COLLAPSING THE BOUNDED WIDTH HIERARCHY FOR INFINITE-DOMAIN CSPS: WHEN SYMMETRIES ARE ENOUGH

ANTOINE MOTTET, TOMÁŠ NAGY, MICHAEL PINSKER, AND MICHAL WRONA

ABSTRACT. We prove that relational structures admitting specific polymorphisms (namely, canonical pseudo-WNU operations of all arities $n \geq 3$) have low relational width. This implies a collapse of the bounded width hierarchy for numerous classes of infinite-domain CSPs studied in the literature. Moreover, we obtain a characterization of bounded width for first-order reducts of unary structures and a characterization of MMSNP sentences that are equivalent to a Datalog program, answering a question posed by Bienvenu et al. In particular, the bounded width hierarchy collapses in those cases as well. Our results extend the scope of theorems of Barto and Kozik characterizing bounded width for finite structures, and show the applicability of infinite-domain CSPs to other fields.

1. INTRODUCTION

Understanding what mathematical structure in problems enables efficient algorithms is one of the most fundamental goals in theoretical computer science; the field therefore strives to achieve such understanding completely for restricted classes of problems that are sufficiently general to contribute to the overall picture. Another core goal is identifying large natural subclasses of the complexity class NP that exhibit a complexity dichotomy, i.e., which contain only problems that are either solvable in polynomial time or NP-complete, and therefore avoid intermediate complexity classes. One class of problems that touches both of the above directions is the class of constraint satisfaction problems (CSPs), which can be cast as homomorphism problems (of a given structure to a fixed structure called a template). Feder and Vardi famously proposed the conjecture that the class of CSPs with a finite template exhibits such a dichotomy \cite{FederVardi98}. With time, this goal evolved from just proving the dichotomy to precisely explaining, as uniformly as possible, the mathematical structure of problems that causes their tractability or hardness. One reason why this goal received considerable attention and was pursued for decades is that CSPs with a finite template hit the sweet spot between generality and structure — they are quite general and include many computational problems of interest, yet they possess sufficient structure that, in a highly non-trivial way, can be extracted and either be used algorithmically or in hardness proofs.

The Feder-Vardi conjecture having been confirmed independently by Bulatov and Zhuk in 2017 \cite{BulatovZhuk17, Zhuk17}, it is natural to look for a more general class of problems where potentially there is still...
sufficient structure to explain all tractability within the class and that possibly exhibits a similar complexity dichotomy. There are several ways to expand the framework, but one of the most natural ways is to allow infinite templates, thereby including many natural problems that can be formulated as CSPs, but not with a finite template. This is the case for linear programming, some reasoning problems in artificial intelligence such as ontology-mediated querying, or even problems as simple to formulate as the digraph acyclicity problem. On the other hand, allowing arbitrary infinite templates would amount to including basically any computational problem [18] and defy the hope for a structural approach as well as for a complexity dichotomy (or even containment in NP); it is therefore necessary to restrict the class of infinite templates under consideration. To achieve the objective of a restricted class that does not suffer from the above deficiencies, properly extends the class of finite-template CSPs, and enriches the latter class by interesting problems, model theory provides us with a useful language and methodology – it has been used to identify a class of infinite templates which share some properties with finite structures such that the corresponding CSPs are always in NP and such that a structural approach to explain tractability within the class is, in principle, achievable. A complexity dichotomy for this class, whose templates are first-order reducts of $k$-homogeneous $\ell$-bounded $\omega$-categorical structures (explained on a high level in Section 1.1 and defined precisely in Section 2), has been conjectured by Bodirsky and Pinsker [32, 11, 10, 5, 6], and a rich mathematical theory to investigate the complexity of problems within the class has evolved over the years (see e.g. [15, 57]).

Methodologically, the key to resolving the Feder-Vardi dichotomy conjecture was the universal-algebraic approach. This approach assigns to each template the set of certain finite-dimensional symmetries, called polymorphisms, and uses them to identify the structure in problem instances which is then used either in polynomial-time algorithms or in hardness proofs. Given how well-developed and powerful this approach is, it is natural to try to lift to the infinite-template framework as many tools from the finite case as possible. The most general results in the area of infinite-template CSPs obtained so far use canonical polymorphisms, which behave like operations on finite sets and which sometimes allow to mimic the universal-algebraic approach to finite-template CSPs. Canonical polymorphisms alone are in general insufficient to explain, say, polynomial-time tractability, and a characterisation of the range of their applicability is missing, but on the positive side their use covers a vast majority of the results obtained in the area. Moreover, even in situations where canonical polymorphisms do not yield a black-box argument by directly lifting results for finite-template CSPs, they prove useful in the design of more elaborate arguments [53, 54].

The two main algorithmic techniques employed in polynomial-time algorithms for finite-template CSPs are local consistency checking and generalised Gaussian elimination. The applicability of pure forms of these techniques is well understood. Again, it is natural to understand the applicability of these techniques for infinite-template CSPs. The present paper is focused on the local consistency checking technique, more precisely on the concept of bounded width. One way to see it is as a generalisation of bounded width resolution for SAT, but adapted to deal with general constraints, rather than just with clauses. Intuitively speaking, local consistency checking consists in propagating local information through a structure so as to infer global information: consider, for example, computing the transitive closure of a relation as deriving global information from local one.

The use of local consistency methods is not limited to constraint satisfaction. Indeed, local consistency checking is also used for such problems as the graph isomorphism problem, where it is known as the Weisfeiler-Leman algorithm. Here, the technique can be used to derive implied constraints that an isomorphism between two graphs has to satisfy so as to narrow down the search space, but local consistency is in fact powerful enough to solve the graph isomorphism problem over any non-trivial minor-closed class of graphs [45]. Notably, the best algorithm for graph isomorphism to date also uses local consistency as a subroutine [3]. Finally, local consistency can be used to solve games involved in formal verification such as parity games and mean-payoff
games [22]. However, unlike in the other applications mentioned above, these algorithms are not efficient since the constructed instances are exponentially large.

One of the reasons for the ubiquity of local consistency methods is that their underlying principles can be described in many different languages, such as the language of category theory [1], the language of finite model theory (by Spoiler-Duplicator games [45] or by homomorphism duality [2]), and logical definability (in Datalog, or infinitary logics with bounded number of variables) [42]. For constraint satisfaction problems over a finite template, the power of local consistency checking can be characterised by conditions on the set of polymorphisms of the template: local consistency checking correctly solves the problem if, and only if, the polymorphisms satisfy these conditions [8, 53, 51]. Moreover, whenever this is the case, then in fact only a very restricted form of local consistency checking is needed [4]. This fact is known as the collapse of the bounded width hierarchy, and it has strong consequences both for complexity and logic. Firstly, the collapse provides efficient algorithms that are able to solve all the CSPs that are solvable by local consistency methods [50], which yields in particular a polynomial-time algorithm solving instances of the uniform CSP where the template is also part of the input. Secondly, this collapse induces collapses in all the areas mentioned at the beginning of this paragraph.

Unlike for finite-template CSPs, an algebraic characterisation of the applicability of local consistency checking is missing for infinite-template CSPs. In fact, even in situations where the complexity of a CSP is captured by the polymorphisms of the template, no purely algebraic description of CSPs solvable by local consistency methods is possible [26, 27, 44]; this is even the case for well-behaved classes of problems such as temporal CSPs [29], which belong to the range of the aforementioned Bodirsky-Pinsker conjecture. These negative results are to be compared with the recent result by Mottet and Pinsker [55] that did provide an algebraic description for several subclasses of infinite-template CSPs within that range.

1.1. Results. In the present paper, we focus on applying the theory of canonical functions to study the power of local consistency checking for CSPs with templates within the Bodirsky-Pinsker conjecture, namely, first-order reducts of \( k \)-homogeneous \( \ell \)-bounded \( \omega \)-categorical structures. Our objective is threefold: firstly, we wish to obtain generic sufficient conditions that imply that local consistency checking solves a given CSP; secondly, we aim at understanding the amount of locality needed for local consistency methods to solve the CSP, as measured by the so-called relational width; and thirdly, we want to show that our sufficient conditions are actually also necessary for certain interesting subclasses. We now give a high-level presentation of our results; precise definitions of all concepts mentioned here can be found in the preliminaries.

In order to achieve the first two objectives, we prove that sufficiently locally consistent instances can be turned into locally consistent instances of a finite-template CSP. Imposing conditions on the canonical polymorphisms of the original template, we can enforce this finite-template CSP to have bounded width. It then has relational width at most \((2,3)\) by [4]; application of this result in turn allows us to obtain a bound on the relational width of the original CSP which only depends on the two simple parameters \( k, \ell \) of the presentation of the template by a \( k \)-homogeneous \( \ell \)-bounded structure. Building on the results for finite templates, we therefore obtain a collapse of the relational width reminiscent of the collapse in the finite case for all structures whose canonical polymorphisms satisfy our conditions.

Before stating this result in full detail, let us comment on the nature of the range of the Bodirsky-Pinsker conjecture and its use in local consistency algorithms. A countable structure is \( \omega \)-categorical if its automorphism group is large in the sense that it acts with finitely many orbits on \( n \)-tuples for all \( n \geq 1 \): that is, the structure has up to automorphisms only finitely many distinct tuples of a fixed finite length. The additional conditions of \( k \)-homogeneity and \( \ell \)-boundedness state that moreover, the orbit of any \( n \)-tuple is uniquely determined by the orbits of its \( k \)-subtuples, and that conversely which combinations of orbits of \( k \)-tuples can appear on an \( n \)-tuple is a local property
(measured by \(\ell\)). A first-order reduct of such a structure is any structure all of whose relations are unions of orbits. Local consistency algorithms, applied to an instance of such a template, keep lists of possible orbits of tuples of variables, reduce them according to the constraints, and then propagate information between these lists locally to make them consistent. By \(\omega\)-categoricity, this procedure terminates (and does so in polynomial time).

**Theorem 1.1.** Let \(k, \ell \geq 1\), and let \(\mathcal{A}\) be a first-order reduct of a \(k\)-homogeneous \(\ell\)-bounded \(\omega\)-categorical structure \(\mathcal{B}\).

- If the clone of \(\text{Aut}(\mathcal{B})\)-canonical polymorphisms of \(\mathcal{A}\) contains pseudo-WNU operations modulo \(\text{Aut}(\mathcal{B})\) of all arities \(n \geq 3\), then \(\mathcal{A}\) has relational width at most \((2k, \max(3k, \ell))\).
- If the clone of \(\text{Aut}(\mathcal{B})\)-canonical polymorphisms of \(\mathcal{A}\) contains pseudo-totally symmetric operations modulo \(\text{Aut}(\mathcal{B})\) of all arities, then \(\mathcal{A}\) has relational width at most \((k, \max(k+1, \ell))\) if \(k > 1\), and \((1, 1)\) if \(k = 1\).

We remark that if \(\mathcal{A}\) is a finite template, then it satisfies the conditions of Theorem 1.1; moreover, the width obtained in the first item of Theorem 1.1 coincides with the width given by Barto’s collapse result from [4], the width obtained in the second item coincides with the width of structures with totally symmetric operations of all arities from [11, 42], and our algebraic conditions coincide with the known conditions for finite templates. We now provide the short argument for the reader familiar with the precise notions of the theorem; others might prefer to skip the remainder of this paragraph, and return to it later. Every finite structure \(\mathcal{A}\) with domain \(\{a_1, \ldots, a_n\}\) is a first-order reduct of the structure \(\mathcal{B} = (\{a_1, \ldots, a_n\}; \{a_1\}, \ldots, \{a_n\})\), which is \(1\)-homogeneous and \(2\)-bounded (and trivially \(\omega\)-categorical). For \(1\)-homogeneity, note first that the automorphism group \(\text{Aut}(\mathcal{B})\) of \(\mathcal{B}\) is trivial. Hence, the orbit of any tuple with respect to \(\text{Aut}(\mathcal{B})\) consists only of the tuple itself, and is clearly determined by the unary singleton relations that hold on the tuple. To see \(2\)-boundedness, note that in order for any structure to embed into \(\mathcal{B}\), it is sufficient to ensure that all elements of that structure are contained in precisely one unary relation, and that no two distinct elements are contained in the same unary relation. Finally, it also follows from the triviality of \(\text{Aut}(\mathcal{B})\) that any polymorphism of \(\mathcal{A}\) is \(\text{Aut}(\mathcal{B})\)-canonical, making this condition void, that any pseudo-WNU operation modulo \(\text{Aut}(\mathcal{B})\) is also a WNU operation, and that any pseudo-totally symmetric operation modulo \(\text{Aut}(\mathcal{B})\) is a totally symmetric operation.

As a corollary of Theorem 1.1 we obtain a collapse of the bounded width hierarchy for first-order reducts of numerous structures studied in the literature [20, 24, 23, 49] where applicability of local consistency methods had already been classified, as well as of unary structures, i.e., structures over a finite unary signature [25]. For the latter class, in order to apply our theorem, we first need to show that all structures in the class with bounded width satisfy the assumptions of the theorem; this is one of our contributions towards the third objective mentioned in the beginning of this section. To this end, we build on recent work by Mottet and Pinsker [55] and expand the use of their smooth approximations to fully suit equational (non-)affineness, which is roughly the algebraic situation imposed by local consistency solvability. The main technical contribution is a new loop lemma that exploits deep algebraic tools from the finite and, assuming the use of canonical functions is unfruitful, allows to obtain the existence of polymorphisms of every arity \(n \geq 3\) and satisfying certain strong symmetry conditions. The study of unary structures is motivated on the one hand by the fact that its first-order reducts are a subclass of the class of first-order reducts of classical structures like \((\mathbb{Z}; +, \leq)\), \((\mathbb{Q}; +, \leq, 1)\), or \((\mathbb{R}; +, \times)\) [25], which are of particular interest in mathematics and computer science [22, 46]. On the other hand, they are also a superclass of the class of finite structures, consisting of templates where single values can be blown up to finite or infinite sets (by means of unary predicates representing them). The class therefore constitutes an excellent testbed for algebraic and combinatorial generalizations, in the spirit of the introduction, of statements about finite structures to structures that are in a sense “just beyond finite.”
Corollary 1.2. Let $\mathbb{A}$ be a structure that has bounded width. If $\mathbb{A}$ is a first-order reduct of:

- the universal homogeneous graph $\mathcal{G}$ or tournament $\mathcal{T}$, or of a unary structure, then $\mathbb{A}$ has relational width at most $(4, 6)$;
- the universal homogeneous $K_n$-free graph $\mathbb{H}_n$, where $n \geq 3$, then at most $(2, n)$;
- $(\mathbb{N}; =)$, the countably infinite equivalence relation $\mathbb{C}^\omega_\omega$ with infinitely many equivalence classes, all of infinite size, or the universal homogeneous partial order $\mathbb{P}$, then at most $(2, 3)$.

In Section 5.2, after our detailed treatment of unary structures, we provide the proof of Corollary 1.2 along with examples showing that most of the bounds from the statement are tight.

Our second contribution (after unary structures) to the third objective of characterizing templates with bounded width concerns templates representing the model-checking problem for Feder and Vardi’s logic MMSNP [42]. Among CSPs definable in this logic, we characterize those with bounded width as precisely those whose appropriate template satisfies the conditions of Theorem 1.3 again using smooth approximations and our new loop lemma. CSPs with a template having bounded width are definable in the declarative language Datalog, which is from a logical point of view the existential positive fragment of least fixpoint logic (see, e.g., [42] for a clear introduction of Datalog in the context of constraint satisfaction). Coupled with our results, this observation allows us to solve the following open problem from [12, 43]. The Datalog-rewritability problem for MMSNP is the problem of deciding, given as input an MMSNP sentence $\Phi$, whether $\neg \Phi$ is equivalent to a Datalog program. This problem appears naturally in the area of ontology-mediated querying, where a certain querying language used in practice and known as (ALC, UCQ) is logically equivalent to the logic MMSNP. The evaluation of such queries is in general a coNP-complete problem. When such a query is Datalog-rewritable, the evaluation of the query becomes a polynomial-time tractable problem.

Theorem 1.3. The Datalog-rewritability problem for MMSNP is decidable, and is 2NExpTime-complete.

1.2. Related results. Local consistency for $\omega$-categorical structures was studied for the first time in [17] where basic notions were introduced and some basic results provided. Bounded width was characterised for first-order reducts of certain $k$-homogeneous $\ell$-bounded structures in [55, 29].

A structure $\mathbb{A}$ has bounded strict width [42] if not only CSP($\mathbb{A}$) is solvable by local consistency methods, but moreover every partial solution of a locally consistent instance can be extended to a total solution. The articles [61] and [60] give the upper bound $(2, \max(3, \ell))$ on the relational width for first-order expansions of some classes of 2-homogeneous, $\ell$-bounded structures under the stronger assumption of bounded strict width; Corollary 1.2 for first-order reducts of $\mathbb{H}_n$ and of $\mathbb{C}^\omega_\omega$ also follows from [61]. The results of [60] were generalized in [56], where the upper bound $(k, \max(k,1,\ell))$ on the relational width was given for first-order expansions of certain classes of $k$-homogeneous, $\ell$-bounded structures with bounded strict width for $k > 2$.

It has been recently shown in [62] that for a large class of $\omega$-categorical structures the existence of an oligopotent quasi near-unanimity polymorphism, which is equivalent to having bounded strict width, and the conditions on canonical polymorphisms in Theorem 1.1 are not the only sufficient conditions for having bounded width. Namely, for certain 2-homogeneous, $\ell$-bounded structures having a chain of directed quasi Jonsson polymorphisms the bound of $(2, \max(3, \ell))$ was obtained for the relational width, which is the same upper bound as in [60].

There are also classes of $k$-homogeneous $\ell$-bounded structures for which it is known that all their first-order reducts with bounded width have relational width at most $(2k, \max(3k, \ell))$, but where this collapse follows neither from Theorem 1.1 nor from any other result mentioned above. For the first-order reducts of $(\mathbb{Q}; <)$, such a collapse follows from [29, 36], and for first-order reducts of certain $k$-uniform hypergraphs with $k \geq 3$, this collapse was proven in [51] using a combination of Theorem 1.1 and additional work. To our knowledge, there is no class of $k$-homogeneous $\ell$-bounded...
structures where the bound on the relational width from the first item of Theorem 1.1 would be known not to apply.

1.3. Organisation of the present article. In Section 2 we provide the basic notions and definitions. The reduction to the finite using canonical functions which leads to the collapse of the bounded width hierarchy, the proof of Theorem 1.1 is given in Section 3. We then extend the algebraic theory of smooth approximations in Section 4 before applying it to first-order reducts of bounded width hierarchy, the proof of Theorem 1.1, is given in Section 3. We then extend the definitions. The reduction to the finite using canonical functions which leads to the collapse of the

2. Preliminaries

2.1. Structures and model-theoretic notions. All relational structures in this article are assumed to be at most countable. For sets B, I, the orbit of a tuple b ∈ B^I under the action of a permutation group G on B is the set \{\alpha(b) | \alpha \in G\}; it is also called the G-orbit of b. A countable relational structure B is \(\omega\)-categorical if its automorphism group Aut(B) is oligomorphic, i.e., for all \(n \geq 1\), the number of orbits of the action of Aut(B) on \(n\)-tuples is finite. For an \(\omega\)-categorical relational structure B and for every \(n \geq 1\) it holds that two \(n\)-tuples of elements of the structure are in the same orbit under Aut(B) if, and only if, they have the same type, i.e., they satisfy the same first-order formulas. The atomic type of an \(n\)-tuple is the set of all atomic formulas satisfied by this tuple; the structure B is homogeneous if any two tuples of the same atomic type belong to the same orbit. For \(k \geq 1\), we say that a relational structure B is \(k\)-homogeneous if for all tuples \(a, b\) of arbitrary finite equal length, if all \(k\)-subtuples of \(a\) and \(b\) are in the same orbit under Aut(B), then \(a\) and \(b\) are in the same orbit under Aut(B). For \(\ell \geq 1\), we say that B is \(\ell\)-bounded if for every finite \(X\), if all substructures \(Y\) of \(X\) of size at most \(\ell\) embed in \(B\), then \(X\) embeds in \(B\). The structure B is universal for a class of finite structures in its signature if it embeds all members of the class. There exist an up to isomorphism unique countable universal homogeneous graph, partial order, tournament, and \(K_n\)-free graph (i.e., for any fixed \(n \geq 3\), a graph not containing any complete graph on \(n\)-vertices), respectively. We say that B is unary if all relations in its signature are unary. A first-order reduct of a structure B is a structure on the same domain whose relations have a first-order definition without parameters in B.

Example 2.1. The relational structure \((\mathbb{Q}; <)\) consisting of the set of rational numbers with the standard linear order has precisely 3 orbits of pairs, namely the orbits \{(a, b) ∈ \mathbb{Q}^2 | a < b\}, \{(a, b) ∈ \mathbb{Q}^2 | a > b\}, \{(a, b) ∈ \mathbb{Q}^2 | a = b\}; in the following, we will denote these orbits by <, >, and Eq, respectively, by abuse of notation. The structure Aut(\(\mathbb{Q}; <)\) is 2-homogeneous since for any \(k \geq 2\), the orbit of a \(k\)-tuple of elements of \(\mathbb{Q}\) is determined by the orbits of its subpairs. It immediately follows that \((\mathbb{Q}; <)\) is also homogeneous and that the automorphism group Aut(\(\mathbb{Q}; <)\) is oligomorphic. Moreover, the structure \((\mathbb{Q}; <)\) is universal for the class of all finite linear orders. Finally, \((\mathbb{Q}; <)\) is 3-bounded since in order to check whether a finite structure \(A = (A; <)\) embeds into \((\mathbb{Q}; <)\), it is enough to check that < is a linear order on A, i.e. that it is a total, antisymmetric, antireflexive, and transitive relation; these properties only depend on the substructures of A of size at most 3. An interesting example of a first-order reduct of \((\mathbb{Q}; <)\) is the structure \((\mathbb{Q}; \text{Betw})\), where Betw is the ternary betweenness relation defined by \{(a, b, c) ∈ \mathbb{Q}^3 | a < b < c or a > b > c\}.

A first-order formula is called primitive positive (pp) if it is built exclusively from atomic formulae, existential quantifiers, and conjunction. A relation is pp-definable in a structure B if it is first-order definable by a pp-formula.

