An Order out of Nowhere: A New Algorithm for Infinite-Domain CSPs

Antoine Mottet
Research Group on Theoretical Computer Science, Hamburg University of Technology, Germany

Tomáš Nagy
Theoretical Computer Science Department, Jagiellonian University, Kraków, Poland

Michael Pinsker
Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Austria

Abstract

We consider the problem of satisfiability of sets of constraints in a given set of finite uniform hypergraphs. While the problem under consideration is similar in nature to the problem of satisfiability of constraints in graphs, the classical complexity reduction to finite-domain CSPs that was used in the proof of the complexity dichotomy for such problems cannot be used as a black box in our case. We therefore introduce an algorithmic technique inspired by classical notions from the theory of finite-domain CSPs, and prove its correctness based on symmetries that depend on a linear order that is external to the structures under consideration. Our second main result is a P/NP-complete complexity dichotomy for such problems over many sets of uniform hypergraphs. The proof is based on the translation of the problem into the framework of constraint satisfaction problems (CSPs) over infinite uniform hypergraphs. Our result confirms in particular the Bodirsky-Pinsker conjecture for CSPs of first-order reducts of some homogeneous hypergraphs. This forms a vast generalization of previous work by Bodirsky-Pinsker (STOC'11) and Bodirsky-Martin-Pinsker-Pongrácz (ICALP'16) on graph satisfiability.

2012 ACM Subject Classification Theory of computation → Logic

Keywords and phrases Constraint Satisfaction Problems, Hypergraphs, Polymorphisms

Digital Object Identifier 10.4230/LIPIcs.ICALP.2024.148

Category Track B: Automata, Logic, Semantics, and Theory of Programming

Related Version *Full Version*: `https://doi.org/10.48550/arXiv.2301.12977` [29]

Funding *Tomáš Nagy*: This research was funded in whole or in part by the Austrian Science Fund (FWF) [P 32337, I 5948]. This research was funded in whole or in part by National Science Centre, Poland 2021/03/Y/ST6/00171. For the purpose of Open Access, the authors have applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.
Michael Pinsker: This research was funded in whole or in part by the Austrian Science Fund (FWF) [P 32337, I 5948]. For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission. This research is also funded by the European Union (ERC, POCOCOP, 101071674). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1 Introduction

In [15], Bodirsky and the third author introduced the computational problem Graph-SAT as a generalization of systematic restrictions of the Boolean satisfiability problem studied by Schaefer [36]. A *graph formula* is a formula formed from the atomic formulas $E(x, y)$ and $x = y$ using negation, conjunction and disjunction, where E is interpreted as the edge relation of a simple undirected graph. Given a finite set Ψ of graph formulas, the *graph satisfiability problem* Graph-SAT(Ψ) gets as an input a finite set $\mathcal{V}$ of variables and a graph formula $\Phi = \phi_1 \wedge \cdots \wedge \phi_n$, where every ϕ_i is obtained from a formula $\psi \in \Psi$ by substituting the variables of ψ by variables from $\mathcal{V}$; the goal is to decide the existence of a graph satisfying Φ. Any instance of a Boolean satisfiability problem can indeed easily be reduced to a problem of this form, roughly by replacing Boolean variables by pairs of variables which are to be assigned vertices in a graph, and by translating the potential Boolean values 0 and 1 into the non-existence or existence of an edge between these two variables. The main result of [15] states that this computational problem is either solvable in polynomial time or is NP-complete. This can be put in contrast with the theorem of Ladner [28] stating that if P $\neq$ NP, then there exist computational problems that are neither solvable in polynomial-time nor NP-complete. Similar *dichotomy theorems* have been established for related problems concerning the satisfaction of constraints by linear orders [8], partially ordered sets [26], tournaments [30], or phylogenetic trees [6]. It is conjectured that such dichotomies defying Ladner's theorem are common; we refer to Section 1.2 and Conjecture 4 for a precise statement.

In order to develop our understanding of such natural generalizations of the classical Boolean satisfiability problem, we consider in this article the complexity of Graph-SAT where graph formulas are replaced by *ℓ-hypergraph formulas* for some fixed $\ell \geq 2$. More precisely, we consider formulas where E is an ℓ-ary symbol denoting the edge relation of an ℓ-uniform undirected hypergraph; in the following, since all our hypergraphs are uniform and undirected, we simply write *ℓ-hypergraph.* The problem ℓ-Hypergraph-SAT is then defined in the same way as the problem Graph-SAT above. We also study the complexity of the natural variant of the ℓ-Hypergraph-SAT problem, investigated in [12] for the special case of graphs, where we ask for the existence of a satisfying hypergraph that belongs to a prescribed set $\mathcal{K}$ of finite ℓ-hypergraphs. This corresponds to imposing structural restrictions on the possible satisfying hypergraph solutions. For example, it is natural to ask for the existence of a solution in the class $\mathcal{K}_r^\ell$ of all finite ℓ-hypergraphs omitting a generalized clique on r vertices. We use the notation ℓ-Hypergraph-SAT($\Psi, \mathcal{K}$) to denote this problem.

Surprisingly, it turns out that ℓ-hypergraph problems behave very differently from the corresponding graph problems, requiring in particular genuinely novel algorithmic methods to handle them. A natural attempt to solve hypergraph satisfiability in polynomial time is to use a generic reduction to constraint satisfaction problems (CSPs) whose domain consists of the hypergraphs with at most ℓ elements [13]. While this reduction always works in the (graph) case of $\ell = 2$ [30], it can happen for $\ell > 2$ that the resulting finite-domain CSP is an NP-complete problem, although the original hypergraph satisfiability problem is solvable in polynomial time. Our main result, Theorem 8, is an algorithm running in polynomial time and solving the ℓ-Hypergraph-SAT($\Psi, \mathcal{K}$) problem under some general algebraic assumptions.

The next example illustrates that in this setting our algorithm is strictly more powerful than the reduction of [13].

▶ **Example 1.** Let $\ell = 3$, let ψ be a formula with 4 free variables that holds precisely for the hypergraphs in Figure 1, and let Ψ be the set consisting of ψ. This is an example where the reduction to the finite from [13] cannot be applied to prove the tractability of ℓ-Hypergraph-SAT(Ψ). However, the problem is solvable in polynomial time, as it can be solved by the algorithm introduced in Section 3.

Building on Theorem 8, our second contribution is a full complexity dichotomy for the problems ℓ-Hypergraph-SAT($\Psi, \mathcal{K}$) where $\mathcal{K}$ is the class $\mathcal{K}_{\text{all}}^\ell$ of all finite ℓ-hypergraphs, or $\mathcal{K}_r^\ell$.

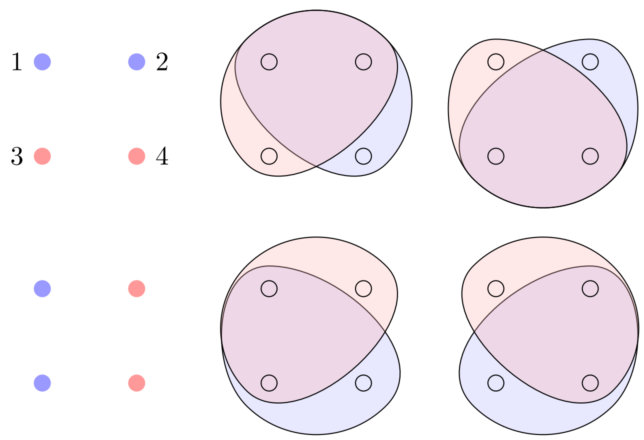

Figure 1 Hypergraphs on at most 4 elements satisfying ψ in Example 1. The vertices are labeled to represent each of the four free variables of ψ (only shown on one of the hypergraphs for readability). The two 3-hypergraphs on the left have 2 vertices, and the color coding denotes vertices that are equal. The other four 3-hypergraphs have four vertices each and precisely two hyperedges.

▶ **Theorem 2.** *Let $\ell \geq 3$, let $\mathcal{K}$ be either the class $\mathcal{K}^\ell_{\text{all}}$ of all finite ℓ-hypergraphs or the class $\mathcal{K}^\ell_r$ for some $r > \ell$, and let Ψ be a set of ℓ-hypergraph formulas. Then ℓ-Hypergraph-SAT$(\Psi, \mathcal{K})$ is either in P, or it is NP-complete. Moreover, given Ψ, one can algorithmically decide which of the cases holds.*

In fact, our polynomial-time algorithm in Theorem 8 solves the hypergraph satisfiability problem for all classes $\mathcal{K}$ of hypergraphs satisfying certain assumptions that we introduce in Section 3. Likewise, our results imply a dichotomy result as in Theorem 2 for every class $\mathcal{K}$ satisfying certain structural assumptions. For more details, see the full version of the article [29].

1.1 Connection to Constraint Satisfaction Problems

The *constraint satisfaction problem* with template $\mathbb{A} = (A; R_1, \dots, R_n)$ is the computational problem CSP$(\mathbb{A})$ of deciding, given an instance with variables $\mathcal{V}$ and constraints $\phi(x_{i_1}, \dots, x_{i_r})$ with $\phi \in \{R_1, \dots, R_n\}$ and $x_{i_1}, \dots, x_{i_r} \in \mathcal{V}$, whether there exists an assignment $f\colon \mathcal{V} \to A$ that satisfies all the constraints.

Note how the problem ℓ-Hypergraph-SAT(Ψ) is similar in nature to a constraint satisfaction problem, where the difference lies in the fact that we are not asking for a labelling of variables to elements of a structure $\mathbb{A}$, but rather for a consistent labelling of ℓ-tuples of variables to a finite set describing all the possible ℓ-hypergraphs on at most ℓ elements. For example, in the case of $\ell = 2$, this set contains 3 elements (for the graph on a single vertex, and the two undirected graphs on 2 vertices), while for $\ell = 3$ this set contains 6 elements (there is one labeled 3-hypergraph on a single element, three on 2 elements, and two on 3 elements).

It was already noticed in [15] that it is possible to design a structure $\mathbb{A}$ (which is necessarily infinite) such that Graph-SAT(Ψ) is equivalent to CSP$(\mathbb{A})$, and this observation also carries out to the hypergraph setting as follows.

