, k, m)-safe using f, and $f' := f$. Thus, $h[b \mapsto a]$ is $(i, k, m-1)$-safe.

If b is at depth 1, then let d_1 and d_2 be the nodes at depth 2 that are children of b. By (S2), it must be the case that $f(d_1)$ and $f(d_2)$ are children of a single node at depth 1 in $\mathfrak{A}_i$ (though $f(d_1)$ could be equal to $f(d_2)$). Select a to be the parent of $f(d_1)$ and $f(d_2)$. This ensures $h[b \mapsto a]$ is a partial rigid homomorphism (using the fact that if $d' \in \boldsymbol{b}$ is in $\mathfrak{B}_i(d_j)$, then $h(d')$ is in $\mathfrak{A}_i(f(d_j))$). (S1) still holds since a is a node at depth 1. (S2) still holds with $f' := f$ since no elements in depth at least 2 have changed, and we have respected siblings. Hence, $h[b \mapsto a]$ is $(i, k, m-1)$-safe.

If b is at the root, then select a to be the root in $\mathfrak{A}_i$. This is a partial rigid homomorphism since depths of elements are preserved by h. (S1) clearly still holds. (S2) holds with $f' := f$. Thus, $h[b \mapsto a]$ is $(i, k, m-1)$-safe. $\square$

This implies that for $i = 2^{2^n-1}$, Duplicator has a winning strategy in the i-round i-width GN-bisimulation game relative to signature ρ_n between the indexed AND/OR trees $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ starting from the empty homomorphism.

5.4. Lower Bounds

Interpolation for GNF. We are now ready to prove Theorem 5.1.

Let $i = 2^{2^n-1}$. Then the indexed AND/OR trees $\mathfrak{A}_i$ and $\mathfrak{B}_i$ defined in Section 5.3 are of depth $2i = 2^{2^n}$. In particular, note that $\mathfrak{A}_i \models \chi_n$ (when viewed as a σ_n structure), but $\mathfrak{B}_i \not\models \chi_n'$ (when viewed as a σ_n' structure).

By Lemma 5.3 (and the observation that the empty partial rigid homomorphism is safe), $\mathfrak{A}_i \sim^{i,i}_{\mathrm{gn}[\rho_n]} \mathfrak{B}_i$. Hence, by Proposition 5.2, $\mathfrak{A}_i$ and $\mathfrak{B}_i$ must agree on all DAG-represented GNF sentences φ over the signature ρ_n such that $|\varphi| \leq i$ and width$(\varphi) \leq i$.

Since $\chi_n \models \chi_n'$, there is a GNF sentence θ_n over the signature ρ_n such that $\chi_n \models \theta_n$ and $\theta_n \models \chi_n'$.

Assume for the sake of contradiction that θ_n is of size at most i (which implies that there is also a DAG representation of θ_n of size at most i). Since $\mathfrak{A}_i \models \chi_n$, it must be the case that $\mathfrak{A}_i \models \theta_n$. Since a DAG representation of θ_n is at most size i and width(θ_n) $\leq i$, $\mathfrak{A}_i$ and $\mathfrak{B}_i$ must agree on θ_n, so $\mathfrak{B}_i \models \theta_n$. But $\mathfrak{B}_i \models \theta_n$ implies $\mathfrak{B}_i \models \chi'_n$, which is a contradiction since the value at the root of $\mathfrak{B}_i$ based on the internal calculations using AOValOne' is 0, not 1 as asserted by χ'_n.

Beth Definability. A similar argument can be used to prove doubly exponential lower bounds on the DAG size of GNF explicit definitions coming from Beth Definability.

THEOREM 5.4. *There is a polynomial function p and a family of* UNF *sentences ϕ_n over $\sigma_n = \rho_n \cup \{\text{AOValOne}\}$ with $|\phi_n| \leq p(n)$ that implicitly define* AOValOne*, and there is no explicit definition ψ_n of* AOValOne *in GNF over ρ_n with size at most 2^{2^n-1}, even when ψ_n is represented via a DAG.*

Indeed, we can implicitly define AOValOne by a UNF formula ϕ_n over signature σ_n expressing: there is AND/OR tree of depth 2^{2^n}, and every AND/OR tree has depth 2^{2^n}, and ConsistentAOValues.

