CSPs with Few Alien Constraints

Peter Jonsson ✉
Department of Computer and Information Science, Linköping University, Linköping, Sweden

Victor Lagerkvist ✉
Department of Computer and Information Science, Linköping University, Linköping, Sweden

George Osipov ✉
Department of Computer and Information Science, Linköping University, Linköping, Sweden

Abstract

The *constraint satisfaction problem* asks to decide if a set of constraints over a relational structure $\mathcal{A}$ is satisfiable (CSP($\mathcal{A}$)). We consider CSP($\mathcal{A} \cup \mathcal{B}$) where $\mathcal{A}$ is a structure and $\mathcal{B}$ is an *alien* structure, and analyse its (parameterized) complexity when at most k alien constraints are allowed. We establish connections and obtain transferable complexity results to several well-studied problems that previously escaped classification attempts. Our novel approach, utilizing logical and algebraic methods, yields an FPT versus pNP dichotomy for arbitrary finite structures and sharper dichotomies for Boolean structures and first-order reducts of $(\mathbb{N}, =)$ (equality CSPs), together with many partial results for general ω-categorical structures.

2012 ACM Subject Classification Theory of computation → Complexity theory and logic; Mathematics of computing → Discrete mathematics

Keywords and phrases Constraint satisfaction, parameterized complexity, hybrid theories

Digital Object Identifier 10.4230/LIPIcs...

Related Version This is an extended version of *CSPs with Few Alien Constraints*, appearing in Proceedings of the 30th Conference on Principles and Practice of Constraint Programming (CP-2024).

Funding The first author is partially supported by the Swedish research council under grant VR-2021-04371. The second author is partially supported by the Swedish research council under grant VR-2022-03214. The first and third author was supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

1 Introduction

The *constraint satisfaction problem* over a structure $\mathcal{A}$ (CSP($\mathcal{A}$)) is the problem of verifying whether a set of constraints over $\mathcal{A}$ admits at least one solution. This problem framework is vast, and, just to name a few, include all Boolean satisfiability problems as well as k-coloring problems, and for infinite domains we may formulate both problems centrally related to model checking first-order formulas and qualitative reasoning. Notable examples where complete complexity dichotomies are known (separating tractable from NP-hard problems) include *all* finite structures [13, 27] and first-order definable relations over well-behaved base structures like $(\mathbb{N}, =)$ and $(\mathbb{Q}, <)$ [2]. While impressive mathematical achievements, these dichotomy results are still somewhat unsatisfactory from a practical perspective since we are unlikely to encounter instances which are based on *purely* tractable constraints. Could it be possible to extend the reach of these powerful theoretical results by relaxing the basic setting so that we may allow greater flexibility than purely tractable constraints while still obtaining something simpler than an arbitrary NP-hard CSP?

We consider this problem in a *hybrid* setting via problems of the form CSP($\mathcal{A} \cup \mathcal{B}$) where $\mathcal{A}$ is a "stable", tractable background structure and $\mathcal{B}$ is an *alien* structure. We focus on the case when CSP($\mathcal{A} \cup \mathcal{B}$) is NP-hard (thus, richer than a polynomial-time solvable problem) but where we have comparably few constraints from the alien structure $\mathcal{B}$. This problem is compatible with the influential framework of *parameterized complexity* which has been used with great effect to study *structurally* restricted problems (e.g., based on tree-width) but where comparably little is known when one simultaneously restricts the allowed constraints.

We begin (in Section 3) by relating the CSP problem with alien constraints to other problems, namely, (1) *model checking*, (2) the problem of checking whether a constraint in a CSP instance is *redundant*, (3) the *implication* problem and (4) the *equivalence* problem. We prove that the latter three problems are equivalent under Turing reductions and provide a general method for obtaining complexity dichotomies for all of these problems via a complexity dichotomy for the CSP problem with alien constraints. Importantly, all of these problems are well-known in their own right, but have traditionally been studied with wildly disparate tools and techniques, but by viewing them under the unifying lens of alien constraints we not only get four dichotomies for the price of one but also open the powerful toolbox based on *universal algebra*. For non-Boolean domains this is not only a simplifying aspect but an absolute necessity to obtain general results. We expand upon the algebraic approach in Section 4 and relate alien constraints to *primitive positive definitions* (pp-definitions) and the important notion of a *core*. As a second general contribution we explore the case when each relation in $\mathcal{B}$ can be defined via an *existential positive formula* over $\mathcal{A}$. This results in a general *fixed-parameter tractable* (FPT) algorithm (with respect to the number of alien constraints) applicable to both finite, and, as we demonstrate later, many natural classes of structures over infinite domains.

In the second half of the paper we attack the complexity of alien constraints more systematically. We begin with structures over finite domains where we obtain a general tractability result by combining the aforementioned FPT algorithm together with the CSP dichotomy theorem [13, 27]. In a similar vein we obtain a general hardness result based on a universal algebraic gadget. Put together this yields a general result: if $\mathcal{A} \cup \mathcal{B}$ is a core (which we may assume without loss of generality) then either CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is FPT, or CSP$_{\leq p}(\mathcal{A} \cup \mathcal{B})$ is NP-hard for some $p \geq 0$, i.e., is *para-NP-hard* (pNP-hard). Thus, from a parameterized complexity view we obtain a complete dichotomy (FPT versus pNP-hardness) for finite-domain structures. However, to also obtain dichotomies for implication, equivalence, and the redundancy problem, we need sharper bounds on the parameter p. We concentrate on two special cases. We begin with Boolean structures in Section 5.2 and obtain a complete classification which e.g. states that $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is FPT if $\mathcal{A}$ is in one of the classical *Schaefer* classes, and give a precise characterization of $\text{CSP}_{\leq p}(\mathcal{A} \cup \mathcal{B})$ for all relevant values of p if $\mathcal{A}$ is not Schaefer. For example, if we assume that $\mathcal{A}$ is Horn, we may thus conclude that $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is FPT for *any* alien Boolean structure $\mathcal{B}$. More generally this dichotomy is sufficiently sharp to also yield dichotomies for implication, equivalence, and redundancy. Compared to the proofs by Schnoor & Schnoor [25] for implication and Böhler [12] for equivalence, we do not use an exhaustive case analysis over Post's lattice.

In Section 6 we consider structures over infinite domains. If we assume that $\mathcal{A}$ and $\mathcal{B}$ are ω-categorical, then we manage to lift the FPT algorithm based on existential positive definability from Section 4 to the infinite setting. Another important distinction is that the notion of a core, and subsequently the common trick of singleton expansion, works differently for ω-categorical languages. Here we follow Bodirsky [2] and use the notion of a *model-complete core*, which means that all n-ary orbits are pp-definable, where an orbit is defined as the action of the automorphism group over a fixed n-ary tuple. This allows us to, for example, prove that $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is FPT whenever $\mathcal{A}$ is an ω-categorical model-complete core and $\text{CSP}(\mathcal{A})$ is in P such that the orbits of the automorphism group of $\mathcal{B}$ are included in the orbits of the automorphism group of $\mathcal{A}$. This forms a cornerstone for the dichotomy for equality languages since the only remaining cases are when $\mathcal{A}$ is 0-valid (meaning that each relation contains a constant tuple) but not Horn (defined similarly to the Boolean domain), and when $\mathcal{B}$ is not 0-valid. The remaining cases are far from trivial, however, and we require the algebraic machinery from Bodirsky et al. [4] which provides a characterization of equality languages in terms of their *retraction* to finite domains. We rely on this description via a recent classification result by Osipov & Wahlström [21]. Importantly, our dichotomy result is sufficiently sharp to additionally obtain complexity dichotomies for the implication, equivalence, and redundancy problems. To the best of our knowledge, these dichotomies are the first of their kind for arbitrary equality languages.

We finish the paper with a comprehensive discussion in Section 7. Most importantly, we have opened up the possibility to systematically study not only alien constraints, but also related problems that have previously escaped complexity classifications. For future research the main open questions are whether (1) sharper results can be obtained for arbitrary finite domains and (2) which further classes of infinite domain structures should be considered.

Proofs of statements marked with ($\star$) can be found in the appendix in the end of the paper.

2 Preliminaries

We begin by introducing the basic terminology and the fundamental problems under consideration. We assume throughout the paper that the complexity classes P and NP are distinct. We let $\mathbb{Q}$ denote the rationals, $\mathbb{N} = \{0, 1, 2, \dots\}$ the natural numbers, $\mathbb{Z} = \{\dots, -2. -1, 0, 1, 2, \dots\}$ the integers, and $\mathbb{Z}_+ = \{1, 2, 3, \dots\}$ the positive integers. For every $c \in \mathbb{Z}_+$, we let $[c] = \{1, 2, \dots, c\}$.

A *parameterized problem* is a subset of $\Sigma^* \times \mathbb{N}$ where Σ is the input alphabet, i.e., an instance is given by $x \in \Sigma^*$ of size n and a natural number k, and the running time of an algorithm is studied with respect to both k and n. The most favourable complexity class is FPT (*fixed-parameter tractable*), which contains all problems that can be decided in $f(k) \cdot n^{O(1)}$ time with f being some computable function. An *fpt-reduction* from a parameterized problem $L_1 \subseteq \Sigma_1^* \times \mathbb{N}$ to $L_2 \subseteq \Sigma_2^* \times \mathbb{N}$ is a function $P : \Sigma_1^* \times \mathbb{N} \to \Sigma_2^* \times \mathbb{N}$ that preserves membership (i.e., $(x, k) \in L_1 \Leftrightarrow P((x, k)) \in L_2$), is computable in $f(k) \cdot |x|^{O(1)}$ time for some computable function f, and there exists a computable function g such that for all $(x, k) \in L_1$, if $(x', k') = P((x, k))$, then $k' \leq g(k)$. It is easy to verify that if L_1 and L_2 are parameterized problems such that L_1 fpt-reduces to L_2 and L_2 is in FPT, then it follows that L_1 is in FPT, too. There are many parameterized classes with less desirable running times than FPT but we focus on pNP-hard problems: a problem is pNP-hard under fpt-reductions if it is NP-hard for some constant parameter value, implying such problems are not in FPT unless P = NP.

We continue by defining *constraint satisfaction problems*. First, a *constraint language* is a (typically finite) set of relations $\mathcal{A}$ over a universe A, and for a relation $R \in \Gamma$ we write $\text{ar}(R) = k$ to denote its arity k. It is sometimes convenient to associate a constraint language with a relational signature, and thus obtaining a *relational structure*: a tuple $(A; \tau, I)$ where A is the *domain*, or *universe*, τ is a relational signature, and I is a function from σ to the set of all relations over D which assigns each relation symbol R a corresponding relation $R^{\mathcal{A}}$ over D. We write $\text{ar}(R)$ for the arity of a relation R, and if $R = \emptyset$ then $\text{ar}(R) = 0$. All structures in this paper are relational and we assume that they have a finite signature unless otherwise stated. Typically, we do not need to make a sharp distinction between relations and the corresponding relation symbols, so we usually simply write $(A; R_1, \ldots, R_m)$, where each R_i is a relation over A, to denote a structure. We also sometimes do not make a sharp distinction between structures and sets of relations when the signature is not important. For arbitrary structures $\mathcal{A}$ and $\mathcal{A}'$ with domains A and A', we let $\mathcal{A} \cup \mathcal{A}'$ denote the structure with domain $A \cup A'$ and containing the relations in $\mathcal{A}$ and $\mathcal{A}'$.

For a constraint language (or structure) $\mathcal{A}$ an instance of the *constraint satisfaction problem* over $\mathcal{A}$ (CSP($\mathcal{A}$)) is then given by $I = (V, C)$ where V is a set of variables and C a set of constraints of the form $R(x_1, \ldots, x_k)$ where $x_1, \ldots, x_k \in V$ and $R \in \mathcal{A}$, and the question is whether there exist a function $f : V \to A$ that satisfies all constraints (a *solution*), i.e., $(f(x_1), \ldots, f(x_k)) \in R$ for all $R(x_1, \ldots, x_k) \in C$. The *CSP dichotomy theorem* says that all finite-domain CSPs are either in P or are NP-complete [13, 27]. Given an instance $I = (V, C)$ of CSP($\mathcal{A}$), we let $\text{Sol}(I)$ be the set of solutions to I. We now define CSPs with alien constraints in the style of Cohen et al. [15].