Fix $\ell \geq 3$ and a class $\mathcal{K}$ of finite ℓ-hypergraphs. Let us assume that $\mathcal{K}$ is an *amalgamation class*: an isomorphism-closed class that is closed under induced sub-hypergraphs and with the property that for any two hypergraphs $\mathbb{H}_1, \mathbb{H}_2 \in \mathcal{K}$ having a common hypergraph $\mathbb{H}$ as intersection, there exists $\mathbb{H}' \in \mathcal{K}$ and embeddings of $\mathbb{H}_1, \mathbb{H}_2$ into $\mathbb{H}'$ which agree on $\mathbb{H}$. A classical result of Fraïssé [22] yields that there exists an infinite limit hypergraph $\mathbb{H}_{\mathcal{K}}$, called the *Fraïssé limit of* $\mathcal{K}$, with the property that the finite induced sub-hypergraphs of $\mathbb{H}_{\mathcal{K}}$ are precisely the hypergraphs in $\mathcal{K}$. Moreover, this limit can be taken to be *homogeneous*, i.e., highly symmetric in a certain precise sense – see Section 2 for precise definitions of these concepts. If Ψ is a set of ℓ-hypergraph formulas, then it defines in $\mathbb{H}_{\mathcal{K}}$ a set of relations that one can view as a CSP template $\mathbb{A}_{\mathcal{K},\Psi}$. It follows that the problem ℓ-Hypergraph-SAT$(\mathcal{K}, \Psi)$ is precisely the same as CSP$(\mathbb{A}_{\mathcal{K},\Psi})$. The assumption that $\mathcal{K}$ is an amalgamation class is rather mild and is for example fulfilled by the classes of interest for Theorem 2, namely by the class $\mathcal{K}^{\ell}_{\mathrm{all}}$ of all finite ℓ-hypergraphs, or for any $\ell < r$ by the class $\mathcal{K}^{\ell}_{r}$. We refer to [1] for a discussion of amalgamation classes of 3-hypergraphs; importantly, while in the case of $\ell = 2$ all such classes are known [27], it seems very difficult to obtain a similar classification in general since there are uncountably-many such classes already for $\ell = 3$. The latter fact obliges us to build on and refine abstract methods rather than relying on the comfort of a classification in our general dichotomy result, which contrasts with the approach for graphs in [12].

Using the reformulation of ℓ-hypergraph problems as constraint satisfaction problems, we show that the border between tractability and NP-hardness in Theorem 2 can be described algebraically by properties of the *polymorphisms* of the structures $\mathbb{A}_{\mathcal{K},\Psi}$, i.e., by the functions preserving all relations of $\mathbb{A}_{\mathcal{K},\Psi}$. This implies, in particular, the above-mentioned decidability of this border. Roughly speaking, the tractable case corresponds to the CSP template enjoying some non-trivial algebraic invariants in the form of polymorphisms, whereas the hard case is characterized precisely by the absence of such invariants.

▶ **Theorem 3.** *Let $\ell \geq 3$, let $\mathcal{K}$ be either the class $\mathcal{K}^{\ell}_{\mathrm{all}}$ of all finite ℓ-hypergraphs or the class $\mathcal{K}^{\ell}_{r}$ for some $r > \ell$, and let Ψ be a set of ℓ-hypergraph formulas. Then precisely one of the following applies.*

1. *The clone of polymorphisms of $\mathbb{A}_{\mathcal{K},\Psi}$ has no uniformly continuous minion homomorphism to the clone of projections $\mathscr{P}$, and* CSP$(\mathbb{A}_{\mathcal{K},\Psi})$ *is in* P.
2. *The clone of polymorphisms of $\mathbb{A}_{\mathcal{K},\Psi}$ has a uniformly continuous minion homomorphism to the clone of projections $\mathscr{P}$, and* CSP$(\mathbb{A}_{\mathcal{K},\Psi})$ *is* NP*-complete.*

The algebraic assumptions in the second item of the theorem correspond to the clone of polymorphisms being trivial in a certain sense (i.e., containing only polymorphisms that imitate the behaviour of projections when restricted to a certain set). For the precise definitions, see [5].

1.2 Related work on constraint satisfaction problems

In the framework of CSPs, it is natural to consider not only classes of finite ℓ-hypergraphs but also classes of different finite structures in a fixed relational signature. If such a class $\mathcal{K}$ is an amalgamation class, then there exists a countably infinite homogeneous structure $\mathbb{B}_{\mathcal{K}}$ whose finite substructures are precisely the structures in $\mathcal{K}$. However, CSP$(\mathbb{B}_{\mathcal{K}})$ is not guaranteed to be contained in the complexity class NP since the class $\mathcal{K}$ does not have to be algorithmically enumerable (as mentioned above, there are uncountably many amalgamation classes of 3-hypergraphs; hence there exists such a $\mathcal{K}$ such that CSP$(\mathbb{B}_{\mathcal{K}})$ is undecidable). A natural way of achieving the algorithmical enumerability of $\mathcal{K}$ is to require that there exists a natural number $b_{\mathcal{K}}$ such that a structure is contained in $\mathcal{K}$ if, and only if, all its substructures of size at most $b_{\mathcal{K}}$ are in $\mathcal{K}$. In this case, we say that $\mathcal{K}$ (or its Fraïssé limit $\mathbb{B}_{\mathcal{K}}$) is *finitely bounded.* For every set Ψ of formulas in the language of the structures at hand, one then gets as in the previous section a structure $\mathbb{A}_{\mathcal{K},\Psi}$ whose domain is the same as $\mathbb{B}_{\mathcal{K}}$ and whose relations are definable in first-order logic from the relations of $\mathbb{B}_{\mathcal{K}}$ – we say that $\mathbb{A}_{\mathcal{K},\Psi}$ is a *first-order reduct* of $\mathbb{B}_{\mathcal{K}}$.

Thus, for every set Ψ of formulas and every finitely bounded amalgamation class $\mathcal{K}$, the generalized satisfiability problem parameterized by Ψ and $\mathcal{K}$ is the CSP of a first-order reduct $\mathbb{A}_{\mathcal{K},\Psi}$ of a finitely bounded homogeneous structure. It is known that the complexity of the CSP over any such template depends solely on the polymorphisms of $\mathbb{A}_{\mathcal{K},\Psi}$ [16]. This motivates the following conjecture generalizing the dichotomy for Graph-SAT which was formulated by Bodirsky and Pinsker in 2011 (see [17]). The modern formulation of the conjecture based on recent progress [2, 3, 5] is the following:

▶ **Conjecture 4.** *Let $\mathbb{A}$ be a CSP template which is a first-order reduct of a finitely bounded homogeneous structure. Then one of the following applies.*

1. *The clone of polymorphisms of $\mathbb{A}$ has no uniformly continuous minion homomorphism to the clone of projections $\mathscr{P}$, and* CSP($\mathbb{A}$) *is in* P.
2. *The clone of polymorphisms of $\mathbb{A}$ has a uniformly continuous minion homomorphism to the clone of projections $\mathscr{P}$, and* CSP($\mathbb{A}$) *is* NP*-complete.*

It follows that Theorem 3 is a special case of Conjecture 4. It is known that if the clone of polymorphisms of any CSP template within the range of Conjecture 4 has a uniformly continuous minion homomorphism to $\mathscr{P}$, then the CSP of such template is NP-hard [5]. Already before Conjecture 4 was introduced, a similar conjecture was formulated by Feder and Vardi [23] for CSPs over templates with finite domains and it was confirmed independently by Bulatov and Zhuk [19, 39, 40] recently. Conjecture 4 itself has been confirmed for many subclasses: for example for CSPs of all structures first-order definable in finitely bounded homogeneous graphs [15, 12], in $(\mathbb{Q};<)$ [8], in any unary structure [14], in the random poset [26], in the random tournament [30], or in the homogeneous branching C-relation [6], in ω-categorical monadically stable structures [18], as well as for all CSPs in the class MMSNP [11], and for CSPs of representations of some relational algebras [9, 10].

1.3 Novelty of the methods and significance of the results

We prove that under the algebraic assumption in item (1) of Theorem 3, $\mathbb{A}_{\mathcal{K},\Psi}$ admits non-trivial symmetries that can be seen as operations acting on the set of *linearly ordered* ℓ-hypergraphs with at most ℓ elements. We moreover know from [34] that the introduction of a linear order "out of nowhere" is unavoidable, in the sense that the symmetries of $\mathbb{A}_{\mathcal{K},\Psi}$ acting on unordered ℓ-hypergraphs can be trivial even if CSP($\mathbb{A}_{\mathcal{K},\Psi}$) is solvable in polynomial time. This is rather surprising (in model-theoretic terms, hypergraphs form a class having the *non-strict order property*, and thus have no ability to encode linear orders) and is to date the only example of such a phenomenon. As a consequence, the aforementioned "reduction to the finite" introduced in [13], which is enough to prove the tractability part of most of the complexity dichotomies mentioned in the previous section, cannot be used in the hypergraph satisfiability setting. In order to prove the tractability part of Theorem 3, we thus introduce new algorithmic techniques inspired by results in the theory of constraint satisfaction problems with *finite* domains, in particular by absorption theory [4] and Zhuk's theory [39, 40]. More precisely, let $\mathcal{I}$ be an instance of CSP($\mathbb{A}_{\mathcal{K},\Psi}$). Our algorithm transforms $\mathcal{I}$ into an equi-satisfiable instance $\mathcal{I}'$ that is sufficiently locally consistent, such that the solution set of a certain relaxation of $\mathcal{I}'$ does not imply any restrictions on the solution set of the whole instance, and that satisfies an additional condition resembling Zhuk's notion of irreducibility [39, 40]. We then prove that any non-trivial instance satisfying those properties has an injective solution. This step is to be compared with the case of *absorbing reductions* in Zhuk's algorithm. The existence of an injective solution can then be checked by the aforementioned reduction to the finite from [13, 14]. In this way, we resolve the trouble with hitherto standard methods pointed out in [34]. A positive resolution of the general Conjecture 4 will likely have to proceed in a similar spirit, albeit at a yet higher level of sophistication. In the case of graphs, the above described algorithm is not necessary since every instance can be immediately reduced to a finite-domain CSP by the black box reduction.

We use the recently developed theory of smooth approximations [30] to prove the dichotomy, i.e., that $\mathbb{A}_{\mathcal{K},\Psi}$ satisfies one of the two items of Theorem 3, for all Ψ. The classification of the complexity of graph-satisfiability problems from [15] used a demanding case distinction over the possible automorphisms groups of the structures $\mathbb{A}_{\mathcal{K},\Psi}$ (where Ψ is a set of graph formulas, and $\mathcal{K}$ is the class of all finite simple undirected graphs) – it was known previously that there are exactly 5 such groups [37]. Our result relies neither on such a classification of the automorphism groups of the structures under consideration, nor on the classification of the hypergraphs of which they are first-order reducts; as mentioned above, no such classification is available. While Thomas [38] obtained a classification of the mentioned automorphism groups for every fixed ℓ and for $\mathcal{K}$ consisting of all finite ℓ-hypergraphs, this number grows with ℓ and makes an exhaustive case distinction impossible. To overcome the absence of such classifications, we rely on the scalability of the theory of smooth approximations, i.e., on the fact that the main results of the theory can be used without knowing the base structures under consideration. This was claimed to be one of the main contributions of this theory; Theorem 3 and its generalization [29, Theorem 21] are the first complexity classification using smooth approximations that truly exemplifies this promise.