Two relational structures A and B are homomorphically equivalent if there exist homomorphisms from A to B and from B to A. We remark that it is easy to see that homomorphically equivalent structures have the same relational width (to be defined later).
A relational structure $\mathcal{A}$ is Ramsey if for any finite substructures $\mathcal{B}, \mathcal{C}$ of $\mathcal{A}$, there exists a substructure $\mathcal{D}$ of $\mathcal{A}$ such that the following holds. For every mapping $\chi$ from the set of all embeddings of $\mathcal{B}$ into $\mathcal{D}$ to $[k]$, there exists an embedding $f$ of $\mathcal{C}$ into $\mathcal{D}$ such that for all embeddings $g_1, g_2$ of $\mathcal{B}$ into $\mathcal{C}$, $\chi(f \circ g_1) = \chi(f \circ g_2)$.

### 2.2. Polymorphisms, clones and identities.

A polymorphism of a relational structure $\mathcal{A}$ is a homomorphism from some finite power of $\mathcal{A}$ to $\mathcal{A}$. The set of all polymorphisms of a structure $\mathcal{A}$ is denoted by $\text{Pol}(\mathcal{A})$; it is a function clone, i.e., a set of finitary operations on a fixed set which contains all projections and which is closed under arbitrary compositions.

If $\mathcal{C}$ is a function clone, then we denote the domain of its functions by $C$, i.e., for every function $f \in \mathcal{C}$, there exists $k \geq 1$ such that $f \in C^{D^k}$; we say that $\mathcal{C}$ acts on $C$. The clone $\mathcal{C}$ also naturally acts (componentwise) on $C^l$ for any $l \geq 1$, and we write $\mathcal{C} \curvearrowright C_l$ for this action. We say that a relation $S \subseteq C^l$ is invariant under $\mathcal{C}$ if for every $f \in \mathcal{C}$ of arity $k \geq 1$, and for all $a_1, \ldots, a_k \in S$, it holds that $f(a_1, \ldots, a_k) \in S$; in this expression, we are referring to the action of $\mathcal{C}$ on $C^k$. If $S$ is invariant, $\mathcal{C}$ acts naturally on $S$ by restriction, and we write $\mathcal{C} \curvearrowright S$ for this action. Finally, if $\sim$ is an invariant equivalence relation on an invariant relation $S$, then $\mathcal{C}$ acts naturally on the classes of $\sim$ (by its action on representatives of the classes); we write $\mathcal{C} \curvearrowright S/\sim$. Any action $\mathcal{C} \curvearrowright S/\sim$ is called a subfactor of $\mathcal{C}$, and we also call the pair $(S, \sim)$ a subfactor. A subfactor $(S, \sim)$ is minimal if $\sim$ has at least two classes and no proper subset of $S$ intersecting at least two $\sim$-classes is invariant under $\mathcal{C}$. For a clone $\mathcal{C}$ acting on a set $X$ and $Y \subseteq X$ we write $(Y)_{\mathcal{C}}$ for the smallest $\mathcal{C}$-invariant subset of $X$ containing $Y$.

Let $n, k \geq 1$, let $A, B$ be sets, and let $\mathcal{G}$ and $\mathcal{G}'$ be permutation groups on $A$ and $B$, respectively. A $k$-ary function $f : A^k \to B$ is $n$-canonical from $\mathcal{G}$ to $\mathcal{G}'$ if for all $a_1, \ldots, a_k \in A^n$ and all $a_1, \ldots, a_k \in \mathcal{G}$ there exists $b \in \mathcal{G}'$ such that $f(a_1, \ldots, a_k) = b \circ f(a_1, a_2, \ldots, a_k)$ (in these expressions, we consider the componentwise action of $f$ on $A^n$ and of $\mathcal{G}, \mathcal{G}'$ on $A^n$ and $B^n$, respectively). We say that $f$ is canonical from $\mathcal{G}$ to $\mathcal{G}'$ if it is $n$-canonical from $\mathcal{G}$ to $\mathcal{G}'$ for every $n \geq 1$. We say that it is diagonally canonical if it satisfies the definition of canonicity in case $a_1 = \cdots = a_k$. If $A = B$, $\mathcal{G} = \mathcal{G}'$, and $f$ is $n$-canonical, canonical, or diagonally canonical from $\mathcal{G}$ to $\mathcal{G}$, we say that $f$ is $n$-canonical, canonical, or diagonally canonical, respectively, with respect to $\mathcal{G}$. A function canonical with respect to $\mathcal{G}$ is also called $\mathcal{G}$-canonical. If $f$ is $n$-canonical with respect to $\mathcal{G}$, it induces an operation on the set $C^n/\mathcal{G}$ of $\mathcal{G}$-orbits of $n$-tuples. If all functions of a function clone $\mathcal{C}$ are $n$-canonical with respect to $\mathcal{G}$, then $\mathcal{C}$ acts on $C^n/\mathcal{G}$ and we write $\mathcal{C}^n/\mathcal{G}$ for this action; if $\mathcal{G}$ is oligomorphic then $\mathcal{C}^n/\mathcal{G}$ is a function clone on a finite set.

**Example 2.1 (continued from p. 5).** The binary minimum operation $\text{min}$ on $\mathbb{Q}$ as well as any function $\text{lex} : \mathbb{Q}^2 \to \mathbb{Q}$ satisfying $\text{lex}(a, b) < \text{lex}(a', b')$ if, and only if, $a < a'$ or $a = a'$ and $b < b'$ are both polymorphisms of $(\mathbb{Q}; <)$. The operation $\text{lex}$ is canonical with respect to $\text{Aut}(\mathbb{Q}; <)$ — using the notation for orbits of pairs under $\text{Aut}(\mathbb{Q}; <)$ introduced above, it acts on these orbits as follows: $\text{lex}(<, <) = \text{lex}(<, \text{Eq}) = \text{lex}(<, >) = \text{lex}(\text{Eq}, <) = <, \text{lex}(>, <) = \text{lex}(>, \text{Eq}) = \text{lex}(>, >) = \text{lex}(\text{Eq}, \text{Eq}) = \text{Eq}$. On the other hand, $\text{min}$ is not canonical with respect to $\text{Aut}(\mathbb{Q}; <)$. Indeed, $0 < 1 < 2 < 3$ but $\text{min}(0, 1) = \text{min}(1, 0)$ and $\text{min}(0, 1) < \text{min}(3, 2)$, and hence the expression $\min(<, >)$ is not well-defined. However, $\min$ is diagonally canonical with respect to $\text{Aut}(\mathbb{Q}; <)$. Indeed, for any tuple $(a_1, \ldots, a_4) \in \mathbb{Q}^4$, the atomic type of $(\text{min}(a_1, a_2), \text{min}(a_3, a_4))$ depends only on the atomic type of $(a_1, \ldots, a_4)$; hence the homogeneity of $(\mathbb{Q}; <)$ implies that if $(b_1, \ldots, b_4) \in \mathbb{Q}^4$ lies in the same $\text{Aut}(\mathbb{Q}; <)$-orbit as $(a_1, \ldots, a_4)$, then the orbits of $(\text{min}(a_1, a_2), \text{min}(a_3, a_4))$ and $(\text{min}(b_1, b_2), \text{min}(b_3, b_4))$ coincide.

Note that for any family of permutation groups on the same set there is a smallest permutation group containing each member of the family, obtained by closing the union over the family under composition. Since any function clone $\mathcal{C}$ is closed under composition, application of this fact to the family of all permutation groups which are contained (as a subset of the unary functions of $\mathcal{C}$) in
\( C \) yields the largest permutation group contained in \( C \) which we denote by \( \mathcal{G}_\varphi \). We say that \( C \) is oligomorphic if \( \mathcal{G}_\varphi \) is oligomorphic. For \( n \geq 1 \), the \( n \)-canonical (canonical) part of \( C \) is the clone of those functions of \( C \) which are \( n \)-canonical (canonical) with respect to \( \mathcal{G}_\varphi \). We write \( C_\text{can} \) for these sets, which form themselves function clones. In particular, we use this notation for polymorphism clones, writing \( \text{Pol}(\mathcal{A})_\text{can} \) and \( \text{Pol}(\mathcal{A})_\text{can} \).

For a set of finitary functions \( \mathcal{F} \) over the same fixed set \( C \) we write \( \overline{\mathcal{F}} \) for the set of those functions \( g \) such that for all finite subsets \( F \) of \( C \), there exists a function in \( \mathcal{F} \) which agrees with \( g \) on \( F \). We say that \( f \) locally interpolates \( g \) modulo \( \mathcal{G} \), where \( f, g \) are \( k \)-ary functions and \( \mathcal{G} \) is a permutation group all of which act on the same domain, if \( g \in \{ \beta \circ f(\alpha_1, \ldots, \alpha_k) \mid \beta, \alpha_1, \ldots, \alpha_k \in \mathcal{G} \} \). Similarly, we say that \( f \) diagonally interpolates \( g \) modulo \( \mathcal{G} \) if \( f \) locally interpolates \( g \) modulo \( \mathcal{G} \) with \( \alpha_1 = \cdots = \alpha_k \). If \( \mathcal{G} \) is the automorphism group of a Ramsey structure in the sense of [14], then every function on its domain locally (diagonally interpolates) a canonical (diagonally canonical) function modulo \( \mathcal{G} \) [33] [31].

**Theorem 2.2** (Consequence of Theorem 5 from [31]). Let \( \mathcal{A} \) be a countable homogeneous Ramsey structure, let \( \mathcal{B} \) be an \( \omega \)-categorical structure, and let \( h : A \to B \) be an arbitrary function. Then \( h \) locally interpolates modulo \( \text{Aut}(\mathcal{A}) \) a function that is canonical from \( \text{Aut}(\mathcal{A}) \) to \( \text{Aut}(\mathcal{B}) \), and it diagonally interpolates modulo \( \text{Aut}(\mathcal{A}) \) a function that is diagonally canonical from \( \text{Aut}(\mathcal{A}) \) to \( \text{Aut}(\mathcal{B}) \).

In particular, if \( f : A^n \to A \) is a polymorphism of a structure \( \mathcal{A} \) with domain \( A \), and \( \mathcal{A} \) is a Ramsey structure, then \( f \) (diagonally) locally interpolates modulo \( \text{Aut}(\mathcal{A}) \) a polymorphism \( g \) of \( \mathcal{A} \) that is (diagonally) canonical with respect to \( \text{Aut}(\mathcal{A}) \). We say that a clone \( \mathcal{D} \) locally interpolates a clone \( C \) modulo a permutation group \( \mathcal{G} \) if for every \( g \in \mathcal{D} \) there exists \( f \in C \) such that \( g \) locally interpolates \( f \) modulo \( \mathcal{G} \).

A clone \( C \) is a model-complete core if its unary functions are equal to \( \overline{\mathcal{G}}_\varphi \), and a structure \( \mathcal{A} \) is a model-complete core if its polymorphism clone is. A finite model-complete core is called a core. For every \( \omega \)-categorical structure \( \mathcal{A} \) there exists an up to isomorphism unique structure that is a model-complete core and that is homomorphically equivalent to \( \mathcal{A} \) by [14]: we call this structure the model-complete core of \( \mathcal{A} \). Since, as remarked earlier, the relational width of a structure is invariant under homomorphic equivalence, one can in this context often work without loss of generality with model-complete cores. In particular, one can then use the following property which distinguishes model-complete cores from other structures: Any \( \omega \)-categorical model-complete core pp-defines all orbits of \( n \)-tuples with respect to its own automorphism group, for all \( n \geq 1 \) [15]. Since for any \( \omega \)-categorical structure \( \mathcal{B} \), pp-definability of a relation is equivalent to its invariance under \( \text{Pol}(\mathcal{B}) \) [28], in the case of model-complete cores we get that all orbits of its automorphism group are invariant under all polymorphisms.

An identity is an equation of terms built from some functional symbols where all variables are implicitly interpreted as being universally quantified. For example, if \( f \) is a \( k \)-ary function symbol for some \( k \geq 1 \), then \( f(x, \ldots, x) = x \) is an identity, and a \( k \)-ary function on some domain \( C \) satisfies the identity \( f(x, \ldots, x) = x \) if this equation holds for all elements \( x \) of \( C \). Functions satisfying this particular identity of the appropriate arity are called idempotent; a function clone is idempotent if all of its functions are. A \( k \)-ary operation \( w \) is called a weak near-unanimity (WNU) operation if it satisfies the set of identities containing an equation for each pair of terms in \( \{ w(x, \ldots, x, y), w(x, y, \ldots, x), \ldots, w(y, x, \ldots, x) \} \). It is called totally symmetric if \( w(x_1, \ldots, x_k) = w(y_1, \ldots, y_k) \) whenever \( \{ x_1, \ldots, x_k \} = \{ y_1, \ldots, y_k \} \). Each set of identities also has a pseudo-variant obtained by composing each term appearing in the identities with a distinct new unary function symbol. For example, a ternary pseudo-WNU operation \( f \) satisfies the identities: \( e_1 \circ f(y, x, x) = e_0 \circ f(x, y, x) \), \( e_3 \circ f(y, x, x) = e_4 \circ f(x, x, y) \) and \( e_5 \circ f(y, x, x) = e_6 \circ f(x, y, x) \). A function clone \( C \) satisfies a given set of identities if the function symbols which appear in the identities can be assigned functions in \( C \) in such a way that all identities in the set are satisfied by the assigned
functions; if \( \mathcal{U} \subseteq \mathcal{C} \) is a set of unary functions, then \( \mathcal{C} \) **satisfies a set of pseudo-identities modulo \( \mathcal{U} \)** if it satisfies the identities in such a way that the new unary function symbols which were added to the original identities to obtain the pseudo-variant are assigned values in \( \mathcal{U} \). For an example for the satisfaction of identities in a polymorphism clone, note that the clone of polymorphisms of \( \mathbb{Q}; < \) satisfies the \( n \)-ary WNU identities for every \( n \geq 2 \), as witnessed by the \( n \)-ary operation \( \min(x_1, \min(\ldots, \min(x_{n-1}, x_n))) \). The function clone of all finitary operations on a countable set which are injective or arise from an injection by the addition of dummy variables is easily seen to be the polymorphism clone of a suitable structure (see e.g. [16] for such a structure); it is not hard to see that it satisfies the pseudo-identity \( e_1 \circ f(x, y) = e_2 \circ f(y, x) \) modulo the set of unary injections, but that it does not contain any commutative operation.

An arity-preserving map \( \xi: \mathcal{C} \to \mathcal{D} \) between function clones is called a **clone homomorphism** if it preserves projections, i.e., it maps every projection in \( \mathcal{C} \) to the corresponding projection in \( \mathcal{D} \), and compositions, i.e., it satisfies \( \xi(f \circ (g_1, \ldots, g_n)) = \xi(f) \circ (\xi(g_1), \ldots, \xi(g_n)) \) for all \( n, m \geq 1 \) and all \( n \)-ary \( f \in \mathcal{C} \) and \( m \)-ary \( g_1, \ldots, g_n \in \mathcal{C} \). An arity-preserving map \( \xi \) is a **minion homomorphism** if \( \xi(f \circ (\pi_{i_1}, \ldots, \pi_{i_n})) = \xi(f) \circ (\pi_{i_1}, \ldots, \pi_{i_n}) \) for all \( n, m \geq 1 \) and all \( n \)-ary \( f \in \mathcal{C} \) and \( m \)-ary projections \( \pi_{i_1}, \ldots, \pi_{i_n} \). We say that a function clone \( \mathcal{C} \) is **equationally trivial** if it has a clone homomorphism to the clone \( \mathcal{P} \) of projections over the two-element domain, and **equationally non-trivial** otherwise.

We also say that \( \mathcal{C} \) is **equationally affine** if it has a clone homomorphism to an **affine clone**, i.e., a clone of affine maps over a finite module. It is known that an idempotent function clone on a finite domain is either equationally affine, or it contains WNU operations of all arities \( n \geq 3 \) ([53] shows a slightly weaker statement; an argument proving the precise statement made here is attributed to E. Kiss in [51, Theorem 2.8]; a more recent proof of this statement can be found in [65]). Similarly, if \( \mathcal{A} \) is an \( \omega \)-categorical model-complete core, then \( \text{Pol}(\mathcal{A})^{\text{can}} \) is either equationally affine, or it contains pseudo-WNU operations modulo \( \text{Aut}(\mathcal{A}) \) of all arities \( n \geq 3 \) (see [32, 55] for the lift of the corresponding result from the finite).

If \( \mathcal{C}, \mathcal{D} \) are function clones and \( \mathcal{D} \) has a finite domain, then a clone or minion homomorphism \( \xi: \mathcal{C} \to \mathcal{D} \) is **uniformly continuous** if for all \( n \geq 1 \) there exists a finite subset \( F \) of \( \mathbb{C}^n \) such that \( \xi(f) = \xi(g) \) for all \( n \)-ary \( f, g \in \mathcal{C} \) which agree on \( F \). If \( \mathcal{C} \) is finite, then clearly every clone or minion homomorphism \( \mathcal{C} \to \mathcal{D} \) is uniformly continuous.

### 2.3. CSP, Relational Width, Minimality

A **CSP instance** over a set \( A \) is a pair \( \mathcal{I} = (\mathcal{V}, \mathcal{C}) \), where \( \mathcal{V} \) is a finite set of variables, and \( \mathcal{C} \) is a finite set of **constraints**; each constraint \( C \in \mathcal{C} \) is a subset of \( A^U \) for some non-empty \( U \subseteq \mathcal{V} \); the set \( U \) is called the **scope** of \( C \). A map \( f: \mathcal{V} \to A \) is a **solution** of a CSP instance \( \mathcal{I} = (\mathcal{V}, \mathcal{C}) \) if for every constraint \( C \in \mathcal{C} \) with scope \( U \subseteq \mathcal{V} \), it holds that \( f|_U \in C \). If \( \mathcal{A} \) is a relational structure with domain \( A \), then we say that \( \mathcal{I} \) is an **instance of CSP(\( \mathcal{A} \))** if every constraint \( C \in \mathcal{C} \) can be viewed as a relation pp-definable in \( \mathcal{A} \) by totally ordering its scope \( U \); more precisely, there exists an enumeration \( u_1, \ldots, u_k \) of the elements of \( U \) and a \( k \)-ary relation \( R \) that is pp-definable in \( \mathcal{A} \) such that for all \( f: U \to A \) we have \( f \in C \iff \langle f(u_1), \ldots, f(u_k) \rangle \in R \). A CSP instance is **non-trivial** if it does not contain any empty constraint; otherwise, it is **trivial**. Given a constraint \( C \subseteq A^U \) and \( K \subseteq U \), the **projection of \( C \) onto \( K \)** is defined by \( C|_K := \{ f|_K : f \in C \} \). We say that a constraint \( C \) of a CSP instance over a set \( A \) is **preserved** by a function \( f: A^n \to A \) if for all \( g_1, \ldots, g_n \in C \) we have \( f(g_1, \ldots, g_n) \in C \).

**Definition 2.3.** Let \( 1 \leq m \leq n \). We say that an instance \( \mathcal{I} \) of CSP(\( \mathcal{A} \)) with variables \( \mathcal{V} \) is \((m, n)\)-**minimal** if both of the following hold:

- every at most \( n \)-element subset of the variable set \( \mathcal{V} \) is contained in the scope of some constraint in \( \mathcal{I} \);
- for every at most \( m \)-element subset of variables \( K \subseteq \mathcal{V} \) and any two constraints \( C_1, C_2 \in \mathcal{I} \) whose scopes contain \( K \), the projections of \( C_1 \) and \( C_2 \) onto \( K \) coincide.

We say that an instance \( \mathcal{I} \) is **\( m \)-minimal** if it is \((m, m)\)-minimal.
Example 2.4. Let us consider the following instance $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ of CSP($\mathbb{Q}; <$). The set $\mathcal{V}$ of variables consists of elements $v_1, v_2, v_3$, and the set $\mathcal{C}$ consists of the constraints $C_{(1, 2)}, C_{(2, 3)}, C_{(3, 1)}$, where $C_{i,j} := \{ f \in \mathbb{Q}^{v_i,v_j} \mid f(v_i) < f(v_j) \}$ for every $(i, j) \in \{(1, 2), (2, 3), (3, 1)\}$. The instance $\mathcal{I}$ has clearly no solution since every assignment $g \in \mathbb{Q}^{v_1,v_2,v_3}$ with the property that $g(v_1) < g(v_2)$ satisfies both $g(v_1) < g(v_2)$ since $g(v_2, v_3) \in C_{(1, 2)}$, and $g(v_1) > g(v_2)$ (by the transitivity of $<$). The instance $\mathcal{I}$ is $(2, 2)$-minimal since every pair of variables is contained in the scope of some constraint, and $C_{i,j} \forall v_i = \mathbb{Q}^{v_i}$, $C_{i,j} \forall v_j = \mathbb{Q}^{v_j}$ for every $(i, j) \in \{(1, 2), (2, 3), (3, 1)\}$.

Let $1 \leq m \leq n$, and let $\mathbb{A}$ be a relational structure. Clearly not every instance $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ of CSP($\mathbb{A}$) is $(m, n)$-minimal. However, every instance $\mathcal{I}$ is equivalent to an $(m, n)$-minimal instance $\mathcal{I}'$ in the sense that $\mathcal{I}$ and $\mathcal{I}'$ have the same set of solutions. In particular we have that if $\mathcal{I}'$ is trivial, then $\mathcal{I}$ has no solutions. Moreover, if $\mathbb{A}$ is $\omega$-categorical, then the instance $\mathcal{I}'$ can be computed in only polynomially many steps in the number of variables and constraints as follows. First introduce a new constraint $A^L$ for every set $L \subseteq \mathcal{V}$ with at most $n$ elements to satisfy the first condition. Then remove orbits with respect to $\text{Aut} (\mathbb{A})$ from the constraints in the instance as long as the second condition is not satisfied.

