September 28, 2023

THE COMPLEXITY OF RESILIENCE PROBLEMS VIA VALUED CONSTRAINT SATISFACTION PROBLEMS

MANUEL BODIRSKY, ŽANETA SEMANIŠINOVÁ, AND CARSTEN LUTZ

ABSTRACT. Valued constraint satisfaction problems (VCSPs) are a large class of computational optimisation problems. If the variables of a VCSP take values from a finite domain, then recent results in constraint satisfaction imply that the problem is in P or NP-complete, depending on the set of admitted cost functions. Here we study the larger class of cost functions over countably infinite domains that have an oligomorphic automorphism group. We present a hardness condition based on a generalisation of pp-constructability as known for (classical) CSPs. We also provide a universal-algebraic polynomial-time tractability condition, based on the concept of fractional polymorphisms.

We apply our general theory to study the computational complexity of resilience problems in database theory (under bag semantics). We show how to construct, for every fixed conjunctive query (and more generally for every union of conjunctive queries), a set of cost functions with an oligomorphic automorphism group such that the resulting VCSP is polynomial-time equivalent to the resilience problem; we only require that the query is *connected* and show that this assumption can be made without loss of generality. For the case where the query is *acylic*, we obtain a complexity dichotomy of the resilience problem, based on the dichotomy for finite-domain VCSPs. To illustrate the utility of our methods, we exemplarily settle the complexity of a (non-acyclic) conjunctive query whose computational complexity remained open in the literature by verifying that it satisfies our tractability condition. We conjecture that for resilience problems, our hardness and tractability conditions match, which would establish a complexity dichotomy for resilience problems for (unions of) conjunctive queries.

1. INTRODUCTION

If $\mathfrak{A}$ is a database and μ is a conjunctive query (or a union of conjunctive queries), then the *resilience* of μ in $\mathfrak{A}$ is the minimum number of tuples that need to be removed from $\mathfrak{A}$ to achieve that $\mathfrak{A}$ does not satisfy μ. For some queries μ, the computational complexity of computing the resilience of a given database is in P for some queries μ and NP-hard for others, but a complete classification has so far been elusive [20, 21, 37]. A natural variation of this problem introduced in [37] is that the input is a *bag database*, that is, it contains each tuple with a multiplicity $k \in \mathbb{N}$.

We present a surprising link between the resilience problem for (unions of) conjunctive queries under bag semantics and a large class of computational optimisation problems called *valued constraint satisfaction problems (VCSPs)*. In a VCSP, we are given a finite set of variables, a finite set of cost functions on those variables, and a threshold u, and the task is to find an assignment to the variables so that the sum of the costs is at most u. The computational complexity of such problems has been studied depending on the admitted cost functions, which we may view as a *valued structure*. A complete classification has been obtained for valued structures with a finite

The first two authors have been funded by the European Research Council (Project POCOCOP, ERC Synergy Grant 101071674) and by the DFG (Project FinHom, Grant 467967530). Views and opinions expressed are however those of the authors 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. The third author was supported by the DFG project LU 1417/3-1 QTEC.

domain, showing that the corresponding VCSPs are in P or NP-hard [11, 32, 34, 45, 46]. There are also some results about VCSPs of valued structures with infinite domains [7, 44].

We show that the resilience problem for every connected conjunctive query (and in fact for every *union* of connected conjunctive queries) can be formulated as a VCSP for a valued structure with an *oligomorphic automorphism group*, i.e., a structure with a countable domain that, for every fixed k, has only finitely many orbits of k-tuples under the action of the automorphism group. This property is important for classical CSPs (which can be seen as VCSPs where all cost functions take values in $\{0, \infty\}$), since it enables to extend and use some tools from finite-domain CSPs (see, e.g., [5]). The complexity classification for the general, not necessarily connected case can be reduced to the connected case. This result also extends to the more general setting where some relations or tuples are *exogenous*, meaning that they may *not* be removed from the database. If the conjunctive query is acyclic, then we even obtain a VCSP for a valued structure with a finite domain and obtain a P versus NP-complete dichotomy from the known dichotomy for such VCSPs.