1.4 Bonus track: local consistency

Our structural analysis of ℓ-hypergraph problems allows us to obtain as an easy consequence a description of the hypergraph satisfiability problems ℓ-Hypergraph-SAT$(\Psi, \mathcal{K})$ that are solvable by local consistency methods, assuming that Ψ contains the atomic formula E. Similar classifications, and general results on the amount of local consistency needed in those cases, had previously obtained for various other problems (including Graph-SAT problems) which can be modeled as CSPs of first-order reducts of finitely bounded homogeneous structures [31, 30].

▶ **Theorem 5.** *Let $r > \ell \geq 3$, let $\mathcal{K}$ be either the class $\mathcal{K}^{\ell}_{\mathrm{all}}$ of all finite ℓ-hypergraphs or the class $\mathcal{K}^{\ell}_{r}$, and let Ψ be a set of ℓ-hypergraph formulas containing $E(x_1, \ldots, x_\ell)$. Then precisely one of the following applies.*

1. *The clone* $\mathrm{Pol}(\mathbb{A})$ *has no uniformly continuous minion homomorphism to the clone of affine maps over a finite module, and* $\mathrm{CSP}(\mathbb{A})$ *has relational width* $(2\ell, \max(3\ell, r))$*.*
2. *The clone* $\mathrm{Pol}(\mathbb{A})$ *has a uniformly continuous minion homomorphism to the clone of affine maps over a finite module.*

1.5 Future work

This work is concerned with the complexity of the decision version of constraint satisfaction problems whose study is motivated by Conjecture 4. A natural variant of such problems is the optimisation version, where one is interested in finding a solution to an instance of the CSP that minimizes the number of unsatisfied constraints. The complexity of such problems (called MinCSPs) has mostly been investigated for finite templates, but recently also in the case of infinite templates falling within the scope of Conjecture 4 from the point of view of exact optimisation and approximation [21, 20, 24], as well as from the point of view of parameterized complexity [33].

Our complexity classification for the decision CSP (Theorem 3) can be seen as a foundation for a systematic structural study of optimisation problems over hypergraphs.

1.6 Organisation of the present article

After introducing a few notions needed for the formulation of the main algorithm in Section 2, we introduce the algorithm and prove its correctness in Section 3. For lack of space, we only present the proof of one of the main correctness arguments (Theorem 13) for the algorithmic part of Theorem 3 and otherwise illustrate the main concepts that we introduce using examples. In Section 4, we give an outline of the proof of Theorem 3. The rest of the proofs can be found in the appendix.

2 Preliminaries

For any $k \geq 1$, we write $[k]$ to denote the set $\{1,\ldots,k\}$. A tuple is called *injective* if its entries are pairwise distinct. In the entire article, we consider only relational structures in a finite signature.

A *primitive-positive* (pp-)formula is a first-order formula built only from atomic formulas, existential quantification, and conjunction. A relation $R \subseteq A^n$ is *pp-definable* in a relational structure $\mathbb{A}$ if there exists a pp-formula $\phi(x_1,\ldots,x_n)$ such that the tuples in R are precisely the tuples satisfying ϕ.

Let $\ell \geq 2$. A structure $\mathbb{H} = (H;E)$ is an *ℓ-hypergraph* if the relation E is of arity ℓ, contains only injective tuples (called *hyperedges*), and is *fully symmetric*, i.e., every tuple obtained by permuting the components of a hyperedge is a hyperedge as well. Given any ℓ-hypergraph $\mathbb{H} = (H;E)$, we write N for the set of all injective ℓ-tuples in H that are not hyperedges, and we call this set the *non-hyperedge* relation.

2.1 CSPs and Relational Width

A *CSP instance* over a set A is a pair $\mathcal{I} = (\mathcal{V},\mathcal{C})$, where $\mathcal{V}$ is a non-empty finite set of variables, and $\mathcal{C}$ is a set of *constraints*; each constraint $C \in \mathcal{C}$ is a subset of A^U for some non-empty $U \subseteq \mathcal{V}$ (U is called the *scope* of C). For a relational structure $\mathbb{A}$, we say that $\mathcal{I}$ is an *instance of* CSP($\mathbb{A}$) if for every $C \in \mathcal{C}$ with scope U, there exists an enumeration $u_1,\ldots,u_k$ of the elements of U and a k-ary relation R of $\mathbb{A}$ such that for all $f\colon U \to A$ we have $f \in C \Leftrightarrow (f(u_1),\ldots,f(u_k)) \in R$. A mapping $s\colon \mathcal{V} \to A$ is a *solution* of the instance $\mathcal{I}$ if we have $s|_U \in C$ for every $C \in \mathcal{C}$ with scope U. Given a constraint $C \subseteq A^U$ and a tuple $\mathbf{v} \in U^k$ for some $k \geq 1$, the *projection of C onto* $\mathbf{v}$ is defined by $\mathrm{proj}_{\mathbf{v}}(C) := \{f(\mathbf{v})\colon f \in C\}$. Let $U \subseteq \mathcal{V}$. We define the *restriction* of $\mathcal{I}$ to U to be an instance $\mathcal{I}|_U = (U,\mathcal{C}|_U)$ where the set of constraints $\mathcal{C}|_U$ contains for every $C \in \mathcal{C}$ the constraint $C|_U = \{g|_U \mid g \in C\}$.

We denote by $\mathrm{CSP}_{\mathrm{Inj}}(\mathbb{A})$ the restriction of CSP($\mathbb{A}$) to those instances of CSP($\mathbb{A}$) where for every constraint C and for every pair of distinct variables u,v in its scope, $\mathrm{proj}_{(u,v)}(C) \subseteq \{(a,b) \in A^2 \mid a \neq b\}$.

▶ **Definition 6.** *Let $1 \leq m \leq n$. We say that an instance $\mathcal{I} = (\mathcal{V},\mathcal{C})$ is* (m,n)-minimal *if both of the following hold:*

- *every non-empty subset of at most n variables in $\mathcal{V}$ is contained in the scope of some constraint in $\mathcal{I}$;*
- *for every at most m-element tuple of variables $\mathbf{v}$ and any two constraints $C_1, C_2 \in \mathcal{C}$ whose scopes contain all variables of $\mathbf{v}$, the projections of C_1 and C_2 onto $\mathbf{v}$ coincide.*

For $m \geq 1$, we say that an instance is *m-minimal* if it is (m, m)-minimal. We say that an instance $\mathcal{I}$ of the CSP is *non-trivial* if it does not contain any empty constraint. Otherwise, $\mathcal{I}$ is *trivial.*

For all $1 \leq m \leq n$ and for every instance $\mathcal{I}$ of a CSP($\mathbb{A}$) for some finite-domain structure $\mathbb{A}$, an (m, n)-minimal instance with the same solution set as $\mathcal{I}$ can be computed from $\mathcal{I}$ in polynomial time. The same holds for any ω-categorical structure $\mathbb{A}$ (see Section 2.2 for the definition of ω-categoricity, and see e.g., Section 2.3 in [32] for a description of the (m, n)-minimality algorithm in this setting). The resulting instance $\mathcal{I}'$ is called the *(m, n)-minimal instance equivalent to* $\mathcal{I}$ and the algorithm that computes this instance is called the *(m, n)-minimality algorithm.* Note that the instance $\mathcal{I}'$ is not necessarily an instance of CSP($\mathbb{A}$). However, $\mathcal{I}'$ is an instance of CSP($\mathbb{A}'$) where $\mathbb{A}'$ is the expansion of $\mathbb{A}$ by all at most n-ary relations pp-definable in $\mathbb{A}$. Moreover, CSP($\mathbb{A}'$) has the same complexity as CSP($\mathbb{A}$).

If $\mathcal{I}$ is m-minimal and $\mathbf{v}$ is a tuple of variables of length at most m, then by definition there exists a constraint of $\mathcal{I}$ whose scope contains all variables in $\mathbf{v}$, and all the constraints who do have the same projection on $\mathbf{v}$. We write $\mathrm{proj}_{\mathbf{v}}(\mathcal{I})$ for this projection, and call it the *projection of $\mathcal{I}$ onto* $\mathbf{v}$.

▶ **Definition 7.** *Let $1 \leq m \leq n$, and let $\mathbb{A}$ be a relational structure. We say that* CSP($\mathbb{A}$) *has relational width (m, n) if every non-trivial (m, n)-minimal instance equivalent to an instance of* CSP($\mathbb{A}$) *has a solution.* CSP($\mathbb{A}$) *has* bounded width *if it has relational width (m, n) for some natural numbers $m \leq n$.*

2.2 Basic model-theoretic definitions

Let $\mathbb{B}$ and $\mathbb{C}$ be relational structures in the same signature. A *homomorphism* from $\mathbb{B}$ to $\mathbb{C}$ is a mapping $f \colon B \to C$ with the property that for every relational symbol R from the signature of $\mathbb{B}$ and for every $\mathbf{b} \in R^{\mathbb{B}}$, it holds that $f(\mathbf{b}) \in R^{\mathbb{C}}$. An embedding of $\mathbb{B}$ into $\mathbb{C}$ is an injective homomorphism $f \colon \mathbb{B} \to \mathbb{C}$ such that f^{-1} is a homomorphism from the structure induced by the image of f in $\mathbb{C}$ to $\mathbb{B}$, and an *isomorphism* from $\mathbb{B}$ to $\mathbb{C}$ is a bijective embedding of $\mathbb{B}$ into $\mathbb{C}$. An endomorphism of $\mathbb{B}$ is a homomorphism from $\mathbb{B}$ to $\mathbb{B}$, an *automorphism* of $\mathbb{B}$ is an isomorphism from $\mathbb{B}$ to $\mathbb{B}$. We denote the set of endomorphisms of $\mathbb{B}$ by End($\mathbb{B}$) and the set of its automorphisms by Aut($\mathbb{B}$).