$\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$

Instance: A natural number k and an instance $I = (V, C_1 \cup C_2)$ of CSP($\mathcal{A} \cup \mathcal{B}$), where (V, C_1) is an instance of CSP($\mathcal{A}$) and (V, C_2) is an instance of CSP($\mathcal{B}$) with $|C_2| \leq k$.
Question: Does there exist a satisfying assignment to I?

Throughout the paper, we assume without loss of generality that the structures $\mathcal{A}$ and $\mathcal{B}$ can be associated with disjoint signatures. The parameter in $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is the *number of alien constraints* (abbreviated #ac). We let $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ denote the $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ problem restricted to a fixed value k of parameter #ac. Note that if CSP($\mathcal{A}$) is not in P, then $\text{CSP}_{\leq 0}(\mathcal{A} \cup \mathcal{B})$ is not in P; moreover, if CSP($\mathcal{A} \cup \mathcal{B}$) is in P, then $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in P. Thus, it is sensible to always require that CSP($\mathcal{A}$) is in P and CSP($\mathcal{A} \cup \mathcal{B}$) is not in P. In many natural cases (e.g., all finite-domain CSPs), CSP($\mathcal{A} \cup \mathcal{B}$) not being polynomial-time solvable implies that CSP($\mathcal{A} \cup \mathcal{B}$) is NP-hard.

A k-ary relation R is said to have a *primitive positive definition* (pp-definition) over a constraint language $\mathcal{A}$ if $R(x_1, \ldots, x_k) \equiv \exists y_1, \ldots, y_{k'} : R_1(\mathbf{x_1}) \wedge \ldots \wedge R_m(\mathbf{x_m})$ where each $R_i \in \mathcal{A} \cup \{=_A\}$ and each $\mathbf{x_i}$ is a tuple of variables over $x_1, \ldots, x_k, y_1, \ldots, y_{k'}$ matching the arity of R_i. Here, and in the sequel, $=_A$ is the equality relation over A, i.e. $\{(a, a) \mid a \in A\}$. If $\mathcal{A}$ is a constraint language, then we let $\langle \mathcal{A} \rangle$ be the inclusion-wise smallest set of relations containing $\mathcal{A}$ closed under pp-definitions.

▶ **Theorem 1** ([18]). *Let $\mathcal{A}$ and $\mathcal{B}$ be structures with the same domain. If every relation of $\mathcal{A}$ has a primitive positive definition in $\mathcal{B}$, then there is a polynomial-time reduction from CSP($\mathcal{A}$) to CSP($\mathcal{B}$).*

When working with problems of the form $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ we additionally introduce the following simplifying notation: $\langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq k}$ denotes the set of all pp-definable relations over $\mathcal{A} \cup \mathcal{B}$ using at most k atoms from $\mathcal{B}$. We now describe the corresponding algebraic objects. An operation $f \colon D^m \to D$ is a *polymorphism* of a relation $R \subseteq D^k$ if, for any choice of m tuples $(t_{11}, \ldots, t_{1k}), \ldots, (t_{m1}, \ldots, t_{mk})$ from R, it holds that $(f(t_{11}, \ldots, t_{m1}), \ldots, f(t_{1k}, \ldots, t_{mk}))$ is in R. An *endomorphism* is a polymorphism with arity one. If f is a polymorphism of R, then we sometimes say that R is *invariant* under f. A constraint language $\mathcal{A}$ has the polymorphism f if every relation in $\mathcal{A}$ has f as a polymorphism. We let $\text{Pol}(\mathcal{A})$ and $\text{End}(\mathcal{A})$ denote the sets of polymorphisms and endomorphisms of $\mathcal{A}$, respectively. If F is a set of functions over D, then $\text{Inv}(F)$ denotes the set of relations over D that are invariant under every function in F. There are close algebraic connections between the operators $\langle \cdot \rangle$, $\text{Pol}(\cdot)$, and $\text{Inv}(\cdot)$. For instance, if $\mathcal{A}$ has a finite domain (or, more generally, if $\mathcal{A}$ is ω-categorical; see below), then we have a Galois connection $\langle \mathcal{A} \rangle = \text{Inv}(\text{Pol}(\mathcal{A}))$ [9, Theorem 5.1].

Polymorphisms enable us to compactly describe the tractable cases of Boolean CSPs.

▶ **Theorem 2** ([24]). *Let $\mathcal{A}$ be a constraint language over the Boolean domain. The problem CSP($\mathcal{A}$) is decidable in polynomial time if $\mathcal{A}$ is invariant under one of the following six operations: (1) the constant unary operation 0 ($\mathcal{A}$ is 0-valid), (2) the constant unary operation 1 ($\mathcal{A}$ is 1-valid), (3) the binary min operation $\sqcap$ ($\mathcal{A}$ is Horn), (4) the binary max operation $\sqcup$ ($\mathcal{A}$ is anti-Horn), (5) the ternary majority operation $M(x, y, z) = (x \sqcap y) \sqcup (x \sqcap z) \sqcup (y \sqcap z)$ ($\mathcal{A}$ is 2-SAT), or (6) the ternary minority operation $m(x, y, z) = x \oplus y \oplus z$ where $\oplus$ is the addition operator in $GF(2)$ ($\mathcal{A}$ is affine). Otherwise, the problem CSP($\mathcal{A}$) is NP-complete.*

A Boolean constraint language that satisfies condition (3), (4), (5), or (6) is called *Schaefer*.

A finite-domain structure $\mathcal{A}$ is a *core* if every $e \in \text{End}(\mathcal{A})$ is a bijection. We let $f(R) = \{(f(t_1), \ldots, f(t_n)) \mid (t_1, \ldots, t_n) \in R\}$ when $f \colon A \to A$ and $R \in \mathcal{A}$. If $e \in \text{End}(\mathcal{A})$ has minimal range, then $e(\mathcal{A}) = \{e(R) \mid R \in \mathcal{A}\}$ is a core and this core is unique up to isomorphism. We can thus speak about *the core* $\mathcal{A}^c$ of $\mathcal{A}$. It is easy to see that CSP($\mathcal{A}$) and CSP($\mathcal{A}^c$) are equivalent under polynomial-time reductions (indeed, even log-space reductions suffice). Another useful equivalence concerns constant relations. Let $\mathcal{A}^+$ denote the structure $\mathcal{A}$ expanded by all unary singleton relations $\{(a)\}$, $a \in A$. If $\mathcal{A}$ is a core, then CSP($\mathcal{A}$) and CSP($\mathcal{A}^+$) are equivalent under polynomial-time reductions [1].

We will frequently consider *ω-categorical* structures. An *automorphism* of a structure $\mathcal{A}$ is a permutation α of its domain A such that both α and its inverse are homomorphisms. The set of all automorphisms of a structure $\mathcal{A}$ is denoted by $\text{Aut}(\mathcal{A})$, and forms a group with respect to composition. The *orbit* of $(a_1, \ldots, a_n) \in A^n$ in $\text{Aut}(\mathcal{A})$ is the set $\{(\alpha(a_1), \ldots, \alpha(a_n)) \mid \alpha \in \text{Aut}(\mathcal{A})\}$. Let $\text{Orb}(\mathcal{A})$ denote the set of orbits of n-tuples in $\text{Aut}(\mathcal{A})$ (for all $n \geq 1$). A structure $\mathcal{A}$ with countable domain is ω-categorical if and only if $\text{Aut}(\mathcal{A})$ is *oligomorphic*, i.e., it has only finitely many orbits of n-tuples for all $n \geq 1$.

Two important classes of ω-categorical structures are *equality languages* (respectively, *temporal languages*) where each relation can be defined as the set of models of a first-order formula over $(\mathbb{N}; =)$ (respectively, $(\mathbb{Q}; <)$). Importantly, $\text{Aut}(\mathcal{A})$ is the full symmetric group if $\mathcal{A}$ is an equality language. A relation in an equality language is said to be *0-valid* if it contains *any* constant tuple. This is justified since if the relation is invariant under one constant operation, then it is invariant under all constant operations.The computational complexity of CSP for equality languages was classified by Bodirsky and Kára [7, Theorem 1]: for any equality language $\mathcal{A}$, CSP($\mathcal{A}$) is solvable in polynomial time if $\mathcal{A}$ is 0-valid or invariant under a binary injective operation, and is NP-complete otherwise.

3 Applications of Alien Constraints

We will now demonstrate how alien constraints can be used for studying the complexity of CSP-related problem: Section 3.1 contains an example where we analyse the complexity of *redundancy*, *equivalence*, and *implication* problems, and we consider connections between the model checking problem and CSPs with alien constraints in Section 3.2. To relate problem complexity we use *Turing reductions*: a problem L_1 is *polynomial-time Turing reducible* to L_2 (denoted $L_1 \leq_T^p L_2$) if it can be solved in polynomial time using an oracle for L_2. Two problems L_1 and L_2 are *polynomial-time Turing equivalent* if $L_1 \leq_T^p L_2$ and $L_2 \leq_T^p L_1$.

3.1 The Redundancy Problem and its Relatives

We will now study the complexity of a family of well-known computational problems. We begin by some definitions. Let $\mathcal{A}$ denote a constraint language and assume that $I = (V, C)$ is an instance of CSP($\mathcal{A}$). We say that a constraint $c \in C$ is *redundant* in I if $\text{Sol}((V, C)) = \text{Sol}((V, C \setminus \{c\}))$. We have the following computational problems.

Redundant($\mathcal{A}$)
Instance: An instance (V, C) of CSP($\mathcal{A}$) and a constraint $c \in C$.
Question: Is c redundant in (V, C)?

Impl($\mathcal{A}$)
Instance: Two instances $(V, C_1), (V, C_2)$ of CSP($\mathcal{A}$).
Question: Does (V, C_1) imply (V, C_2), i.e., is it the case that $\text{Sol}((V, C_1)) \subseteq \text{Sol}((V, C_2))$?

Equiv($\mathcal{A}$)
Instance: Two instances $(V, C_1), (V, C_2)$ of CSP($\mathcal{A}$).
Question: Is it the case that $\text{Sol}((V, C_1)) = \text{Sol}((V, C_2))$?

Before we start working with alien constraints, we exhibit a close connection between Redundant($\cdot$), Equiv($\cdot$), and Impl($\cdot$).

▶ **Lemma 3.** *Let $\mathcal{A}$ be a structure. The problems* Equiv($\mathcal{A}$)*,* Impl($\mathcal{A}$)*, and* Redundant($\mathcal{A}$) *are polynomial-time Turing equivalent.*

Proof. We show that (1) Equiv($\mathcal{A}$) $\leq_T^p$ Impl($\mathcal{A}$), (2) Impl($\mathcal{A}$) $\leq_T^p$ Redundant($\mathcal{A}$), and (3) Redundant($\mathcal{A}$) $\leq_T^p$ Equiv($\mathcal{A}$).

(1). Let $((V, C_1), (V, C_2))$ be an instance of Equiv($\mathcal{A}$). We need to check whether $\text{Sol}((V, C_1)) = \text{Sol}((V, C_2))$. This is true if and only if the two Impl instances $((V, C_1), (V, C_2))$ and $((V, C_2), (V, C_1))$ are yes-instances.

(2). Let $((V, C_1), (V, C_2))$ be an instance of Impl($\mathcal{A}$). For each constraint $c \in C_2$, first check whether C_1 implies $\{c\}$ by (a) checking if $c \in C_1$, in which case C_1 trivially implies $\{c\}$, (b) if not, then check whether c is redundant in $C_1 \cup \{c\}$, in which case we answer yes, and otherwise no. If C_1 implies $\{c\}$ for every $c \in C_2$ then C_1 implies C_2 and we answer yes, and otherwise no.

(3). Let $I = ((V, C), c)$ be an instance of REDUNDANT($\mathcal{A}$). It is obvious that I is a yes-instance if and only if the instance $((V, C), (V, C \setminus \{c\}))$ is a yes-instance of EQUIV($\mathcal{A}$). ◀

Next, we show how the complexity of REDUNDANT($\mathcal{A}$) can be analysed by exploiting CSPs with alien constraints. If R is a k-ary relation over domain D, then we let $\bar{R}$ denote its *complement*, i.e. $\bar{R} = D^k \setminus R$.