To see that $\mathcal{I}'$ is an instance of CSP($\mathbb{A}$), observe that the relation $A^L$ is pp-definable in $\mathbb{A}$ for every $L \subseteq \mathcal{V}$ and hence, the instance created from $\mathcal{I}$ by adding the new constraints is still an instance of CSP($\mathbb{A}$). Moreover, when the algorithm removes orbits from some constraint, then the new constraint is obtained by a pp-definition from two old constraints and hence, it is pp-definable in $\mathbb{A}$. Whence, the instance $\mathcal{I}'$ is an instance of CSP($\mathbb{A}$).

The algorithm just described introduces at most $\sum_{i=1}^{n} \binom{|\mathcal{V}|}{i}$ new constraints. Let $p$ denote the maximum of $n$ and the maximal arity of the relations of $\mathbb{A}$. Then the algorithm uses only relations of arity at most $p$. Note that by the $\omega$-categoricity of $\mathbb{A}$, every relation pp-definable in $\mathbb{A}$ is a finite union of orbits of tuples with respect to $\text{Aut} (\mathbb{A})$. Hence, since in every step the algorithm removes at least one orbit from some constraint and since the number of orbits that can be removed is polynomial in $|\mathcal{V}| + |\mathcal{C}|$ by the discussion above, the number of steps of the algorithm is polynomial in $|\mathcal{V}| + |\mathcal{C}|$.

Definition 2.5. Let $1 \leq m \leq n$. A relational structure $\mathbb{A}$ has relational width $(m, n)$ if every non-trivial $(m, n)$-minimal instance of $\mathbb{A}$ has a solution. $\mathbb{A}$ has bounded width if it has relational width $(m, n)$ for some $m, n$.

Example 2.4 (continuing from p. 10). Let us consider the instance $\mathcal{I}$ from Example 2.4. The algorithm above which produces a $(2, 3)$-minimal instance $\mathcal{I}'$ equivalent to the instance $\mathcal{I}$ proceeds as follows. It first introduces a new constraint $C_{(1,2,3)} := \mathbb{Q}^{v_1}$ so as to satisfy the first item of Definition 2.3. In order to satisfy the second item, the algorithm compares projections of constraints to subsets of $\mathcal{V}$ containing at most 2 elements. By comparing the projections of $C_{(1,2,3)}$ and of $C_{(1,2)}$ to $\{v_1, v_2\}$, it removes from $C_{(1,2,3)}$ all maps $f$ which do not satisfy $f(v_1) < f(v_2)$. By considering the projections of $C_{(1,2,3)}$ and $C_{(2,3)}$ to $\{v_2, v_3\}$ and of $C_{(1,2,3)}$ and $C_{(3,1)}$ to $\{v_3, v_1\}$, it finally removes all maps from $C_{(1,2,3)}$. Now, comparing the projections of $C_{(1,2,3)}$ and of $C_{(i,j)}$ for every $(i, j) \in \{(1, 2), (2, 3), (3, 1)\}$, the algorithm removes also all maps from $C_{(1,2)}$, $C_{(2,3)}$, and from $C_{(3,1)}$. The resulting instance $\mathcal{I}'$ is then $(2, 3)$-minimal but it contains only empty constraints, and is therefore in particular trivial. In fact, it can be shown that $(\mathbb{Q}; <)$ has relational width precisely $(2, 3)$.

The following statement gives an overview of conditions characterizing relational structures with bounded width that are used in several proofs later. Let $p \geq 2$ be a prime number and let $R_0$ and $R_1$ be the relations defined by $\{(x, y, z) \in \mathbb{Z}_p \mid x + y + z = i \text{ mod } p\}$ for $i \in \{0, 1\}$. We say that a structure $\mathbb{B}$ has a pp-construction in a structure $\mathbb{A}$ if $\mathbb{B}$ is homomorphically equivalent to a
structure with domain $A^n$, for some $n \geq 1$, whose relations are pp-definable in $\mathbb{A}$ (for this purpose, a $k$-ary relation on $A^n$ is regarded as a $kn$-ary relation on $A$).

**Theorem 2.6.** Let $\mathbb{A}$ be an $\omega$-categorical relational structure. All the implications $(i) \Rightarrow (j)$ and $(i) \iff (i')$ in the list below hold for $1 \leq i \leq j \leq 3$.

- $(1)$ $\mathbb{A}$ has relational width at most $(2,3)$.
- $(2)$ $\mathbb{A}$ has bounded width.
- $(2')$ The class of finite structures that do not have a homomorphism to $\mathbb{A}$ is definable by a Datalog program.
- $(3)$ $\text{Pol}(\mathbb{A})$ does not admit a uniformly continuous minion homomorphism to an affine clone.
- $(3')$ $\mathbb{A}$ does not pp-construct $(\mathbb{Z}_p; R_0, R_1)$ for any prime number $p$.

If $\mathbb{A}$ is finite, then all these conditions are equivalent, and moreover equivalent to the following two statements.

- $(4)$ The expansion $\mathbb{B}$ of the core of $\mathbb{A}$ by all unary singleton relations is such that $\text{Pol}(\mathbb{B})$ is not equationally affine.
- $(5)$ $\text{Pol}(\mathbb{A})$ contains WNU operations of all arities $n \geq 3$.

The implication $(1)$ to $(2)$ is trivial, the equivalence of $(2)$ and $(2')$ follows from [42] Theorem 23 and [55] Corollary 1. $(2)$ implies $(3)$ by [52] and [9], and the equivalence of $(3)$ and $(3')$ is a folklore consequence of [9]. For finite structures, the implication from $(4)$ to $(1)$ was proven in [4], (3) implies $(4)$ by [9], and (4) is equivalent to (5) by the results from [53, 51] (in fact, a slightly weaker statement is formulated there; the precise statement made here is attributed to E. Kiss in [51, Theorem 2.8], and another proof can be found in [65]).

**Theorem 2.7 (H1, 42).** Let $\mathbb{A}$ be a finite relational structure. $\mathbb{A}$ has relational width $(1,1)$ if, and only if, $\text{Pol}(\mathbb{A})$ contains totally symmetric operations of all arities.

2.4. **Smooth Approximations.** We are going to apply the fundamental theorem of smooth approximations [55] to lift an action of a function clone to a larger clone.

**Definition 2.8.** (Smooth approximations) Let $A$ be a set, $n \geq 1$, and let $\sim$ be an equivalence relation on a subset $S$ of $A^n$. We say that an equivalence relation $\eta$ on some set $S'$ with $S \subseteq S'$ **approximates** $\sim$ if the restriction of $\eta$ to $S$ is a (possibly non-proper) refinement of $\sim$. We call $\eta$ an **approximation** of $\sim$.

For a permutation group $\mathcal{G}$ acting on $A$ and leaving $\eta$ as well as each $\sim$-class invariant, we say that the approximation $\eta$ is **smooth** if each equivalence class $C$ of $\sim$ intersects some equivalence class $C'$ of $\eta$ such that $C \cap C'$ contains a $\mathcal{G}$-orbit.

**Theorem 2.9** (The fundamental theorem of smooth approximations [55]). Let $\mathcal{C} \subseteq \mathcal{D}$ be function clones on a set $A$, and let $\mathcal{G}$ be a permutation group on $A$ such that $\mathcal{D}$ locally interpolates $\mathcal{C}$ modulo $\mathcal{G}$. Let $\sim$ be a $\mathcal{C}$-invariant equivalence relation on $S \subseteq A$ with $\mathcal{G}$-invariant classes and finite index, and $\eta$ be a $\mathcal{D}$-invariant smooth approximation of $\sim$ with respect to $\mathcal{G}$. Then there exists a uniformly continuous minion homomorphism from $\mathcal{D}$ to $\mathcal{C} \rtimes S/\sim$.

3. **Collapses in the Relational Width Hierarchy**

We now introduce a construction for reducing CSPs over structures in the scope of Theorem [1.1] to CSPs over finite sets. The idea is to create from an instance another instance whose variables are subsets of variables which take orbits as values, of which there is only a finite number thanks to oligomorphicity. This will enable us to prove Theorem [1.1].

**Definition 3.1.** Let $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ be a CSP instance over a set $A$. Let $\mathcal{G}$ be a permutation group on $A$, let $k \geq 1$, and for every $K \in \binom{V}{k}$, let $O^k_{\mathcal{G}}$ be the set of $K$-orbits, i.e., orbits of maps $f: K \to A$
under the natural action of $\mathcal{G}$. Let $\mathcal{O}_{\mathcal{G},k}^\mathcal{V} := \bigcup_{K \in \mathcal{V}} \mathcal{O}_{\mathcal{G}}^K$ and let $\mathcal{I}_{\mathcal{G},k}$ be the following instance over $\mathcal{O}_{\mathcal{G},k}^\mathcal{V}$:

$\mathcal{O}_{\mathcal{G},k}^\mathcal{V}$:

- The variable set of $\mathcal{I}_{\mathcal{G},k}$ is the set $\binom{\mathcal{V}}{k}$ of $k$-element subsets of $\mathcal{V}$. Every variable $K$ of $\mathcal{I}_{\mathcal{G},k}$ is meant to take a value in $\mathcal{O}_{\mathcal{G}}^K$.
- For every constraint $C \subseteq A^U$ in $\mathcal{I}$, $\mathcal{I}_{\mathcal{G},k}$ contains the constraint $C_{\mathcal{G},k} \subseteq (\mathcal{O}_{\mathcal{G},k}^\mathcal{V})^{\binom{\mathcal{V}}{k}}$ defined by

$$C_{\mathcal{G},k} := \left\{ g: \left( \binom{U}{k} \right) \to \mathcal{O}_{\mathcal{G},k}^\mathcal{V} \mid \exists f \in C \; \forall K \in \left( \binom{U}{k} \right) \left( f \mid K \in g(K) \right) \right\}.$$ 

That is, the set of $k$-element subsets of $U$ is the scope of a constraint which allows precisely those assignments of orbits to these subsets which are naturally induced by the assignments allowed by $C$ for the variables in $U$. Note that for every $K \in \binom{\mathcal{V}}{k}$, for every constraint $C$ whose scope contains $K$ and for every $g \in C_{\mathcal{G},k}$ we have $g(K) \in \mathcal{O}_{\mathcal{G}}^K$.

Observe that if $\mathcal{I}$ is non-trivial, then so is $\mathcal{I}_{\mathcal{G},k}$.

**Lemma 3.2.** Let $1 \leq a \leq b$. If $\mathcal{I}$ is $(ak,bk)$-minimal, then $\mathcal{I}_{\mathcal{G},k}$ is $(a,b)$-minimal.

**Proof.** Let $b' \leq b$, and let $K_1, \ldots, K_{b'} \in \binom{\mathcal{V}}{b'}$. Note that $U := \bigcup K_i$ has size at most $bk$, and therefore there exists a set $W$ with $U \subseteq W \subseteq \mathcal{V}$ and a constraint $C \subseteq A^W$ in $\mathcal{I}$ since $\mathcal{I}$ is $(ak,bk)$-minimal. The scope of the associated constraint $C_{\mathcal{G},k}$ is $\binom{W}{k}$, which contains $K_1, \ldots, K_{b'}$. Hence, the first item of Definition 2.3 is satisfied.

For the second item, let $a' \leq a$, let $K := \{K_1, \ldots, K_{a'}\}$, and let $C_{\mathcal{G},k} \subseteq (\mathcal{O}_{\mathcal{G},k}^\mathcal{V})^{\binom{\mathcal{V}}{k}}$, $D_{\mathcal{G},k} \subseteq (\mathcal{O}_{\mathcal{G},k}^\mathcal{V})^{\binom{\mathcal{V}}{k}}$ be two constraints whose scopes contain $K$. Then $\bigcup K$ is contained in the scope of the associated $C \subseteq A^U$ and $D \subseteq A^{W'}$ and has size at most $ak$, so that by $(ak,bk)$-minimality of $\mathcal{I}$, the projections of $C$ and $D$ onto $\bigcup K$ coincide. Thus for every $g \in C_{\mathcal{G},k}$, there exists by definition an $f \in C$ such that $f\mid K_i \in g(K_i)$ for all $i$, and by the previous sentence there exists $f' \in D$ such that $f'\mid K_i \in g(K_i)$ for all $i$. Thus, $g\mid K$ is in the projection of $D_{\mathcal{G},k}$ to $K$. The argument is symmetric, showing that the projections of $C_{\mathcal{G},k}$ and $D_{\mathcal{G},k}$ to $K$ coincide. \hfill \Box

Note that for every solution $h$ of $\mathcal{I}$, the map $\chi_h: \binom{\mathcal{V}}{k} \to \mathcal{O}_{\mathcal{G},k}^\mathcal{V}$ defined by $K \mapsto \{ah\mid K \mid a \in \mathcal{G}\}$ defines a solution to $\mathcal{I}_{\mathcal{G},k}$. The next lemma proves that every solution to $\mathcal{I}_{\mathcal{G},k}$ is of the form $\chi_h$ for some solution $h$ of $\mathcal{I}$, provided that $\mathcal{I}$ is $(k,\ell)$-minimal and that $\mathcal{G} = \text{Aut}(\mathbb{B})$ for some $k$-homogeneous $\ell$-bounded structure $\mathbb{B}$.

**Lemma 3.3.** Let $1 \leq k < \ell$ if $k > 1$, and if $k = 1$, suppose that $\ell \geq 1$. Let $\mathbb{B}$ be a $k$-homogeneous $\ell$-bounded structure, let $\mathcal{A}$ be a first-order reduct of $\mathbb{B}$, and let $\mathcal{I}$ be a $(k,\ell)$-minimal instance of CSP($\mathcal{A}$). Then every solution to $\mathcal{I}_{\text{Aut}(\mathbb{B}),k}$ lifts to a solution of $\mathcal{I}$, i.e., it is of the form $\chi_h$ for some solution $h$ of $\mathcal{I}$.

**Proof.** Let $h: \binom{\mathcal{V}}{k} \to \mathcal{O}_{\mathcal{G},k}^\mathcal{V}$ be a solution to $\mathcal{I}_{\text{Aut}(\mathbb{B}),k}$. Recall that $h(K)$ is a $K$-orbit for any $K \in \binom{\mathcal{V}}{k}$, and one can therefore restrict $h(K)$ to any $L \subseteq K$ by setting $h(K)|_L := \{f|_L \mid f \in h(K)\}$. Note that since $\mathcal{I}$ is $k$-minimal, we have $h(K)|_{K \cap K'} = h(K')|_{K \cap K'}$ for all $K, K' \in \binom{\mathcal{V}}{k}$.

We now define an equivalence relation $\sim$ on $\mathcal{V}$. Suppose first that $k = 1$. Then every orbit of elements of $\mathbb{B}$ under the action of $\text{Aut}(\mathbb{B})$ must be a singleton (for any orbit with two distinct elements $a, b$, the pairs $(a, a)$ and $(a, b)$ would not be in the same orbit but their subtupes of length one would be, so that $\mathbb{B}$ would not be 1-homogeneous). In that case, we identify $\mathcal{O}_{\mathcal{G},k}^\mathcal{V}$ with the domain $B$ itself, and set $x \sim y$ if and only if $h(\{x\}) = h(\{y\})$; that is, $\sim$ is essentially the kernel of $h$.

Suppose next that $k \geq 2$, and set $x \sim y$ if and only if there is $K \in \binom{\mathcal{V}}{k}$ containing $x, y$ such that $h(K)|_{\{x,y\}}$ consists of constant maps. One could equivalently ask that this holds for all $K$...
containing $x, y$ by 2-minimality, and it then follows that this is indeed an equivalence relation by $(2, 3)$-minimality of $I$. Moreover, $h$ descends to $(V/\sim)_{K}$: if $K' = \{v_1, \ldots, v_k\}$ is a $k$-element set, define $\tilde{h}(K') := h(\{v_1, \ldots, v_k\})$. The definition of $h$ does not depend on the choice of representatives, by the very definition of $\sim$.

Define a finite structure $\mathbb{C}$ with domain $V/\sim$ in the signature of $\mathbb{B}$ as follows. Let $\mathbb{V}/\sim = \{[v_1], \ldots, [v_n]\}$. We define $\mathbb{C}$ such that the relations holding on the tuple $([v_1], \ldots, [v_n])$ in $\mathbb{C}$ are the same as the relations holding on an arbitrary tuple $(b_1, \ldots, b_n) \in \mathbb{B}^n$ that satisfies the following. For every $m \leq k$ and for all $[v_1], \ldots, [v_m] \sim$ pairwise different, the atomic types of $(b_1, \ldots, b_m)$ and $(g([v_1]), \ldots, g([v_m]))$ agree for all $g \in h(K)$ and $K \supseteq \{v_1, \ldots, v_m\}$. This construction does not depend on the choice of the tuple $(b_1, \ldots, b_n)$ by the $k$-homogeneity of $\mathbb{B}$ and is well-defined by the consistency of the assignment given by $h$.

Finally, note that all substructures of $\mathbb{C}$ of size at most $\ell$ embed into $\mathbb{B}$. Indeed, let $m \leq \ell$ and let $\mathbb{L}$ be an $m$-element substructure of $\mathbb{C}$, and let $L' \subseteq V$ be an $m$-element set containing one representative for each element of $\mathbb{L}$. By $(k, \ell)$-minimality of $\mathbb{I}$, there exists $C \subseteq A^\ell$ in $\mathbb{I}$, and a corresponding constraint $C_{\text{Aut}(\mathbb{B}), k}$ of $\mathbb{I}_{\text{Aut}(\mathbb{B}), k}$. Thus, $h|[L'] \in C_{\text{Aut}(\mathbb{B}), k}$, so that there exists $g \in C$ such that for all $K \in (V/\sim), g|K \in h(K)$. Thus $g$ corresponds to an embedding of every $k$-element substructure of $\mathbb{L}$ into $\mathbb{B}$, and since $\mathbb{B}$ is $k$-homogeneous, $g$ is an embedding of $\mathbb{L}$ into $\mathbb{B}$. Finally, since $\mathbb{B}$ is $\ell$-bounded, it follows that there exists an embedding $e$ of $\mathbb{C}$ into $\mathbb{B}$.

It remains to check that the composition of $e$ with the canonical projection $\mathbb{V} \rightarrow \mathbb{V}/\sim$ is a solution to $\mathbb{I}$, which is trivial since the relations of $\mathcal{A}$ are unions of orbits under $\text{Aut}(\mathbb{B})$. 

For every finite set $\mathbb{V}$ and for every $K \in (V/\sim)$, every operation $f$ that is canonical with respect to a permutation group $\mathcal{G}$ induces an operation on the set of orbits of $K$-tuples under $\mathcal{G}$. We denote this operation by $f^K\mathcal{G}$. Finally, we denote by $f^K\mathcal{G}, f^K$ the union of $f^K\mathcal{G}$ for all $K \in (V/\sim)$ and we call it a multisorted operation, i.e., this operation is defined only on tuples where all elements belong to the same $O^K$ for some fixed $K \in (V/\sim)$. We say that a multisorted operation $f^K\mathcal{G}$ is a multisorted WNU if $f^K\mathcal{G}$ satisfies the WNU identities for every $K \in (V/\sim)$. Similarly, we define multisorted totally symmetric operations. We remark that we could have avoided the use of multisorted functions by using $K$-tuples of elements of $\mathbb{V}$ instead of $K$-element subsets of $\mathbb{V}$ in Definition [14], however, encoding the original instance this way would have added considerable redundancy and made the proof more technical.

Note that in the above situation, for every constraint $C$ of an instance $\mathcal{I} = (\mathcal{V}, C)$ it makes formally sense to ask whether $f^K\mathcal{G}, f^K$ preserves $C_{\mathcal{G}, k}$. Indeed, let $f$ be $n$-ary. Whenever $K$ is a $k$-element subset of the scope of $C$, and hence a variable in the scope of $C_{\mathcal{G}, k}$, then for all $g_1, \ldots, g_n \in C_{\mathcal{G}, k}$ we have $g_1(K), \ldots, g_n(K) \in O^K$; hence $f^K\mathcal{G}$ can be applied to these values, and doing so for all variables in the scope of $C_{\mathcal{G}, k}$ altogether yields a function from this scope to $O^K$ which can be an element of $C_{\mathcal{G}, k}$ or not.

**Lemma 3.4.** Let $f$ be a polymorphism of $\mathcal{A}$ that is canonical with respect to $\mathcal{G}$. Every constraint in $\mathcal{I}_{\mathcal{G}, k}$ is preserved by $f^K\mathcal{G}, f^K$.

**Proof.** Let $n$ be the arity of $f$ and let $C \subseteq A^\ell$ be a constraint in $\mathbb{I}$. In particular, since $\mathcal{I}$ is an instance of CSP$(\mathcal{A})$, and $f$ is a polymorphism of $\mathcal{A}$, we have that $C$ is preserved by $f$.

Let $C_{\mathcal{G}, k}$ be the constraint in $\mathcal{I}_{\mathcal{G}, k}$ given by $C$, and let $g_1, \ldots, g_n \in C_{\mathcal{G}, k}$. By definition, for every $i \in \{1, \ldots, n\}$ there is $g_i' \in C$ such that for all $K \in (V/\sim)$, $g_i'|K \in \mathbb{G}(K)$. Note that $f(g_1'|K, \ldots, g_n'|K) = f(g_1', \ldots, g_n'|K)$, so that $f(g_1', \ldots, g_n'|K) \in f^K\mathcal{G}(g_1(K), \ldots, g_n(K))$. Since $f(g_1', \ldots, g_n') \in C$, it follows that $f^K\mathcal{G}(g_1, \ldots, g_n)$ is in $C_{\mathcal{G}, k}$. 

\[ \square \]
Let Theorem 1.1. The proof is identical to the proof of Corollary 3.5, using Theorem 2.7 instead of Theorem 2.6.

Corollary 3.5. Suppose that $|V| \geq k$, and that $\mathcal{A}$ has for every $m \geq 3$ a polymorphism $f_m$ that is canonical with respect to $\mathcal{G}$ and such that $(f_m)^{\mathcal{G},k}_{\mathcal{G}}$ is a multisorted WNU operation. If $\mathcal{I}_{\mathcal{G},k}$ is non-trivial and $(2,3)$-minimal, then it has a solution.

Proof. Let $\mathcal{D}$ be the finite structure with domain $O^V_{\mathcal{G},k}$ which contains for every constraint $C_{\mathcal{G},k}$ in $\mathcal{I}_{\mathcal{G},k}$ the corresponding relation. More precisely, for every constraint $C_{\mathcal{G},k}$ with $C \subseteq A^V$ for some $U \subseteq V$, $\mathcal{D}$ contains the relation $R_C := \{(g(K))_{K \in L^U_k} \mid g \in C_{\mathcal{G},k}\}$.