Let $\ell \geq 2$, and let $\mathcal{K}$ be an isomorphism-closed class of finite ℓ-hypergraphs. We say that $\mathcal{K}$ is an *amalgamation class* if the following two conditions are satisfied: It is closed under induced substructures, and for any ℓ-hypergraphs $\mathbb{H}, \mathbb{H}_1, \mathbb{H}_2 \in \mathcal{K}$ and for any embeddings f_i of $\mathbb{H}$ into $\mathbb{H}_i$ ($i \in \{1, 2\}$), there exists an ℓ-hypergraph $\mathbb{H}'$ and embeddings g_i of $\mathbb{H}_i$ into $\mathbb{H}'$ ($i \in \{1, 2\}$) such that $g_1 \circ f_1 = g_2 \circ f_2$. We write $\overrightarrow{\mathcal{K}}$ for the class which contains for every ℓ-hypergraph $\mathbb{H}$ from $\mathcal{K}$ all ordered ℓ-hypergraphs obtained by linearly ordering $\mathbb{H}$.

A relational structure $\mathbb{B}$ is *homogeneous* if every isomorphism between finite induced substructures of $\mathbb{B}$ extends to an automorphism of $\mathbb{B}$. The class of finite substructures of a homogeneous structure $\mathbb{B}$ is an amalgamation class; and conversely, for every amalgamation class $\mathcal{K}$ there exists a homogeneous structure $\mathbb{B}_{\mathcal{K}}$ whose finite induced substructures are exactly the structures in $\mathcal{K}$ (see e.g. [25] for this as well as the other claims in this section). The structure $\mathbb{B}_{\mathcal{K}}$ is called the Fraïssé limit of $\mathcal{K}$. The *universal homogeneous ℓ-hypergraph* is the Fraïssé limit of $\mathcal{K}^{\ell}_{\mathrm{all}}$.

A *first-order reduct* of a structure $\mathbb{B}$ is a structure $\mathbb{A}$ on the same domain whose relations are definable over $\mathbb{B}$ by first-order formulas without parameters. Recall that for any amalgamation class $\mathcal{K}$ and for any set Ψ of ℓ-hypergraph formulas, $\mathbb{A}_{\mathcal{K},\Psi}$ denotes the first-order reduct of the Fraïssé limit $\mathbb{H}_{\mathcal{K}}$ of $\mathcal{K}$ whose relations are defined by the formulas in Ψ. We remark that if $\mathbb{B}$ is the Fraïssé limit of a finitely bounded class, then every first-order formula is equivalent to one without quantifiers.

A countable relational structure is *ω-categorical* if its automorphism group has finitely many orbits in its componentwise action on n-tuples of elements for all $n \geq 1$. This is equivalent to saying that there are only finitely many relations of any fixed arity $n \geq 1$ that are first-order definable from $\mathbb{A}$. Every first-order reduct of a finitely bounded homogeneous structure is ω-categorical.

2.3 Polymorphisms

A *polymorphism* of a relational structure $\mathbb{A}$ is a function from A^n to A for some $n \geq 1$ which *preserves* all relations of $\mathbb{A}$, i.e., for every such relation R of arity m and for all tuples $(a_1^1, \ldots, a_m^1), \ldots, (a_1^n, \ldots, a_m^n) \in R$, it holds that $(f(a_1^1, \ldots, a_1^n), \ldots, f(a_m^1, \ldots, a_m^n)) \in R$. We also say that a polymorphism of $\mathbb{A}$ preserves a constraint $C \subseteq A^U$ if for all $g_1, \ldots, g_n \in C$, it holds that $f \circ (g_1, \ldots, g_n) \in C$. The set of all polymorphisms of a structure $\mathbb{A}$, denoted by $\mathrm{Pol}(\mathbb{A})$, 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. Every relation that is pp-definable in a relational structure $\mathbb{A}$ is preserved by all polymorphisms of $\mathbb{A}$.

Let $S \subseteq R \subseteq A^n$ be relations pp-definable in a structure $\mathbb{A}$. We say that S *is a binary absorbing subuniverse of* R *in* $\mathbb{A}$ if there exists a binary operation $f \in \mathrm{Pol}(\mathbb{A})$ such that for every $\mathbf{s} \in S, \mathbf{r} \in R$, we have that $f(\mathbf{s}, \mathbf{r}), f(\mathbf{r}, \mathbf{s}) \in S$. In this case, we write $S \trianglelefteq_{\mathbb{A}} R$, and we say that f *witnesses* the binary absorption.

Let $\mathbb{A}$ be a relational structure, and let $\mathscr{G} = \mathrm{Aut}(\mathbb{A})$ be the group of its automorphisms. For $n \geq 1$, a k-ary operation f defined on the domain of $\mathbb{A}$ is *n-canonical* with respect to $\mathbb{A}$ if for all $\mathbf{a}_1, \ldots, \mathbf{a}_k \in A^n$ and all $\alpha_1, \ldots, \alpha_k \in \mathscr{G}$, there exists $\beta \in \mathscr{G}$ such that $f(\mathbf{a}_1, \ldots, \mathbf{a}_k) = \beta \circ f(\alpha_1(\mathbf{a}_1), \ldots, \alpha_k(\mathbf{a}_k))$. A function f that is n-canonical with respect to $\mathbb{A}$ for all $n \geq 1$ is called *canonical* with respect to $\mathbb{A}$. In particular, f induces an operation on the set $A^n/\mathscr{G}$ of orbits of n-tuples under $\mathscr{G}$ for every $n \geq 1$. In our setting, we are interested in operations that are canonical with respect to a homogeneous ℓ-hypergraph $\mathbb{H}$ or to a homogeneous linearly ordered ℓ-hypergraph $(\mathbb{H}, <)$. In this case, an operation canonical with respect to $\mathbb{H}$ can simply be seen as an operation on labeled ℓ-hypergraphs with at most n elements, while an operation canonical with respect to $(\mathbb{H}, <)$ can be seen as an operation on labeled ℓ-hypergraphs with at most n elements which carry a weak linear order.

3 Polynomial-Time Algorithms From Symmetries

In this section, we fix $\ell \geq 3$ and a finitely bounded class $\mathcal{K}$ of ℓ-hypergraphs such that $\overrightarrow{\mathcal{K}}$ is an amalgamation class. We write $(\mathbb{H}, <)$ for the Fraïssé limit of $\overrightarrow{\mathcal{K}}$, I_n for the set of injective n-tuples of elements from H for any $n \geq 1$, I for I_ℓ, and $b_{\mathbb{H}}$ for an integer witnessing that $\mathcal{K}$ is finitely bounded. We also fix a first-order reduct $\mathbb{A}$ of $\mathbb{H}$. We say that $\mathbb{A}$ *admits an injective linear symmetry* if it has a ternary injective polymorphism m which is canonical with respect to $(\mathbb{H}, <)$, and which has the property that for any $\mathbf{a}, \mathbf{b} \in I$, the orbits under $\mathrm{Aut}(\mathbb{H})$ of $m(\mathbf{a}, \mathbf{a}, \mathbf{b}), m(\mathbf{a}, \mathbf{b}, \mathbf{a}), m(\mathbf{b}, \mathbf{a}, \mathbf{a})$ and $\mathbf{b}$ agree. Note that in this case, m induces an operation on the set $\{E, N\}$ of orbits of injective ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$. We say that m *acts as a minority operation on* $\{E, N\}$ since the second condition on m can be equivalently written as $m(X, X, Y) = m(X, Y, X) = m(Y, X, X) = Y$ for all $X, Y \in \{E, N\}$.

We prove the following.

▶ **Theorem 8.** *Let $\ell \geq 3$, let $\mathcal{K}$ be a finitely bounded class of ℓ-hypergraphs such that $\overrightarrow{\mathcal{K}}$ is an amalgamation class. Let $\mathbb{A}$ be a first-order reduct of the Fraïssé limit $\mathbb{H}$ of $\mathcal{K}$. Suppose that $I \trianglelefteq_{\mathbb{A}} H^\ell$, and that $\mathbb{A}$ admits an injective linear symmetry or is such that $\mathrm{CSP}_{\mathrm{Inj}}(\mathbb{A})$ has bounded width. Then $\mathrm{CSP}(\mathbb{A})$ is solvable in polynomial time.*

If $\mathrm{CSP}_{\mathrm{Inj}}(\mathbb{A})$ has bounded width, $\mathrm{CSP}(\mathbb{A})$ has bounded width as well by general principles, and is therefore in particular solvable in polynomial time (see [29, Section 3.2]). In the rest of this section, we will therefore focus on the case when $\mathbb{A}$ admits an injective linear symmetry.

Let $\mathbb{A}$ be a first-order reduct of $\mathbb{H}$ admitting an injective linear symmetry. Set $p_1(x, y) := m(x, y, y)$. It follows that p_1 is canonical with respect to $(\mathbb{H}, <)$ and that it acts as the first projection on $\{E, N\}$, i.e., it satisfies for any $\mathbf{a}, \mathbf{b} \in I$ that the orbit of $p_1(\mathbf{a}, \mathbf{b})$ under $\mathrm{Aut}(\mathbb{H})$ is equal to the orbit of $\mathbf{a}$. Moreover, by composing p_1 with a suitable endomorphism of $\mathbb{H}$, we can assume that $p_1(y, x)$ acts lexicographically on the order, i.e., $p_1(x, y) < p_1(x', y')$ if $y < y'$ or $y = y'$ and $x < x'$ (for more details, see [29, Section 5.2]).

In the remainder of this section, we present an algorithm solving $\mathrm{CSP}(\mathbb{A})$ in polynomial time, given that $\mathbb{A}$ has among its polymorphisms operations p_1 and m with the properties derived above. Before giving the technical details, we give here an overview of the methods we employ. Let $\mathcal{I}$ be an instance of $\mathrm{CSP}(\mathbb{A})$. Our algorithm transforms $\mathcal{I}$ into an equi-satisfiable instance $\mathcal{I}'$ that is sufficiently minimal, such that the solution set of a certain relaxation of $\mathcal{I}'$ is subdirect on all projections to an ℓ-tuple $\mathbf{v}$ of pairwise distinct variables (i.e., for every tuple $\mathbf{a}$ in this projection, this relaxation of $\mathcal{I}'$ has a solution where the variables from $\mathbf{v}$ are assigned values from $\mathbf{a}$), and that additionally satisfies a condition which we call *inj-irreducibility*, inspired by Zhuk's notion of irreducibility [39, 40]. We then prove that any non-trivial instance satisfying those properties has an injective solution. This step is to be compared with the case of *absorbing reductions* in Zhuk's algorithm, and in particular with Theorem 5.5 in [40], in which it is proved that any sufficiently minimal and irreducible instance that has a solution also has a solution where an arbitrary variable is constrained to belong to an absorbing subuniverse. Since in our setting I is an absorbing subuniverse of H^ℓ in $\mathbb{A}$, this fully establishes a parallel between the present work and [40]. The algorithm that we are going to introduce will work with infinite sets which are however always unions of orbits of ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$. $\mathrm{Aut}(\mathbb{H})$ is oligomorphic, i.e., it has only finitely many orbits in its action on H^k for every $k \geq 1$; in particular, there are only finitely many orbits of ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$, whence we can represent every union of such orbits by listing all orbits included in this union.