▶ **Theorem 4.** (⋆) *Let $\mathcal{A}$ be a structure with domain A. If CSP($\mathcal{A}$) is not in* P*, then* REDUNDANT($\mathcal{A}$) *is not in* P*. In particular,* REDUNDANT($\mathcal{A}$) *is* NP*-hard (under polynomial-time Turing reductions) whenever CSP($\mathcal{A}$) is* NP*-hard. Otherwise,* REDUNDANT($\mathcal{A}$) *is in* P *if and only if for every relation $R \in \mathcal{A}$, CSP$_{\leq 1}(\mathcal{A} \cup \{\bar{R}\})$ is in* P*.*

Combining Theorem 4 with the forthcoming complexity classification of Boolean CSPs with alien constraints (Theorem 14) shows that Boolean REDUNDANT($\mathcal{A}$) is in P if and only if $\mathcal{A}$ is Schaefer. We have not found this result in the literature but we view it as folklore since it follows from other classification results (start from [12] or [25] and transfer the results to REDUNDANT($\mathcal{A}$) with the aid of Lemma 3). However, we claim that our proof is very different when compared to the proofs in [12] and [25]): Böhler et al. use a lengthy case analysis while Schnoor & Schnoor in addition uses the so-called weak base method, which scales poorly since not much is known about this construction for non-Boolean domains. We do not claim that our proof is superior, but we do not see how to generalize the classifications by Böhler et al. and Schnoor & Schnoor to larger (in particular infinite) domains since they are fundamentally based on Post's classification of Boolean clones. Such a generalization, on the other hand, is indeed possible with our approach. We demonstrate in Section 6.2 that we can obtain a full understanding of the complexity of CSPs with alien constraints for equality languages. This result carries over to REDUNDANT($\cdot$) via Theorem 4, implying that we have a full complexity classification of REDUNDANT($\cdot$) for equality languages. This result can immediately be transferred to IMPL($\cdot$) and EQUIV($\cdot$) by Lemma 3.

3.2 Model Checking

We follow [20] and view the *model checking* problem as follows: given a logic $\mathscr{L}$, a structure $\mathcal{A}$, and a sentence ϕ of $\mathscr{L}$, decide whether $\mathcal{A} \models \phi$. The main motivation for this problem is its connection to databases [26]. From the CSP perspective, we consider a slightly reformulated version: given an instance $I = (V, C)$ of CSP($\mathcal{A}$) and a formula ϕ with free variables in V, we ask if there is a tuple in Sol(I) that satisfies ϕ. If ϕ can be expressed as an instance I' of CSP($\mathcal{B}$) for some structure $\mathcal{B}$, then this is the same thing as if asking whether $I \cup I'$ has a solution or not. In the model-checking setting, we want to check whether ϕ is true in all solutions of I. If $\neg\phi$ can be expressed as an instance I' of CSP($\mathcal{B}$) for some structure $\mathcal{B}$, then we are done: every solution to I satisfies ϕ if and only if CSP($I \cup I'$) is not satisfiable, and this clarifies the connection with CSPs with alien constraints. For instance, one may view IMPL($\mathcal{A}$) (and consequently the underlying CSP$_{\leq 1}(\mathcal{A} \cup \bar{R})$ problems by Lemma 3 and Theorem 4) as the model checking problem restricted to queries that are $\mathcal{A}$-sentences constructed using the operators $\forall$ and $\vee$. Naturally, one wants the ability to use more complex queries such as (1) queries extended with other relations, i.e. queries constructed over an expanded structure, or (2) queries that are built using other logical connectives.

In both cases, it makes sense to study the fixed-parameter tractability of CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ with parameter #ac since the query is typically much smaller than the structure $\mathcal{A}$. The connection is quite obvious in the first case (one may view #ac as measuring how "complex" the given query is) while it is more hidden in the second case. Let us therefore consider the negation operator. From a logical perspective, one may view a constraint $\bar{R}(x_1, \ldots, x_k)$ as the formula $\neg R(x_1, \ldots, x_k)$. Needless to say, the relation $\bar{R}$ is often not pp-definable in a structure $\mathcal{A}$ containing R but it may be existential positive definable in $\mathcal{A}$. Assume that the preconditions of the example hold and that CSP($\mathcal{A}$) is in P. We know that $\bar{R}$ has an existential positive definition in $\mathcal{A}$ for every $R \in \mathcal{A}$. Let $\bar{\mathcal{A}} = \{\bar{R} \mid R \in \mathcal{A}\}$ and consider the problem $\text{CSP}_{\leq}(\mathcal{A} \cup \bar{\mathcal{A}})$. The forthcoming Theorem 15 is applicable so this problem is in FPT parameterized by #ac. Now, the corresponding model checking problem is to decide if $\mathcal{A} \models \phi$ where ϕ is an $\mathcal{A}$-sentence constructed using the operators $\forall$ and $\vee$ and where we are additionally allowed to use negated relations $\neg R(x_1, \ldots, x_m)$. It follows that this problem is in FPT parameterized by the number of negated relations.

4 General Tools for Alien Constraints

We analyze the complexity of $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$, starting in Section 4.1 by investigating which of the classic algebraic tools are applicable to the alien constraint setting, and continuing in Section 4.2 by presenting a general FPT result. We will use these observations for proving various results but also for obtaining a better understanding of alien constraints.

4.1 Alien Constraints and Algebra

First, we have a straightforward generalization of Theorem 1 in the alien constraint setting.

▶ **Theorem 5.** ($\star$) *Let $\mathcal{A}$ and $\mathcal{B}$ be two structures with disjoint signatures. There exists a polynomial time many-one reduction f from $\text{CSP}_{\leq}(\mathcal{A}^* \cup \mathcal{B}^*)$ to $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ for any finite $\mathcal{A}^* \subseteq \langle \mathcal{A} \rangle$ and $\mathcal{B}^* \subseteq \langle \mathcal{A} \cup \mathcal{B} \rangle$. If $I = (V, C, k)$ is an instance of $\text{CSP}_{\leq}(\mathcal{A}^* \cup \mathcal{B}^*)$ and $f(I) = (V', C', k')$, then k' only depends on k, $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{B}^*$, so f is an fpt-reduction.*

This claim is, naturally, in general not true for $\text{CSP}_{\leq k}(\mathcal{A}^* \cup \mathcal{B})$ for finite $\mathcal{A}^* \subseteq \langle \mathcal{A} \cup \mathcal{B} \rangle$. The idea underlying Theorem 5 can be used in many different ways and we give one example.

▶ **Proposition 6.** *If $\mathcal{A}, \mathcal{B}$ are structures and $R \in \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq 1}$, then $\text{CSP}_{\leq k}(\mathcal{A} \cup (\mathcal{B} \cup \{R\}))$ is polynomial-time reducible to $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$.*

We proceed by relating $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ to the important idea of reducing to a core (recall Section 2). Recall that $\mathcal{A}^c$ denotes the (unique up to isomorphism) core of a finite-domain structure $\mathcal{A}$. For two structures $\mathcal{A} \cup \mathcal{B}$ we similarly write $(\mathcal{A} \cup \mathcal{B})^c$ for the core. Specifically, if $e \in \text{End}(\mathcal{A} \cup \mathcal{B})$ has minimal range, then the core consists of $\{e(R) \mid R \in \mathcal{A}\} \cup \{e(R) \mid R \in \mathcal{B}\}$ of the same signature as $\mathcal{A}$ and $\mathcal{B}$, and the problem $\text{CSP}_{\leq}((\mathcal{A} \cup \mathcal{B})^c)$ is thus well-defined.

▶ **Theorem 7.** ($\star$) *Let $\mathcal{A}$ and $\mathcal{B}$ be two structures over a finite universe A. Then $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ and $\text{CSP}_{\leq}((\mathcal{A} \cup \mathcal{B})^c)$ are interreducible under both polynomial-time and fpt reductions.*

In general, it is *not* possible to reduce from $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ to $\text{CSP}_{\leq k}(\mathcal{A}^c \cup \mathcal{B})$ or from $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ to $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B}^c)$. This can be seen as follows. Consider the Boolean relation $R(x_1, x_2, x_3) \equiv x_1 = x_2 \vee x_2 = x_3$, and let $\mathcal{A} = \{R\}$, $\mathcal{B} = \{\neq\}$. Then, $\text{CSP}_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is NP-hard (see e.g. Exercise 3.24 in [14]) so $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is pNP-hard. However, $\mathcal{A}$ is 0-valid, so $\mathcal{A}^c = \{\{(0, 0, 0)\}\}$, implying that $\text{CSP}_{\leq}(\mathcal{A}^c \cup \mathcal{B})$ is in P.

4.2 Fixed-Parameter Tractability

We present an algorithm in this section that underlies many of our fixed-parameter tractability results and it is based on a particular notion of definability. The *existential fragment* of first-order logic consists of formulas that only use the operations negation, conjunction, disjunction, and existential quantification, while the *existential positive* fragment additionally disallows negation. We emphasize that it is required that the equality relation is allowed in existential (positive) definitions. We can view existential positive in a different way that is easier to use in our algorithm. Let $\mathcal{A}$ be a structure with domain A and assume that $R \subseteq A^m$ is defined via a existential positive definition over $\mathcal{A}$, i.e., $R(x_1, \ldots, x_m) \equiv \exists y_1, \ldots, y_n \colon \ \phi(x_1, \ldots, x_m, y_1, \ldots, y_n)$ where ϕ is a quantifier-free existential positive $\mathcal{A}$-formula. Since ϕ can be written in disjunctive normal form without introducing negation or quantifiers, it follows that R is a finite union of relations in $\langle \mathcal{A} \rangle$.

▶ **Theorem 8.** *Assume the following.*

1. *$\mathcal{A}, \mathcal{B}$ are structures with the same domain A,*
2. *every relation in $\mathcal{B}$ is existential positive definable in $\mathcal{A}$, and*
3. *CSP($\mathcal{A}$) is in P.*

Then CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT *parameterized by* #ac.

Proof. Assume $\mathcal{B} = \{A; B_1, \ldots, B_m\}$. Condition 2. implies that B_i, $i \in [m]$, is a finite union of relations $B_i = R_i^1 \cup \cdots \cup R_i^{c_i}$ where $R_i^1, \ldots, R_i^{c_i}$ are in $\langle \mathcal{A} \rangle$. Let the structure $\mathcal{A}^*$ contain the relations in $\mathcal{A} \cup \{R_i^j \mid i \in [m] \text{ and } j \in [c_i]\}$. Clearly, $\mathcal{A}^*$ has a finite signature and the problem CSP($\mathcal{A}^*$) is in P by Theorem 1 since every relation in $\mathcal{A}^*$ is a member of $\langle \mathcal{A} \rangle$. Let $b = \max\{c_i \mid i \in [m]\}$.

Let $((V, C), k)$ denote an arbitrary instance of CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$). The satisfiability of (V, C) can be checked via the following procedure. If C contains no $\mathcal{B}$-constraint, then check the satisfiability of (V, C) with the polynomial-time algorithm for CSP($\mathcal{A}$). Otherwise, pick one constraint $c = B_i(x_1, \ldots, x_q)$ with $B_i \in \mathcal{B}$ and check recursively the satisfiability of the following instances:

$$(V, (C \setminus \{c\}) \cup \{R_i^1(x_1, \ldots, x_q)\}), \ldots, (V, (C \setminus \{c\}) \cup \{R_i^{c_i}(x_1, \ldots, x_q)\}).$$

If at least one of the instances is satisfiable, then answer "yes" and otherwise "no". This is clearly a correct algorithm for CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$.

We continue with the complexity analysis. Note that the leaves in the computation tree produced by the algorithm are CSP($\mathcal{A}^*$) instances and they are consequently solvable in polynomial time. The depth of the computation tree is at most k (since (V, C) contains at most k $\mathcal{B}$-constraints) and each node has at most b children. Thus, the problem can be solved in $b^k \cdot \text{poly}(|I|)$ time. We conclude that CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT parameterized by #ac since b is a fixed constant that only depends on the structures $\mathcal{A}$ and $\mathcal{B}$. ◀

5 Finite-Domain Languages

This section is devoted to CSPs over finite domains. We begin in Section 5.1 by studying how the definability of constants affect the complexity of finite-domain CSPs with alien constraints, and we use this as a cornerstone for a parameterized FPT versus pNP dichotomy result for of CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$. We show a sharper result for Boolean structures in Section 5.2.

5.1 Parameterized Dichotomy

We begin with a simplifying result. For a finite set A, let $\mathcal{C}_A$ be the structure whose relations are the constants over A.