Let $m \geq 3$ be arbitrary. We will now prove that Pol($\mathcal{D}$) contains a WNU operation of arity $m$. Since $|V| \geq k$, there exists $K' \in \binom{V}{k}$; let us fix such $K'$, and let $O \in O^V_{\mathcal{G},k}$ be arbitrary. Let $f := (f_m)^{\mathcal{G},k}_{\mathcal{G}}$, and let us define an $m$-ary function $f'$: $(O^V_{\mathcal{G},k})^m \rightarrow O^V_{\mathcal{G},k}$ as follows. We define $f'$ to be equal to $f$ everywhere, where $f$ is defined, and we define $f'(O_1, \ldots, O_m) := O$ for every $(O_1, \ldots, O_m) \in (O^V_{\mathcal{G},k})^m$, where $f$ is not defined. Note that $f'$ is a WNU operation, and since $f$ preserves all constraints of $\mathcal{I}_{\mathcal{G},k}$ by Lemma 3.4, so does $f'$. It follows that $f'$ preserves all relations $R_C$ of $\mathcal{D}$, whence it is a polymorphism of $\mathcal{D}$.

Finally, the implication from item (5) to item (1) in Theorem 2.6 yields that $\mathcal{I}_{\mathcal{G},k}$ has relational width at most $(2,3)$. Since $\mathcal{I}_{\mathcal{G},k}$ is an instance of CSP($\mathcal{D}$) by definition, and since it is non-trivial and $(2,3)$-minimal by the assumption, $\mathcal{I}_{\mathcal{G},k}$ has a solution. $\square$

In order to prove the second item of Theorem 1.1 we need also the following version of Corollary 3.5.

Corollary 3.6. Suppose that $|V| \geq k$, and that $\mathcal{A}$ has for every $m \geq 1$ a polymorphism $f_m$ that is canonical with respect to $\mathcal{G}$ and such that $(f_m)^{\mathcal{G},k}_{\mathcal{G}}$ is a multisorted totally symmetric operation. If $\mathcal{I}_{\mathcal{G},k}$ is non-trivial and $(1,1)$-minimal, then it has a solution.

Proof. The proof is identical to the proof of Corollary 3.5 using Theorem 2.7 instead of Theorem 2.6. $\square$

Finally, this allows us to prove Theorem 1.1 from the introduction.

Theorem 1.1. Let $k, \ell \geq 1$, and let $\mathcal{A}$ be a first-order reduct of a $k$-homogeneous $\ell$-bounded $\omega$-categorical structure $\mathcal{B}$.

- If the clone of Aut($\mathcal{B}$)-canonical polymorphisms of $\mathcal{A}$ contains pseudo-WNU operations modulo Aut($\mathcal{B}$) of all arities $m \geq 3$, then $\mathcal{A}$ has relational width at most $(2k, \max(3k, \ell))$.
- If the clone of Aut($\mathcal{B}$)-canonical polymorphisms of $\mathcal{A}$ contains pseudo-totally symmetric operations modulo Aut($\mathcal{B}$) of all arities, then $\mathcal{A}$ has relational width at most $(k, \max(k+1, \ell))$ if $k > 1$, and $(1,1)$ if $k = 1$.

Proof. Suppose that the assumption of the first item of Theorem 1.1 is satisfied. Let $\mathcal{I}$ be a non-trivial $(2k, \max(3k, \ell))$-minimal instance of CSP($\mathcal{A}$), and let $\mathcal{I}_{\text{Aut}(\mathcal{B}),k}$ be the associated instance from Definition 3.1. Note that $\mathcal{I}_{\text{Aut}(\mathcal{B}),k}$ is a CSP instance over a finite set by the $\omega$-categoricity of $\mathcal{B}$. Note moreover that any $(2k, \max(3k, \ell))$-minimal non-trivial instance with less than $k$ variables admits a solution. Hence, we may assume that $\mathcal{I}$ has at least $k$ variables. By Lemma 3.2, $\mathcal{I}_{\text{Aut}(\mathcal{B}),k}$ is a $(2,3)$-minimal instance, and it is non-trivial by definition and since $\mathcal{I}$ has at least $k$ variables. Corollary 3.5 yields that $\mathcal{I}_{\text{Aut}(\mathcal{B}),k}$ admits a solution. Since $\mathcal{I}$ is $(2k, \max(3k, \ell))$-minimal, it is also $(k, \ell)$-minimal, and hence this solution lifts to a solution of $\mathcal{I}$ by Lemma 3.3. Thus, $\mathcal{A}$ has relational width at most $(2k, \max(3k, \ell))$.

Suppose now that the assumption in the second item is satisfied. If $k = 1$, then the orbits of elements under Aut($\mathcal{B}$) are singletons, whence $\mathcal{B}$ is a finite structure and the result follows.
immediately from Theorem 2.7. Otherwise, by the same reasoning as above but using Corollary 3.6 instead of Corollary 3.5 given a \((k, \max(k, \ell))-\)minimal instance \(\mathcal{I}\), the associated instance \(\mathcal{I}_{\text{Aut}(B),k}\) is \((1, 1))-\)minimal and therefore has a solution. Since \(\mathcal{I}\) is \((k, \max(k, \ell))-\)minimal, and if \(k = 2\), then it is also \((2, 3))-\)minimal, this solution lifts to a solution of \(\mathcal{I}\) by Lemma 3.3.

4. A NEW LOOP LEMMA FOR SMOOTH APPROXIMATIONS

We refine the algebraic theory of smooth approximations from \([55]\). Building on deep algebraic results from \([7]\) on finite idempotent algebras that are equationally non-trivial, we lift some of the theory from binary symmetric relations to cyclic relations of arbitrary arity.

4.1. The loop lemma.

**Definition 4.1.** Let \(R\) be a relation of arity \(m \geq 1\), let \(1 \leq n \leq m\), and let \(i_1, \ldots, i_n\) be pairwise distinct elements from \(\{1, \ldots, m\}\). We denote by \(\text{proj}_{\{i_1, \ldots, i_n\}}(R)\) the projection of \(R\) to the coordinates \((i_1, \ldots, i_n)\), i.e., \(\text{proj}_{\{i_1, \ldots, i_n\}}(R) = \{(b_{i_1}, \ldots, b_{i_n}) \mid (b_1, \ldots, b_m) \in R\}\).

The **linkedness congruence** of a binary relation \(R \subseteq A \times B\) is the equivalence relation \(\lambda_R\) on \(\text{proj}_2(R)\) defined by \((b, b') \in \lambda_R\) if there are \(k \geq 0\) and \(a_0, \ldots, a_{k-1} \in A\) and \(b = b_0, \ldots, b_k = b' \in B\) such that \((a_i, b_i) \in R\) and \((a_i, b_{i+1}) \in R\) for all \(i \in \{0, \ldots, k - 1\}\). We say that \(R\) is **linked** if it is non-empty and \(\lambda_R\) relates any two elements of \(\text{proj}_2(R)\). If \(R\) is of arity greater than two, then we define linkedness as before viewing \(R\) as a binary relation between \(\text{proj}_{\{1, \ldots, m-1\}}(R)\) and \(\text{proj}_{\{m\}}(R)\).

If \(A\) is a set and \(m \geq 2\), then we call a relation \(R \subseteq A^m\) **cyclic** if it is invariant under cyclic permutations of the components of its tuples, i.e., for every cyclic permutation \(\sigma\) of \(\{1, \ldots, m\}\), and for every \((a_1, \ldots, a_m) \in R\), it holds that \((a_{\sigma(1)}, \ldots, a_{\sigma(m)}) \in R\). The **support** of a relation \(R\) is the set of all elements of \(A\) which appear in some tuple of \(R\). Note that if \(R\) is cyclic, then its support equals the projection of \(R\) to any of its arguments.

If \(R\) is invariant under an oligomorphic group action on \(A \times B\), then there is an upper bound on the length \(k\) to witness \((b, b') \in \lambda_R\), and therefore \(\lambda_R\) is pp-definable from \(R\); in particular, it is invariant under any function clone acting on \(A \times B\) and preserving \(R\).

**Definition 4.2.** Let \(\mathcal{G}\) be a permutation group on a set \(A\). A **pseudo-loop with respect to \(\mathcal{G}\)** is a tuple of elements of \(A\) all of whose components belong to the same \(\mathcal{G}\)-orbit \([58][10][11]\). If \(\mathcal{G}\) contains only the identity function, then a pseudo-loop is called a **loop**.

Our next goal is to prove that any cyclic linked relation on a finite domain which is invariant under an equationally non-trivial idempotent function clone contains a loop. This result will be used to obtain a generalization of the second loop lemma of smooth approximations \([55]\) Theorem 11. In order to refer to results on finite algebras more easily, we will use the language of algebras rather than clones in the proof. We need the following definitions.

**Definition 4.3.** Let \(A\) be an algebra and let \(B\) be a subuniverse of \(A\), i.e., a subset of the universe of \(A\) closed under all operations of \(A\). We say that \(B\) is an **absorbing subuniverse of \(A\)** (or \(B\) absorbs \(A\)) if there exists a term operation \(f\) of arity \(n \geq 2\) such that for any \(j \in \{1, \ldots, n\}\) and for any \((a_1, \ldots, a_n) \in A^n\) with \(a_i \in B\) for all \(i \neq j\), \(f(a_1, \ldots, a_n) \in B\).

If \(B\) is an absorbing subuniverse of \(A\) and no proper subuniverse of \(B\) absorbs \(A\), we call \(B\) a **minimal absorbing subuniverse** of \(A\) and write \(B \ll A\).

We will use the following results from \([7]\).

**Proposition 4.4** (Proposition 2.15 in \([7]\)). Let \(A_1, A_2\) be algebras on finite domains whose polymorphism clones are equationally non-trivial, and let \(R\) be a subalgebra of \(A_1 \times A_2\) with \(\text{proj}_{\{1\}}(R) = A_1, \text{proj}_{\{2\}}(R) = A_2\). Then both of the following hold.
(1) If \( R \) is linked, \( B_1 \ll A_1, B_2 \ll A_2 \), and \( (B_1 \times B_2) \cap R \neq \emptyset \), then \( B_1 \times B_2 \ll R \).

(2) If \( R \) is linked, \( B_1 \ll A_1 \) or \( B_1 \ll A_2, B_2 \ll A_2, b_1 \in B_1, \) and \( b_2 \in B_2 \), then \( b_1 \) and \( b_2 \) can be linked via \( a_0, \ldots, a_k \) for some \( k \geq 1 \), where \( a_i \) is contained in a minimal absorbing subalgebra of \( A_1 \) or \( A_2 \) for every \( i \in \{0, \ldots, k\} \).

We say that a digraph has algebraic length 1 if it contains a path from some element to itself whose numbers of forward arcs and backward arcs differ by 1.

**Theorem 4.5** (Theorem 3.6 in [7]). Let \( A \) be an algebra on a finite domain with equationally non-trivial polymorphism clone, and let \( \mathcal{G} = (A; H) \) be a digraph of algebraic length 1 such that every vertex has an outgoing and an incoming edge and such that \( H \) is a subuniverse of \( A^2 \). Then \( \mathcal{G} \) contains a loop. Moreover, if there exists an absorbing subuniverse \( I \) of \( A \) which is contained in a weak component of \( \mathcal{G} \) of algebraic length 1, then the loop can be found in a minimal absorbing subuniverse of \( A \).

**Theorem 4.6** (Consequence of the proof of Theorem 4.2 in [7]). Let \( \mathcal{C} \) be an idempotent function clone on a finite domain that is equationally non-trivial. Then any \( \mathcal{C} \)-invariant cyclic linked relation on its domain contains a loop.

Proof. Let \( R \) as in the statement be given, and denote its arity by \( m \); we may assume \( m \geq 2 \). Given \( 1 \leq i \leq m \), we set \( R_i := \text{proj}_{(\ell, \ldots, m)}(R) \); moreover, we set \( R_{(i, j)} := \text{proj}_{(i, j)}(R) \) for all distinct \( i, j \) with \( 1 \leq i, j \leq m \).

We denote the support of \( R \) by \( A \). Note that for all \( i \in \{1, \ldots, m-1\} \) we have that \( R_{i+1} \) is linked when viewed as a binary relation between \( R_i \) and \( A \). We give the short argument showing linkedness for the convenience of the reader. Let \( a, b \in A \) be arbitrary; since they are linked in \( R_m \), there exist \( k \geq 1 \) and \( a = c_0, \ldots, c_{2k} = b \) such that \((c_{2j+1}, c_{2j}) \in R_m \) and \((c_{2j+1}, c_{2j+2}) \in R_m \) for all \( 0 \leq j < k \). For all such \( j \), we have \( c_{2j+1} \in R_{m-1} \); moreover, \( \text{proj}_{(\ell, \ldots, m-1)}(c_{2j+1}) \in R_{m-\ell} \) for all \( 1 \leq \ell \leq m-1 \) by the cyclicity of \( R \). Hence, these elements prove linkedness of \( a, b \) in \( R_{i+1} \) for all \( 1 \leq i \leq m-1 \).

As a consequence, \( R_{(1, i)} \) is linked for all \( i \in \{2, \ldots, m\} \). Indeed, for arbitrary \( a, b \in A \), if \( a = c_0, \ldots, c_{2k} = b \) witness the linkedness of \( a, b \) in \( R_i \), then \( c_0, \text{proj}_{(1)}(c_1), c_2, \ldots, \text{proj}_{(1)}(c_{2k-1}), c_{2k} \) witness the linkedness of \( a, b \) in \( R_{(1, i)} \).

Let \( A \) be the algebra whose domain is the support \( A \) of \( R \) and whose fundamental operations are the restrictions of the functions of the clone \( \mathcal{C} \) to that domain, i.e., \( A \) is the algebra \((A; \{f|_A \mid f \in \mathcal{C}\})\); note that this is well-defined since \( A \), as the projection of \( R \) to any coordinate, is invariant under the functions of \( \mathcal{C} \). Let \( R \) be the subalgebra of \( A^m \) with domain \( R \). Similarly, we write \( I_i \) for the subalgebra of \( I_i \) on the domain \( R_i \), for all \( 1 \leq i \leq m \). Following the proof of Theorem 4.2 in [7], we prove by induction on \( i \in \{1, \ldots, m\} \) that the following properties hold:

1. There exists \( I \ll A \) such that \( I^i \ll R_i \).
2. For all \( I_1, \ldots, I_i \ll A \) such that \( R_i \cap (I_1 \times \cdots \times I_i) \neq \emptyset \), we have \( R_i \cap (I_1 \times \cdots \times I_i) \ll I_i \).

The case \( i = 1 \) is trivial, since \( R_1 = A \). The case \( i = m \) gives us a constant tuple in \( R \).

We prove property (2) for \( i+1 \). Let \( I_1, \ldots, I_{i+1} \ll A \) be such that \( R_{i+1} \cap (I_1 \times \cdots \times I_{i+1}) \neq \emptyset \). By the induction hypothesis, \( R_i \cap (I_1 \times \cdots \times I_i) \ll R_i \). By the argument above, \( R_{i+1} \) is linked as a relation between \( R_i \) and \( A \). Thus, the assumptions of item (1) of Proposition 4.2 are satisfied, and we have \( R_{i+1} \cap (I_1 \times \cdots \times I_{i+1}) \ll R_{i+1} \).

We prove property (1) for \( i+1 \). Define a directed graph \( H \) on \( R_i \) by setting

\[
H := \{(a_1, a_i, (a_2, \ldots, a_{i+1})) \mid (a_1, \ldots, a_{i+1}) \in R_{i+1}\}.
\]

Let \( I \ll A \) be such that \( I^i \ll R_i \), which exists by the induction hypothesis. We show that:

- \( I^i \) is a subset of a (weak) connected component of \( H \);
- this connected component has algebraic length 1.
For the first item, let $X := \{ x \mid \exists a_1, \ldots, a_i \in I \,((a_1, \ldots, a_i, x) \in R_{i+1})\}$, which is an absorbing subuniverse of $A$. Let $X_1 \subseteq X$ be a minimal absorbing subuniverse of $A$. Then since $(I^i \times X_1) \cap R_{i+1} \neq \emptyset$ property (2) gives us $(I^i \times X_1) \subseteq R_{i+1}$. Reiterating this idea, we find minimal absorbing subuniverses $X_2, \ldots, X_i$ of $A$ such that for all $1 \leq j \leq i$ we have that $I^{i-j+1} \times X_1 \times \cdots \times X_j$ is contained in $R_{i+1}$. Now pick an arbitrary tuple $(a_1, \ldots, a_i) \in I^i$, and an arbitrary tuple $(x_1, \ldots, x_i) \in X_1 \times \cdots \times X_i$. Then there is a path in $H$ from $(a_1, \ldots, a_i)$ to $(x_1, \ldots, x_i)$ by the above, proving the first item.

For the second item, let $E := \{(x, y) \mid \exists c_2, \ldots, c_i \in I \,((x, c_2, \ldots, c_i, y) \in R_{i+1})\}$. Let $V_1, V_2$ be the projection of $E$ onto its first and second coordinate, respectively; then $E$ is a relation between $V_1$ and $V_2$. We have that $E$ is an absorbing subuniverse of the algebra $R_{(1,i+1)}$ induced by $R_{(1,i+1)}$ in $A^2$. By assumption on $R$, we have that $R_{(1,i+1)}$ is linked. Therefore, $E$ is linked, and $V_1$ and $V_2$ are absorbing subuniverses of $A$. Note that $I \subseteq V_1$ and $I \subseteq V_2$. Let $b \in I$ be arbitrary. Then there exist $k \geq 0$ and $c_0, \ldots, c_{2k+1}$ such that $(c_2j, c_{2j+1}) \in E$ for all $0 \leq j \leq k$ and $(c_{2(j+1)}, c_{2j+1}) \in E$ for all $0 \leq j \leq k - 1$ and such that $c_0 = b = c_{2k+1}$. These elements can be assumed to lie in minimal absorbing subuniverses of $A$ by item (2) of Proposition 4.4. We then have, by property (2), $(c_{2j}, \ldots, b, c_{2(j+1)}) \in R_{i+1}$ for all $0 \leq j \leq k$ and $(c_{2(j+1)}, \ldots, b, c_{2(j+1)}) \in R_{i+1}$ for all $0 \leq j \leq k - 1$. This gives a path of algebraic length 1 in $H$ from $(b, \ldots, b)$ to itself. This proves the second item.

Since $R$ is linked, it follows that every vertex of $H$ has an outgoing and an incoming edge. Moreover, $H$ has algebraic length 1, and is invariant under $C$, which is equationally non-trivial. Hence, by Theorem 4.5 there exists a loop in $H$ which lies in a minimal absorbing subuniverse $K$ of $R_i$. By the definition of $H$, this loop is a constant tuple $(a, \ldots, a)$. By projecting $K$ on the first component, we obtain a minimal absorbing subuniverse $J \ll A$; since $a \in J$, we have that $J^{i+1} \cap R_{i+1} \neq \emptyset$. By property (2), we get that $J^{i+1} \ll R_{i+1}$, so that property (1) holds.

The following is the previously announced generalization of [55, Theorem 11] from binary symmetric relations to arbitrary cyclic relations.

**Proposition 4.7.** Let $n \geq 1$, and let $D$ be an oligomorphic function clone on a set $A$ which is a model-complete core. Let $C \subseteq D^{\text{can}}$ be such that $C^n / \mathcal{G}_D$ is equationally non-trivial. Let $(S, \sim)$ be a minimal subfactor of the action $C^n$ with $\mathcal{G}_D$-invariant $\sim$-classes. Then for every $D$-invariant cyclic relation $R$ with support $(S)_{\mathcal{G}_D}$ one of the following holds:

1. The linkedness congruence of $R$ is a $D$-invariant approximation of $\sim$.
2. $R$ contains a pseudo-loop with respect to $\mathcal{G}_D$.

**Proof.** Let $R$ be given, and denote its arity by $m$. Assuming that (1) does not hold, we prove (2).

Denote by $O$ the set of orbits of $m$-tuples under the action of $\mathcal{G}_D$ thereon. Let $R'$ be the relation obtained by considering $R$ as a relation on $O$, i.e.,

$$R' := \{(O_1, \ldots, O_m) \in O^m \mid R \cap (O_1 \times \cdots \times O_m) \neq \emptyset\}.$$  

Thus, $R'$ is an $m$-ary cyclic relation with support $S' \subseteq O$, and $R'$ contains a loop if and only if $R$ satisfies (2).

By assumption, the action $C^n / \mathcal{G}_D$ is equationally non-trivial; moreover, it is idempotent since $D$ is a model-complete core. Note also that $R'$, and in particular $S'$, are preserved by this action. It is therefore sufficient to show that $R'$ is linked and apply Theorem 4.6.

Recall that we consider $R$ also as a binary relation between $\text{proj}_{m-1} (R)$ and $(S)_{\mathcal{G}_D}$; similarly, we consider $R'$ as a binary relation between $\text{proj}_{m-1} (R')$ and $S'$. By the oligomorphicity of $D$, the linkedness congruence $\lambda_R$ of $R$ is invariant under $D$.

By our assumption that (1) does not hold, there exist $c, d \in S$ which are not $\sim$-equivalent and such that $\lambda_R (c, d)$ holds; otherwise, $\lambda_R$ would be an approximation of $\sim$. This implies that the
orbits \( O_c, O_d \) of \( c, d \) are related via \( \lambda_{R'} \). By the minimality of \( (S, \sim) \), we have that \( (S)_D = \{(c, d)\}_D \).

Since \( D \) is a model-complete core, it preserves the \( \mathcal{G}_D \)-orbits, and it follows that any tuple in \( (S)_D = \{(c, d)\}_D \) is \( \lambda_{R'} \)-related to a tuple in the orbit of \( c \). Hence, \( \lambda_{R'} = (S')^2 \), and thus \( R' \) is linked. Theorem 4.6 therefore implies that \( R' \) contains a loop, and hence \( R \) contains a pseudo-loop with respect to \( \mathcal{G}_D \), which is what we had to show. \( \square \)

The following is a generalization of Lemma 14 in [55] from binary relations and functions to relations and functions of higher arity. It is used in Section 5 to obtain characterization of bounded width for first-order reducts of unary structures and of structures related to the logic MMSNP.