We now show that the structure defined by the formula from Example 1 has polymorphisms satisfying our assumptions on p_1 and on m, and hence it falls into the scope of this section.

▶ **Example 1** (continued). Let $R \subseteq H^4$ be the relation defined by ψ, and let $\mathbb{A} := (H; R)$. We define the canonical behaviour of a binary injection (i.e., a binary injective function) p_1 with respect to $(\mathbb{H}, <)$ as follows. We require that p_1 acts as the first projection on $\{E, N\}$, if O is a non-injective orbit of triples under $\mathrm{Aut}(\mathbb{H})$, and P is an injective orbit, then we require that $p_1(O, P) = p_1(P, O) = P$. Finally, for two non-injective orbits O_1, O_2 of triples under $\mathrm{Aut}(\mathbb{H}, <)$ such that $p_1(O_1, O_2)$ needs to be injective, we require that $p_1(O_1, O_2) = E$ if the minimum of any triple in O_1 appears only once in this triple, and $p_1(O_1, O_2) = N$ otherwise. Now, we take a ternary injection m' canonical with respect to $(\mathbb{H}, <)$ which behaves like a minority on $\{E, N\}$, and we define $m := m'(p_1(x, p_1(y, z)), p_1(y, p_1(z, x)), p_1(z, p_1(x, y)))$. It is easy to see that p_1 and m preserve R, hence $\mathbb{A}$ satisfies the assumptions from Theorem 8.

Let $\sim$ denote the 6-ary relation containing the tuples $(\mathbf{a}, \mathbf{b})$ where $\mathbf{a}, \mathbf{b}$ are triples that are in the same orbit under $\mathrm{Aut}(\mathbb{H})$. Note that the relation T defined by

$$(x_1, x_2, x_3, x_4) \in T :\Longleftrightarrow R(x_1, x_2, x_3, x_4) \wedge (x_1, x_3, x_2) \sim (x_4, x_2, x_3)$$

is preserved by all polymorphisms of $\mathbb{A}$ that are canonical with respect to $\mathbb{H}$ and that it contains precisely those tuples of the form (a,a,b,b) and (a,b,a,b) for arbitrary $a \neq b$. It can be seen (e.g., from [7]) that $\mathrm{Pol}(H;T)$ only contains *essentially unary* operations, of the form $(x_1,\ldots,x_n) \mapsto \alpha(x_i)$ for arbitrary permutations α of H, and therefore the polymorphisms of $\mathbb{A}$ that are canonical with respect to $\mathbb{H}$ are also essentially unary. It follows that the finite-domain CSP used in the reduction from [13] is NP-complete.

3.1 Finitisation of instances

Let $\mathbb{A}$ be a first-order reduct of $\mathbb{H}$. Let $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ be an instance of $\mathrm{CSP}(\mathbb{A})$. In this section, we always assume that the variable set $\mathcal{V}$ is equipped with an arbitrary linear order; this assumption is however inessential and only used to formulate the statements and proofs in a more concise way. We denote by $[\mathcal{V}]^\ell$ the set of injective increasing ℓ-tuples of variables from $\mathcal{V}$. Given any instance $\mathcal{I}$ of $\mathrm{CSP}(\mathbb{A})$, consider the following CSP instance $\mathcal{I}_{\mathrm{fin}}$ over the set $\mathcal{O}$ of orbits of ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$, called the *finitisation of* $\mathcal{I}$:

- The variable set of $\mathcal{I}_{\mathrm{fin}}$ is the set $[\mathcal{V}]^\ell$.
- For every constraint $C \subseteq A^U$ in $\mathcal{I}$, $\mathcal{I}_{\mathrm{fin}}$ contains the constraint C' containing the maps $g \colon [U]^\ell \to \mathcal{O}$ such that there exists $f \in C$ satisfying $f(\mathbf{v}) \in g(\mathbf{v})$ for every $\mathbf{v} \in [U]^\ell$.

This instance corresponds to the instance $\mathcal{I}_{\mathrm{Aut}(\mathbb{H}),\ell}$ from [32, Definition 3.1], with the difference that there the ℓ-element subsets of $\mathcal{V}$ were used as variables, and the domain consisted of orbits of maps. However, the translation between the two definitions is straightforward. Note that if a mapping $f \colon \mathcal{V} \to A$ is a solution of $\mathcal{I}$, then the mapping $h \colon [\mathcal{V}]^\ell \to \mathcal{O}$, where $h(\mathbf{v})$ is the orbit of $f(\mathbf{v})$ under $\mathrm{Aut}(\mathbb{H})$ for every $\mathbf{v} \in [\mathcal{V}]^\ell$ is a solution of $\mathcal{I}_{\mathrm{fin}}$.

Let $\mathcal{J} = (S, \mathcal{C})$ be an instance over the set $\mathcal{O}$ of orbits of ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$, e.g., $\mathcal{J} = \mathcal{I}_{\mathrm{fin}}$ for some $\mathcal{I}$. The *injectivisation* of $\mathcal{J}$, denoted by $\mathcal{J}^{(\mathrm{inj})}$, is the instance obtained by removing from all constraints all maps taking some value outside the two injective orbits E and N.

Let $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ be an instance of $\mathrm{CSP}(\mathbb{A})$; the *injective finitisation of* $\mathcal{I}$ is the instance $(\mathcal{I}_{\mathrm{fin}})^{(\mathrm{inj})}$. Let $S \subseteq [\mathcal{V}]^\ell$. The *injective finitisation of* $\mathcal{I}$ *on* S is the restriction of the injective finitisation of $\mathcal{I}$ to S. For any constraint $C \in \mathcal{C}$, the corresponding constraint in the injective finitisation of $\mathcal{I}$ is called the injective finitisation of C. Note that if $\mathbb{A}$ admits an injective linear symmetry, then for any instance $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ of $\mathrm{CSP}(\mathbb{A})$ and for any $S \subseteq [\mathcal{V}]^\ell$, the injective finitisation of $\mathcal{I}$ on S is solvable in polynomial time. This follows from Lemma 3.4 in [32] and from the dichotomy theorem for finite-domain CSPs [39, 40, 19].

Let $\mathbb{A}$ be a first-order reduct of $\mathbb{H}$ preserved by m and by p_1. We can assume that $\mathbb{A}$ has among its relations all unions of orbits of ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$ that are preserved by p_1 and by the ternary injection m. Otherwise, we expand $\mathbb{A}$ by these finitely many relations and we prove that the CSP of this expanded structure is solvable in polynomial time. Note that in particular, every orbit of ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$ is a relation of $\mathbb{A}$. Moreover, we suppose that $\mathbb{A}$ has the property that for every instance $\mathcal{I}$ of $\mathrm{CSP}(\mathbb{A})$, the $(2\ell, \max(3\ell, b_{\mathbb{H}}))$-minimal instance equivalent to $\mathcal{I}$ is again an instance of $\mathrm{CSP}(\mathbb{A})$. This can be achieved without loss of generality since it is enough to expand $\mathbb{A}$ by finitely many pp-definable relations, which are also preserved by m and p_1. Note that if $\mathcal{I}$ is a $(2\ell, \max(3\ell, b_{\mathbb{H}}))$-minimal instance of $\mathrm{CSP}(\mathbb{A})$, then its injective finitisation $\mathcal{I}_{\mathrm{fin}}$ is $(2,3)$-minimal by [32, Lemma 3.2]; in particular, $\mathcal{I}_{\mathrm{fin}}$ is *cycle consistent*, i.e., it satisfies one of the basic consistency notions used in Zhuk's algorithm [39]. Moreover, if $\mathcal{I}_{\mathrm{fin}}$ is $(2,3)$-minimal, then for any solution $h \colon [\mathcal{V}]^\ell \to \mathcal{O}$ of $\mathcal{I}_{\mathrm{fin}}$, any mapping $f \colon \mathcal{V} \to A$ with $f(\mathbf{v}) \in h(\mathbf{v})$ for every $\mathbf{v} \in [\mathcal{V}]^\ell$ is a solution of $\mathcal{I}$ by [32, Lemma 3.3].

Let $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ be an instance of CSP($\mathbb{A}$), and let $C \in \mathcal{C}$. Since $\mathbb{A}$ is preserved by m, there exists a set of linear equations over $\mathbb{Z}_2$ associated with the injective finitisation of C. By abuse of notation, we write every linear equation as $\sum_{\mathbf{v}\in S} X_{\mathbf{v}} = P$, where $P \in \{E, N\}$ and $S \subseteq [\mathcal{V}]^\ell$ is a set of injective ℓ-tuples of variables from the scope of C. In these linear equations, we identify E with 1 and N with 0, so that e.g. $E + E = N$ and $N + E = E$. Using this notation, the canonical behaviour of the function m on $\{E, N\}$ can be written as $m(X, Y, Z) = X + Y + Z$ which justifies the notion of $\mathbb{A}$ admitting linear symmetries.

For an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ of CSP($\mathbb{A}$), we define an instance $\mathcal{I}_{\text{eq}} = (\mathcal{V}, \mathcal{C}_{\text{eq}})$ of the equality-CSP (i.e., CSP over structures first-order definable over $(H; =)$) over the same base set H corresponding to the closure of the constraints under the full symmetric group on H. Formally, for every constraint $C \in \mathcal{C}$, the corresponding constraint $C_{\text{eq}} \in \mathcal{C}_{\text{eq}}$ contains all functions αh for all $h \in C$ and $\alpha \in \mathrm{Sym}(H)$. Since $\mathbb{A}$ is preserved by a binary injection, the constraints of $\mathcal{I}_{\text{eq}}$ are preserved by the same or indeed any binary injection and hence, its CSP has relational width $(2, 3)$ by the classification of equality CSPs [7].

Let $\mathcal{I}$ be an ℓ-minimal instance of CSP($\mathbb{A}$), let $\mathbf{v} \in [\mathcal{V}]^\ell$, and let $R \subseteq \mathrm{proj}_{\mathbf{v}}(\mathcal{I})$ be an ℓ-ary relation from the signature of $\mathbb{A}$. Let $\mathcal{I}^{\mathbf{v}\in R}$ be the instance obtained from $\mathcal{I}$ by replacing every constraint C containing all variables from $\mathbf{v}$ by $\{g \in C \mid g(\mathbf{v}) \in R\}$.