▶ **Lemma 9.** (⋆) *Let $\mathcal{A}$ be a structure over a domain A. For every $\mathcal{C} \subseteq \mathcal{C}_A$, CSP($\mathcal{A} \cup \mathcal{C}$) is polynomial-time reducible to $CSP_{\leq|\mathcal{C}|}(\mathcal{A} \cup \mathcal{C})$.*

Lemma 9 together with the basic algebraic results from Section 4.1 allows us to prove the following result that combines a more easily formulated fixed-parameter result (compared to Theorem 8) with a powerful hardness result.

▶ **Theorem 10.** (⋆) *Let $\mathcal{A}, \mathcal{B}$ be structures with finite domain D. Assume that $\mathcal{A} \cup \mathcal{B}$ is a core. If CSP($\mathcal{A} \cup \mathcal{C}_A$) is in* P*, then $CSP_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in* FPT *with parameter* #ac*. Otherwise, $CSP_{\leq p}(\mathcal{A} \cup \mathcal{B})$ is* NP*-hard for some p that only depends on the structures $\mathcal{A}$ and $\mathcal{B}$.*

Proof. We provide a short proof sketch, the full proof is in Appendix E. Using the dichotomy of finite domain CSPs [13, 27], we first assume CSP($\mathcal{A} \cup \mathcal{C}_D$) is in P. One can prove that every tuple over D is pp-definable over $\mathcal{A} \cup \mathcal{C}_D$ and then that each relation in $\mathcal{B}$ is existential positive definable over $\mathcal{A} \cup \mathcal{C}_D$. We can now apply Theorem 8, and $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT.

For the NP-hard case, we assume CSP($\mathcal{A} \cup \mathcal{C}_D$) is NP-hard and construct a polynomial-time reduction from CSP($\mathcal{A} \cup \mathcal{C}_D$) to $\text{CSP}_{\leq p}(\mathcal{A} \cup \mathcal{B})$. We use the endomorphisms of $\mathcal{A} \cup \mathcal{B}$ to construct a pp-definable relation E which allow us to simulate the constant relations, and a reduction to $\text{CSP}_{\leq 1}(\mathcal{A} \cup \{E\})$ to establish the claim via Lemma 9 and Theorem 5. ◀

Theorem 10 has broad applicability. Let us, for instance, consider a structure $\mathcal{A}$ with finite domain A and containing a finite number of relations from Inv(f) where $f \colon A^m \to A$ is idempotent ($f : A^m \to D$ is *idempotent* if $f(a, \ldots, a) = a$ for all $a \in A$.) If CSP($\mathcal{A}$) is in P, then CSP($\mathcal{A} \cup \mathcal{C}_A$) is in P since constant relations are invariant under f. Hence, $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT parameterized by #ac for *every* finite structure $\mathcal{B}$ with domain A by Theorem 10. Idempotent functions that give rise to polynomial-time solvable CSPs are fundamental and well-studied in the literature; see e.g. the survey by Barto et al. [1].

Via Theorem 7 we obtain the following parameterized complexity dichotomy separating problems in FPT from pNP-hard problems.

▶ **Corollary 11.** *Let $\mathcal{A}, \mathcal{B}$ be structures over the finite domain A. Then, $CSP_{\leq}(\mathcal{A} \cup \mathcal{B})$ is either in* FPT *or* pNP*-hard (in parameter* #ac*).*

Proof. Let $e \in \text{End}(\mathcal{A} \cup \mathcal{B})$ have minimal range and let $\mathcal{A}' = \{e(R) \mid R \in \mathcal{A}\}$ and $\mathcal{B}' = \{R \mid R \in \mathcal{B}\}$ be the two components of the core $(\mathcal{A} \cup \mathcal{B})^c$, and let $A' = \{e(a) \mid a \in A\}$ be the resulting domain. The problems $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ and $\text{CSP}_{\leq}(\mathcal{A}' \cup \mathcal{B}')$ are fpt-interreducible by Theorem 7. The problem CSP($\mathcal{A}' \cup \mathcal{C}_{A'}$) is either in P or is NP-hard by the CSP dichotomy theorem [13, 27]. In the first case, $\text{CSP}_{\leq}(\mathcal{A}' \cup \mathcal{B}')$ (and thus $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$) is in FPT with parameter #ac. Otherwise, $\text{CSP}_{\leq}(\mathcal{A}' \cup \mathcal{B}')$ is pNP-hard, and the fpt-reduction from $\text{CSP}_{\leq}(\mathcal{A}' \cup \mathcal{B}')$ to $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ establishes pNP-hardness for the latter. ◀

Corollary 11 must be used with caution: it does not imply that $\text{CSP}_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is NP-hard and results such as Theorem 4 may not be applicable. This encourages the refinement of coarse complexity results based on Theorem 10. We use Boolean relations as an example of this in the next section.

5.2 Classification of Boolean Languages

We present a complexity classification of $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ when $\mathcal{A}$ and $\mathcal{B}$ are Boolean structures (Theorem 14). We begin with two auxiliary results and we define relations $c_0 = \{(0)\}$ and $c_1 = \{(1)\}$.

▶ **Lemma 12.** ($\star$) *Let $\mathcal{A}$ be a Boolean structure where $c_0 \in \langle \mathcal{A} \rangle$. If an n-ary Boolean $R \neq \emptyset$ is not 0-valid then $c_1 \in \langle \mathcal{A} \cup \{R\} \rangle_{\leq 1}$.*

We say that a Boolean relation R is *invariant under complement* if it is invariant under the operation $\{0 \mapsto 1, 1 \mapsto 0\}$. This is equivalent to $(t_1, \ldots, t_k) \in R$ if and only if $(1 - t_1, \ldots, 1 - t_k) \in R$.

▶ **Lemma 13.** ($\star$) *Let $\mathcal{A}$ be a Boolean structure with finite signature. If $\mathcal{A}$ is invariant under complement, then CSP$(\mathcal{A} \cup \{c_0, c_1\})$ is polynomial-time reducible to CSP$_{\leq 1}(\mathcal{A} \cup \{\neq\})$.*

We are now ready for analysing the complexity of CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ when $\mathcal{A}$ and $\mathcal{B}$ are Boolean structures. We use a simplifying concept: a *0/1*-pair (R_0, R_1) contains two Boolean relations where R_0 is 0-valid but not 1-valid and R_1 is 1-valid but not 0-valid.

▶ **Theorem 14.** *Let $\mathcal{A}$ and $\mathcal{B}$ be Boolean structures such that CSP$(\mathcal{A})$ is in* P *and CSP$(\mathcal{A} \cup \mathcal{B})$ is* NP*-hard. Then the following holds.*

1. *If $\mathcal{A}$ is Schaefer, then CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in* FPT *with parameter* #ac.
2. *If (i) $\mathcal{A}$ is not Schaefer, (ii) $\mathcal{A}$ is both 0- and 1-valid, (iii) $\mathcal{B}$ contains a 0/1-pair, and (iv) $\mathcal{B}$ is 0- or 1-valid, then CSP$_{\leq 2}(\mathcal{A} \cup \mathcal{B})$ is* NP*-hard and CSP$_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is in* P.
3. *Otherwise, CSP$_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is* NP*-hard.*

Proof. Assume $\mathcal{A}$ is Schaefer and let $\mathcal{A}^+ = \mathcal{A} \cup \{c_0, c_1\}$. The structure $\mathcal{A}^+$ is clearly a core and $\mathcal{A}^+ \cup \mathcal{B}$ is a core, too. The problem CSP$(\mathcal{A}^+)$ is in P by Theorem 2 so Theorem 10 implies that CSP$_{\leq}(\mathcal{A}^+ \cup \mathcal{B})$ (and naturally CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$) is in FPT parameterized by #ac. Since CSP$(\mathcal{A})$ is in P, we know from Theorem 2 that $\mathcal{A}$ is 0-valid, 1-valid or Schaefer. We assume henceforth that $\mathcal{A}$ is 0-valid and not Schaefer; the other case is analogous. If $\mathcal{B}$ is 0-valid, then CSP$(\mathcal{A} \cup \mathcal{B})$ is trivially in P and this is ruled out by our initial assumptions. We assume henceforth that $\mathcal{B}$ is not 0-valid and consider two cases depending on whether c_0 is pp-definable in $\mathcal{A}$ or not.

Case 1. c_0 is pp-definable in $\mathcal{A}$. We know that CSP$(\mathcal{A} \cup \{c_0, c_1\})$ is NP-hard by Theorem 2 since $\mathcal{A}$ is not Schaefer. We can thus assume that CSP$(\mathcal{A} \cup \{c_1\})$ is NP-hard. Lemma 9 implies that CSP$_{\leq 1}(\mathcal{A} \cup \{c_1\})$ is NP-hard. The relation c_1 is in $\langle \mathcal{A} \cup \mathcal{B} \rangle_{\mathcal{B} \leq 1}$ by Lemma 12 so we conclude that CSP$_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is NP-hard.

Case 2. c_0 is not pp-definable in $\mathcal{A}$. This implies that every relation in $\mathcal{A}$ is simultaneously 0- and 1-valid. To see this, assume to the contrary that $\mathcal{A}$ contains a relation that is not 1-valid. Then, $x = 0 \Leftrightarrow R(x, \ldots, x)$ and c_0 is pp-definable in $\mathcal{A}$. This implies that $\mathcal{B}$ contains (a) a relation that is not invariant under any constant operation or (b) every relation is closed under a constant operation and $\mathcal{B}$ contains a 0/1-pair. Note that if (a) and (b) does not hold, then $\mathcal{B}$ is invariant under a constant operation and CSP$(\mathcal{A} \cup \mathcal{B})$ is trivially in P.

Case 2(a). There is a a relation R in $\mathcal{B}$ that is not invariant under any constant operation, i.e. $(0, \ldots, 0) \notin R$ and $(1, \ldots, 1) \notin R$. The relation R has arity $a \geq 2$. Let t be the tuple in R that contains the maximal number b of 0:s. Clearly, $b < a$. We assume that the arguments are permuted so that t begins with b 0:s and continues with $a - b$ 1:s. Consider the pp-defintion

$$S(x, y) \equiv R(\underbrace{x, \ldots, x}_{b \text{ occ.}}, \underbrace{y, \ldots, y}_{a-b \text{ occ.}}).$$

There are two possibilities: either $S(x, y) \Leftrightarrow x = 0 \wedge y = 1$ or $S(x, y) \Leftrightarrow x \neq y$. In the first case we are done since CSP$(\mathcal{A} \cup \{c_0, c_1\})$ is NP-hard (recall that $\mathcal{A}$ is not Schaefer) and CSP$_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is easily seen to be NP-hard by Lemma 9. Let us consider the second case.

If $\mathcal{A}$ is invariant under complement, then $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is NP-hard by Lemma 13. If $\mathcal{A}$ is not invariant under complement, then we claim that c_0 and c_1 can be pp-defined with the aid of $\neq$. Arbitrarily choose a relation T in $\mathcal{A}$ that contains a tuple $t = (t_1, \ldots, t_a)$ such that $(1 - t_1, \ldots, 1 - t_a) \notin T$—note that t cannot be a constant tuple since both $(0, \ldots, 0)$ and $(1, \ldots, 1)$ are in T. Assume that t contains b 0:s and that the arguments are permuted so that t begins with b 0:s followed by $a - b$ 1:s. Consider the pp-definition

$$U(x, y) \equiv x \neq y \wedge T(\underbrace{x, \ldots, x}_{b \text{ occ.}}, \underbrace{y, \ldots, y}_{a-b \text{ occ.}}).$$

The relation U contains the single tuple $(0, 1)$. We know that $\mathrm{CSP}(\mathcal{A} \cup \{c_0, c_1\})$ is NP-hard (recall that $\mathcal{A}$ is not Schaefer) and Lemma 9 implies that $\mathrm{CSP}_{\leq 2}(\mathcal{A} \cup \{c_0, c_1\})$ is NP-hard, too. It is now easy to see that $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is NP-hard via the definition of U.

Case 2(b). Every relation in $\mathcal{B}$ is closed under at least one constant operation and $\mathcal{B}$ contains a 0/1-pair (R_0, R_1). Since $\mathcal{A}$ is both 0- and 1-valid, it follows that $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ is in P. The constant relations c_0 and c_1 are pp-definable in $\{R_0, R_1\}$ since $x = 0 \Leftrightarrow R_0(x, \ldots, x)$ and $x = 1 \Leftrightarrow R_1(x, \ldots, x)$. This implies with the aid of Lemma 9 that $\mathrm{CSP}_{\leq 2}(\mathcal{A} \cup \mathcal{B})$ is NP-hard since $\mathcal{A}$ is not Schaefer. ◀

Theorem 14 carries over to Boolean REDUNDANT(·), EQUIV(·) and IMPL(·) by Lemma 3 combined with Theorem 4, so these problems are in P if and only if $\mathcal{A}$ is Schaefer (case 2. in Theorem 14 is not applicable when analysing these problems since it requires $|\mathcal{B}| \geq 2$). Otherwise, they are NP-complete under polynomial-time Turing reductions. The *meta-problem* for Boolean CSPs with alien constraints is decidable, i.e., there is an algorithm that decides for Boolean structures $\mathcal{A}, \mathcal{B}$ whether $\mathrm{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in case 1., 2., or 3. of Theorem 14. This is obvious since we have polymorphism descriptions of the Schaefer languages.