**Lemma 4.8.** Let \( D \) be a clone of polymorphisms of an \( \omega \)-categorical relational structure on a set \( A \), and assume that \( D \) is a model-complete core. Let \( n \geq 1 \), and let \( \sim \) be an equivalence relation on a set \( S \subseteq A^n \) with \( \mathcal{G}_D \)-invariant classes. Let \( m \geq 1 \), and let \( P \) be an \( m \)-ary relation on \( (S)_D \).

Suppose that every \( m \)-ary \( D \)-invariant cyclic relation \( R \) on \( (S)_D \) which contains a tuple in \( P \) with components \( m \)-invariant with respect to \( \mathcal{G}_D \).

Then there exists an \( m \)-ary \( f \in D \) such that for all \( a_1, \ldots, a_m \in A^n \) we have that if the tuple \( (f(a_1, \ldots, a_m), f(a_2, \ldots, a_m, a_1), \ldots, f(a_m, a_1, \ldots, a_{m-1})) \) is in \( P \), then it intersects at most one \( \sim \)-class.

**Proof.** The proof is similar to the proof of Lemma 14 in [55]. Fix \( m \geq 2 \). Call an \( m \)-tuple \((a_1, \ldots, a_m)\) of elements of \( A^n \) troublesome if there exists an \( m \)-ary \( r \in D \) such that

\[
(b_1, \ldots, b_m) := (r(a_1, \ldots, a_m), r(a_2, \ldots, a_m, a_1), \ldots, r(a_m, a_1, \ldots, a_{m-1}))
\]

is in \( P \) and has components in at least two \( \sim \)-classes. We first show that for any tuple \((a_1, \ldots, a_m) \in (A^n)^m \), there exists an \( m \)-ary function \( g \in D \) such that

\[
(g(a_1, \ldots, a_m), \ldots, g(a_m, a_1, \ldots, a_{m-1}))
\]

is not troublesome. We then use this fact to obtain a sequence \((f^i)_{i \in \mathbb{N}}\) of operations from \( D \) satisfying the same property on subsets of \( A^n \) of increasing size, and finally obtain a function satisfying the conclusion of the lemma by a standard compactness argument.

Note that if \((a_1, \ldots, a_m)\) is not troublesome and \( h \in D \), then \((h(a_1, \ldots, a_m), \ldots, h(a_m, a_1, \ldots, a_{m-1}))\) is not troublesome either. Moreover, if \( a_1, \ldots, a_m \) are in the same orbit under \( \mathcal{G}_D \), then \((a_1, \ldots, a_m)\) is not troublesome: since \( D \) is a model-complete core, \((h(a_1, \ldots, a_m), \ldots, h(a_m, a_1, \ldots, a_{m-1}))\) are also in the same orbit under \( \mathcal{G}_D \) for any \( m \)-ary \( h \in D \); since the classes of \( \sim \) are closed under \( \mathcal{G}_D \), our claim follows.

For each troublesome tuple \((a_1, \ldots, a_m)\), the smallest \( D \)-invariant relation containing the set

\[
\{(b_1, \ldots, b_m), (b_m, b_1, \ldots, b_{m-1})\},
\]

where \((b_1, \ldots, b_m)\) is defined via a witnessing function \( r \) as above, is cyclic, contains a tuple in \( P \) with components in at least two \( \sim \)-classes, and its support is contained in \((S)_D \). Hence, it contains a pseudo-loop with respect to \( \mathcal{G}_D \) by our assumptions. This implies that there exists an \( m \)-ary \( g \in D \) such that the entries of the tuple

\[
(g(b_1, \ldots, b_m), \ldots, g(b_m, b_1, \ldots, b_{m-1}))
\]

all belong to the same \( \mathcal{G}_D \)-orbit. The function

\[
s(x_1, \ldots, x_m) := g(r(x_1, \ldots, x_m), \ldots, r(x_m, x_1, \ldots, x_{m-1}))
\]

thus has the property that the entries of

\[
(s(a_1, \ldots, a_m), \ldots, s(a_m, a_1, \ldots, a_{m-1}))
\]

all lie in the same \( \mathcal{G}_D \)-orbit.

In conclusion, for every tuple \((a_1, \ldots, a_m)\) of elements of \( A^n \) – troublesome or not – there exists an \( m \)-ary function \( d \) in \( D \) such that \((d(a_1, \ldots, a_m), \ldots, d(a_m, a_1, \ldots, a_{m-1}))\) is not troublesome: if
\((a_1, \ldots, a_m)\) is troublesome one can take \(d\) to be the operation \(s\) we just described; if \((a_1, \ldots, a_m)\) is not troublesome then the first projection works.

Let \((a_1^i, \ldots, a_m^i)\) in be an enumeration of all \(m\)-tuples of elements of \(A^n\). We build by induction on \(i \in \mathbb{N}\) an operation \(f^i \in \mathcal{D}\) such that

\[
(f^i(a_1^1, \ldots, a_m^1), \ldots, f^i(a_m^i, a_1^i, \ldots, a_{m-1}^i))
\]

is not troublesome for any \(j < i\). For \(i = 0\) there is nothing to show, so suppose that \(f^i\) is built. Let \(d \in \mathcal{D}\) be an \(m\)-ary operation making the tuple

\[
(f^i(a_1^i, \ldots, a_m^i), \ldots, f^i(a_m^i, a_1^i, \ldots, a_{m-1}^i))
\]

not troublesome, in the manner of the previous paragraph. Then setting

\[
f^{i+1}(x_1, \ldots, x_m) := d(f^i(x_1, \ldots, x_m), \ldots, f^i(x_m, x_1, \ldots, x_{m-1}))
\]

clearly has the desired property for the tuple \((a_1^i, \ldots, a_m^i)\), and also for all tuples \((a_1^i, \ldots, a_m^i)\) with \(j < i\) by the remark in the first paragraph.

By a standard compactness argument, the sequence \((f^i)_{i \in \mathbb{N}}\) yields the desired function \(f\). This follows directly from Lemma 4 in [31]. We provide a different argument for the convenience of the reader. More precisely, let us construct the following infinite, finitely branching tree. For \(i > 0\), let \(A_i\) be the union of the entries of all tuples \(a_1^j, \ldots, a_m^j\) with \(j \leq i\). Define an equivalence relation \(\sim_i\) on the \(k\)-ary functions in \(\mathcal{D}\) by \(f \sim_i g\) if there exists \(\alpha \in \mathcal{G}_\mathcal{D}\) such that \(\alpha f|_{A_i} = g|_{A_i}\). The tree contains a node for every \(i > 0\) and for every equivalence class \([f]_i\) of some \(k\)-ary \(f \in \mathcal{D}\) with respect to \(\sim_i\) with the property that \((f(a_1^1, \ldots, a_m^1), \ldots, f(a_m^i, a_1^i, \ldots, a_{m-1}^i))\) is not troublesome for every \(j \leq i\). A node \([g]_{i+1}\) is a child of \([f]_i\) if \(g \sim_i f\). Since \(\mathcal{G}_\mathcal{D}\) is the automorphism group of an \(\omega\)-categorical structure, every \(\sim_i\) has only finitely many equivalence classes, and thus every \([f]_i\) has finite degree. Since \(\mathcal{G}_\mathcal{D}\) is oligomorphic, there are only finitely many such orbits \(O^i\) for every \(i > 0\), whence this tree is finitely branching, and it is clearly infinite. The tree contains a path of arbitrary large length, as the sequence \([f^1]_1, [f^2]_1, \ldots, [f^i]_1\) is a path of length \(i\) for any \(i\). König’s tree lemma then yields that there exists an infinite branch \(([f_i]_1)_{i \in \mathbb{N}}\). Up to replacing \(f_i\) by \(\alpha_i f_i\) for some suitably chosen \(\alpha_i \in \mathcal{G}_\mathcal{D}\), one can assume without loss of generality that \(f_{i+1}|_{A_i} = f_i|_{A_i}\) holds for all \(i \in \mathbb{N}\). Then \(f = \bigcup [f_i]_{A_i}\) is well-defined and is in \(\mathcal{D}\). Moreover, it satisfies the conclusion of the lemma.

\[\square\]

5. Applications: Collapses of the bounded width hierarchies for some classes of infinite structures

We now apply the algebraic results of Section 4 and the theory of smooth approximations to obtain a characterisation of bounded width for CSPs of first-order reducts of unary structures and for CSPs in MMSNP. Moreover, we obtain a collapse of the bounded width hierarchy for such CSPs and for CSPs of many other structures studied in the literature.

5.1. Unary Structures. We are going to prove the following characterization of bounded width for first-order reducts of unary structures.

**Theorem 5.1.** Let \(\mathcal{A}\) be a first-order reduct of a unary structure, and assume that \(\mathcal{A}\) is a model-complete core. Then one of the following holds:

- The clone of \(\text{Aut}(\mathcal{A})\)-canonical polymorphisms of \(\mathcal{A}\) is not equationally affine, or equivalently, it contains pseudo-WNU operations modulo \(\text{Aut}(\mathcal{A})\) of all arities \(n \geq 3\);
- \(\text{Pol}(\mathcal{A})\) has a uniformly continuous minion homomorphism to an affine clone.

In the first case, the stated equivalence follows from Section 2.2 and \(\mathcal{A}\) has relational width at most \((4, 6)\) by Theorem 1.1 and in the second case it does not have bounded width by Theorem 2.6. We remark that by [25, Lemma 6.7], the model-complete core of a first-order reduct of a unary
structure is also a first-order reduct of a unary structure. Thus, Theorem 5.1 gives a characterization of bounded width for all first-order reducts of unary structures.

The two items of Theorem 5.1 are invariant under expansions of \( A \) by a finite number of constants (for the first item, this follows exactly as the preservation of the pseudo-Siggers identity under adding constants in [10, 11], for the second item, [9] Theorem 1.8 yields that the existence of a uniformly continuous minion homomorphism from \( \text{Pol}(A) \) to an arbitrary clone \( C \) is equivalent to the existence of such a homomorphism from the polymorphism clone of an expansion of \( A \) by a finite number of constants.) Thus, by Proposition 6.8 in [25], one can assume in the following that \( A \) is a first-order expansion of \( (\mathbb{N}; V_1, \ldots, V_r) \) (and not only a first-order reduct thereof, i.e., the relations \( V_1, \ldots, V_r \) are present in \( A \)) where \( V_1, \ldots, V_r \) form a partition of \( \mathbb{N} \) in which every set is either a singleton or infinite. Such partitions were called stabilized partitions in [25], and we shall also call the structure \((\mathbb{N}; V_1, \ldots, V_r)\) a stabilized partition.

We will use the following fact which states that \( \text{Pol}(A) \) locally interpolates \( \text{Pol}(A)^{\text{can}} \).

**Lemma 5.2** (Proposition 6.5 in [25]). Let \( A \) be a first-order expansion of a stabilized partition \((\mathbb{N}; V_1, \ldots, V_r)\). For every \( f \in \text{Pol}(A) \) there exists \( g \in \text{Pol}(A)^{\text{can}} \) which is locally interpolated by \( f \) modulo \( \text{Aut}(A) \).

The following will allow us to assume, in most proofs, the presence of functions in \( \text{Pol}(A) \) which are injective on every set in the stabilized partition. This is the analogue to the efforts to obtain binary injections in [55].

**Lemma 5.3** (Subset of the proof of Proposition 6.6 [25]). Let \( A \) be a first-order expansion of a stabilized partition \((\mathbb{N}; V_1, \ldots, V_r)\), and assume it is a model-complete core. If \( \text{Pol}(A) \) has no continuous clone homomorphism to \( \mathcal{P} \), then it contains operations of all arities whose restrictions to \( V_i \) are injective for all \( 1 \leq i \leq r \).

**Proof.** We show by induction that for all \( 1 \leq j \leq r \), there exists a binary operation in \( \text{Pol}(A) \) whose restriction to each of \( V_1, \ldots, V_j \) is injective; higher arity functions with the same property are then obtained by nesting the binary operation. For the base case \( j = 1 \), observe that the disequality relation \( \neq \) is preserved on \( V_1 \) since \( A \) is a model-complete core; together with the restriction of \( \text{Pol}(A) \) to \( V_1 \) being equationally non-trivial, we then obtain an operation which acts as an essential function on \( V_1 \). This in turn easily yields a function that acts as a binary injection on \( V_1 \) – see e.g. [19]. For the induction step, assuming the statement holds for \( 1 \leq j < r \), we show the same for \( j + 1 \). By the induction hypothesis, there exist binary functions \( f, g \in \text{Pol}(A) \) such that the restriction of \( f \) to each of the sets \( V_1, \ldots, V_j \) is injective, and the restriction of \( g \) to \( V_{j+1} \) is injective. If the restriction of \( f \) to \( V_{j+1} \) depends only on its first or only on its second variable, then it is injective in that variable since disequality is preserved on \( V_{j+1} \), and hence either the function \( f(g(x, y), f(x, y)) \) or the function \( f(f(x, y), g(x, y)) \) has the desired property. If on the other hand the restriction of \( f \) to \( V_{j+1} \) depends on both variables, then the same argument as in the base case yields a function which is injective on \( V_{j+1} \), and this function is still injective on each of the sets \( V_1, \ldots, V_j \).

**Proposition 5.4.** Let \( A \) be a first-order expansion of a stabilized partition \((\mathbb{N}; V_1, \ldots, V_r)\), and assume it is a model-complete core. Suppose that \( \text{Pol}(A) \) contains operations of all arities whose restrictions to \( V_i \) are injective for all \( 1 \leq i \leq r \). Then the following are equivalent:

- \( \text{Pol}(A)^{\text{can}} \) is equationally affine;
- \( \text{Pol}(A)^{\text{can}} \cap \mathbb{N}/\text{Aut}(A) \) is equationally affine.

**Proof.** Trivially, if \( \text{Pol}(A)^{\text{can}} \cap \mathbb{N}/\text{Aut}(A) \) is equationally affine, then so is \( \text{Pol}(A)^{\text{can}} \).

For the other direction, assume that \( \text{Pol}(A)^{\text{can}} \cap \mathbb{N}/\text{Aut}(A) \) is not equationally affine. We will show that \( \text{Pol}(A)^{\text{can}} \) contains pseudo-WNU operations of all arities \( \geq 3 \), whence it is not equationally affine. Since \( \text{Pol}(A)^{\text{can}} \cap \mathbb{N}/\text{Aut}(A) \) is not equationally affine, \( \text{Pol}(A)^{\text{can}} \) contains for
every $k \geq 3$ a $k$-ary operation $g_k$ whose action on $N/\text{Aut}(\mathbb{A})$ is a WNU operation by the implication from \cite{4} to \cite{5} in Theorem 2.6 and since $N/\text{Aut}(\mathbb{A})$ is a finite set. Fix for all such $k \geq 3$ an operation $f_k$ of arity $k$ whose restriction to $V_i$ is injective for all $i \in \{1, \ldots, r\}$, and consider the operation $$h_k(x_1, \ldots, x_k) := f_k(g_k(x_1, \ldots, x_k), g_k(x_2, \ldots, x_k, x_1), \ldots, g_k(x_k, x_1, \ldots, x_{k-1}).$$ Then evaluating $h_k(a, b, \ldots, b)$ for arbitrary $a, b \in N$, all the arguments of $f_k$ belong to the same set $V_i$, by the fact that $g_k$ acts on $N/\text{Aut}(\mathbb{A})$ as a weak near-unanimity operation. Since $\mathbb{A}$ is an expansion of $(N; V_1, \ldots, V_r)$, we obtain that $h_k(a, b, \ldots, b)$ belongs to $V_i$. The same is true for any permutation of the tuple $(a, b, \ldots, b)$, so that $h_k$ acts as a WNU operation on $N/\text{Aut}(\mathbb{A})$. By the injectivity of $f_k$ when restricted to $V_i$, it also follows that $h_k$ acts as a WNU operation on $N^2/\text{Aut}(\mathbb{A})$; hence, it is a pseudo-WNU operation since it preserves the equivalence of orbits of 2-tuples under $\text{Aut}(N; V_1, \ldots, V_r)$ and since the stabilized partition is 2-homogeneous. Taking for every $h_k$ an operation in $\text{Pol}(\mathbb{A})^{\text{can}}$ locally interpolated by $h_k$ modulo $\text{Aut}(\mathbb{A})$ which exists by Lemma 5.2, we see that $\text{Pol}(\mathbb{A})^{\text{can}}$ contains pseudo-WNU operations of all arities $\geq 3$, and hence it is not equationally affine.

Before finally proving Theorem 5.1, we state the following standard result from \cite{59} that allows us to derive, under the assumption that $\text{Pol}(\mathbb{A})^{\text{can}} \cap A/\text{Aut}(\mathbb{A})$ is equationally affine, the existence of an invariant subset $S \subseteq A$ and an equivalence relation $\sim$ on $S$ with finitely many classes such that $\text{Pol}(\mathbb{A})^{\text{can}} \cap S/\sim$ is itself an affine clone.

**Proposition 5.5** (Consequence of Proposition 3.1 from \cite{59}). Let $C$ be an idempotent function clone on a finite domain that is equationally affine. Then there exists a minimal subfactor $(S, \sim)$ of $C$ such that $C$ acts on the $\sim$-classes as an affine clone.

**Proof of Theorem 5.1.** Let $\mathbb{A}$ as in Theorem 5.1 be given; by the remark following that theorem, we may without loss of generality assume that $\mathbb{A}$ is a first-order expansion of a stabilized partition $(N; V_1, \ldots, V_r)$. Assume henceforth that $\text{Pol}(\mathbb{A})^{\text{can}}$ is equationally affine; we show that $\text{Pol}(\mathbb{A})$ has a uniformly continuous minion homomorphism to an affine clone.

If $\text{Pol}(\mathbb{A})$ has a continuous clone homomorphism to $\mathcal{P}$, then we are done. Assume therefore the contrary; then by Lemma 5.3 $\text{Pol}(\mathbb{A})$ contains for all $k \geq 2$ a $k$-ary operation whose restriction to $V_i$ is injective for all $1 \leq i \leq r$. In particular, Proposition 5.4 applies, and thus $\text{Pol}(\mathbb{A})^{\text{can}} \cap N/\text{Aut}(\mathbb{A})$ is equationally affine. Proposition 5.5 yields that there exists a minimal subfactor $(S, \sim)$ of $\text{Pol}(\mathbb{A})^{\text{can}}$ such that $\text{Pol}(\mathbb{A})^{\text{can}}$ acts on the $\sim$-classes as an affine clone.

Let $R$ be any $\text{Pol}(\mathbb{A})$-invariant cyclic relation with support $\langle S \rangle_{\text{Pol}(\mathbb{A})}$, containing a tuple with components in pairwise distinct $\text{Aut}(\mathbb{A})$-orbits and which intersects at least two $\sim$-classes. By Proposition 4.7 $R$ either gives rise to a $\text{Pol}(\mathbb{A})$-invariant approximation of $\sim$, or it contains a pseudo-loop with respect to $\text{Aut}(\mathbb{A})$. In the first case, the presence of the tuple required above implies smoothness of the approximation: if $t \in R$ is such a tuple, $c \in \langle S \rangle_{\text{Pol}(\mathbb{A})}$ appears in $t$, and $d \in \langle S \rangle_{\text{Pol}(\mathbb{A})}$ belongs to the same $\text{Aut}(\mathbb{A})$-orbit as $c$, then there exists an element of $\text{Aut}(\mathbb{A})$ which sends $c$ to $d$ and fixes all other elements of $t$. Hence, $c$ and $d$ are linked in $R$, and the entire $\text{Aut}(\mathbb{A})$-orbit of $c$ is contained in a class of the linkedness relation of $R$. Thus, $\text{Pol}(\mathbb{A})$ admits a uniformly continuous minion homomorphism to an affine clone by Theorem 2.9.

Hence we may assume that for any $R$ as above the second case holds. We now show that this leads to a contradiction, finishing the proof of Theorem 5.1. We apply Lemma 1.8 with $n = 1$, $m \geq 2$ arbitrary and $P$ the set of $m$-tuples with entries in pairwise distinct $\text{Aut}(\mathbb{A})$-orbits within $\langle S \rangle_{\text{Pol}(\mathbb{A})}$, i.e., $P = \{(a_1, \ldots, a_m) \in (\langle S \rangle_{\text{Pol}(\mathbb{A})})^n \mid \forall 1 \leq i < j \leq m, \forall \alpha \in \text{Aut}(\mathbb{A}), \alpha(a_i) \neq a_j\}$. We obtain an $m$-ary function $f \in \text{Pol}(\mathbb{A})$ with the property that the tuple $$(f(a_0, \ldots, a_{m-1}), \ldots, f(a_1, \ldots, a_{m-1}, a_0))$$ intersects at most one $\sim$-class whenever it has entries in pairwise distinct $\text{Aut}(\mathbb{A})$-orbits, for all $a_0, \ldots, a_{m-1} \in S$. Let $(\mathbb{A}, <)$ be the expansion of $\mathbb{A}$ by a linear order that is convex with respect
to the partition \( V_1, \ldots, V_r \) and dense and without endpoints on every infinite set of the partition. The structure \((\mathbb{A}, <)\) is a Ramsey structure, since \(\text{Aut}(\mathbb{A}, <)\) is isomorphic as a permutation group to the action of the product \(\prod_{i=1}^r \text{Aut}(V_i; <)\), and each of the groups of the product is either trivial or the automorphism group of a Ramsey structure \[\text{[47]}\]. By Theorem \[\text{[2.2]}\] we may assume that \(f\) is diagonally canonical with respect to \(\text{Aut}(\mathbb{A}, <)\). Let \(a, a' \in A^m\) be so that \(a_i, a'_i\) belong to the same orbit with respect to \(\text{Aut}(\mathbb{A})\) for all \(1 \leq i \leq m\). Then there exists \(\alpha \in \text{Aut}(\mathbb{A}, <)\) such that \(\alpha(a) = a'\), and hence \(f(a)\) and \(f(a')\) lie in the same \(\text{Aut}(\mathbb{A})\)-orbit by diagonal canonicity; hence \(f\) is \(1\)-canonical with respect to \(\text{Aut}(\mathbb{A})\).