We call an ℓ-minimal instance of CSP($\mathbb{A}$) *eq-subdirect* if for every $\mathbf{v} \in [\mathcal{V}]^\ell$ and for every non-injective orbit $O \subseteq \mathrm{proj}_{\mathbf{v}}(\mathcal{I})$ under $\mathrm{Aut}(\mathbb{H})$, the instance $(\mathcal{I}^{\mathbf{v}\in O})_{\text{eq}}$ has a solution. Note that by ℓ-minimality and since all constraints of the instance are preserved by a binary injection, the instance $(\mathcal{I}^{\mathbf{v}\in O})_{\text{eq}}$ has a solution for every injective orbit $O \subseteq \mathrm{proj}_{\mathbf{v}}(\mathcal{I})$ under $\mathrm{Aut}(\mathbb{H})$. Indeed, any injective mapping from $\mathcal{V}$ to H is a solution of $(\mathcal{I}^{\mathbf{v}\in O})_{\text{eq}}$.

▶ **Example 9.** Let $\mathbb{H}$ be the universal homogeneous ℓ-hypergraph. Let $\mathbf{u} = (u_1, \ldots, u_\ell), \mathbf{v} = (v_1, \ldots, v_\ell)$ be disjoint ℓ-tuples of variables, and let $\mathcal{V}$ be the set of all variables contained in these tuples. We define a CSP instance $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ over the set H as follows. Let $\mathbf{u}' = (u_2, \ldots, u_\ell, u_1)$. We set $\mathcal{C}$ to contain two constraints C, C' such that C contains all mappings $f \colon \mathcal{V} \to H$ such that $f(\mathbf{u})$ and $f(\mathbf{v})$ belong to the same orbit under $\mathrm{Aut}(\mathbb{H})$, and C' contains all mappings $f \colon \mathcal{V} \to H$ such that $f(\mathbf{u}')$ and $f(\mathbf{v})$ belong to the same orbit. Note that the constraints C, C' are preserved by any function which is canonical with respect to $\mathbb{H}$. It is easy to see that $\mathcal{I}$ is non-trivial and ℓ-minimal, but it is not eq-subdirect. Indeed, for any non-injective and non-constant mapping g from the set of variables of $\mathbf{u}$ to H, it holds that $g(\mathbf{u})$ and $g(\mathbf{u}')$ are contained in different orbits under $\mathrm{Aut}(\mathbb{H})$.

We can obtain an eq-subdirect instance out of an ℓ-minimal instance in polynomial time by the algorithm introduced in [29, Section 3.1]. The algorithm successively shrinks for every $\mathbf{v} \in [\mathcal{V}]^\ell$ the projection $\mathrm{proj}_{\mathbf{v}}(\mathcal{I})$ to the union of those orbits $O \subseteq \mathrm{proj}_{\mathbf{v}}(\mathcal{I})$ under $\mathrm{Aut}(\mathbb{H})$ for which the instance $(\mathcal{I}^{\mathbf{v}\in O})_{\text{eq}}$ has a solution; it stops when no more orbits can be removed from $\mathrm{proj}_{\mathbf{v}}(\mathcal{I})$ for any $\mathbf{v} \in [\mathcal{V}]^\ell$.

Note that for any $1 \leq m \leq n$ and any instance $\mathcal{I}$ of CSP($\mathbb{A}$), we can compute an instance that is both eq-subdirect and (m, n)-minimal and that has the same solution set as $\mathcal{I}$ in polynomial time. Indeed, it is enough to repeat the above-mentioned algorithm and the (m, n)-minimality algorithm until no orbits under $\mathrm{Aut}(\mathbb{H})$ are removed from any constraint.

3.2 Inj-irreducibility

For any $\mathbf{a} \in H^\ell$, we write $O(\mathbf{a})$ for the orbit of $\mathbf{a}$ under $\mathrm{Aut}(\mathbb{H})$ and $O_<(\mathbf{a})$ for the orbit of $\mathbf{a}$ under $\mathrm{Aut}(\mathbb{H}, <)$. Recall that being canonical with respect to $(\mathbb{H}, <)$, the function p_1 acts naturally on orbits under $\mathrm{Aut}(\mathbb{H}, <)$; we can therefore abuse the notation and write $p_1(O, P)$ for orbits O, P under $\mathrm{Aut}(\mathbb{H}, <)$. We will say that a non-injective orbit O of ℓ-tuples under $\mathrm{Aut}(\mathbb{H})$ is:

- *deterministic* if for every $\mathbf{a} \in O$, there exists $\alpha \in \mathrm{Aut}(\mathbb{H})$ such that $p_1(O_<(\alpha(\mathbf{a})), O_<(\mathbf{e})) = p_1(O_<(\alpha(\mathbf{a})), O_<(\mathbf{n}))$, where $\mathbf{e}, \mathbf{n}$ are arbitrary strictly increasing ℓ-tuples of elements of $\mathbb{H}$ with $\mathbf{e} \in E$, $\mathbf{n} \in N$,
- *non-deterministic* otherwise.

For a tuple $\mathbf{a}$ contained in a deterministic orbit O, we call any $\alpha \in \mathrm{Aut}(\mathbb{H})$ with the property that $p_1(O_<(\alpha(\mathbf{a})), E) = p_1(O_<(\alpha(\mathbf{a})), N)$ for the strictly increasing ordering in the second coordinate *deterministic for* $\mathbf{a}$. Note that for any $\beta \in \mathrm{Aut}(\mathbb{H}, <)$, $\beta\alpha \in \mathrm{Aut}(\mathbb{H})$ is deterministic for $\mathbf{a}$ as well since p_1 is canonical with respect to $(\mathbb{H}, <)$.

Let $\mathcal{J} = (\mathcal{V}, \mathcal{C})$ be a CSP instance over a set B. A sequence $v_1, C_1, v_2, \ldots, C_k, v_{k+1}$, where $k \geq 1$, $v_i \in \mathcal{V}$ for every $i \in [k+1]$, $C_i \in \mathcal{C}$ for every $i \in [k]$, and v_i, v_{i+1} are contained in the scope of C_i for every $i \in [k]$, is called a *path* in $\mathcal{J}$. We say that two elements $a, b \in B$ are *connected* by a path $v_1, C_1, v_2, \ldots, C_k, v_{k+1}$ if there exists a tuple $(c_1, \ldots, c_{k+1}) \in B^{k+1}$ such that $c_1 = a, c_{k+1} = b$, and such that $(c_i, c_{i+1}) \in \mathrm{proj}_{(v_i, v_{i+1})}(C_i)$ for every $i \in [k]$. Suppose that $\mathcal{J}$ is 1-minimal. The *linkedness congruence* on $\mathrm{proj}_v(\mathcal{J})$ is the equivalence relation λ on $\mathrm{proj}_v(\mathcal{J})$ defined by $(a, b) \in \lambda$ if there exists a path $v_1, C_1, v_2, \ldots, C_k, v_{k+1}$ from a to b in $\mathcal{J}$ such that $v_1 = v_{k+1} = v$. Note that for a finite relational structure $\mathbb{B}$, for a $(2,3)$-minimal instance $\mathcal{J} = (\mathcal{V}, \mathcal{C})$ of $\mathrm{CSP}(\mathbb{B})$, and for any $v \in \mathcal{V}$, the linkedness congruence λ on $\mathrm{proj}_v(\mathcal{J})$ is a relation pp-definable in $\mathbb{B}$. Indeed, it is easy to see that the binary relation containing precisely the pairs $(a, b) \in B^2$ that are connected by a particular path in $\mathcal{I}$ is pp-definable in $\mathbb{B}$. If we concatenate all paths that connect two elements $(a, b) \in \lambda$, the resulting path connects every pair $(a, b) \in \lambda$ since by the $(2,3)$-minimality of $\mathcal{J}$, every path from v to v connects c to c for every $c \in \mathrm{proj}_v(\mathcal{J})$. It follows that λ is pp-definable.

▶ **Definition 10.** *Let $\mathbb{A}$ be a first-order reduct of $\mathbb{H}$, and let $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ be a non-trivial ℓ-minimal instance of* $\mathrm{CSP}(\mathbb{A})$*. We call $\mathcal{I}$* inj-irreducible *if for every set $S \subseteq [\mathcal{V}]^\ell$, one of the following holds for the instance $\mathcal{J} = \mathcal{I}_{fin}|_S$:*
- *$\mathcal{J}^{(\mathrm{inj})}$ has a solution,*
- *for some $\mathbf{v} \in S$, $\mathrm{proj}_{\mathbf{v}}(\mathcal{J})$ contains the two injective orbits and the linkedness congruence on $\mathrm{proj}_{\mathbf{v}}(\mathcal{J})$ does not connect them,*
- *for some $\mathbf{v} \in S$, the linkedness congruence on $\mathrm{proj}_{\mathbf{v}}(\mathcal{J})$ links an injective orbit to a non-deterministic orbit.*

▶ **Example 11.** We illustrate the concept of inj-irreducibility on the following instance. Let $\ell = 3$, and let $\mathbb{H}$ be the universal homogeneous ℓ-hypergraph. Let $a \neq b \in H$ be arbitrary; we call the orbits $O(a, a, b), O(a, b, a), O(b, a, a)$ under $\mathrm{Aut}(\mathbb{H})$ *half-injective*. Let us define an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C})$ over the set H as follows. Above, we identified E with 1 and N with 0 in the linear equations associated to injective finitisations of constraints. In this example, we identify also all half-injective orbits under $\mathrm{Aut}(\mathbb{H})$ with 1. Hence, we can write, e.g., $E + O(a, a, b) = 0$.

Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ be increasing triples of pairwise disjoint variables, and set $\mathcal{V}$ to be the union of all variables of these tuples. We define $\mathcal{C}$ to be a set consisting of two constraints, C_0 and C_1, defined as follows. For $i \in \{0, 1\}$, we set C_i to contain all mappings $f \in H^{\mathcal{V}}$ such that both of the following hold:
- either $O(f(\mathbf{v}_1)) + O(f(\mathbf{v}_2)) + O(f(\mathbf{v}_3)) = i$, or $O(f(\mathbf{v}_1)) = O(f(\mathbf{v}_2)) = O(f(\mathbf{v}_3)) = \{(a, a, a) \mid a \in H\}$,
- $f(x) \neq f(y)$ for all $x, y \in \mathcal{V}$ belonging to different triples from $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

We show that the constraints C_0, C_1 are preserved by a binary injection p_1 and a ternary injection m that are both canonical with respect to $\mathbb{H}$. To this end, we define the canonical behaviours of p_1 and m on the orbits of triples under $\mathrm{Aut}(\mathbb{H})$ as follows. We set p_1 to act as the first projection on the numbers associated to the respective orbits, and to satisfy $p_1(O(a,a,a),P) = p_1(P,O(a,a,a)) = P$ for an arbitrary $a \in H$ and for an arbitrary orbit P. Note that these assumptions together with the requirement that p_1 is an injection uniquely determine the behaviour of p_1 – e.g., $p_1(P,N) = E$ for an arbitrary half-injective orbit P. We set m to act as an idempotent minority on the numbers associated to the orbits, and to act as $p_1(x, p_1(y,z))$ on the orbits where the action is not determined by the previous condition. It is easy to verify that the constraints C_0 and C_1 are preserved by both p_1 and m, and hence $\mathcal{I}$ is an instance of $\mathrm{CSP}(\mathbb{A})$ for some first-order reduct $\mathbb{A}$ of $\mathbb{H}$ which falls into the scope of this section.