6 Infinite-Domain Languages

We focus on infinite-domain CSPs in this section. We begin Section 6.1 by discussing certain problems when CSPs with alien constraints are generalized to infinite domains. Our conclusion is that restricting ourselves to ω-categorical structures is a viable first step: ω-categorical structures constitute a rich class of CSPs and we can generalize at least some of the machinery from Section 5 to this setting. We demonstrate this in Section 6.2 where we obtain a complete complexity classification for equality languages.

6.1 Orbits and Infinite-Domain CSPs

It is not straightforward to tranfer the results in Section 5 to the infinite-domain regime. First, let us consider Theorem 8. In contrast to finite domains, relations in $\mathcal{B}$ may not be finite unions of relations in $\langle \mathcal{A} \rangle$ or, equivalently, not being definable with an existential positive formula. Second, let us consider Theorem 10: the proof is based on structures expanded with symbols for each domain value and this leads to problematic structures with infinite signatures. The proof is also based on the assumption that CSPs are either polynomial-time solvable or NP-complete, and this is no longer true [5]. It is thus necessary to restrict our attention to some class of structures with sufficiently pleasant properties. A natural choice is ω-categorical structures that allows us to reformulate Theorem 8 as follows.

▶ **Theorem 15.** ($\star$) *Assume the following.*

1. *$\mathcal{A}, \mathcal{B}$ are structures with the same countable (not necessarily infinite) domain A,*
2. *$\mathcal{A}$ and $\mathcal{B}$ are ω-categorical,*
3. *every relation in* $\mathrm{Orb}(\mathcal{B})$ *is existential primitive definable in* $\langle \mathcal{A} \rangle$*, and*
4. *CSP($\mathcal{A}$) is in* P

Then CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT *parameterized by* #ac.

▶ **Example 16.** Results related to Theorem 15 have been presented in the literature. Recall that RCC5 and RCC8 are spatial formalism with binary relations that are disjunctions of certain basic relations [23]. Li et al. [19] prove that if $\mathcal{A}$ is a polynomial-time solvable RCC5 or RCC8 constraint language containing all basic relations, then Redundant($\mathcal{A}$) is in P. This immediately follows from combining Theorem 4 and Theorem 15 since RCC5 and RCC8 can be represented by ω-categorical constraint languages [3, 11] and every RCC5/RCC8 relation is existential primitive definable in the structure of basic relations by definition. This result can be generalized to a much larger class of relations in the case of RCC5 since the orbits of k-tuples are pp-definable in the structure of basic relations [6, Proposition 35].

A general hardness result based on the principles behind Theorem 10 does not seem possible in the infinite-domain setting, even for ω-categorical structures. The hardness proof in Theorem 10 utilizes variables given fixed values and a direct generalization would lead to groups of variables that together form an orbit of an n-tuple. Such gadgets behave very differently from variables given fixed values: in particular, they do not admit a result similar to Lemma 9. Thus, hardness results needs to be constructed in other ways.

We know from Section 4.1 that CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ and CSP$_{\leq}((\mathcal{A} \cup \mathcal{B})^c)$ are the same when $\mathcal{A}$ and $\mathcal{B}$ has the same finite domain. We now consider a generalisation of cores to infinite domains from Bodirsky [2]: an ω-categorical structure $\mathcal{A}$ with countable domain is a *model-complete core* if every relation in $\mathrm{Orb}(\mathcal{A})$ is pp-definable in $\mathcal{A}$. There is an obvious infinite-domain analogue of Theorem 7: if $\mathcal{A}' \cup \mathcal{B}'$ is the model-complete core of $\mathcal{A} \cup \mathcal{B}$ (where $\mathcal{A}, \mathcal{B}$ are ω-categorical structures over a countable domain A), then CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ polynomial-time reduces to CSP$_{\leq}(\mathcal{A}' \cup \mathcal{B}')$. Model-complete cores share many other properties with cores, too. With this said, it is interesting to understand model-complete cores in the context of CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$, simply because they are so well-studied and exhibit useful properties. We merely touch upon this subject by making an observation that we use in Section 6.2.

▶ **Lemma 17.** ($\star$) *Let $\mathcal{A}$ and $\mathcal{B}$ denote ω-categorical structures with a countable domain A. Assume that $\mathcal{A}$ is a model-complete core and CSP($\mathcal{A}$) is in* P*. Then, CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in* FPT *parameterized by* #ac *for every structure $\mathcal{B}$ such that* $\mathrm{Orb}(\mathcal{B}) \subseteq \mathrm{Orb}(\mathcal{A})$.

6.2 Classification of Equality Languages

We present a complexity classification of CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ for equality languages $\mathcal{A}$, $\mathcal{B}$. Essentially, there are two interesting cases: when $\mathcal{A}$ is Horn, and when $\mathcal{A}$ is 0-valid and not Horn. In the former case, CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT parameterized by #ac, while in the second case it is pNP-hard. It turns out that the ability to pp-define the arity-c disequality relation, where c depends only on $\mathcal{A}$, using at most k alien constraints, determines the complexity. A dichotomy for Redundant($\cdot$), Impl($\cdot$), and Equiv($\cdot$) follows: these problems are either in P or NP-hard under polynomial-time Turing reductions.

Recall that CSP($\mathcal{A}$) for a finite equality constraint language $\mathcal{A}$ is in P if $\mathcal{A}$ is 0-valid or preserved by a binary injective operation, and NP-hard otherwise, and that the automorphism group for equality languages is the symmetric group Σ on $\mathbb{N}$, i.e. the set of permutations on $\mathbb{N}$. It is easy to see that an orbit of a k-tuple $(a_1, \ldots, a_k)$ is pp-definable in $\{=, \neq\}$. For instance, the orbit of $(0,0,1,2)$ is defined by $O(x_1,x_2,x_3,x_4) \equiv x_1 = x_2 \wedge x_2 \neq x_3 \wedge x_2 \neq x_4 \wedge x_3 \neq x_4$. Observe that $\neq$ is invariant under every binary injective operation, so if $\mathcal{A}$ is Horn, then $\neq \in \langle \mathcal{A} \rangle$ and every orbit of n-tuples under Σ is pp-definable in $\mathcal{A}$. Thus, $\mathcal{A}$ is a model-complete core as pointed out in Section 6.1. Lemma 17 now implies the following.

▶ **Corollary 18.** *Let $\mathcal{A}$ and $\mathcal{B}$ be equality languages. If $\mathcal{A}$ is Horn, then CSP$_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in* FPT *parameterized by* #ac.

Thus, we need to classify the complexity of $\mathrm{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ for every k, where $\mathcal{A}$ is 0-valid and not Horn, and $\mathcal{B}$ is not 0-valid. We will rely on results about the complexity of singleton expansions of equality constraint languages. Let $\mathcal{A}$ be a constraint language over the domain $\mathbb{N}$. By $\mathcal{A}_c^+$ we denote the expansion of $\mathcal{A}$ with c singleton relations, i.e. $\mathcal{A}_c^+ = \mathcal{A} \cup \{\{1\}, \dots, \{c\}\}$. The complexity of $\mathrm{CSP}(\mathcal{A}_c^+)$ for equality constraint languages $\mathcal{A}$ and all constants c was classified by Osipov & Wahlström [21, Section 7], building on the detailed study of polymorphisms of equality constraint languages by Bodirsky et al. [4].

The connection between $\mathrm{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ and $\mathrm{CSP}(\mathcal{A}_c^+)$ is the following. In one direction, we can augment every instance of $\mathrm{CSP}(\mathcal{A})$ with c fresh variables $z_1, \dots, z_c$ and, assuming k is large enough and $\mathcal{B}$ is not 0-valid, use $\mathcal{B}$-constraints to ensure that $z_1, \dots, z_c$ attain distinct values in every satisfying assignment. Given that $\mathcal{A}$ is invariant under every permutation of $\mathbb{N}$, we can now treat $z_1, \dots, z_c$ as constants, e.g. as $1, \dots, c$, and transfer hardness results from the singleton expansion to our problem. In the other direction, if the relation $\mathrm{NEQ}_{c+1} \notin \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq k}$, then every satisfiable instance of $\mathrm{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ has a solution with range $[c]$, and $\mathcal{A}_c^+$ is tractable: indeed, a satisfiable instance without such a solution would be a pp-definition of $\mathrm{NEQ}_{c'}$ for some $c' > c$. These connections are formalized in Lemmas 23 and 24. We will leverage the following hardness result.

▶ **Lemma 19** (Follows from Theorem 54 in [21]). *Let $\mathcal{A}$ be a finite equality language. If $\mathcal{A}$ is not Horn, then CSP$(\mathcal{A}_c^+)$ is* NP*-hard for some $c = c(\mathcal{A})$.*

Our main tool for studying singleton expansions are *retractions*.

▶ **Definition 20.** *Let $\mathcal{A}$ be an equality language. An operation $f \colon \mathbb{N} \to [c]$ is a* retraction of $\mathcal{A}$ to $[c]$ *if f is an endomorphism of $\mathcal{A}$ where $f(i) = i$ for all $i \in [c]$. If $\mathcal{A}$ admits a retraction f to $[c]$, then we say that $\mathcal{A}$* retracts to $[c]$*, and $\mathcal{A}_f$ is* a retract (of $\mathcal{A}$ to $[c]$).

We obtain a useful characterization of retracts.

▶ **Lemma 21.** *Let $\mathcal{A}$ be an equality language and f be a retraction from $\mathcal{A}$ to $[c]$. Then $f(R) = R \cap [c]^{\mathrm{ar}(R)}$ for all $R \in \mathcal{A}$.*

Proof. First, observe that $f(R) \subseteq R \cap [c]^{\mathrm{ar}(R)}$: indeed, f is an endomorphism, so $f(R) \subseteq R$, and $f(R) \subseteq [c]^{\mathrm{ar}(R)}$ because the range of f is $[c]$. Moreover, we have $R \cap [c]^{\mathrm{ar}(R)} \subseteq f(R)$ because f is constant on $[c]$, so it preserves every tuple in $[c]^{\mathrm{ar}(R)}$. ◀

The finite-domain language $\{R \cap [c]^{\mathrm{ar}(R)} : R \in \mathcal{A}\}$ is called a *c-slice of $\mathcal{A}$* in [21, Section 7]. Lemma 21 shows that a c-slice of $\mathcal{A}$ is the retract $\mathcal{A}_f$ under any retraction f from $\mathcal{A}$ to $[c]$. Note that the definition of the c-slice does not depend on f, so we can talk about *the retract of $\mathcal{A}$ to* $[c]$. We will use this fact implicitly when transferring results from Theorem 57 in [21].

▶ **Lemma 22** (Follows from Theorem 57 in [21]). *Let $\mathcal{A}$ be an equality language that is 0-valid and not Horn, and let c be a positive integer. Then exactly one of the following holds:*
- *$\mathcal{A}$ does not retract to $[c]$, and CSP$(\mathcal{A}_c^+)$ is* NP*-hard.*
- *$\mathcal{A}$ retracts to $[c]$, and CSP$(\mathcal{A}_c^+)$ is* NP*-hard for all $c \geq 2$.*

- *$\mathcal{A}$ retracts to $[c]$, and both CSP(Δ_c^+) for the retract Δ and CSP($\mathcal{A}_c^+$) are in* P.

Let $\text{NEQ}_r = \{(t_1, \ldots, t_r) \in \mathbb{N}^r : |\{t_1, \ldots, t_r\}| = r\}$, i.e. the relation that contains every tuple of arity r with all entries distinct.