Assume for the sake of contradiction that there is a GNF formula $\psi_n(x)$ explicitly defining AOValOne over signature ρ_n, and such that $|\psi_n(x)| \leq 2^{2^n-1}$ (so there is a DAG representation with this size bound as well). Let $i = 2^{2^n-1}$ and let a, b be the roots of the structures $\mathfrak{A}_i$ and $\mathfrak{B}_i$ over signature σ_n defined in Section 5.3. These structures satisfy the implicit definition ϕ_n. We have $\mathfrak{A}_i \models \text{AOValOne}\, a$, so $\mathfrak{A}_i \models \psi_n(a)$. But $\mathfrak{A}_i, a \sim^{k,m}_{\text{gn}[\rho_n]} \mathfrak{B}_i, b$ for $i = k = m = 2^{2^n-1}$ by Lemma 5.3 (and observing that $a \mapsto b$ and $b \mapsto a$ are (i, k, m)-safe since a and b are the roots). This means $\mathfrak{B}_i \models \psi_n(b)$ by Proposition 5.2, so $\mathfrak{B}_i \models \text{AOValOne}\, b$, a contradiction.

Interpolation for GF. A variant of this construction inspired by Grädel [1999] can be used to get a lower bound for the size of interpolants of GF validities.

THEOREM 5.5. *There is a polynomial function p and a family of* GF *sentences $\chi_n \models \chi'_n$ with relations of unbounded arity (bounded arity, respectively) such that $|\chi_n| + |\chi'_n| \leq p(n)$, and there is no GNF interpolant θ_n for $\chi_n \models \chi'_n$ with size at most 2^{2^n} (size at most 2^n, respectively), even when θ_n is represented by a DAG.*

In Grädel [1999], a family of GF sentences φ_n using equality and using relations of unbounded arity are defined such that every model $\mathfrak{A}$ of φ_n contains a binary tree of depth 2^{2^n}. This is described in detail in the proof of Theorem 4.4 in Grädel [1999] (where it is used to show that satisfiability is 2ExpTime-hard for GF).

We can adapt this construction to define indexed AND/OR trees with small (polynomial in n) size GF formulas. A technical difference in this approach is that a node in the AND/OR tree is specified by a pair of elements in the structure rather than a single element, and equality is used in the GF formulas.

As before, this allows us to prove a doubly exponential lower bound on the size of GNF interpolants (even represented as a DAG) for GF validities, but it relies on relations of unbounded arity.

For the single exponential lower bound for relations of bounded arity, we can modify the construction in a different way. Instead of attaching index trees to each node in the AND/OR tree, we can introduce new unary relations $\text{IndBit}_0, \ldots, \text{IndBit}_n$, and require that for every node in the AND/OR tree, the index value (when we view these new predicates as n bits of a binary number) corresponds to the depth of this node in the AND/OR tree. This allows us to define indexed AND/OR trees of depth 2^n in GF, which can be used to prove an exponential lower bound on the size of GNF interpolants for GF validities with relations of bounded arity.

6. PRESERVATION THEOREMS

In this section, we consider some preservation theorems for GNF and GF that follow from the constructive interpolation results of Theorem 4.10. We remark that the results in this section hold when preservation and equivalence is considered over finite structures only, as well as over general structures.

We first consider the analog of the Łoś-Tarski theorem. A formula φ is *preserved under extensions* if $\mathfrak{A} \subseteq \mathfrak{B}$ and $\mathfrak{A} \models \varphi$ together imply $\mathfrak{B} \models \varphi$. The classical Łoś-Tarski theorem states that every FO formula preserved under extensions has an equivalent *existential FO formula*, that is, a formula using only existential quantification once negations are pushed inside. It was shown in Bárány et al. [2013b] that every GNF formula preserved under extensions has an equivalent *existential GNF formula*, which is a GNF formula where no quantifier is in the scope of a negation symbol.

Otto [2000] showed a connection between relativized interpolation and the Łoś-Tarski theorem. We adapt this technique, using relativized GNF interpolation to get a bound on the size of an equivalent existential GNF formula. This contrasts with the result of Dawar et al. [2007] that for general first-order formulas there is no elementary bound for the Łoś-Tarski theorem.