It immediately follows that all the half-injective orbits under $\mathrm{Aut}(\mathbb{H})$ are deterministic since $p_1(O_<(\mathbf{c}),E) = p_1(O_<(\mathbf{c}),N) = E$ for any $\mathbf{c} \in H^3$ contained in a half-injective orbit; the orbit P of the constant tuples is non-deterministic since $p_1(P,E) = E$, and $p_1(P,N) = N$. It is also easy to see that $\mathcal{I}$ is non-trivial and $(6,9)$-minimal. Moreover, $\mathcal{I}$ is not inj-irreducible. Indeed, setting $S := \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, the linkedness congruence on $\mathrm{proj}_{\mathbf{v}_1}(\mathcal{I}_{\mathrm{fin}}|_S)$ connects precisely all injective and half-injective orbits, and the injective finitisation of $\mathcal{I}$ on S does not have a solution.

Note that if $\mathbb{A}$ is the structure from Example 11, its CSP can be solved by the reduction to the finite from [13]. For simplicity, we choose to illustrate the concepts that we have just introduced on this example rather than on an example where the canonical behaviour of the functions p_1 and m depends on the linear order. However, Example 1 provides us with a structure admitting linear symmetries where the canonical behaviour of any polymorphism satisfying the assumptions on m or on p_1 depends on the additional linear order.

▶ **Lemma 12.** *Let C be a constraint of an instance of $\mathrm{CSP}(\mathbb{A})$ which contains an injective mapping, and let S be a set of variables appearing together in an unsplittable linear equation associated with the injective finitisation of C. Then for every $g \in C$, either $g(\mathbf{v})$ is in an injective or deterministic orbit for all $\mathbf{v} \in S$, or $g(\mathbf{v})$ is in a non-deterministic orbit for all $\mathbf{v} \in S$.*

▶ **Theorem 13.** *Let $\mathbb{A}$ be a first-order reduct of $\mathbb{H}$ that admits linear symmetries. Let $\mathcal{I}$ be a $(2\ell, \max(3\ell, b_{\mathbb{H}}))$-minimal, inj-irreducible instance of $\mathrm{CSP}(\mathbb{A})$ with variables $\mathcal{V}$ such that for every distinct $u, v \in \mathcal{V}$, $\mathrm{proj}_{(u,v)}(\mathcal{I}) \cap I_2 \neq \emptyset$. Then $\mathcal{I}$ has an injective solution.*

Proof. Note that if $\mathcal{I}$ has fewer than ℓ variables, it has an injective solution by the assumption on binary projections of $\mathcal{I}$ and since all constraints of $\mathcal{I}$ are preserved by the binary injection p_1. Let us therefore suppose that $\mathcal{I}$ has at least ℓ variables. Let us assume for the sake of contradiction that $\mathcal{I}$ does not have an injective solution. Let $\mathcal{J}$ be $\mathcal{I}_{\mathrm{fin}}$, and let $\mathcal{C}$ be the set of its constraints. By assumption, $\mathcal{J}^{(\mathrm{inj})}$ does not have a solution. Note that $\mathcal{J}^{(\mathrm{inj})}$ corresponds to a system of linear equations over $\mathbb{Z}_2$, which is therefore unsatisfiable. In case this system can be written as a diagonal block matrix, there exists a set $S \subseteq [\mathcal{V}]^\ell$ of variables such that the system of equations associated with the injectivisation of $\mathcal{L} := \mathcal{J}|_S = (S, \mathcal{C}')$ corresponds to a minimal unsatisfiable block. By definition, this means that $\mathcal{L}^{(\mathrm{inj})}$ is unsatisfiable. The instance $\mathcal{L}$ has the property that for every non-trivial partition of S into parts S_1, S_2, there exists an unsplittable equation associated with the injectivisation of a constraint $C \in \mathcal{C}'$ which contains variables from both S_1 and S_2.

Since $\mathcal{I}$ is inj-irreducible, there exists $\mathbf{v} \in S$ such that the two injective orbits are elements of $\mathrm{proj}_{\mathbf{v}}(\mathcal{L})$ and are not linked, or some injective orbit in $\mathrm{proj}_{\mathbf{v}}(\mathcal{L})$ is linked to a non-deterministic orbit in $\mathrm{proj}_{\mathbf{v}}(\mathcal{L})$.

In the first case, we note that for *all* $\mathbf{w} \in S$ such that $\mathrm{proj}_{\mathbf{w}}(\mathcal{L})$ contains the two injective orbits, the two injective orbits are not linked. Indeed, suppose that there exists $\mathbf{w} \in S$ such that $E, N \in \mathrm{proj}_{\mathbf{w}}(\mathcal{L})$ are linked, i.e., there exists a path $\mathbf{v}_1 = \mathbf{w}, C_1, \ldots, C_k, \mathbf{v}_{k+1} = \mathbf{w}$ in $\mathcal{L}$ connecting E and N. Since $\mathcal{I}$ is $(2\ell, \max(3\ell, b_{\mathbb{H}}))$-minimal, Lemma 3.2 in [32] yields that $\mathcal{J}$ and hence also $\mathcal{L}$ is $(2,3)$-minimal. In particular, there exists $C \in \mathcal{C}'$ containing in its scope both $\mathbf{v}$ and $\mathbf{w}$. Let $O_1, O_2 \in \{E, N\}$ be disjoint such that there exist $g_1, g_2 \in C$ with $g_1(\mathbf{v}) = O_1, g_1(\mathbf{w}) = E, g_2(\mathbf{v}) = O_2, g_2(\mathbf{w}) = N$. It follows that the path $\mathbf{v}, C, \mathbf{w} = \mathbf{v}_1, C_1, \ldots, C_k, \mathbf{v}_{k+1} = \mathbf{w}, C, \mathbf{v}$ connects O_1 with O_2 in $\mathrm{proj}_{\mathbf{v}}(\mathcal{J})$, a contradiction. Let $g \colon S \to \{E, N\}$ be defined as follows. For a fixed $\mathbf{v} \in S$, let $g(\mathbf{v})$ be an arbitrary element of $\mathrm{proj}_{\mathbf{v}}(\mathcal{L}^{(\mathrm{inj})})$. Next, for $\mathbf{w} \in S$, define $g(\mathbf{w})$ to be the unique injective orbit O such that there exists a constraint $C \in \mathcal{C}'$ containing both $\mathbf{v}$ and $\mathbf{w}$ in its scope and such that there exists $g' \in C$ with $g'(\mathbf{v}) = g(\mathbf{v})$ and $g'(\mathbf{w}) = O$. This g is a solution of $\mathcal{L}^{(\mathrm{inj})}$, a contradiction.

Thus, it must be that a non-deterministic orbit in $\mathrm{proj}_{\mathbf{v}}(\mathcal{L})$ is linked to an injective orbit in $\mathrm{proj}_{\mathbf{v}}(\mathcal{L})$. Hence, there exists a path in $\mathcal{L}$ from $\mathbf{v}$ to $\mathbf{v}$ and connecting an injective orbit to a non-deterministic one. Moreover, up to composing this path with additional constraints, one can assume that this path goes through all the variables in S. This follows by the $(2,3)$-minimality of $\mathcal{J}$. Define a partition of S where $\mathbf{w} \in S_1$ if the first time that $\mathbf{w}$ appears in the path, the element associated with $\mathbf{w}$ is in an injective orbit, and $\mathbf{w} \in S_2$ otherwise. Since the system of unsplittable equations associated with $\mathcal{L}^{(\mathrm{inj})}$ cannot be decomposed as a diagonal block matrix, some constraint $C \in \mathcal{C}'$ gives an equation in that system containing $\mathbf{u}_1 \in S_1$ and $\mathbf{u}_2 \in S_2$. Thus, there exists $g \in C$ with $g(\mathbf{u}_1)$ injective, and $g(\mathbf{u}_2)$ non-deterministic. This contradicts Lemma 12. ◀

3.3 Establishing inj-irreducibility

We introduce a polynomial-time algorithm which produces, given an instance $\mathcal{I}$ of $\mathrm{CSP}(\mathbb{A})$, an instance $\mathcal{I}'$ of $\mathrm{CSP}(\mathbb{A})$ that is either inj-irreducible or trivial and that has a solution if, and only if, $\mathcal{I}$ has a solution. It uses the fact that the injective finitisation of an instance $\mathcal{I}$ of $\mathrm{CSP}(\mathbb{A})$ on S is solvable in polynomial time for any set $S \subseteq [\mathcal{V}]^\ell$. This follows from the fact that the constraints of the injective finitisation of $\mathcal{I}$ are preserved by a ternary minority by Lemma 3.4 from [32].

We give a brief description of the algorithm. It gradually ensures that the instance is $(2\ell, \max(3\ell, b_{\mathbb{H}}))$-minimal, eq-subdirect, and so that for no distinct variables $u, v \in \mathcal{V}$, it holds that $\mathrm{proj}_{(u,v)}(\mathcal{I}) = \{(a,a) \mid a \in H\}$. If the instance satisfies these assumptions, we consider for every $\mathbf{u} \in [\mathcal{V}]^\ell$ every partition $\{E^1_{\mathbf{u}}, \ldots, E^s_{\mathbf{u}}\}$ on $\mathrm{proj}_{\mathbf{u}}(\mathcal{I})$ with pp-definable classes satisfying that $E^1_{\mathbf{u}}$ contains no non-deterministic orbit. We find a subset $S \subseteq [\mathcal{V}]^\ell$ such that for every $\mathbf{w} \in S$, the set $\{E^1_{\mathbf{w}}, \ldots, E^s_{\mathbf{w}}\}$ defined by $E^i_{\mathbf{w}} := \{\mathbf{a} \in H^\ell \mid \exists \mathbf{b} \in H^\ell \colon (\mathbf{b}, \mathbf{a}) \in \mathrm{proj}_{(\mathbf{u},\mathbf{w})}(\mathcal{I})\}$ for every $i \in [s]$ forms a partition on $\mathrm{proj}_{\mathbf{w}}(\mathcal{I})$ with the property that $E^1_{\mathbf{w}}$ contains no non-deterministic orbit, and such that this partition cannot be extended to any other tuple in $[\mathcal{V}]^\ell$. For every such partition, the algorithm checks if the injective finitisation of $\mathcal{I}$ on S has a solution; if not, we constrain every $\mathbf{w} \in S$ to not take any value from $E^1_{\mathbf{w}}$. It is *a priori* unclear that adding these constraints on the one hand yields an instance of $\mathrm{CSP}(\mathbb{A})$, and on the other hand that we do not transform a satisfiable instance into an unsatisfiable one in this way. The following technical result is that, in fact, this does not happen.