▶ **Lemma 23.** ($\star$) *Let $\mathcal{A}$ and $\mathcal{B}$ be equality languages and $c \in \mathbb{Z}_+$. If $\text{NEQ}_{c+1} \notin \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq k}$, then every satisfiable instance of $CSP_{\leq k}(\mathcal{A} \cup \mathcal{B})$ has a solution whose range is in $[c]$.*

▶ **Lemma 24.** ($\star$) *Let $\mathcal{A}$, $\mathcal{B}$ be two equality constraint languages, and let $c \in \mathbb{Z}_+$ be an integer. CSP($\mathcal{A}_c^+$) is polynomial-time reducible to $CSP_{\leq k}(\mathcal{A} \cup \mathcal{B})$ whenever $\text{NEQ}_c \in \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq k}$.*

We are ready to present the classification.

▶ **Theorem 25.** *Let $\mathcal{A}$ and $\mathcal{B}$ be equality languages such that CSP($\mathcal{A}$) is in* P *and CSP($\mathcal{A} \cup \mathcal{B}$) is* NP*-hard.*

1. *If $\mathcal{A}$ is Horn, $CSP_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in* FPT *parameterized by* #ac.
2. *If $\mathcal{A}$ is not Horn, $CSP_{\leq}(\mathcal{A} \cup \mathcal{B})$ is* pNP*-hard parameterized by* #ac. *Moreover, there exists an integer $c = c(\mathcal{A})$ such that $CSP_{\leq k}(\mathcal{A} \cup \mathcal{B})$ is in* P *whenever $\text{NEQ}_c \notin \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq k}$, and is* NP*-hard otherwise.*

Proof. CSP($\mathcal{A}$) is in P so $\mathcal{A}$ is Horn or 0-valid. If $\mathcal{A}$ is Horn, then Corollary 18 applies, proving part 1 of the theorem. Suppose $\mathcal{A}$ is 0-valid and not Horn. By applying Lemma 19 to $\mathcal{A}$, we infer that there is a minimum positive integer c such that CSP($\mathcal{A}_c^+$) is NP-hard. Since $\mathcal{A}$ is 0-valid, we have $c \geq 2$. Using Lemma 24, we can reduce CSP($\mathcal{A}_c^+$) to $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ in polynomial time whenever $\text{NEQ}_c \in \langle \mathcal{A} \cup \mathcal{B} \rangle_{\mathcal{B} \leq k}$, proving that the latter problem is NP-hard. Observe that $\mathcal{B}$ is not 0-valid because CSP($\mathcal{A} \cup \mathcal{B}$) is NP-hard, so $\neq \in \langle \mathcal{B} \rangle$ and $\text{NEQ}_c \in \langle \mathcal{A} \cup \mathcal{B} \rangle_{\mathcal{B} \leq k}$ for some finite $k \leq \binom{c}{2}$. This show the pNP-hardness result in part 2.

To complete the proof of part 2, it suffices to show that we can solve $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ in polynomial time whenever $\text{NEQ}_c \notin \langle \mathcal{A} \cup \mathcal{B} \rangle_{\mathcal{B} \leq k}$. To this end, observe that, by the choice of c, if $c' < c$, then CSP($\mathcal{A}_{c'}^+$) is in P. Then, by Lemma 22, $\mathcal{A}$ retracts to the finite domain $[c']$, and the retract Δ is such that CSP($\Delta_{c'}^+$) is in P. We will use the algorithm for CSP($\Delta_{c'}^+$) in our algorithm for $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ that works for all k such that $\text{NEQ}_c \notin \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq k}$

Let I be an instance of $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$. Since $\text{NEQ}_c \notin \langle \mathcal{A} \cup \mathcal{B} \rangle_{\mathcal{B} \leq k}$, Lemma 23 implies that I is satisfiable if and only if it admits a satisfying assignment with range $[c-1]$. Let X be the set of variables in I that occur in the scopes of the alien constraints. Note that $|X| \in O(k)$. Enumerate all assignments $\alpha : X \to [c-1]$, and check if it satisfies all $\mathcal{B}$-constraints in I. If not, reject it, otherwise remove the $\mathcal{B}$-constraints and add unary constraints $x = \alpha(x)$ for all $x \in X$ instead. This leads to an instance of CSP(Δ_{c-1}^+), which is solvable in polynomial time. If we obtain a satisfiable instance for some α, then accept I, and otherwise reject it. Correctness follows by Lemma 23 and the fact that the algorithm considers all assignments from X to $[c]$. We make $2^{O(k)}$ calls to the algorithm for CSP(Δ_{c-1}^+), where k is a fixed constant, and each call runs in polynomial time. This completes the proof. ◀

Theorem 14 implies that $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is pNP-hard if and only if $\text{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ is NP-hard for some k, and it is in FPT parameterized by #ac otherwise. Theorem 25 now implies a dichotomy for Redundant($\cdot$), Impl($\cdot$), and Equiv($\cdot$) over finite equality languages.

▶ **Theorem 26.** ($\star$) *Let $\mathcal{A}$ be a finite equality language. Then* Redundant($\mathcal{A}$), Impl($\mathcal{A}$), *and* Equiv($\mathcal{A}$) *are either in* P *or* NP*-hard (under polynomial-time Turing reductions).*

Algebraically characterizing the exact borderline between tractable and hard cases of the problem seems difficult. In particular, given a 0-valid non-Horn equality language $\mathcal{A}$, answering whether $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \bar{\mathcal{A}})$ is in P, i.e. whether $\mathrm{NEQ}_c \in \langle \mathcal{A} \cup \bar{R} \rangle_{\leq 1}$ for some $R \in \mathcal{A}$ and large enough c, requires a deeper understanding of such languages. However, one can show that the answer to this, and even a more general question is decidable.

▶ **Proposition 27.** *($\star$) There is an algorithm that takes two equality constraint languages $\mathcal{A}$ and $\mathcal{B}$ and outputs minimum $k \in \mathbb{N} \cup \{\infty\}$ such that $CSP_{\leq k}(\mathcal{A} \cup \mathcal{B})$ is* NP*-hard.*

7 Discussion

We have focused on structures with finite signatures in this paper. This is common in the CSP literature since relational structures with infinite signature cause vexatious representational issues. It may, though, be interesting to look at structures with infinite signatures, too. Zhuk [28] observes that the complexity of the following problem is open: given a system of linear equations mod 2 and a single linear equation mod 24, find a satisfying assignment over the domain $\{0, 1\}$. The equations have unbounded arity so this problem can be viewed as a $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \mathcal{B})$ problem where $\mathcal{A}, \mathcal{B}$ have infinite signatures. This question is thus not directly answered by Theorem 14. Second, let us also remark that when considering $\mathrm{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$, we have assumed that both $\mathcal{A}$ and $\mathcal{B}$ are taken from some nice "superstructure". For example, in the equality language case we assume that both structures are first-order reducts of $(\mathbb{N}; =)$. One could choose structures more freely and, for example, let $\mathcal{A}$ be an equality language and $\mathcal{B}$ a finite-domain language. This calls for modifications of the underlying theory since (for instance) the algorithm that Theorem 8 is based on breaks down.

For finite domains we obtained a *coarse* parameterized dichotomy for $\mathrm{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ separating FPT from pNP-hardness. Sharper results providing the exact borderline between P and NP-hardness for the pNP-hard cases are required for classifying implication, equivalence, and redundancy. Via Theorem 7 and Theorem 10 the interesting case is when $\mathrm{CSP}(\mathcal{A})$ is in P, $\mathcal{A} \cup \mathcal{B}$ is core but $\mathcal{A}$ is not core. This question may be of independent algebraic interest and could be useful for other problems where the core property is not as straightforward as in the CSP case. For example, in *surjective* CSP we require the solution to be surjective, and this problem is generally hardest to analyze when the template is not a core [8].

Any complexity classification of the first-order reducts of a structure includes by necessity a classification of equality CSPs. Thus, our equality language classification lay the foundation for studying first-order reducts of more expressive structures. A natural step is to study *temporal languages*, i.e. first-order reducts of $(\mathbb{Q}; <)$. Our classification of equality constraint languages relies on the work in [4] via [21], who studied the clones of polymorphisms of equality constraint languages in more detail. One important result, due to Haddad & Rosenberg [16], is that after excluding several easy cases, every equality constraint language we end up with is only closed under operations with range $[c]$ for some constant c. Then, pp-defining the relation NEQ_{c+1} brings us into pNP-hard territory. Similar characterizations of the polymorphisms for reducts of other infinite structures, e.g. $(\mathbb{Q}; <)$, would imply corresponding pNP-hardness results, and this appear to be a manageable way forward.

References

1 Libor Barto, Andrei A. Krokhin, and Ross Willard. Polymorphisms, and how to use them. In *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*, pages 1–44. 2017.

2 Manuel Bodirsky. *Complexity of Infinite-Domain Constraint Satisfaction*. Cambridge University Press, 2021.

3 Manuel Bodirsky and Hubie Chen. Qualitative temporal and spatial reasoning revisited. *Journal of Logic and Computation*, 19(6):1359–1383, 2009.

4 Manuel Bodirsky, Hubie Chen, and Michael Pinsker. The reducts of equality up to primitive positive interdefinability. *The Journal of Symbolic Logic*, 75(4):1249–1292, 2010. `doi:10.2178/jsl/1286198146`.

5 Manuel Bodirsky and Martin Grohe. Non-dichotomies in constraint satisfaction complexity. In *Proc. 35th International Colloquium on Automata, Languages and Programming (ICALP-2008)*, pages 184 –196, 2008.

6 Manuel Bodirsky and Peter Jonsson. A model-theoretic view on qualitative constraint reasoning. *Journal of Artificial Intelligence Research*, 58:339–385, 2017.

7 Manuel Bodirsky and Jan Kára. The complexity of equality constraint languages. *Theory of Computing Systems*, 43(2):136–158, 2008.

8 Manuel Bodirsky, Jan Kára, and Barnaby Martin. The complexity of surjective homomorphism problems - a survey. *Discrete Applied Mathematics*, 160(12):1680–1690, 2012.

9 Manuel Bodirsky and Jaroslav Nešetřil. Constraint satisfaction with countable homogeneous templates. In *Proc. 17th International Workshop on Computer Science Logic (CSL-2003)*, pages 44–57, 2003.

10 Manuel Bodirsky, Michael Pinsker, and Todor Tsankov. Decidability of definability. *The Journal of Symbolic Logic*, 78(4):1036–1054, 2013.

11 Manuel Bodirsky and Stefan Wölfl. RCC8 is polynomial on networks of bounded treewidth. In *Proc. 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, pages 756–761, 2011.

12 Elmar Böhler, Edith Hemaspaandra, Steffen Reith, and Heribert Vollmer. Equivalence and isomorphism for Boolean constraint satisfaction. In *Proc. 16th International Workshop on Computer Science Logic (CSL-2002)*, pages 412–426, 2002.

13 Andrei A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *Proc. 58th IEEE Annual Symposium on Foundations of Computer Science (FOCS-2017)*, pages 319–330, 2017.

14 Hubie Chen. A rendezvous of logic, complexity, and algebra. *ACM SIGACT News*, 37(4):85–114, 2006.

15 David A. Cohen, Peter Jeavons, Peter Jonsson, and Manolis Koubarakis. Building tractable disjunctive constraints. *Journal of the ACM*, 47(5):826–853, 2000.

16 Lucien Haddad and Ivo G. Rosenberg. Finite clones containing all permutations. *Canadian Journal of Mathematics*, 46(5):951–970, 1994.

17 Wilfrid Hodges. *A shorter model theory*. Cambridge University Press, Cambridge, 1997.

18 Peter G. Jeavons. On the algebraic structure of combinatorial problems. *Theoretical Computer Science*, 200:185–204, 1998.

19 Sanjiang Li, Zhiguo Long, Weiming Liu, Matt Duckham, and Alan Both. On redundant topological constraints. *Artificial Intelligence*, 225:51–76, 2015.

20 Florent R. Madelaine and Barnaby Martin. On the complexity of the model checking problem. *SIAM Journal on Computing*, 47(3):769–797, 2018.

21 George Osipov and Magnus Wahlström. Parameterized complexity of equality MinCSP. *arXiv preprint arXiv:2305.11131*, 2023. This is the report version of a paper that appears in Proc. 31st Annual European Symposium on Algorithms (ESA-2023), pp. 86:1-86:17.

22 George Osipov and Magnus Wahlström. Parameterized complexity of equality MinCSP. *arXiv preprint arXiv:2305.11131*, 2023. This is the report version of a paper that appeared at the 31st Annual European Symposium on Algorithms (ESA-2023).

23 David A. Randell, Zhan Cui, and Anthony G. Cohn. A spatial logic based on regions and connection. In *Proc. 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR-1992)*, pages 165–176, 1992.