COROLLARY 6.1 (ANALOG OF ŁOŚ-TARSKI THEOREM). *Let φ be a GNF formula over signature σ with $|\varphi| = n$. If φ is preserved under extensions, then we can construct a DAG representation of an equivalent existential GNF formula φ' such that $|\varphi'| \leq 2^{2^{p(n)}}$ for some polynomial function p independent of φ.*

PROOF. We follow Otto [2000]. Let $\mathbb{U} := \{U^1, U^2\}$ for distinct unary predicates U^1, U^2 not occurring in σ. Since φ is preserved under extensions, we have

$$\neg \exists y . (U^1 y \wedge \neg U^2 y) \wedge \bigwedge_{t \in \boldsymbol{x} \cup \boldsymbol{e}} U^1 t \wedge \varphi^1 \models \varphi^2,$$

where $\mathrm{free}(\varphi) = \boldsymbol{x}$, $\mathrm{con}(\varphi) = \boldsymbol{e}$, and φ^i denotes the relativization of φ to U^i. For readability, here and in the following similar arguments, the equality guard for the negation of the first conjunct in the antecedent has been omitted.

Let N be the size of this validity, which is linear in n. By Theorem 4.10, the previous GNF validity admits a $\mathbb{U}$-relativized GNF interpolant θ over signature $\sigma \cup \mathbb{U}$ with a DAG representation of size at most $2^{2^{p(N)}}$ for some polynomial function p independent of φ.

Since U^1 does not occur on both sides of the validity, U^1 cannot appear at all in θ. Since U^2 does not occur negatively on the left-hand side of the original validity, U^2 cannot appear negatively in θ. This means that every existential quantifier in θ must occur positively (since every quantifier must be relativized by U^2, and U^2 can only occur positively).

Replacing every occurrence of some $U^2 z$ in θ by $z = z$ results in a GNF formula φ' over σ that is equivalent to φ, and such that any quantifier occurs within the scope of an even number of negations. By pushing the negations to the atomic level (and potentially duplicating guards to cover any negated atoms), we have a doubly exponential size DAG representation of an existential GNF formula. □

Similarly, we can use the Lyndon interpolation result to prove a relationship between monotonicity and positivity. We say a formula φ is *monotone* if it is preserved when a structure is modified via adding tuples to the interpretation of some relation. We say a formula φ is *positive* if all relation symbols (except possibly equality) occur positively (within the scope of an even number of negations). The classical result states that monotone FO formulas are equivalent to positive FO formulas. We can prove an effective version of this result for GNF. We can also refine this result in the

domain-independent case (a formula is *domain independent* if it depends only on the interpretations of relation symbols; see Abiteboul et al. [1995]).

COROLLARY 6.2 (MONOTONE=POSITIVE). *Let φ be a GNF formula over signature σ with $|\varphi| = n$ such that φ is monotone. Then we can construct a DAG representation of an equivalent positive GNF formula φ' such that $|\varphi'| \leq 2^{2^{p(n)}}$ for some polynomial function p independent of φ. Furthermore, if φ is domain independent, then φ' can be taken to be positive and in existential GNF.*

PROOF. Since φ is monotone, we have

$$\bigwedge_{R \in \sigma} \neg \exists \boldsymbol{y}. (R^1 \boldsymbol{y} \wedge \neg R^2 \boldsymbol{y}) \wedge \varphi^1 \models \varphi^2,$$

where φ^i denotes φ with each $R \in \sigma$ replaced with R^i. Applying Theorem 4.10 yields an interpolant only using relations of the form R^2 for $R \in \sigma$. Moreover, these relations must occur positively. Replacing each R^2 with R, yields φ'.

The preceding argument is modified in the domain-independent case by using additional predicates for the domains. More precisely, we relativize the antecedent and consequent of the preceding entailment by fresh unary predicates U^1 and U^2, respectively; and we add $\bigwedge_{R \in \sigma; i,j \in \{1,2\}} \neg \exists \boldsymbol{x}. (\bigwedge_{x \in \boldsymbol{x}} U^i x \wedge R^1 \boldsymbol{x} \wedge \neg \bigwedge_{x \in \boldsymbol{x}} U^j x)$ as conjunct to the antecedent. Applying Theorem 4.10 as before, and replacing each R^2 with R, as well as $U^2 z$ by $z = z$, yields the desired formula φ'. □

Guarded Fragment. We now turn to effective preservation for GF. Recall that Andréka et al. [1998] prove a Łoś-Tarski theorem for GF, but with no bound. Here we state a constructive version for sentences.