Applying Lemma \[\text{[5.2]}\] we obtain a canonical function \(g \in \text{Pol}(\mathbb{A})^{\text{can}}\) which acts like \(f\) on \(\mathbb{N}/\text{Aut}(\mathbb{A})\). The property of \(f\) stated above then implies for \(g\) that \(g(a_0, \ldots, a_{m-1}) \sim g(a_1, \ldots, a_{m-1}, a_0)\) for all \(a_0, \ldots, a_{m-1} \in S\) such that the values \(g(a_0, \ldots, a_{m-1}), \ldots, g(a_{m-1}, a_0, \ldots, a_{m-2})\) lie in pairwise distinct \(\text{Aut}(\mathbb{A})\)-orbits.

By the choice of \((S, \sim)\) we have that \(\text{Pol}(\mathbb{A})^{\text{can}}\) acts on \(S/\sim\) by affine functions over a finite module. We use the symbols \(+, \cdot\) for the addition and multiplication in the corresponding ring, and also \(+\) for the addition in the module and \(\cdot\) for multiplication of elements of the module with elements of the ring. We denote by \(1\) the multiplicative identity of the ring, by \(-1\) its additive inverse, and identify their powers in the additive group with the non-zero integers.

The domain of the module is \(S/\sim\), and we denote the identity element of its additive group by \([a_0]_\sim\). Pick an arbitrary element \([a_1]_\sim \neq [a_0]_\sim\) from \(S/\sim\), and let \(m \geq 2\) be its order in the additive group of the module, i.e., the minimal positive number such that \(m \cdot [a_1]_\sim = [a_0]_\sim\). For \(i \in \{2, \ldots, m-1\}\), let \(a_i\) be an arbitrary element such that \([a_i]_\sim = i \cdot [a_1]_\sim\). Let \(g \in \text{Pol}(\mathbb{A})^{\text{can}}\) be the \(m\)-ary operation obtained in the preceding paragraph. If the values \(g(a_0, \ldots, a_{m-1})\) and \(g(a_1, \ldots, a_{m-1}, a_0)\) lie in pairwise distinct \(\text{Aut}(\mathbb{A})\)-orbits, then (computing indices modulo \(m\)) we have that \(g([a_0]_\sim, \ldots, [a_{m-1}]_\sim), \ldots, g([a_{m-1}]_\sim, \ldots, [a_{m-1}]_\sim)\) are all equal. If on the other hand they do not, then \(g([a_k]_\sim, \ldots, [a_{k+m-1}]_\sim) = g([a_{k+j}]_\sim, \ldots, [a_{k+j+m-1}]_\sim)\) for some \(0 \leq k < m\) and \(1 \leq j < m\). Hence, in either case we may assume the latter equation holds. By assumption, \(g\) acts on \(S/\sim\) as an affine map, i.e., as a map of the form \((x_0, \ldots, x_{m-1}) \mapsto \sum_{i=0}^{m-1} c_i \cdot x_i\), where \(c_0, \ldots, c_{m-1}\) are elements of the ring which sum up to \(1\). We compute (with indices to be read modulo \(m\))

\[
[a_0]_\sim = g([a_{k+j}]_\sim, \ldots, [a_{k+j+m-1}]_\sim) + (-1) \cdot g([a_k]_\sim, \ldots, [a_{k+m-1}]_\sim)
= \sum_{i=0}^{m-1} c_i \cdot [a_{k+j+i}]_\sim + (-1) \cdot \sum_{i=0}^{m-1} c_i \cdot [a_{k+i}]_\sim
= \sum_{i=0}^{m-1} c_i \cdot (k+i+j) \cdot [a_1]_\sim + (-1) \cdot \sum_{i=0}^{m-1} c_i \cdot (k+i) \cdot [a_1]_\sim
= \left(\sum_{i=0}^{m-1} c_i\right) \cdot j \cdot [a_1]_\sim = j \cdot [a_1]_\sim.
\]

But \(j \cdot [a_1]_\sim \neq [a_0]_\sim\) since the order of \([a_1]_\sim\) equals \(m > j\), a contradiction. \(\square\)

5.2. **Proof of Corollary [1.2]**. Before proceeding with the proof of Corollary [1.2] we summarize the currently known characterizations of structures with bounded width.

**Corollary 5.6** ([Theorems 39 and 57 in [53], Theorems 21 and 67 in [49], Theorem 8.1 in [34], Theorem 1 in [19], Theorem 5.1]). Let \(\mathbb{A}\) be a model-complete core structure that has bounded width. Let \(\mathcal{C}\) be the clone of polymorphisms of \(\mathbb{A}\) that are \(\text{Aut}(\mathbb{A})\)-canonical. If \(\mathbb{A}\) is a first-order reduct of:

- the universal homogeneous graph or tournament, or of a unary structure, then \(\mathcal{C}\) contains pseudo-WNU operations modulo \(\text{Aut}(\mathbb{A})\) of all arities \(n \geq 3\).
• \((\mathbb{N};=)\), the universal homogeneous \(K_n\)-free graph \(\mathbb{H}_n\), where \(n \geq 3\), or the countably infinite equivalence relation with infinitely many equivalence classes all of infinite size, or the universal homogeneous partial order \(\mathbb{P}\), but not of \((\mathbb{Q},<)\), then \(\mathcal{G}\) contains pseudo-totally symmetric operations modulo \(\text{Aut}(\mathcal{A})\) of all arities.

Now, we can finally prove Corollary 1.2 from Section 1.

**Corollary 1.2.** Let \(\mathcal{A}\) be a structure that has bounded width. If \(\mathcal{A}\) is a first-order reduct of:

- the universal homogeneous graph \(\mathcal{G}\) or tournament \(\mathcal{T}\), or of a unary structure, then \(\mathcal{A}\) has relational width at most \((4,6)\);
- the universal homogeneous \(\mathcal{G}\)-free graph \(\mathbb{H}_n\), where \(n \geq 3\), then at most \((2,n)\);
- \((\mathbb{N};=)\), the countably infinite equivalence relation \(\mathbb{C}_\omega\) with infinitely many equivalence classes, all of infinite size, or the universal homogeneous partial order \(\mathbb{P}\), then at most \((2,3)\).

**Proof.**

For the first item, let \(\mathcal{A}\) be a first-order reduct of \(\mathcal{G}\) or \(\mathcal{T}\). Then \(\mathcal{A}\) has a model-complete core by [13]; this model-complete core has the same relational width as \(\mathcal{G}\). Moreover, by Lemma 17 and Lemma 46 in [55], this model-complete core is again a first-order reduct of \(\mathcal{G}\) or \(\mathcal{T}\), respectively, or a one-element structure. In the latter case, \(\mathcal{A}\) has relational width \((1,1)\); hence, we may assume that \(\mathcal{A}\) is itself a model-complete core. It then follows from Corollary 5.6 that \(\mathcal{A}\) has, for every \(n \geq 3\), an \(\text{Aut}(\mathcal{A})\)-canonical polymorphism of arity \(n\) that is a pseudo-WNU operation modulo \(\text{Aut}(\mathcal{A})\).

Since both \(\mathcal{G}\) and \(\mathcal{T}\) are 2-homogeneous and 3-bounded, our claim follows from our Theorem 1.1.

By Lemma 6.7 in [25], the set of first-order reducts of unary structures is closed under taking model-complete cores. Let therefore \(\mathcal{A}\) be a first-order reduct of a unary structure \(\mathcal{B}\) which is a model-complete core. Then \(\mathcal{A}\) is also a first-order reduct of a unary structure \(\mathcal{C} = (C; V_1, \ldots, V_r)\), where \(V_i\) either contains only one element or is infinite and where \(V_i \cap V_j = \emptyset\) for every \(1 \leq i, j \leq r\) with \(i \neq j\). Moreover, \(\mathcal{C}\) is 2-homogeneous since the orbit of every tuple under \(\text{Aut}(\mathcal{B})\) is determined by the universal relations and equalities holding on pairs of its elements. We claim it is 2-bounded. Indeed, every finite structure in the signature of \(\mathcal{C}\) embeds into \(\mathcal{C}\) unless it contains an element contained in two relations \(V_i, V_j\) with \(1 \leq i, j \leq r, i \neq j\), or it contains two distinct elements contained in a relation \(V_i\) for some \(1 \leq i \leq r\) with \(|V_i^C| = 1\). Hence, by appeal to Theorem 1.1 and Corollary 5.6 in the present paper, our claim holds for this class as well.

If \(\mathcal{A}\) is a first-order reduct of \(\mathbb{H}_n\), \((\mathbb{N};=)\) or \(\mathbb{C}_\omega\) that has bounded width and is a model-complete core, then it has \(\text{Aut}(\mathcal{A})\)-canonical pseudo-totally symmetric polymorphisms modulo \(\text{Aut}(\mathcal{A})\) of all arities, by Corollary 5.6 above. Since \(\mathbb{H}_n\) is 2-homogeneous and \(n\)-bounded, and since both \((\mathbb{N};=)\) and \(\mathbb{C}_\omega\) are 2-homogeneous and 3-bounded, the claimed bound follows.

Finally, a first-order reduct of \(\mathcal{P}\) with bounded width is either homomorphically equivalent to a first-order reduct of \((\mathbb{Q};<)\) or it satisfies the algebraic condition in the second item of Theorem 1.1 by Corollary 5.6. In the latter case we are done by Theorem 1.1 in the former we appeal to the syntactical characterization of first-order reducts of \((\mathbb{Q};<)\). Indeed, such a structure has bounded width if and only if it is definable by a conjunction of so-called Ord-Horn clauses [29]. It then follows by [36] that a first-order reduct of \((\mathbb{Q};<)\) with bounded width has relational width at most \((2,3)\). The result for \(\mathcal{P}\) follows.

The following example shows that for some of the structures under consideration, the bounds on relational width provided by Corollary 1.2 are tight.

**Example 5.7.** To show the tightness of the bound in the case of the universal homogeneous graph \(\mathcal{G} = (G; E)\), we exhibit a first-order reduct \(\mathcal{A}\) such that for all \(i \leq j\) with \(i \leq 4\), if \(1 \leq i < 4\) or \(1 \leq j < 6\), then there exists a non-trivial, \((i,j)\)-minimal instance of \(\text{CSP}(\mathcal{A})\) that has no solution. Let \(N := (G^2 \setminus E) \cap \neq\). Consider the first-order reduct \(\mathcal{A} := (G; R_e, R_d)\) of \(\mathcal{G}\), where \(R_e := \{(a, b, c, d) \in G^4 \mid E(a, b) \land E(c, d)\} \) and \(R_d := \{(a, b, c, d) \in G^4 \mid E(a, b) \land N(c, d)\} \) and \(R_e := \{(a, b, c, d) \in G^4 \mid E(a, b) \land N(c, d)\} \).
It can be seen that $A$ has bounded width, so that Corollary 1.2 implies that $A$ has relational width at most (4, 6). The instance $I = \{(v_1, \ldots, v_4), \{C_{\sigma}, C_{\tau}\}\}$, where $C_{\sigma} := \{f \in A^{v_1, \ldots, v_4} | f(v_1, \ldots, v_4) \in R_{\sigma}\}$, $C_{\tau} := \{f \in A^{v_1, \ldots, v_4} | f(v_1, \ldots, v_4) \in R_{\tau}\}$, is non-trivial, and $(i, j)$-minimal for all $i \leq j$ with $1 \leq i < 4$ since all variables are contained in the scopes of both constraints and the projection of any of the constraints to any subset of $\{v_1, \ldots, v_4\}$ containing at most 3 elements is equal to the projection of $\{f \in A^{v_1, \ldots, v_4} | f(v_1) \neq f(v_2), f(v_3) \neq f(v_4)\}$ to this subset. However, $I$ has no solution. Moreover, the $(4, 5)$-minimal instance equivalent to the instance $J = \{(v_1, \ldots, v_6), \{D_1, D_2, D_3\}\}$, where $D_1 := \{f \in A^{v_1, \ldots, v_4} | f(v_1, \ldots, v_4) \in R_{\tau}\}$, $D_2 := \{f \in A^{v_3, \ldots, v_6} | f(v_3, \ldots, v_6) \in R_{\sigma}\}$, and $D_3 := \{f \in A^{v_1, v_2, v_3, v_6} | f(v_1, v_2, v_3, v_6) \in R_{\tau}\}$, is non-trivial and has no solution. It follows that the exact relational width of $A$ is (4, 6).

The tightness of the bound for first-order reducts of the universal homogeneous tournament can be shown similarly. However, it is an open question whether the bound is tight for first-order reducts of unary structures (for more details, see [57]).

The bound on relational width provided in the second item of Corollary 1.2 is tight. Indeed, let $n \geq 3$, let $H_n := (H_n; E)$ be the universal homogeneous $K_n$-free graph, let $N := ((H_n)^2; E) \cap \neq$, and let $A := (H_n; E, N)$. $A$ is preserved by canonical pseudo-totally symmetric operations modulo $\text{Aut}(H_n)$ of all arities and therefore has relational width at most $(2, n)$ by Theorem 1.1 (setting $B := H_n$). But the non-trivial, $(2, n-1)$-minimal instance $I = \{(v_1, \ldots, v_n), \{C_{i,j} | 1 \leq i \neq j \leq n\}\}$, where $C_{i,j} := \{f \in A^{v_1, v_j} | (f(v_1), f(v_j)) \in E\}$ for every $1 \leq i \neq j \leq n$, has no solution; moreover, the instance $J = \{(v_1, v_2), \{(f(v_1), f(v_2)) \in E\}, \{f \in A^{v_1, v_2} | (f(v_1), f(v_2)) \in E\}\}$ is non-trivial, $(1, j)$-minimal for every $j \geq 1$ and has no solution either.

For the structures from the third item of Corollary 1.2, the tightness of the bound can be shown similarly.

5.3. MMSNP. Monotone Monadic SNP (MMSNP) is a fragment of existential second order logic that was discovered by Feder and Vardi in their seminal paper [42]. MMSNP is defined as the class of formulas of the form

$$
\Phi := \exists M_1 \ldots \exists M_r \forall x_1 \ldots \forall x_k \bigwedge_i \neg (\alpha_i \land \beta_i)
$$

where $M_1, \ldots, M_r$ are unary predicates, each $\alpha_i$ is a conjunction of positive atoms in a signature $\tau$ not containing the equality (called the \textit{input signature}) and each $\beta_i$ is a conjunction of positive or negative existentially quantified unary predicates. We say that $\Phi$ is \textit{connected} if every conjunction $\alpha_i$ is connected, that is, no $\alpha_i$ can be written as the conjunction of two non-empty formulas which do not share any variables. A simple syntactic procedure shows that every $\Phi$ is equivalent to a disjunction of connected MMSNP sentences, which can be computed in exponential time.

As in the case of graphs, we call a structure \textit{connected} if it is not the disjoint union of two of its proper substructures. Let $\tau$ be a relational signature, let $\sigma$ be a unary signature whose relations are called the \textit{colors}, and let $\mathcal{F}$ be a finite set of finite connected $(\tau \cup \sigma)$-structures, with the property that every element of any structure in $\mathcal{F}$ has exactly one color. We call $\mathcal{F}$ a \textit{colored obstruction set} in the following. The problem FPP($\mathcal{F}$) takes as input a $\tau$-structure $\mathcal{X}$ and asks whether there exists a $(\tau \cup \sigma)$-expansion $\mathcal{X}^*$ of $\mathcal{X}$ whose vertices are all colored with exactly one color and such that for every $\mathcal{F} \in \mathcal{F}$, there exists no homomorphism from $\mathcal{F}$ to $\mathcal{X}^*$. We say in this case that the coloring $\mathcal{X}^*$ is $\mathcal{F}$-\textit{free} or $\mathcal{F}$-\textit{obstruction-free}. It can be seen that connected MMSNP and FPP are equivalent: for every connected MMSNP sentence $\Phi$, there exists a set $\mathcal{F}$ such that $\mathcal{X} \models \Phi$ if and only if $\mathcal{X}$ is a yes-instance to FPP($\mathcal{F}$), for any $\tau$-structure $\mathcal{X}$. Conversely, given any $\mathcal{F}$, there exists a connected MMSNP sentence $\Phi$ as above. We call $\mathcal{F}$ a \textit{colored obstruction set associated with $\Phi$}.

Every connected MMSNP sentence $\Phi$ has an equivalent normal form $\Psi$ that can be computed from $\mathcal{F}$ in double exponential time [21] Lemma 4.4]. It is shown in [21] Definition 4.12 that for
every MMSNP sentence $\Phi$ in normal form, there exists an $\omega$-categorical structure $A_\Phi$ (denoted in [21] by $C_\Phi^\omega$) such that for any $\tau$-structure $X$ it is true that $X \models \Phi$ if and only if $X$ admits a homomorphism to $A_\Phi$ (which can equivalently be taken to be injective) [21 Lemma 4.13].

Additionally, $\Phi$ is in strong normal form if any identification of two existentially quantified predicates in $\Phi$ yields an inequivalent sentence. Finally, we say that $\Phi$ is precolored if for every symbol $M \in \sigma$, there is an associated unary symbol $P_M \in \tau$, and $\Phi$ contains the conjunct $\neg (P_M(x) \land M'(x))$ for every color $M' \in \sigma \setminus \{M\}$. Note that any precolored sentence in normal form is automatically in strong normal form (identifying two unary predicates $M, M'$ would yield a sentence $\Psi$ containing the conjunct $\neg (P_M(x) \land M(x))$, and $\Phi$ and $\Psi$ can be separated by a 1-element structure whose sole vertex is in $P_M$). Any sentence $\Phi$ has a standard precoloration obtained by adding for every symbol $M \in \sigma$ a predicate $P_M$ to $\tau$ and for every $M' \in \sigma \setminus \{M\}$ the conjunct $\neg (P_M(x) \land M'(x))$ to $\Phi$. The colored obstruction set of the standard precoloration of $\Phi$ consists of the obstruction set for $\Phi$ together with one-element obstructions $F$ whose sole vertex belongs to $P_M$ and $M'$ in $F$ and all other relational symbols are interpreted as empty relations.

A useful (but imprecise) parallel between MMSNP and finite-domain CSPs is the following. Connected MMSNP sentences in normal form, in strong normal form, and in precolored normal form have the same relationship as do finite structures, finite cores, and expansions of cores by all singleton unary relations.

When $\Phi$ is in strong normal form or precolored, $A_\Phi$ can additionally be chosen to have the following properties:

(i) If $\Phi$ is precolored, then the orbits of the elements of $A_\Phi$ under $\text{Aut}(A_\Phi)$ correspond to the colors of $\Phi$ and coincide with the interpretations of the corresponding symbols in $\tau$ [21 Lemma 6.2]. In particular, the action of $\text{Pol}(A_\Phi)^\text{can}_1$ on $\text{Aut}(A_\Phi)$-orbits of elements is idempotent [21 Proposition 7.2].

(ii) The expansion of $A_\Phi$ by a generic linear order $<$ is a Ramsey structure [21 Corollary 5.9]. By Theorem [2.2] every $f \in \text{Pol}(A_\Phi)$ locally interpolates an operation $g \in \text{Pol}(A_\Phi)^\text{can}_1$, and every $f$ diagonally interpolates an operation $f'$ that is diagonally canonical with respect to $\text{Aut}(A_\Phi, <)$.

We finally solve the Datalog-rewritability problem for MMSNP and prove that a precolored connected sentence $\Phi$ in normal form is equivalent to a Datalog program if and only if the action of $\text{Pol}(A_\Phi)^\text{can}_1$ on $\text{Aut}(A_\Phi)$-orbits of elements is not equationally affine.

We will need the following results from Ramsey theory.

**Theorem 5.8 (Consequence of Theorem 2.11 from [40]).** Let $\Phi$ be an MMSNP sentence in normal form. Let $A^*$ be the expansion of $A_\Phi$ by all pp-definable relations. Then there exists a countably infinite homogeneous structure $(\mathbb{H}, <)$ whose finite substructures are exactly the finite substructures of structures satisfying $\Phi$ and expanded by all pp-definable relations and by an arbitrary linear order. Moreover, $(\mathbb{H}, <)$ is Ramsey.

**Proof.** See Appendix A in [21] for the translation of Theorem 2.11 from [40] into the MMSNP framework. $\square$

The following proposition is proved in [21 Lemma 7.5] in the case where $m = 2$. We give the proof for the convenience of the reader.

**Proposition 5.9.** Let $\Phi$ be a precolored MMSNP sentence in normal form and let $m \geq 1$. There exist self-embeddings $e_1, \ldots, e_m$ of $A_\Phi$ such that $(e_{i_1}(a_1), \ldots, e_{i_m}(a_m))$ and $(e_{j_1}(b_1), \ldots, e_{j_m}(b_m))$ are in the same orbit under $\text{Aut}(A_\Phi, <)$ provided that:

- $a_k$ and $b_k$ are in the same color for all $k \in \{1, \ldots, m\}$
- $a_k$ and $a_\ell$ are in distinct colors for all $k \neq \ell$,
- $\{i_1, \ldots, i_m\} = \{j_1, \ldots, j_m\} = \{1, \ldots, m\}$. 


Proof. Let \((\mathbb{H}, <)\) be the Ramsey structure from Theorem 5.8. There exists a homomorphism \(h: \mathbb{H} \to A^*\) since the reduct of \(\mathbb{H}\) to the input signature of \(\Phi\) by definition and hence, it admits a homomorphism to \(A^*\); this homomorphism also preserves the additional relations in the signature of \(\mathbb{H}\) and \(A^*\) since these relations are obtained by pp-definitions from relations in the input signature of \(\Phi\). Theorem 2.2 yields that \(h\) can moreover be assumed to be canonical from \(\text{Aut}(\mathbb{H}, <)\) to \(\text{Aut}(A^*, <)\).

Let \(\mathbb{B}\) be the expansion of \(\{1, \ldots, m\} \times A\), the disjoint union of \(m\) copies of \(A\), by all pp-definable relations. Endow \(\mathbb{B}\) with a linear order that is convex with respect to the colors. Since \(\Phi\) is connected, and since \(A\) satisfies \(\Phi\), we have that \(\mathbb{B}\) satisfies \(\Phi\). All the finite substructures of \((\mathbb{B}, <)\) then embed into \((\mathbb{H}, <)\), and by compactness there exists an embedding \(e'\) of \((\mathbb{B}, <)\) into \((\mathbb{H}, <)\). Let \(e_i(x) := (h \circ e')(i, x)\).