▶ **Theorem 14.** *The instance $\mathcal{I}'$ produced by the procedure* InjIrreducibility *in [29, Section 3.1] is an instance of* $\mathrm{CSP}(\mathbb{A})$ *and it has a solution if, and only if, the original instance has a solution. Moreover, $\mathcal{I}'$ is either trivial or inj-irreducible.*

4 Final Arguments for the Complexity Dichotomy

Finally, we briefly sketch the proof of Theorem 3. If the algebraic assumptions in the second item are met, it is known that $\mathrm{CSP}(\mathbb{A})$ is NP-complete [5]. On the other hand, if I is a binary absorbing subuniverse of H^ℓ and $\mathbb{A}$ either admits an injective linear symmetry or is such that $\mathrm{CSP}_{\mathrm{Inj}}(\mathbb{A})$ has bounded width, then $\mathrm{CSP}(\mathbb{A})$ is polynomial-time solvable by Theorem 8. It therefore remains to prove that the assumptions in the first item of Theorem 3 (i.e., the clone of polymorphisms of $\mathbb{A}$ being non-trivial) imply that $\mathbb{A}$ falls into the scope of Theorem 8.

By [35], if some polymorphism of $\mathbb{A}$ acts on $\{E, N\}$ in a non-trivial way, then $\mathbb{A}$ either admits an injective linear symmetry or is such that $\mathrm{CSP}_{\mathrm{Inj}}(\mathbb{A})$ has bounded width. Moreover, by results from [30, Proposition 25], the polymorphisms of $\mathbb{A}$ being non-trivial imply that $\mathbb{A}$ contains a binary injection witnessing that I is a binary absorbing subuniverse of H^ℓ. On the other hand, if the action of the polymorphisms of $\mathbb{A}$ on $\{E, N\}$ is trivial, then we can use the theory of smooth approximations [30] to show that the polymorphisms of $\mathbb{A}$ are trivial and $\mathbb{A}$ thus falls into the scope of the second item of Theorem 3.

References

1. Reza Akhtar and Alistair H. Lachlan. On countable homogeneous 3-hypergraphs. *Archive for Mathematical Logic*, 34:331–344, 1995. doi:10.1007/BF01387512.
2. Libor Barto, Michael Kompatscher, Miroslav Olšák, Trung Van Pham, 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. doi:10.1109/LICS.2017.8005128.
3. Libor Barto, Michael Kompatscher, Miroslav Olšák, Trung Van Pham, and Michael Pinsker. Equations in oligomorphic clones and the constraint satisfaction problem for ω-categorical structures. *Journal of Mathematical Logic*, 19(02):1950010, 2019. doi:10.1142/S0219061319500107.
4. 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.
5. Libor Barto, Jakub Opršal, and Michael Pinsker. The wonderland of reflections. *Israel Journal of Mathematics*, 223(1):363–398, 2018. doi:10.1007/s11856-017-1621-9.
6. Manuel Bodirsky, Peter Jonsson, and Trung Van Pham. The complexity of phylogeny constraint satisfaction problems. *ACM Transactions on Computational Logic*, 18(3), 2017. An extended abstract appeared in the conference STACS 2016. doi:10.1145/3105907.
7. 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). doi:10.1007/S00224-007-9083-9.
8. Manuel Bodirsky and Jan Kára. The complexity of temporal constraint satisfaction problems. *Journal of the ACM*, 57(2), 2010. An extended abstract appeared in the Proceedings of the Symposium on Theory of Computing (STOC). doi:10.1145/1667053.1667058.
9. Manuel Bodirsky and Simon Knäuer. Network satisfaction for symmetric relation algebras with a flexible atom. *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, 35(7):6218–6226, 2021. doi:10.1609/AAAI.V35I7.16773.
10. Manuel Bodirsky and Simon Knäuer. The complexity of network satisfaction problems for symmetric relation algebras with a flexible atom. *J. Artif. Intell. Res.*, 75:1701–1744, 2022. doi:10.1613/JAIR.1.14195.
11. Manuel Bodirsky, Florent R. Madelaine, and Antoine Mottet. A universal-algebraic proof of the complexity dichotomy for monotone monadic SNP. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science – LICS'18*, 2018. doi:10.1145/3209108.3209156.

12 Manuel Bodirsky, Barnaby Martin, Michael Pinsker, and András Pongrácz. Constraint satisfaction problems for reducts of homogeneous graphs. *SIAM Journal on Computing*, 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`.

13 Manuel Bodirsky and Antoine Mottet. Reducts of finitely bounded homogeneous structures, and lifting tractability from finite-domain constraint satisfaction. In *Proceedings of the 31th Annual ACM/IEEE Symposium on Logic in Computer Science – LICS'16*, 2016. `doi:10.1145/2933575.2934515`.

14 Manuel Bodirsky and Antoine Mottet. A dichotomy for first-order reducts of unary structures. *Logical Methods in Computer Science*, 14(2), 2018. `doi:10.23638/LMCS-14(2:13)2018`.

15 Manuel Bodirsky and Michael Pinsker. Schaefer's theorem for graphs. *Journal of the ACM*, 62(3):19:1–19:52, 2015. A conference version appeared in the Proceedings of the Symposium on Theory of Computing (STOC 2011), pages 655–664. `doi:10.1145/2764899`.

16 Manuel Bodirsky and Michael Pinsker. Topological Birkhoff. *Transactions of the American Mathematical Society*, 367:2527–2549, 2015. `doi:10.1090/S0002-9947-2014-05975-8`.

17 Manuel Bodirsky, Michael Pinsker, and András Pongrácz. Projective clone homomorphisms. *Journal of Symbolic Logic*, 86(1):148–161, 2021. `doi:10.1017/jsl.2019.23`.

18 Bertalan Bodor. CSP dichotomy for ω-categorical monadically stable structures. PhD dissertation, Institute of Algebra, Technische Universität Dresden, 2021. URL: `https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-774379`.

19 Andrei A. Bulatov. A dichotomy theorem for nonuniform CSPs. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 319–330. IEEE Computer Society, 2017. `doi:10.1109/FOCS.2017.37`.

20 Moses Charikar, Venkatesan Guruswami, and Rajsekar Manokaran. Every permutation CSP of arity 3 is approximation resistant. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15-18 July 2009*, pages 62–73. IEEE Computer Society, 2009. `doi:10.1109/CCC.2009.29`.

21 Vaggos Chatziafratis and Konstantin Makarychev. Phylogenetic CSPs are approximation resistant. *CoRR*, abs/2212.12765, 2022. `doi:10.48550/arXiv.2212.12765`.

22 Roland Fraïssé. Une hypothèse sur l'extension des relations finies et sa vérification dans certaines classes particulières (deuxième partie). *Synthese*, 16(1):34–46, 1966. URL: `http://www.jstor.org/stable/20114493`.

23 Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal on Computing*, 28(1):57–104, 1998. `doi:10.1137/S0097539794266766`.

24 Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar. Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM J. Comput.*, 40(3):878–914, 2011. `doi:10.1137/090756144`.

25 Wilfrid Hodges. *Model theory*, volume 42 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 1993.

26 Michael Kompatscher and Trung Van Pham. A complexity dichotomy for poset constraint satisfaction. *IfCoLog Journal of Logics and their Applications (FLAP)*, 5(8):1663–1696, 2018. A conference version appeared in the Proceedings of the 34th Symposium on Theoretical Aspects of Computer Science (STACS 2017), pages 47:1–47:12. URL: `https://www.collegepublications.co.uk/downloads/ifcolog00028.pdf`.

27 Alistair H. Lachlan and Robert E. Woodrow. Countable ultrahomogeneous undirected graphs. *Transactions of the American Mathematical Society*, 262, 1980. `doi:10.1090/S0002-9947-1980-0583847-2`.

28 Richard Emil Ladner. On the structure of polynomial time reducibility. *J. ACM*, 22(1):155–171, January 1975. `doi:10.1145/321864.321877`.

29 Antoine Mottet, Tomáš Nagy, and Michael Pinsker. An order out of nowhere: a new algorithm for infinite-domain csps, 2023. arXiv:2301.12977.

30 Antoine Mottet and Michael Pinsker. Smooth approximations and CSPs over finitely bounded homogeneous structures. In *Proceedings of the 37th Annual ACM/IEEE Symposium on Logic in Computer Science – LICS'22*, 2022. doi:10.1145/3531130.3533353.

31 Antoine Mottet, Tomáš Nagy, Michael Pinsker, and Michał Wrona. Smooth approximations and relational width collapses. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 138:1–138:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.ICALP.2021.138.

32 Antoine Mottet, Tomáš Nagy, Michael Pinsker, and Michał Wrona. When symmetries are enough: collapsing the bounded width hierarchy for infinite-domain CSPs. arxiv:2102.07531, 2022. doi:10.48550/arXiv.2102.07531.

33 George Osipov and Magnus Wahlström. Parameterized Complexity of Equality MinCSP. In Inge Li Gørtz, Martin Farach-Colton, Simon J. Puglisi, and Grzegorz Herman, editors, *31st Annual European Symposium on Algorithms, ESA 2023, September 4-6, 2023, Amsterdam, The Netherlands*, volume 274 of *LIPIcs*, pages 86:1–86:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICS.ESA.2023.86.

34 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, 2022. IEEE Computer Society. doi:10.1109/ISMVL52857.2022.00019.

35 Emil L. Post. The two-valued iterative systems of mathematical logic. *Annals of Mathematics Studies*, 5, 1941.

36 Thomas J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the tenth annual ACM Symposium on Theory of Computing (STOC '78)*, pages 216–226, New York, NY, USA, 1978. Association for Computing Machinery. doi:10.1145/800133.804350.

37 Simon Thomas. Reducts of the random graph. *The Journal of Symbolic Logic*, 56(1):176–181, 1991. doi:10.2307/2274912.

38 Simon Thomas. Reducts of random hypergraphs. *Annals of Pure and Applied Logic*, 80(2):165–193, 1996. doi:10.1016/0168-0072(95)00061-5.

39 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.

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