THEOREM 6.3. *Let φ be a GF sentence over signature σ with $|\varphi| = n$. If φ is preserved under extensions, then we can construct an equivalent existential GF sentence φ' such that $|\varphi'| \leq 2^{2^{2^{2^{p(n)}}}}$ for a polynomial function p independent of φ.*

PROOF. By Corollary 6.1, we can find a DAG representation of some GNF existential χ' that is equivalent to φ and is of doubly exponential size. We can convert χ' to χ'' that is a union of χ_i'', where each χ_i'' is a "GNCQ"—a conjunctive query with atomic negation, where every negated atom is guarded. Further, each GNCQ χ_i'' within χ'' is at most exponential in the size of χ', and the overall size of χ'' (as a formula) is at most doubly exponential in the size of χ'. Let A_i be the set of formulas in GF of the form $\exists \boldsymbol{y}. \bigwedge_j A_{ij}$, where $\bigwedge A_{ij}$ is obtained by identifying variables in χ_i'' and then adding on at most $2^{|\varphi|}$ additional positive atoms. We claim that χ' is equivalent to $\bigvee_i \bigvee_{Q \in A_i} Q$. Since the set of indices i is doubly exponential in the size of χ' and each A_i has size doubly exponential in χ', this gives the desired bound. Clearly, each formula in A_i implies χ_i'' and hence implies φ. On the other hand, consider a model M satisfying φ. Given any model M there is another structure M^* agreeing with M on all GF sentences that has a guarded tree decomposition. Indeed, the guarded unraveling of M (see, e.g., Andréka et al. [1998]) gives such an M^*. M^* must satisfy φ and therefore must satisfy some χ_i'' via a homomorphism h. Let C be the image of h and Q_C be the conjunctive query that describes the restriction of M^* to the bags containing C—such a formula can be written in GF by constructing it inductively from the leaves of the tree up. The size of Q_C is at most the size of χ_i'' times the maximal size of bags—the latter being a factor that is at most exponential in the size of the schema. Note that every element $c \in C$ corresponds to exactly one variable y_c of Q_C. We let Q_C' extend Q_C by adding, for every negated atom in χ_i'', the result of replacing each variable x_i by $y_{h(x_i)}$. Then Q_C' also holds in M^* and $Q_C' \in$ GF. Hence, Q_C' holds in M and is in A_i, which completes the argument. □

We get similar bounds for monotone domain-independent GF sentences.

THEOREM 6.4. *Let φ be a GF sentence over signature σ with $|\varphi| = n$. If φ is monotone and domain independent, then we can construct an equivalent sentence φ' that is positive and is in existential GF, and such that $|\varphi'| \leq 2^{2^{2^{2^{p(n)}}}}$ for a polynomial function p independent of φ.*

PROOF. By Corollary 6.2, φ is equivalent to a positive and existential GNF sentence χ' whose DAG representation has doubly exponential size. As previously, we can convert this to a union $\chi'' = \bigvee_i \chi''_i$, with χ''_i in this case being a conjunctive query, but possibly including guarded inequalities. As before, this results in a doubly exponential blowup in the overall size, but only a singly exponential blowup in the size of each CQ.

Again letting A_i be the set of formulas in GF of the form $\exists \boldsymbol{y}. \bigwedge_j A_{ij}$, where $\bigwedge A_{ij}$ is obtained by identifying variables in χ''_i and then adding on at most $2^{|\varphi|}$ additional positive atoms, we can show that χ' is equivalent to $\bigvee_i \bigvee_{Q \in A_i} Q$. □

7. CONCLUSIONS

We have shown that GNF is a rich logic that admits interpolation and preservation theorems with elementary complexity, in sharp contrast to full first-order logic. In doing this, we developed the mosaic method for use in interpolation, and have gained insight into the flexibility of the mosaic method.

In future work, we will look at revisiting the applications of GNF interpolation to databases (as in Bárány et al. [2013a]) in light of these results. For these applications, the mosaic method does not seem like an attractive proof system to use because of the high levels of nondeterminism involved. It would be interesting to develop a more traditional proof system for GNF that would allow interpolants of doubly exponential DAG size to be extracted.