To check that these self-embeddings satisfy the required properties, let \(a_1, b_1, \ldots, a_m, b_m\) and \(i_1, j_1, \ldots, i_m, j_m\) be as in the statement. Note that since \(a_k\) and \(b_k\) are in the same color for all \(k\), they are in the same orbit in \(A\) by the fact that \(\Phi\) is precolored \([1]\). They in particular satisfy the same primitive positive formulas, meaning that \(e'\) maps \((i_k, a_k)\) and \((j_k, b_k)\) to elements of \(\mathbb{H}\) that satisfy the same atomic formulas, and that are therefore in the same orbit as \(\mathbb{H}\) is homogeneous. Note moreover that by definition of the order on \(\mathbb{B}\), \((i_k, a_k) < (i_l, a_l)\) if and only if \((j_k, b_k) < (j_l, b_l)\), and that no other atomic relation holds within the tuples \(((i_1, a_1), \ldots, (i_m, a_m))\) and \(((j_1, b_1), \ldots, (j_m, b_m))\). Thus, the required tuples are in the same orbit in \((\mathbb{H}, <)\), by the homogeneity of \((\mathbb{H}, <)\). Since \(h\) is canonical, we obtain that their image under \(h\) is in the same orbit in \((A^*, <)\). \(\square\)

Lemma 5.10. Let \((S, \sim)\) be a subfactor of \(\text{Pol}(A^*_1)\) with \(\text{Aut}(A^*_1)\)-invariant \(\sim\)-classes. Let \(m \geq 2\), and let \(f \in \text{Pol}(A^*_1)\) be such that for all \(a_1, \ldots, a_m \in A\), if the entries of the tuple \((f(a_1, \ldots, a_m), f(a_2, \ldots, a_m, a_1), \ldots, f(a_m, a_1, \ldots, a_{m-1}))\) all belong to different colors, then the tuple intersects at most one \(\sim\)-class. Let \(O_0, \ldots, O_{m-1} \in S\) be pairwise distinct orbits under \(\text{Aut}(A^*_1)\). There exists \(g \in \text{Pol}(A^*_1)\) that is locally interpolated by \(f\) and that satisfies

\[(*) \quad g(O_k, \ldots, O_{k+m-1}) \sim g(O_{j+k}, \ldots, O_{j+k+m-1})\]

for some \(0 \leq k < m\) and \(1 \leq j < m\) (where the indices are computed modulo \(m\)).

Proof. Recall that by \([ii]\) the expansion of \(A^*_1\) by a generic linear order is a Ramsey structure. By Theorem 2.2 \(f\) diagonally interpolates a function \(g \in \text{Pol}(A^*_1)\) with the same properties and which is diagonally canonical with respect to \(\text{Aut}(A^*_1, <)\), and without loss of generality we can therefore assume that \(f\) is itself diagonally canonical.

Let \(e_0, \ldots, e_{m-1}\) be self-embeddings of \(A\) with the properties stated in Proposition 5.9. Consider \(f'(x_0, \ldots, x_{m-1}) := f(e_0x_0, \ldots, e_{m-1}x_{m-1})\), and note that \(f'\) is 1-canonical when restricted to \(m\)-tuples where all entries are in pairwise distinct orbits. Let \(g \in \text{Pol}(A^*_1)\) be a function that is diagonally interpolated by \(f'\) and which is diagonally canonical with respect to \(\text{Aut}(A^*_1, <)\). In particular \(g \in \text{Pol}(A^*_1)\) and \(g(O_k, \ldots, O_{k+m-1})\) and \(f'(O_k, \ldots, O_{k+m-1})\) are in \(S\) and \(\sim\)-equivalent for all \(k\).

As in the proof of Theorem 5.11 there are suitable \(0 \leq k < m\) and \(1 \leq j < m\) such that

\[f(e_kO_k, \ldots, e_{k+m-1}O_{k+m-1}) \sim f(e_{k+j}O_{k+j}, \ldots, e_{k+j+m-1}O_{j+k+m-1})\]
holds, where indices are computed modulo $m$. Then
\[
g(O_k, \ldots, O_{k+m-1}) \sim f(e_0 O_k, \ldots, e_{m-1} O_{k+m-1})
\]
\[\sim f(e_k O_k, \ldots, e_{k+m-1} O_{k+m-1})
\]
\[\sim f(e_{k+j} O_{k+j}, \ldots, e_{k+j+m-1} O_{k+j+m-1})
\]
where the equivalences marked (‡) hold by the fact that $f$ is diagonally canonical with respect to $\text{Aut}(A_\Phi, <)$ and by Proposition 5.9.

Let $\Phi$ be an MMSNP sentence in normal form with input signature $\tau$ and let $\mathcal{F}$ be the associated colored obstruction set. Note that $A_\Phi$ satisfies $\Phi$, so that there exists a coloring of $A_\Phi$ that is $\mathcal{F}$-free. Fix such a coloring $A_\Phi^*$. For each $\tau$-reduct $\mathcal{F}^r$ of a structure $\mathcal{F} \in \mathcal{F}$ of size $k$, fix an enumeration of the domain and assume without loss of generality that the domain is $\{1, \ldots, k\}$. Let $B_\Phi$ be the finite structure whose domain is the set of colors of $\Phi$, with one $k$-ary relation $R_{\mathcal{F}^r}$ for each $\tau$-reduct of an obstruction $\mathcal{F} \in \mathcal{F}$ of size $k$, and where the interpretation of $R_{\mathcal{F}^r}$ in $B_\Phi$ contains all tuples $(C_1, \ldots, C_k)$ of colors such that there exists a homomorphism $h: \mathcal{F}^r \to A_\Phi$ and such that $h(i)$ is in the color $C_i$ in $A_\Phi^*$ for all $i$.

Let $A$ be a relational structure, and let $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ be an instance of CSP($A$). The canonical database of $\mathcal{I}$ is the relational structure $\mathcal{D}$ in the signature of $A$ such that for every relational symbol $R$ of arity $k \geq 1$, a tuple $(v_1, \ldots, v_k) \in \mathcal{V}^k$ is contained in $R^\mathcal{D}$ if $\mathcal{C}$ contains the constraint $\{f \in A^{(v_1, \ldots, v_k)} | f(v_1, \ldots, v_k) \in R^A\}$.

The following theorem gives a characterization of Datalog-rewritability for precolored normal forms. The proof is similar to that of Theorem 5.1.

**Theorem 5.11.** Let $\Phi$ be a precolored connected MMSNP $\tau$-sentence in normal form, and let $\ell$ be the maximum size of a structure in a minimal colored obstruction set associated with $\Phi$. The following are equivalent:

1. $\neg \Phi$ is equivalent to a Datalog program;
2. $\text{Pol}(A_\Phi)$ does not have a uniformly continuous minion homomorphism to an affine clone;
3. The action of $\text{Pol}(A_\Phi)^\text{can}_1$ on $\text{Aut}(A_\Phi)$-orbits of elements is not equationally affine;
4. $\text{Pol}(B_\Phi)$ is not equationally affine;
5. $A_\Phi$ has relational width at most $(2, \max(3, \ell))$.

**Proof.** We prove the implications (1)$\Rightarrow$(2)$\Rightarrow$(3)$\Rightarrow$(4)$\Rightarrow$(5)$\Rightarrow$(1).

(1) implies (2) by the implication from (2) to (3) in Theorem 2.6.

(2) implies (3). We do the proof by contraposition. The proof is essentially the same as in the case of reducts of unary structures (Theorem 5.1). Suppose that $\text{Pol}(A_\Phi)^\text{can}_1 \sim A_\Phi/\text{Aut}(A_\Phi)$ is equationally affine and let $(S, \sim)$ be as given by Proposition 5.5.

Let $m \geq 2$ and let $R$ be an $m$-ary cyclic relation invariant under $\text{Pol}(A_\Phi)$ and containing a tuple $(a_1, \ldots, a_m)$ whose entries are pairwise distinct. By Proposition 4.7, either the linkedness congruence of $R$ defines an approximation of $\sim$, or $R$ contains a pseudoloop modulo $\text{Aut}(A_\Phi)$. In the first case, the approximation is smooth by the same argument as in the proof of Theorem 5.1 and we obtain a uniformly continuous minion homomorphism from $\text{Pol}(A_\Phi)$ to an affine clone by Theorem 2.9.

So let us assume that for all $m \geq 2$, every such relation $R$ contains a pseudoloop. By applying Lemma 4.8 with $P$ being the set of $m$-tuples whose entries belong to pairwise distinct colors, we obtain a polymorphism $f$ such that for all $a_1, \ldots, a_m$, if $f(a_1, \ldots, a_m), \ldots, f(a_m, a_0, \ldots, a_{m-1})$ are in pairwise distinct colors, then they intersect at most one $\sim$-class. As in the proof of Theorem 5.1, pick an arbitrary $a_1 \in S$ such that $[a_1]_\sim$ is not the zero element of the module $S/\sim$. Let $m \geq 2$ be
its order, and let \( O_i \) be the orbit of \( i \cdot [a]_\sim \), for \( i \in \{0, 1, \ldots, m-1\} \). By Lemma 5.10 we obtain \( g \in \Pol(\mathcal{A}_\Phi)^\text{can}_1 \) such that

\[
g(O_{k, \ldots, O_{k+m-1}}) \sim g(O_{j+k}, \ldots, O_{j+k+m-1})
\]

for some \( k \in \{0, \ldots, m-1\} \) and \( j \in \{1, \ldots, m-1\} \). The same computation as in Theorem 5.1 then gives a contradiction and concludes the proof.

(3) implies (4). Since \( \Phi \) is precolored, orbits under \( \Aut(\mathcal{A}_\Phi) \) corresponds to the colors of \( \Phi \), i.e., to the domain of \( \mathbb{B}_\Phi \) (by \([1]\)). Then under the bijection between colors and orbits, \( \Pol(\mathcal{A}_\Phi)^\text{can}_1 \cap \mathcal{A}_\Phi/\Aut(\mathcal{A}_\Phi) \) is a subset of \( \Pol(\mathbb{B}_\Phi) \). Thus if the smaller clone is not equationally affine, the larger clone is also not.

(4) implies (5). Let \( F \) be a minimal colored obstruction set associated with \( \Phi \). Suppose that \( \Pol(\mathbb{B}_\Phi) \) is not equationally affine. Then \( \Pol(\mathbb{B}_\Phi) \) does not have a minion homomorphism to an affine clone and moreover, all colors (i.e., elements of the domain of \( \mathbb{B}_\Phi \)) are pp-definable in \( \mathbb{B}_\Phi \) since \( \Phi \) is precolored. Hence, by the implication from \([1]\) to \([1]\) in Theorem 2.6, \( \mathbb{B}_\Phi \) has relational width at most \((2, 3)\). It is proven in \([21]\) Theorem 7.3 that the following is a reduction from \( \text{CSP}(\mathcal{A}_\Phi) \) to \( \text{CSP}(\mathbb{B}_\Phi) \): given an instance \( I = (\mathcal{V}, \mathcal{C}) \) of \( \text{CSP}(\mathcal{A}_\Phi) \), define an instance \( \mathcal{J} \) of \( \text{CSP}(\mathbb{B}_\Phi) \) with the same variables and which, for every obstruction \( F \in F \) of size \( k \geq 1 \) and for every homomorphism \( h \) from \( F^\tau \) to the canonical database of \( \mathcal{I} \), contains the constraint \( \{ f \in B^{h(1), \ldots, h(k)}_\Phi \mid (f(h(1)), \ldots, f(h(k))) \in R^F \} \). If \( \mathcal{I} \) is equivalent to a non-trivial \((2, \max(3, \ell))\)-minimal instance, then \( \mathcal{J} \) is equivalent to a \((2, 3)\)-minimal instance. It follows that \( \mathcal{A}_\Phi \) has relational width at most \((2, \max(3, \ell))\).

(5) implies (1). By the implication from \([2]\) to \([2]\) in Theorem 2.6, (5) implies that the class of finite structures that do not have a homomorphism to \( \mathcal{A}_\Phi \) is definable by a Datalog program. This class is by construction the class of finite models of \( \neg \Phi \), which proves (1). \( \square \)

The bound from item (5) of Theorem 5.11 can be reached for any \( \ell \geq 3 \) which in particular implies that the relational width of structures \( \mathcal{A}_\Phi \) associated to precolored MMSNP sentences in normal form is not uniformly bounded. Indeed, let \( \ell \geq 3 \), let \( \tau \) consist of a single binary relational symbol \( E \), and let \( \Phi \) be the MMSNP sentence whose colored obstruction set consists of a complete graph on \( \ell \) vertices, all of which belong to a single color. Then the structure \( \mathcal{A}_\Phi \) is the universal homogeneous \( K_\ast \)-free graph, and we showed in Section 5.2 that the exact relational width of \( \mathcal{A}_\Phi \) is \((2, \ell)\).

Note that the characterization in Theorem 5.11 only holds for precolored sentences in normal form. We show below how to characterize Datalog also for arbitrary MMSNP sentences in normal form. The steps involved in the proof to go from normal form to precolored normal form reflect a similar situation as for finite domain CSPs, when going from arbitrary finite structures to their cores: given a finite structure \( \mathcal{A} \) and its core \( \mathbb{B} \), we have that \( \text{CSP}(\mathcal{A}) \) is in Datalog if and only if \( \Pol(\mathbb{B}) \) is not equationally affine if and only if \( \Pol(\mathcal{A}) \) does not admit a minion homomorphism to an affine clone (see Theorem 2.6).

On the algebraic side, we first show in Proposition 5.12 that the existence of a uniformly continuous minion homomorphism from \( \Pol(\mathcal{A}_\Phi) \) or \( \Pol(\mathcal{A}_\Theta) \) to an affine clone does not depend on whether we consider an MMSNP formula in normal form \( \Phi \) or the standard precoloration \( \Theta \) of an equivalent formula in strong normal form. The same step was shown in \([20] [21]\) for the P/NP-complete dichotomy, with \( \mathcal{P} \) replacing affine clones. We then show in Lemma 5.13 that \( \mathbb{B}_\Theta \) is the expansion of the core of \( \mathbb{B}_\Phi \) by singleton unary relations, which allows us to use the characterization of bounded width in Theorem 2.6. Finally we show in Lemma 5.14 that \( \neg \Theta \) is equivalent to a Datalog program if and only if \( \neg \Phi \) is.

**Proposition 5.12.** Let \( \Phi \) be an MMSNP sentence in normal form, let \( \Psi \) be an equivalent formula in strong normal form, and let \( \Theta \) be the standard precoloration of \( \Psi \). There is a uniformly continuous
minion homomorphism from \( \text{Pol}(\mathbb{A}_{\Phi}) \) to an affine clone if and only if there is a uniformly continuous
minion homomorphism from \( \text{Pol}(\mathbb{A}_{\Phi}) \) to an affine clone.

**Proof.** Since \( \Phi \) and \( \Psi \) are equivalent, \( \mathbb{A}_{\Phi} \) and \( \mathbb{A}_{\Psi} \) are homomorphically equivalent (this can be seen using a compactness argument, cf. [21, Theorem 5.15]). Therefore, there exist uniformly continuous
minion homomorphisms \( \text{Pol}(\mathbb{A}_{\Phi}) \to \text{Pol}(\mathbb{A}_{\Phi}) \) and \( \text{Pol}(\mathbb{A}_{\Psi}) \to \text{Pol}(\mathbb{A}_{\Phi}) \).

It is shown in [21 Theorem 6.9] that \( \text{Pol}(\mathbb{A}_{\Phi}) \) has a uniformly continuous
minion homomorphism to \( \text{Pol}(\mathbb{A}_{\Psi}) \) and that \( \text{Pol}(\mathbb{A}_{\Psi}, \neq) \) has a uniformly continuous
minion homomorphism to \( \text{Pol}(\mathbb{A}_{\Phi}) \). Thus, it suffices to show that if \( \text{Pol}(\mathbb{A}_{\Psi}, \neq) \) has a uniformly continuous
minion homomorphism to an affine clone, then so does \( \text{Pol}(\mathbb{A}_{\Phi}) \).

Let \( p \geq 2 \) be prime and let \( R_0 \) and \( R_1 \) be the relations from item (3) in Theorem 2.6. By the equivalence of items (3) and (3') in Theorem 2.6, it is enough to show that if \( (\mathbb{A}_{\Psi}, \neq) \) pp-constructs \( (\mathbb{Z}_p; R_0, R_1) \), then so does \( \mathbb{A}_{\Phi} \).

Suppose that \( (\mathbb{Z}_p; R_0, R_1) \) has a pp-construction in \( (\mathbb{A}_{\Psi}, \neq) \). Thus, there is \( n \geq 1 \) and pp-formulas \( \phi_0(x, y, z), \phi_1(x, y, z) \) defining relations \( S_0, S_1 \) such that \( (\mathbb{A}^n; S_0, S_1) \) and \( (\mathbb{Z}_p; R_0, R_1) \) are homomorphically equivalent; we take \( n \) to be minimal with the property that such pp-formulas exist. Let \( x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n), z = (z_1, \ldots, z_n) \). Since \( R_0 \) and \( R_1 \) are totally symmetric relations (i.e., the order of the entries in a tuple does not affect its membership into any of \( R_0 \) or \( R_1 \)), we can assume that \( S_0 \) and \( S_1 \) are, too, and that the formulas pp-defining them are syntactically invariant under permutation of the blocks of variables \( x, y, \) and \( z \).

We first claim that \( \phi_i \) does not contain any inequality atom \( x_r \neq y_r \) for \( r \in \{1, \ldots, n\} \) (so that by symmetry, also \( y_r \neq z_r \) and \( x_r \neq z_r \) do not appear). Let \( h: (\mathbb{Z}_p; R_0, R_1) \to (\mathbb{A}^n; S_0, S_1) \) be a homomorphism. Since \( (0, 0, 0) \in R_0 \), we have that \( (h(0), h(0), h(0)) \) satisfies \( \phi_0 \), and therefore the listed inequality atoms cannot appear. The same holds for \( \phi_1 \), by considering \( (h(0), h(0), h(1)) \) and its permutations.

Moreover, the only possible equality atoms in \( \phi_0 \) are of the form \( u_r = v_r \) for some \( r \in \{1, \ldots, n\} \) and \( u, v \in \{x, y, z\} \). Indeed, if \( \phi_0 \) contains \( u_r = v_r \) for \( r \neq s \), then by symmetry of \( \phi_0 \) we obtain by transitivity that \( \phi_0 \) entails \( x_r = x_s \). One could obtain a pp-construction with smaller dimension by adding all the equalities \( x_r = x_s, y_r = y_s, z_r = z_s \) to \( \phi_0 \) and \( \phi_1 \) and existentially quantifying the \( s \)th coordinate of the two formulas. The same reasoning applies to \( \phi_1 \).

Let \( \psi_1 \) be the formula obtained from \( \phi_1 \) by removing the possible inequality literals, and let \( T_i \) be defined by \( \psi_i \) in \( \mathbb{A}_{\Phi} \). We claim that \( (\mathbb{A}^n; T_0, T_1) \) and \( (\mathbb{A}^n; S_0, S_1) \) are homomorphically equivalent, which concludes the proof. Since \( \phi_i \) implies \( \psi_i \), we have that \( (\mathbb{A}^n; S_0, S_1) \) is a (non-induced) substructure of \( (\mathbb{A}^n; T_0, T_1) \), and therefore it homomorphically maps to \( (\mathbb{A}^n; T_0, T_1) \) by the identity map. For the other direction, we prove the result by compactness and show that every finite substructure \( \mathbb{B} \) of \( (\mathbb{A}^n; T_0, T_1) \) has a homomorphism to \( (\mathbb{A}^n; S_0, S_1) \). Let \( b^1, \ldots, b^m \) be the elements of \( \mathbb{B} \), where \( b^i = (b^i_1, \ldots, b^i_n) \). Let \( C \) be the \( \tau \)-structure over at most \( n \cdot m \) elements \( \{c^i_r \mid i \in \{1, \ldots, m\}, r \in \{1, \ldots, n\} \} \) such that:

- \( c^i_r \) and \( c^i_s \) are taken to be equal if, and only if, \( b^i_r \) and \( b^i_s \) are connected by a sequence of equalities coming from \( \phi_0 \) and \( \phi_1 \). By the claims above, this is only possible if \( r = s \).
- the relations of \( C \) are defined by pulling back the relations from \( \mathbb{A}_{\Psi} \) under the map \( \pi: c^i_j \mapsto b^i_j \).

Note that \( \pi \) is a homomorphism \( \mathbb{C} \to \mathbb{A}_{\Phi} \), and therefore \( \mathbb{C} \) admits an injective homomorphism \( g \) to \( \mathbb{A}_{\Phi} \). Let \( c^i = (c^i_1, \ldots, c^i_n) \) for \( i \in \{1, \ldots, n\} \). We claim that if \( (b^i_r, b^j_r, b^k_r) \in T_0 \) then \( (g(c^i), g(c^j), g(c^k)) \in S_0 \). Indeed, suppose that \( (b^i_r, b^j_r, b^k_r) \) satisfies \( \psi_0 \). Then by construction \( (c^i, c^j, c^k) \) satisfies \( \psi_0 \) in \( \mathbb{C} \), and thus \( (g(c^i), g(c^j), g(c^k)) \) satisfies \( \psi_0 \) in \( \mathbb{A}_{\Phi} \). Moreover, by injectivity of \( g \), we have \( g(c^i) \neq g(c^j) \) as long as \( r \neq s \). Consider any inequality atom in \( \phi_0 \). By our first claim, it can only be of the form \( x_r \neq y_s \) for some \( r \neq s \), and therefore it is satisfied by \( (g(c^i), g(c^j), g(c^k)) \). Thus, \( (g(c^i), g(c^j), g(c^k)) \) satisfies \( \phi_0 \). The same reasoning for \( \phi_1 \) shows that \( g \) induces a homomorphism \( \mathbb{B} \to (\mathbb{A}^n; S_0, S_1) \) by mapping \( b^i \) to \( g(c^i) \).
Lemma 5.13. Let $\Phi$ be an MMSNP sentence in normal form, let $\Psi$ be an MMSNP sentence in strong normal form equivalent to $\Phi$, and let $\Theta$ be the standard precoloration of $\Psi$. Then $B_\Theta$ is the expansion of the core of $B_\Phi$ by all unary singleton relations.