24 Thomas J. Schaefer. The complexity of satisfiability problems. In *Proc. 10th Annual ACM Symposium on Theory of Computing (STOC-1978)*, pages 216–226, 1978.

25 Henning Schnoor and Ilka Schnoor. Partial polymorphisms and constraint satisfaction problems. In *Complexity of Constraints - An Overview of Current Research Themes [Result of a Dagstuhl Seminar]*, pages 229–254. Springer, 2008.

26 Moshe Y. Vardi. The complexity of relational query languages (extended abstract). In *Proc. 14th Annual ACM Symposium on Theory of Computing (STOC-1982)*, pages 137–146, 1982.

27 Dmitriy Zhuk. A proof of the CSP dichotomy conjecture. *Journal of the ACM*, 67(5):30:1–30:78, 2020.

28 Dmitriy Zhuk. Constraint satisfaction problem: what makes the problem easy. In *Proc. International Conference of Mathematicians 2022 (ICM-2022)*, pages 1530–1553, 2022.

APPENDIX

A Proof of Theorem 4

Proof. Let $I = (V, C)$ be an arbitrary instance of CSP($\mathcal{A}$) with domain A.

CSP($\mathcal{A}$) is not in P. We show that REDUNDANT($\mathcal{A}$) is not in P. Choose a relation $R \in \mathcal{A}$ of arity $p > 0$ that satisfies $\emptyset \subsetneq R \subsetneq A^p$. Note that $\mathcal{A}$ must contain at least one such relation R since otherwise we can trivially determine whether an instance is a yes-instance or not, and this contradicts that CSP($\mathcal{A}$) is not in P. Let $(t_1, \ldots, t_p)$ be an arbitrary tuple in R. We construct another instance $I' = (V', C')$ such that a certain constraint $c \in C'$ is redundant in I' if and only if I is not satisfiable.

1. Introduce p fresh variables $y_1, \ldots, y_p$ and define $V' = V \cup \{y_1, y_2, ..., y_p\}$.
2. Define the constraint $c = R(y_1, y_2, ..., y_p)$ and let $C' = C \cup \{c\}$.

These steps describe a polynomial time reduction from the CSP($\mathcal{A}$) instance I to the REDUNDANT($\mathcal{A}$) instance (I', c). We prove that I is a yes-instance if and only if (I', c) is a no-instance.

If I is satisfiable, then there exists a satisfying assignment $f : V \to A$ that satisfies all constraints in C. We show that I' is satisfiable by extending the assignment f to $f' : V' \to A$: let $f'(x) = f(x)$ when $x \in V$ and $f'(y_i) = t_i$, $i \in [p]$. Note that $\text{Sol}((V', C')) \neq \text{Sol}((V', C' \setminus \{c\}))$ since $R \subsetneq D^p$ so c is not a redundant constraint in I'.

If I is not satisfiable, then I' is not satisfiable since $C \subseteq C'$. Thus, $\text{Sol}((V', C')) = \text{Sol}((V', C' \setminus \{c\}))$ and (I', c) is a yes-instance of REDUNDANT($\mathcal{A}$).

We conclude that this is a polynomial-time Turing reduction and the lemma follows. Note that REDUNDANT(Γ) is NP-hard (under polynomial-time Turing reductions) whenever CSP(Γ) is NP-hard.

CSP($\mathcal{A}$) is in P. We show that REDUNDANT($\mathcal{A}$) is in P if and only if for every relation $R \in \mathcal{A}$, $\text{CSP}_{\leq 1}(\mathcal{A} \cup \{\bar{R}\})$ is in P.

Right-to-left direction. Assume $\text{CSP}_{\leq 1}(\mathcal{A} \cup \{\bar{R}\})$ is in P for every $R \in \mathcal{A}$. For an instance $I = ((V, C), c)$ of REDUNDANT($\mathcal{A}$), let $c = R(x_1, \ldots, x_k)$ and define $\bar{c} = \bar{R}(x_1, \ldots, x_k)$. Observe that $I' = (V, (C \setminus \{c\}) \cup \bar{c})$ an instance of $\text{CSP}_{\leq 1}(\mathcal{A} \cup \bar{R})$ and check whether it is satisfiable. We claim that I is a no-instance if and only I' is satisfiable. Indeed, I is a no-instance if and only if $\text{Sol}(V, C \setminus \{c\}) \neq \text{Sol}(V, C)$. Clearly, $\text{Sol}(V, C) \subseteq \text{Sol}(V, C \setminus \{c\})$, so I is a no-instance if and only if there is an assignment α that satisfies $C \setminus \{c\}$ and does not satisfy c. Note that such an assignment α satisfies $I' = (V, (C \setminus \{c\}) \cup \bar{c})$, so it exists if and only if I' is satisfiable.

Left-to-right direction. Assume that REDUNDANT($\mathcal{A}$) is in P. We show $\text{CSP}_{\leq 1}(\mathcal{A} \cup \bar{R})$ is in P as well. Let $I = (V, C)$ be an instance of the former problem, where $c = \bar{R}(x_1, \ldots, x_k)$ is in C, and let $\bar{c} = R(x_1, \ldots, x_k)$. Observe that $I' = (V, (C \setminus \{c\}) \cup \bar{c}, \bar{c})$ is an instance of REDUNDANT($\mathcal{A}$), and check whether it is a yes-instance. We claim that I is satisfiable if and only if I' is a no-instance. Indeed, I' is no-instance if and only if $\text{Sol}(V, (C \setminus \{c\}) \cup \bar{c}) \neq \text{Sol}(V, C \setminus \{c\})$. Clearly, $\text{Sol}(V, (C \setminus \{c\}) \cup \bar{c}) \subseteq \text{Sol}(V, C \setminus \{c\})$, so I' is a no-instance if and only if there exists an assignment α that satisfies $C \setminus \{c\}$ and does not satisfy $\bar{c}$. Note that such an assignment α satisfies both $(C \setminus \{c\})$ and c, and hence satisfies $I = (V, C)$, so α exists if and only if I is satisfiable. ◀

B Proof of Theorem 5

Proof. We only sketch the proof since the details are very similar to the classical reduction for CSPs in Theorem 1. The structures $\mathcal{A}$ and $\mathcal{B}$ have finite signatures so we can (without loss of generality) assume that we have access to the following information: (1) the pp-definitions in $\mathcal{A}$ for the relations in $\mathcal{A}^* \setminus \mathcal{A}$, and (2) for every $R \in \mathcal{B}^* \setminus \mathcal{B}$, a pp-definiton of R in $\mathcal{A} \cup \mathcal{B}$ with k_R $\mathcal{B}$-constraints.

Let $I = (V, C, k)$ be an arbitrary instance of $\text{CSP}_{\leq}(\mathcal{A}^* \cup \mathcal{B}^*)$. We begin by replacing each $(\mathcal{A}^* \setminus \mathcal{A})$-constraint by its precomputed pp-definition in $\mathcal{A}$. This does not increase the parameter. We similarly replace every $(\mathcal{B}^* \setminus \mathcal{B})$-constraint by its pp-definition over $\mathcal{A} \cup \mathcal{B}$. There are at most k such constraints in C, and each of them is replaced by at most k_R constraints over $\mathcal{B}$ for a fixed constant k_R. This reduction is obviously correct and can be computed in polynomial time. The bound on the parameter follows since k_R only depends on the chosen pp-definition over the fixed and finite language $\mathcal{A} \cup \mathcal{B}$. ◀

C Proof of Theorem 7

Proof. Let e be an endomorphism with minimal range in $\text{End}(\mathcal{A} \cup \mathcal{B})$, let $\mathcal{A}' = \{e(R) \mid R \in \mathcal{A}\}$ and $\mathcal{B}' = \{e(R) \mid R \in \mathcal{B}\}$, of the same signature as $\mathcal{A}$ and $\mathcal{B}$. First, let (V, C, k) be an instance of $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$. For each constraint $R(\mathbf{x}) \in C$ we simply replace it by $e(R)(\mathbf{x})$. It is then easy to verify, and well-known, that the resulting instance is satisfiable if and only if (V, C) is satisfiable. Furthermore, observe that if (1) $R \in \mathcal{A}$ then $e(R) \in \mathcal{A}'$, and (2) if $R \in \mathcal{B}$ then $e(R) \in \mathcal{B}'$. Hence, (V, C) has k alien constraints $R_1(\mathbf{x}_1), \ldots, R_k(\mathbf{x}_k)$ then the new instance has k alien constraints $e(R_1)(\mathbf{x}_1), \ldots, e(R_k)(\mathbf{x}_k)$, too. Hence, it is an fpt-reduction.

The other direction is similar: let $(V, C_1 \cup C_2, k)$ be an instance of $\text{CSP}_{\leq}(\mathcal{A}' \cup \mathcal{B}')$. For each constraint $e(R)(\mathbf{x}) \in C_1$ we replace it by $R(\mathbf{x})$ for $R \in \mathcal{A}$, and for each constraint $e(R)(\mathbf{x}) \in C_2$ we replace it by $R(\mathbf{x})$ for $R \in \mathcal{B}$. Clearly, the number of alien constraints remains unchanged, and the reduction is an fpt-reduction which exactly preserves #ac. ◀

D Proof of Lemma 9

Proof. Let (V, C) be an instance of $\text{CSP}(\mathcal{A} \cup \mathcal{C})$. Pick $c \in \mathcal{C}$ and consider the set of constraints $C^c = \{c(x) \mid c \in C\}$. Pick an arbitrary $c(v) \in C^c$ and consider the instance (V', C') obtained by (1) identifying v' with v for any $c(v') \in C^c$ throughout the instance and (2) replacing C^c from the set of constraints with the single constraint $c(v)$. If we repeat this for every $c \in \mathcal{C}$ we obtain an instance of $\text{CSP}_{\leq |\mathcal{C}|}(\mathcal{A} \cup \mathcal{C})$ which is satisfiable if and only if (V, C) is satisfiable. ◀

E Proof of Theorem 10

Proof. We use the fact that every structure with finite domain has a CSP that is either polynomial-time solvable or NP-hard [13, 27]. Assume that $\text{CSP}(\mathcal{A} \cup \mathcal{C}_D)$ is in P. First, we claim that every tuple over D is pp-definable over $\mathcal{A} \cup \mathcal{C}_D$. Thus, let $n \geq 1$ and pick $t = (d_1, \ldots, d_n) \in D^n$. It follows that $\{t\}(x_1, \ldots, x_n) \equiv \{d_1\}(x_1) \wedge \ldots \wedge \{d_n\}(x_n)$ since each $\{d_i\} \in \mathcal{C}_D$. Second, pick an n-ary relation $R = \{t_1, \ldots, t_m\} \in \mathcal{B}$. Since each $\{t_i\} \in \langle \mathcal{A} \cup \mathcal{C}_D \rangle$, R is a finite union of relations in $\langle \mathcal{A} \cup \mathcal{C}_D \rangle$, and every relation in $\mathcal{B}$ is existential positive definable over $\mathcal{A} \cup \mathcal{C}_D$. We conclude that Theorem 8 is applicable and that $\text{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT parameterized by #ac.

For the second statement, we assume that $\mathrm{CSP}(\mathcal{A} \cup \mathcal{C}_D)$ is NP-hard. We show that there is a polynomial-time reduction from $\mathrm{CSP}(\mathcal{A} \cup \mathcal{C}_D)$ to $\mathrm{CSP}_{\leq p}(\mathcal{A} \cup \mathcal{B})$ for some p that only depends on $\mathcal{B}$. First, let $D = \{a_1, \ldots, a_d\}$ and consider the relation $E = \{(e(a_1), \ldots, e(a_d)) \mid e \in \mathrm{End}(\mathcal{A} \cup \mathcal{B})\}$, i.e., the set of endomorphisms of $\mathcal{A}$ viewed as a d-ary relation. It is known that $E \in \langle \mathcal{A} \cup \mathcal{B} \rangle$ [1, proof of Theorem 17] since $\mathcal{A} \cup \mathcal{B}$ is a core. Let $I = (V, C)$ be an instance of $\mathrm{CSP}(\mathcal{A} \cup \mathcal{C}_D)$. By Lemma 9 we can without loss of generality assume that I is an instance of $\mathrm{CSP}_{\leq d}(\mathcal{A} \cup \mathcal{C}_D)$, and we will produce a polynomial-time reduction to $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \{E\})$ which is sufficient to prove the claim under Theorem 5.