The bounds for interpolation are tight for GNF and UNF, in terms of the size of the DAG representation of the output. We leave open the question of tightness of our bounds for formula size and for preservation. We also do not know the status of many other preservation theorems for GNF, such as preservation under surjective homomorphism, and the status of monotone=positive for GF formulas that are not domain independent.

REFERENCES

Serge Abiteboul, Richard Hull, and Victor Vianu. 1995. *Foundations of Databases*. Addison-Wesley.

Hajnal Andréka, István Németi, and Johan van Benthem. 1998. Modal languages and bounded fragments of predicate logic. *Journal of Philosophical Logic* 27, 3 (1998), 217–274.

Jeremy Avigad. 2003. Eliminating definitions and Skolem functions in first-order logic. *ACM Transactions on Computational Logic* 4, 3 (2003).

Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider (Eds.). 2003. *The Description Logic Handbook*. Cambridge.

Vince Bárány, Michael Benedikt, and Pierre Bourhis. 2013a. Access patterns and integrity constraints revisited. In *ICDT*.

Vince Bárány, Michael Benedikt, and Balder ten Cate. 2013b. Rewriting guarded negation queries. In *MFCS*.

Vince Bárány, Balder ten Cate, and Luc Segoufin. 2011. Guarded negation. In *ICALP*.

Michael Benedikt, Balder ten Cate, and Michael Vanden Boom. 2014. Effective interpolation and preservation in guarded logics. In *CSL-LICS*.

William Craig. 1957. Three uses of the Herbrand-Gentzen theorem in relating model theory and proof theory. *Journal of Symbolic Logic* 22, 3 (1957), 269–285.

Giovanna D'Agostino and Giacomo Lenzi. 2015. Bisimulation quantifiers and uniform interpolation for guarded first order logic. *Theoretical Computer Science* 563 (2015), 75–85.

Anuj Dawar, Martin Grohe, Stephan Kreutzer, and Nicole Schweikardt. 2007. Model theory makes formulas large. In *ICALP*.

Melvin Fitting. 1996. *First-Order Logic and Automated Reasoning*. Springer.

Harvey Friedman. 1976. The complexity of explicit definitions. *Advances in Mathematics* 20, 1 (1976), 18–29.

Erich Grädel. 1999. On the restraining power of guards. *Journal of Symbolic Logic* 64, 4 (1999), 1719–1742.

Eva Hoogland. 2000. *Definability and Interpolation: Model-Theoretic Investigations*. Ph.D. dissertation. University of Amsterdam.

Eva Hoogland, Maarten Marx, and Martin Otto. 1999. Beth definability for the guarded fragment. In *LPAR*.

Roger C. Lyndon. 1959. An interpolation theorem in the predicate calculus. *Pacific Journal of Mathematics* 9 (1959), 129–142.

Maarten Marx. 2007. Queries determined by views: Pack your views. In *PODS*.

Maarten Marx, Szabolcs Mikulás, and Mark Reynolds. 2000. The mosaic method for temporal logics. In *TABLEAUX*.

Ken McMillan. 2004. Applications of Craig interpolation to model checking. In *CSL*.

Szabolcs Mikulás. 1998. Taming first-order logic. *Journal of the IGPL* 6, 2 (1998), 305–316.

Nobuyoshi Motohashi. 1984. Equality and Lyndon's interpolation theorem. *Journal of Symbolic Logic* 49, 1 (1984), 123–128.

Alan Nash, Luc Segoufin, and Victor Vianu. 2010. Views and queries: Determinacy and rewriting. *ACM Transactions on Database Systems* 35, 3 (2010).

Istvan Németi. 1986. *Free Algebras and Decidability in Algebraic Logic*. Ph.D. dissertation. Hungarian Academy of Sciences, Budapest.

Martin Otto. 2000. An interpolation theorem. *Bulletin of Symbolic Logic* 6, 4 (2000), 447–462.

Balder ten Cate, Enrico Franconi, and Inanc Seylan. 2011. Beth definability in expressive description logics. In *IJCAI*.

Balder ten Cate and Luc Segoufin. 2011. Unary negation. In *STACS*.