Proof. Note that since $\Phi$ and $\Psi$ are equivalent, the structures $A_\Phi$ and $A_\Psi$ are homomorphically equivalent. The homomorphisms realizing the equivalence can be taken to be 1-canonical without loss of generality (by (ii), and thus they induce a homomorphic equivalence between $B_\Phi$ and $B_\Psi$. Moreover, $B_\Psi$ is a core: if $h$ is an endomorphism of $B_\Phi$ such that $h(M) = h(M')$ for some colors $M \neq M'$, one can identify in $\Psi$ the colors according to $h$ and obtain an equivalent sentence, contradicting the fact that $\Psi$ is in strong normal form. Recall that the standard precoloration of $\Psi$ is obtained by adding 1-element obstructions for each pair of colors $M \neq M'$. Since in $A_\Theta$ the relation $P_M$ equals the color $M$ (by (i)), these obstructions yield in $B_\Theta$ unary relations containing a single element. Thus, $B_\Theta$ is the expansion of the core of $B_\Phi$ by all unary singleton relations. \square

Lemma 5.14. Let $\Phi$ be an MMSNP sentence in strong normal form and let $\Theta$ be its standard precoloration. If $\neg \Theta$ is equivalent to a Datalog program, so is $\neg \Phi$.

Proof. Note that if $X$ is a structure in the signature of $\Phi$, it satisfies $\Phi$ if and only if its expansion by the relations $P_M$ for every color $M$ interpreted as empty relations satisfies $\Theta$. Thus, we obtain that $\neg \Phi$ is equivalent to a Datalog program obtained by taking a program for $\neg \Theta$ and removing all rules involving the extra predicates $P_M$. \square

Corollary 5.15. Let $\Phi$ be an MMSNP sentence in normal form. Then $\neg \Phi$ is equivalent to a Datalog program if, and only if, $Pol(B_\Phi)$ does not admit a minion homomorphism to an affine clone.

Proof. Let $\Psi$ be an MMSNP sentence in strong normal form equivalent to $\Phi$, and let $\Theta$ be the standard precoloration of $\Psi$. By Lemma 5.13 we have that $B_\Theta$ is the expansion of the core of $B_\Phi$ by all singleton unary relations.

If $Pol(B_\Phi)$ does not admit a minion homomorphism to an affine clone, then $Pol(B_\Theta)$ is not equationally affine by the implication from (3) to (1) in Theorem 2.6 and by Theorem 5.11 $\neg \Theta$ is equivalent to a Datalog program. By Lemma 5.14 we obtain that $\neg \Phi$ is equivalent to a Datalog program.

Conversely, suppose that $Pol(B_\Phi)$ has a minion homomorphism to an affine clone. Then $Pol(B_\Theta)$ is equationally affine by the implication from (1) to (3) in Theorem 2.6, therefore by Theorem 5.11 there is a uniformly continuous minion homomorphism from $Pol(A_\Theta)$ to an affine clone. By Proposition 5.12 there is a uniformly continuous minion homomorphism from $Pol(A_\Phi)$ to an affine clone. Since $A_\Phi$ and $A_\Psi$ are homomorphically equivalent, we obtain a uniformly continuous minion homomorphism from $Pol(A_\Phi)$ to an affine clone. Finally, by the implication from (2) to (3) in Theorem 2.6, this implies that $A_\Phi$ does not have bounded width, i.e., $\neg \Phi$ is not equivalent to a Datalog program by the same reasoning as in the proof of (5) implies (1) in Theorem 5.11. \square

This finally allows us to obtain Theorem 1.3 from the introduction.

Theorem 1.3. The Datalog-rewritability problem for MMSNP is decidable, and is $2NExpTime$-complete.

Proof. Let $\Phi$ be an MMSNP sentence, which is equivalent to a disjunction $\Phi_1 \lor \cdots \lor \Phi_p$ of connected MMSNP sentences, and this decomposition can be computed in exponential time [20, Proposition 3.2]. Each $\Phi_i$ has size polynomial in $\Phi$. Moreover, if $p$ is minimal then $\neg \Phi$ is equivalent to a Datalog program if and only if every $\neg \Phi_i$ is equivalent to a Datalog program (see, e.g., Proposition 3.3 in [21], for a proof of a similar fact). Such a minimal set $\{\Phi_1, \ldots, \Phi_p\}$ of sentences can be computed in exponential time, given $\Phi$ as input. After having computed any set $\{\Phi_1, \ldots, \Phi_p\}$ whose disjunction is equivalent to $\Phi$, it suffices to iterate the following procedure: for any $i, j \in \{1, \ldots, p\}$, check...
whether $\Phi_i$ implies $\Phi_j$ (by Theorem 5.15 in [21], this problem is in NP). If $\Phi_i$ implies $\Phi_j$ for some $i \neq j$, then remove $\Phi_i$ and continue. Otherwise we claim $\{\Phi_1, \ldots, \Phi_p\}$ is minimal. Indeed, no $\Phi_i$ distinct from $\Phi_j, \ldots, \Phi_{j_k}$ implies any disjunction $\Phi_{j_1} \lor \cdots \lor \Phi_{j_k}$: by taking for each $j \in \{j_1, \ldots, j_k\}$ a structure $X_j$ witnessing that $\Phi_i$ does not imply $\Phi_j$ (i.e., $X_j$ satisfies $\Phi_i$ but not $\Phi_{j_j}$), then the disjoint union $X_{j_1} \cup \cdots \cup X_{j_k}$ witnesses that $\Phi_i$ does not imply the disjunction.

Fix $i \in \{1, \ldots, p\}$. Let $\Psi_i$ be a normal form associated with $\Phi_i$, which can be computed in double exponential time, and where $\Psi$ itself can be of size doubly exponential in the size of $\Phi_i$. By Corollary 5.15 $\neg \Psi_i$ is equivalent to a Datalog program if, and only if, $\text{Pol}(\mathbb{B}_{\Phi_i})$ does not admit a minion homomorphism to an affine clone. Deciding this property is in NP [39 Corollary 6.8]. We obtain overall a 2NExpTime algorithm. The complexity lower bound is Theorem 18 in [35].

\[ \text{References} \]

[1] Samson Abramsky, Anuj Dawar, and Pengming Wang. The pebbling comonad in finite model theory. In 32nd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2017, Reykjavik, Iceland, June 20-23, 2017, pages 1–12. IEEE Computer Society, 2017. doi:10.1109/LICS.2017.8005129

[2] Albert Atserias, Andrei A. Bulatov, and Víctor Dalmau. On the power of $\omega$-categorical structures. Logical Methods in Computer Science, 8(1), 2012. doi:10.2168/LMCS-8(1:7)2012

[3] Libor Barto, Jakub Opršal, and Michael Pinsker. The equivalence of two dichotomy conjectures for infinite domain constraint satisfaction problems. In Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science – LICS’17, 2017. Preprint arXiv:1612.07551.

[4] Libor Barto, Michael Kompatscher, Miroslav Olšák, Trung Van Pham, and Michael Pinsker. Equations in oligomorphic clones and the constraint satisfaction problem for $\omega$-categorical structures. Journal of Mathematical Logic, 19(2):#1950010, 2019.

[5] Libor Barto and Marcin Kozik. Absorbing subalgebras, cyclic terms, and the constraint satisfaction problem. Log. Methods Comput. Sci., 8(1), 2012. doi:10.2168/LMCS-8(1:7)2012

[6] Libor Barto and Marcin Kozik. Constraint satisfaction problems solvable by local consistency methods. J. ACM, 61(1):3:1–3:19, 2014. doi:10.1145/2556646

[7] Libor Barto, Jakub Opršal, and Michael Pinsker. The wonderland of reflections. Israel Journal of Mathematics, 223(1):363–398, 2018.

[8] Libor Barto and Michael Pinsker. The algebraic dichotomy conjecture for infinite domain constraint satisfaction problems. In Proceedings of the 31th Annual IEEE Symposium on Logic in Computer Science – LICS’16, pages 615–622, 2016. Preprint arXiv:1602.04353.

[9] Libor Barto and Michael Pinsker. Topology is irrelevant. SIAM Journal on Computing, 49(2):365–393, 2020.

[10] Meghyn Bienvenu, Balder ten Cate, Carsten Lutz, and Frank Wolter. Ontology-based data access: A study through disjunctive datalog, CSP, and MMSNP. ACM Trans. Database Syst., 39(4):33:1–33:44, 2014. doi:10.1145/2661643

[11] Manuel Bodirsky. Cores of Countably Categorical Structures. Logical Methods in Computer Science, Volume 3, Issue 1, January 2007. doi:10.2168/LMCS-3(1:2)2007

[12] Manuel Bodirsky. Ramsey classes: Examples and constructions. In Surveys in Combinatorics, London Mathematical Society Lecture Note Series 424, Cambridge University Press, 2015. Invited survey article for the British Combinatorial Conference: ArXiv:1502.05146.

[13] Manuel Bodirsky. Complexity of Infinite-Domain Constraint Satisfaction. Lecture Notes in Logic. Cambridge University Press, 2021. doi:10.1017/9781108737594

[14] Manuel Bodirsky, Hubie Chen, and Michael Pinsker. The reducets of equality up to primitive positive interdefinability. Journal of Symbolic Logic, 75(4):1249–1292, 2010.

[15] Manuel Bodirsky and Víctor Dalmau. Datalog and constraint satisfaction with infinite templates. J. Comput. Syst. Sci., 79(1):79–100, 2013. A conference version appeared in the Proceedings of the 34th Symposium on Theoretical Aspects of Computer Science (STACS 2006), pages 646–659. doi:10.1016/j.jcss.2012.05.012
[18] Manuel Bodirsky and Martin Grohe. Non-dichotomies in constraint satisfaction complexity. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7-11, 2008, Proceedings, Part II - Track B: Logic, Semantics, and Theory of Programming & Track C: Security and Cryptography Foundations, volume 5126 of Lecture Notes in Computer Science, pages 184–196. Springer, 2008. doi:10.1007/978-3-540-70583-3_16

[19] Manuel Bodirsky and Jan Kára. The complexity of equality constraint languages. Theory of Computing Systems, 3(2):136–158, 2008. A conference version appeared in the proceedings of Computer Science Russia (CSR'06).

[20] Manuel Bodirsky, Florent R. Madelaine, and Antoine Mottet. A universal-algebraic proof of the complexity dichotomy for Monotone Monadic SNP. In Anuj Dawar and Erich Grädel, editors, Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018, pages 105–114. ACM, 2018. doi:10.1145/3209108.3209156

[21] Manuel Bodirsky, Florent R. Madelaine, and Antoine Mottet. A Proof of the Algebraic Tractability Conjecture for Monotone Monadic SNP. SIAM Journal on Computing, 50(4):1359–1409, 2021. doi:10.1137/19M128466X

[22] Manuel Bodirsky, Barnaby Martin, and Antoine Mottet. Discrete temporal constraint satisfaction problems. J. ACM, 65(2):9:1–9:41, 2018. doi:10.1145/3154832

[23] Manuel Bodirsky, Barnaby Martin, Michael Pinsker, and András Pongrácz. Constraint satisfaction problems for reducts of homogeneous graphs. SIAM J. Comput., 48(4):1224–1264, 2019. A conference version appeared in the Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, pages 119:1–119:14. doi:10.1137/16M1082974

[24] Manuel Bodirsky and Antoine Mottet. Reducts of finitely bounded homogeneous structures, and lifting tractability from finite-domain constraint satisfaction. In Martin Grohe, Eric Koskinen, and Natarajan Shankar, editors, Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS ’16, New York, NY, USA, July 5-8, 2016, pages 623–632. ACM, 2016. doi:10.1145/2933575.2934515

[25] Manuel Bodirsky and Antoine Mottet. A dichotomy for first-order reducts of unary structures. Log. Methods Comput. Sci., 14(2), 2018. doi:10.23638/LMCS-14(2:13)2018

[26] Manuel Bodirsky, Antoine Mottet, Miroslav Olšák, Jakub Opršal, Michael Pinsker, and Ross Willard. Topology is relevant (in a dichotomy conjecture for infinite-domain constraint satisfaction problems). In 34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019, Vancouver, BC, Canada, June 24-27, 2019, pages 1–12. IEEE, 2019. doi:10.1109/LICS.2019.8785883

[27] Manuel Bodirsky, Antoine Mottet, Miroslav Olšák, Jakub Opršal, Michael Pinsker, and Ross Willard. ω-categorical structures avoiding height 1 identities. Transactions of the AMS, 374:327–350, 2021. doi:10.1090/tran/8179

[28] Manuel Bodirsky and Jaroslav Nešetřil. Constraint satisfaction with countable homogeneous templates. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7-11, 2008, Proceedings, Part II - Track B: Logic, Semantics, and Theory of Programming & Track C: Security and Cryptography Foundations, volume 5126 of Lecture Notes in Computer Science, pages 184–196. Springer, 2008. doi:10.1007/978-3-540-70583-3_16

[29] Manuel Bodirsky and Jaroslav Nešetřil. Constraint satisfaction for Monotone Monadic SNP. SIAM Journal on Computing, 50(4):1359–1409, 2021. doi:10.1137/19M128466X

[30] Pierre Bourhis and Carsten Lutz. Containment in monadic disjunctive datalog, MMSNP, and expressive description logics. In Chitta Baral, James P. Delgrande, and Frank Wolter, editors, Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016, pages 207–216. AAAI Press, 2016. URL: http://www.aaai.org/ocs/index.php/KR/KR16/paper/view/12847
34 ANTOINE MOTTET, TOMÁŠ NAGY, MICHAEL PINSKER, AND MICHAL WRONA

[57] Michael Pinsker. Current challenges in infinite-domain constraint satisfaction: Dilemmas of the infinite sheep. In 2022 IEEE 52nd International Symposium on Multiple-Valued Logic (ISMVL), pages 80–87, Los Alamitos, CA, USA, may 2022. IEEE Computer Society. URL: https://doi.ieeecomputersociety.org/10.1109/ISMVL52857.2022.00019, doi:10.1109/ISMVL52857.2022.00019

[58] Michael Pinsker, Pierre Gillibert, and Julius Jonušas. Pseudo-loop conditions. Bulletin of the London Mathematical Society, 51(5):917–936, 2019.

[59] Matthew A. Valeriote. A subalgebra intersection property for congruence distributive varieties. Canadian Journal of Mathematics, 61(2):451–464, 2009. doi:10.4153/CJM-2009-023-2.

[60] Michał Wrona. On the relational width of first-order expansions of finitely bounded homogeneous binary cores with bounded strict width. In Holger Hermanns, Lijun Zhang, Naoki Kobayashi, and Dale Miller, editors, LICS ’20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020, pages 958–971. ACM, 2020. doi:10.1145/3373718.3394781

[61] Michał Wrona. Relational width of first-order expansions of homogeneous graphs with bounded strict width. In Christophe Paul and Markus Bläser, editors, 37th International Symposium on Theoretical Aspects of Computer Science, STACS 2020, March 10-13, 2020, Montpellier, France, volume 154 of LIPIcs, pages 39:1–39:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.STACS.2020.39

[62] Michał Wrona. Quasi directed jonsson operations imply bounded width (for fo-expansions of symmetric binary cores with free amalgamation), 2024. arXiv:2402.16540

[63] Dmitriy Zhuk. A proof of CSP dichotomy conjecture. In Chris Umans, editor, 58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017, pages 331–342. IEEE Computer Society, 2017. doi:10.1109/FOCS.2017.38

[64] Dmitriy Zhuk. A proof of the CSP dichotomy conjecture. J. ACM, 67(5):30:1–30:78, 2020. doi:10.1145/3402029

[65] Dmitriy Zhuk. Strong subalgebras and the constraint satisfaction problem. J. Multiple Valued Log. Soft Comput., 36(4-5):455–504, 2021.

HAMBURG UNIVERSITY OF TECHNOLOGY, RESEARCH GROUP ON THEORETICAL COMPUTER SCIENCE, GERMANY
Email address: antoine.mottet@tuhh.de
URL: http://amottet.github.io/~mottet/

INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, FG ALGEBRA, TU WIEN, AUSTRIA AND THEORETICAL COMPUTER SCIENCE DEPARTMENT, JAGIELLONIAN UNIVERSITY, POLAND

Email address: tomas.nagy@email.com
URL: https://nagyto.github.io/

INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, FG ALGEBRA, TU WIEN, AUSTRIA
Email address: marula@gmx.at
URL: http://dmg.tuwien.ac.at/pinsker/

THEORETICAL COMPUTER SCIENCE DEPARTMENT, JAGIELLONIAN UNIVERSITY, POLAND
Email address: michal.wrona@uj.edu.pl
URL: https://www.tcs.uj.edu.pl/wrona
Response to comments

We would like to thank all reviewers for their thorough reading of the paper and for all their comments. We have implemented them, please find our responses to all particular comments below.

Comments from the Editor

• Thank you for doing a good job incorporating the referees’ comments. The introduction, in particular, is much improved all the way up to the last paragraph before Section 1.1 (in that paragraph, ω-categoricity pops up without introduction, intuition, or context). I suggest that you re-write that paragraph and carefully check all of Section 1.1 to improve the accessibility. The referees, in the previous round, made some specific suggestions on improving Section 1.1., which you incorporated, but I believe that more can and should be done there. In particular, it would be good to have remarks, which involve technical details or notions that an outsider is not expected to know, clearly designated as such.

We have adjusted Section 1.1 as suggested, we have added some remarks for experts and we have expanded the motivation for the main notions.

• ω-categoricity appears twice in Section 1.1. Once in line 2, where you say that this what you want to study, but, for any outsider, this does not match the rest of the intro. The second time is in the sentence before Thm. 1.2, where the reference to this notion again does not work well with the rest of the intro. You need to either say more about omega-categoricity and put it in the context or else collect all remarks about it in one place and clearly designate it as intended for experts.

We added an explanation of the notion of ω-categoricity.

• Theorem 1.1 needs to be given some context - at the moment, the only motivation for it is an open question asked elsewhere, but of course not every question asked elsewhere is a good motivation. Also, this theorem and the paragraph preceding it look rather disconnected from the rest of the intro.

We added an explanation for the origin of the open problem and the applications in ontology-mediated querying.

• You give some good short intuition behind the technical notions such as homogeneity and boundedness - please do the same for the notion of a first-order reduct.

We gave also some intuition behind the notion of a first-order reduct.

• The technical conditions appearing in Theorem 1.2 can be related to the finite-domain case - the polymorphisms appearing in both cases are the same, but certain technical adjustments must be made for the infinite case. At the moment, too little help is given to an outsider trying to make sense of the statement of the theorem.

We clarified the relation to the finite-domain case.
Comments from Reviewer #1

• The first use of the term “unary structure” after the abstract is in page 3. It wouldn’t hurt to reference to [24] at that point, just like you give a reference to [41] for MMSNP in the same sentence. The next occurrence of the term “unary structure” is in page 4 as “unary structures mentioned above”. This sounds as it is referring to the finite structures with all unary singleton relations of the preceding paragraph, when what is actually meant is any finite or infinite structure over a finite unary vocabulary. This confusion is aggravated by the fact that in that reference [24] appears after “numerous structures studied in the literature” and not after “as well as of the unary structures mentioned above” as it should. Finally, the fact that the class of unary structures is a subclass of the class of first-order reducts of \((\mathbb{Z}; +, \leq)\), \((\mathbb{Q}; +, \leq, 1)\) or \((\mathbb{R}; +, \times)\) is not entirely obvious (at least for Presburger \((\mathbb{Z}; +, \leq)\)). Please refer again to [24] where an explanation of this can be found, or give your own explanation if it is shorter.

We corrected the references and clarified what we mean by unary structures.

Comments from Reviewer #4

• Provide explicit statements for all outsourced results mentioned in point 1 above.

  For all the instances mentioned by the reviewer, we provided explicit statements or we explained or changed the respective parts.

• Provide a brief proof overview before technical and long proofs, e.g. in Lemma 4.6, Proposition 5.4, and Theorem 5.8.

  We added overviews of the proofs.

• In the instances of wordy text mentioned above, add an “i.e.” to elaborate.

  We added explanation by symbolic notation to all instances mentioned above.

• P.3: 1st section sign: “On the one hand.. on the other hand” is typically used to contrast two things

  The formulation was corrected.

• P.4 1st section sign: “its algorithmic meaning there is not guaranteed”. I really don’t get what you mean by the “algorithmic meaning” of something.

  We changed the formulation, emphasizing that \(\omega\)-categoricity is a natural sufficient condition for the consistency algorithm to terminate.

• P. 8, Section 2.3: You define CSP instances but you never define what a solution of an instance is.

  The definition was added.

• P.13: “Finally this allows us to prove Theorem 1.2 from the introduction”. State Theorem 1.2 once more, as the proof refers to the “first item” and the “second item” from the same theorem and the reader is forced to go back and forth.

  We added the statement of the theorem.

• P.13, Def 4.1: I suppose you want \(i_1, \ldots, i_n\) to be increasing, no?

  The definition makes sense also for non-increasing tuples, so we kept it as is.
• P.14, Def. 4.3: I believe you mean “\(f(a_1, \ldots, a_n) \in B\)”.
  
  Corrected.

• p.17, before section 5: “Since \(D\) is a clone of polymorphisms, it is topologically closed”. This is not totally trivial.
  
  *We have replaced the topological argument by a proof that assumes less knowledge about clones of polymorphisms of \(\omega\)-categorical structures.*

• You mention Ramsey structures several times e.g. p.7, 19, 23, 24, so you might as well provide a definition.
  
  *The definition was added on page 7.*

• P.22, 2nd section sign: “As for graphs” \(\rightarrow\) “As in the case of graphs”
  
  Corrected.

• P.22: Define “nct”
  
  *This was a typo, we corrected it.*
This figure "universe.jpg" is available in "jpg" format from:

http://arxiv.org/ps/2102.07531v3