Let $v_1, \ldots, v_d \in V$ such that $c_i(v_i) \in C$, i.e., the variables being enforced constant values via the constraints in $\mathcal{C}_D$. We remove the constraints $c_1(v_1), \ldots, c_d(v_d)$ and replace them with $E(v_1, \ldots, v_d)$. We claim that the resulting instance (V, C') is satisfiable if and only if (V, C) is satisfiable. First, assume that $f \colon V \to D$ is a satisfying assignment to (V, C). We see that $f(v_i) = c_i$ for each $i \in [d]$ and thus that $(f(v_1) \ldots, f(v_d)) \in E$. For the other direction, assume that $g \colon V \to D$ is a satisfying assignment to (V, C') and consider the function defined by $\pi(a_i) = g(v_i)$ for every $i \in [d]$. Clearly, $(\pi(v_1), \ldots, \pi(v_d)) \in E$, and it follows that $\pi \in \mathrm{Aut}(\mathcal{A} \cup \mathcal{B})$. Since $\mathrm{Aut}(\mathcal{A} \cup \mathcal{B})$ is an automorphism group it follows that $\pi^{-1} \in \mathrm{Aut}(\mathcal{A} \cup \mathcal{B})$, too, and the function $h(x) = \pi^{-1}(g(x))$ then gives us the required satisfying assignment. ◀

F Proof of Lemma 12

Proof. By assumption, $c_0 \in \langle \mathcal{A} \rangle$, and to simplify the notation we assume that $c_0 \in \mathcal{A}$. This can be done without loss of generality since in the pp-definition below we can replace any occurrence of c_0 by its pp-definition. Fix a tuple $(a_1, \ldots, a_n) \in R$ which is not constantly 0. This is possible since $R \neq \emptyset$ and since R is not 0-valid. We then use the definition $c_1(x) \equiv \exists y \colon c_0(y) \wedge R(x_1, \ldots, x_n)$ where $x_i = x$ if $a_i = 1$ and $x_i = y$ if $a_i = 0$. ◀

G Proof of Lemma 13

Proof. Let (V, C) denote an instance of $\mathrm{CSP}(\mathcal{A} \cup \{c_0, c_1\})$. Assume (without loss of generality by Lemma 12) that the constant relations c_0 and c_1 appear at most one time, respectively, in C and that they restrict the variables z_0 and z_1 as follows: $c_0(z_0)$ and $c_1(z_1)$. Let (V, C') denote the instance of $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \{\neq\})$ where $C' = (C \setminus \{c_0(z_0), c_1(z_1)\}) \cup \{z_0 \neq z_1\}$. It is not difficult to verify that (V, C') is satisfiable if and only if (V, C) is satisfiable since $\mathcal{A}$ is invariant under complement. ◀

H Proof of Theorem 15

Proof. Condition 3. says that every relation in $\mathrm{Orb}(\mathcal{B})$ is a finite union of relations in $\langle \mathcal{A} \rangle$ (as pointed out in Section 4.2). Condition 2. together with the well-known characterization of ω-categorical structures by Engeler, Svenonius, and Ryll-Nardzewski [17, Theorem 6.3.1] imply that every relation in $\mathcal{B}$ is a finite union of relations in $\langle \mathcal{A} \rangle$. We can now apply Theorem 8. ◀

I Proof of Lemma 17

Proof. The structure $\mathcal{B}$ is a model-complete core so every relation in $\mathrm{Orb}(\mathcal{A})$ is pp-definable in $\mathcal{A}$. Pick an arbitrary relation $R \in \mathcal{B}$. The structure $\mathcal{B}$ is ω-categorical so R is a finite union of relations in $\mathrm{Orb}(\mathcal{B})$. We have assumed that $\mathrm{Orb}(\mathcal{B}) \subseteq \mathrm{Orb}(\mathcal{A})$ so R is existential positive definable in $\mathcal{A}$. The result follows from Theorem 15. ◀

J Proof of Lemma 23

Proof. We prove the contrapositive: if there is a satisfiable instance of $\mathrm{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ with every satisfying assignment taking at least c values, then $\mathcal{A} \cup \mathcal{B}$ admits a pp-definition of NEQ_c with k constraints from $\mathcal{B}$. We will use the fact that for every d,

$$\mathrm{NEQ}_c(x_1, \ldots, x_c) \equiv \exists x_{c+1}, \ldots, x_{c+d} \colon \; \mathrm{NEQ}_d(x_1, \ldots, x_{c+d}),$$

so it is enough to pp-define a relation $\mathrm{NEQ}_{c'}$ with $c' \geq c$ to prove the lemma.

Consider a satisfiable instance I of $\mathrm{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ as a quantifier-free primitive-positive formula $\phi(x_1, \ldots, x_n)$. Note that I contains at most k constraints from $\mathcal{B}$. Let α be a satisfying assignment to I with minimum range, and assume without loss of generality that the range is $[c]$ for some $c \in \mathbb{Z}_+$. We claim that $I' = \phi(y_{\alpha(x_1)}, \ldots, y_{\alpha(x_n)})$ is a pp-definition of NEQ_c. First, note that every injective assignment satisfies I'. Moreover, every satisfying assignment to I' also satisfy I, so it must take at least r values (i.e. be injective) by the choice of α. Finally, note that I' contains at most k constraints from $\mathcal{B}$, hence it is an instance of $\mathrm{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$. ◀

K Proof of Lemma 24

Proof. Let I be an instance of $\mathrm{CSP}(\mathcal{A}^+)$. We construct an equivalent instance I' of $\mathrm{CSP}_{\leq k}(\mathcal{A} \cup \mathcal{B})$ starting with all constraints in I except for the applications of singleton relations, i.e. unit assignments. Assume without loss of generality that I does not contain two contradicting unit assignments. To simulate c constants, create variables $x_1, \ldots, x_c$ and add the pp-definitions of $\mathrm{NEQ}_c(x_1, \ldots, x_c)$ to I'. This requires k applications of $\mathcal{B}$-constraints. Now, replace every variable v in I' such that the constraint $v = i$ is in I with the new variable x_i. Clearly, the reduction requires polynomial time. The correctness follows since we are using a pp-definition to simulate relation NEQ_c, and it can be verified using Theorem 5. ◀

L Proof of Theorem 26

Proof. The problems under consideration are equivalent under polynomial-time Turing reductions by Lemma 3. By Theorem 4, Redundant($\mathcal{A}$) is in P if and only if $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \bar{\mathcal{A}})$ is in P, where $\bar{\mathcal{A}} = \{\bar{R} : R \in \mathcal{A}\}$ is the language of complements of $\mathcal{A}$-relations. Clearly, if $\mathcal{A}$ is neither Horn nor 0-valid, then even $\mathrm{CSP}_{\leq 0}(\mathcal{A} \cup \bar{\mathcal{A}})$ is NP-hard, implying that Redundant($\mathcal{A}$) is coNP-hard as pointed out after Lemma ??. If $\mathcal{A}$ is Horn, then then $\mathrm{CSP}_{\leq}(\mathcal{A} \cup \mathcal{B})$ is in FPT parameterized by #ac so $\mathrm{CSP}_{\leq 1}(\mathcal{A} \cup \bar{\mathcal{A}})$ is in P, and hence Redundant($\mathcal{A}$) is in P. If $\mathcal{A}$ is 0-valid and not Horn, then $\bar{\mathcal{A}}$ is not 0-valid and $\mathrm{CSP}(\mathcal{A} \cup \bar{\mathcal{A}})$ is NP-hard. Now, Case 2 of Theorem 25 applies. ◀

M Proof of Proposition 27

Proof. We will assume that the relations are represented by their defining formulas. This way, we can use the results of [10] immediately. We can also test inclusion of a tuple in a relation compute a representative set of tuples, i.e. a set such that every tuple in the relation is isomorphic to one member of this set.

We first check whether $\mathcal{A}$ and $\mathcal{B}$ are 0-valid and whether they are Horn. For the first, check whether the all-0 tuple is in the relation. For the second, recall from [7, Lemma 8] that a relation is Horn if and only if it is closed under any binary injective operation. Choose an arbitrary binary injective function f and check that, for every pair of tuples in the representative set, the result of applying f to them componentwise is also in the relation. To see that this is sufficient, consider an equality relation R, two arbitrary tuples $a, b \in R$ and their representatives a', b', i.e. tuples in the representative set such that $a_i = a_j \iff a'_i = a'_j$ and $b_i = b_j \iff b'_i = b'_j$. Then $(a_i, b_i) = (a_j, b_j) \iff (a'_i, b'_i) = (a'_j, b'_j)$, so $f(a', b') \in R \implies f(a, b) \in R$. If $\mathcal{A}$ is Horn or both $\mathcal{A}$ and $\mathcal{B}$ are 0-valid, then $k = \infty$ by Corollary 18. Otherwise, $k < \infty$. If $\mathcal{A}$ is neither Horn nor constant, then $\mathrm{CSP}(\mathcal{A})$ is NP-hard, and $k = 0$.

The case we are left with is when $\mathcal{A}$ is constant and not Horn, while $\mathcal{B}$ is not constant. By Lemma 22, there exists $c \in \mathbb{N}$ such that $\mathrm{CSP}(\mathcal{A}^+_c)$ is NP-hard, and $\mathrm{CSP}(\mathcal{A}^+_{c'})$ is in P for all $c' < c$. We show that c can be computed. Note that $\mathrm{CSP}(\mathcal{A}^+_1)$ is in P because every instance is satisfiable by a constant assignment. Now consider $c = 2$. By Theorem 54 in [22] and Lemma 21, $\mathrm{CSP}(\mathcal{A}^+_2)$ is in P if the 2-slice of $\mathcal{A}$ is preserved by an affine operation, and NP-hard otherwise. We can compute the 2-slice and check whether it is closed under an affine operation in polynomial time. If $\mathrm{CSP}(\mathcal{A}^+_2)$ is NP-hard, then $k = 1$ because $\mathrm{NEQ}_2 \in \langle \emptyset \cup \mathcal{B} \rangle_{\leq k}$. Otherwise, proceed to $c \geq 3$. Again, using Theorem 54 in [22] and Lemma 21, we have that $\mathrm{CSP}(\mathcal{A}^+_c)$ for $c \geq 3$ is in P if the c-slice of $\mathcal{A}$ is trivial (contains only empty or complete relations), and NP-hard otherwise. This can also be checked in polynomial time.

Now that c is determined, by Theorem 25, k is the minimum integer such that $\mathrm{NEQ}_c \in \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq k}$. Note that $k \leq \binom{c}{2}$ since $\mathrm{NEQ}_2 \in \langle \emptyset \cup \mathcal{B} \rangle_{\leq k}$. We can find minimum k by considering every value $1 \leq t \leq \binom{c}{2}$ in increasing order and checking whether $\mathrm{NEQ}_c \in \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq t}$. Thus, it remains to show that pp-definability of NEQ_c in $\mathcal{A} \cup \mathcal{B}$ with at most t constraints from $\mathcal{B}$ is decidable. To see this, we can view a pp-definition as a relation $R \in \langle \mathcal{A} \cup \mathcal{B} \rangle_{\leq t}$ such that the projection of R onto first c indices is NEQ_c. Furthermore, $R(x_1, \ldots, x_n) \equiv R_{\mathcal{A}}(x_1, \ldots, x_n) \wedge R_{\mathcal{B}}(x_1, \ldots, x_n)$, where $R_{\mathcal{A}} \in \langle \mathcal{A} \rangle$ and $R_{\mathcal{B}} \in \langle \emptyset \cup \mathcal{B} \rangle_{\leq t}$. Note that $R_{\mathcal{B}}$ can only depend on $\ell \leq r(\mathcal{B}) \cdot t$ arguments, where $r(\mathcal{B})$ is the maximum arity of a relation in $\mathcal{B}$, which is constant. The relation $R_{\mathcal{A}}$ projected onto these ℓ arguments is an equality relation of arity ℓ. We can guess ℓ, enumerate all equality relations $R'_{\mathcal{A}}$ of arity ℓ pp-definable in $\mathcal{A}$ using [10] and enumerate all relations $R'_{\mathcal{B}}$ in $\langle \mathcal{B} \rangle$ definable using t constraints, and check whether $R'_{\mathcal{A}}(x_1, \ldots, x_c) \wedge R'_{\mathcal{B}}(x_1, \ldots, x_c)$ projected onto $x_1, \ldots, x_c$ is NEQ_c. This completes the proof. ◀