This leads us to initiating the systematic study of VCSPs of countably infinite valued structures with an oligomorphic automorphism group. In particular, we develop a notion of *expressive power* which is based on *primitive positive definitions* and other complexity-preserving operators, inspired by the techniques known from VCSPs over finite domains. We use the expressive power to obtain polynomial-time reductions between such VCSPs, which we use to formulate a hardness condition for them and conjecture that for VCSPs that stem from resilience problems this hardness condition is necessary and sufficient, unless P = NP.

We also present an algebraic condition for valued structures that implies that the induced VCSP is in P, based on the concept of *fractional polymorphisms*, which generalise classical polymorphisms, a common tool for proving tractability of CSPs. To prove membership in P, we use a reduction to finite-domain VCSPs which can be solved by a linear programming relaxation technique. We conjecture that the resulting algorithm solves all resilience problems that are in P. We demonstrate the utility of our algebraic tractability condition by applying it to a concrete conjunctive query for which the computational complexity of resilience has been stated as an open problem in the literature [21] (Section 8.5).

Related Work. The study of resilience problems was initiated in [20]. The authors obtain a P versus NP-complete dichotomy for the class of self-join-free conjunctive queries, i.e., queries in which each relation symbol occurs at most once. In a subsequent paper [21], several results are obtained for conjunctive queries with self-joins of a specific form, while the authors also state a few open problems of similar nature that cannot be handled by their methods. In the latest article [37], Gatterbauer and Makhija present a unified approach to resilience problems based on integer linear programming, which works both under bag semantics and under set semantics. However, the new complexity results in [37] again concern only self-join free queries. Our approach is independent from self-joins and hence allows to study conjunctive queries that were difficult to treat before.

We stress that VCSPs of countable valued structures with an oligomorphic automorphism group greatly surpass resilience problems. For example, many computational problems in the recently very active area of *fixed parameter tractability for graph separation problems* can be formulated as VCSPs of appropriate countable valued structures with an oligomorphic automorphism group. In particular, this applies to the *directed feedback edge set problem*, the *directed symmetric multicut problem*, and many other problems with recent breakthrough results concerning FPT algorithms where the parameter is the number of edges in the graph that needs to be removed [29,30]. Several of these problems such as the *multicut problem* and the *coupled min cut problem* can even be formulated as VCSPs over a finite domain. Our notion of expressive power (and fractional polymorphisms) can be used to also study the parametrised complexity of all of these problems.

Outline. The article is organised from the general to the specific, starting with VCSPs in full generality (Section 2), then focussing on valued structures with an oligomorphic automorphism group (Section 3), for which our notion of expressive power (Section 4) leads to polynomial-time reductions. Our general hardness condition, which also builds upon the notion of expressive power, is presented in Section 5. To study the expressive power and to formulate general polynomial-time tractability results, we introduce the concept of *fractional polymorphisms* in Section 6 (they are probability distributions over operations on the valued structure). We take inspiration from the theory of VCSPs for finite-domain valued structures, but apply some non-trivial modifications that are specific to the infinite-domain setting (because the considered probability distributions are over uncountable sets). We then present a general polynomial-time tractability result (Theorem 7.17) which is phrased in terms of fractional polymorphisms. Section 8 applies the general theory to resilience problems. We illustrate the power of our approach by settling the computational complexity of a resilience problem for a concrete conjunctive query from the literature (Section 8.5). Section 9 closes with open problems for future research.

2. Preliminaries

2.1. **Valued Structures.**

The set $\{0, 1, 2, \dots\}$ of natural numbers is denoted by $\mathbb{N}$, the set of rational numbers is denoted by $\mathbb{Q}$, the set of non-negative rational numbers by $\mathbb{Q}_{\geq 0}$ and the set of positive rational numbers by $\mathbb{Q}_{>0}$. We use analogous notation for the set of real numbers $\mathbb{R}$ and the set of integers $\mathbb{Z}$. We also need an additional value ∞; all we need to know about ∞ is that

- $a < \infty$ for every $a \in \mathbb{Q}$,
- $a + \infty = \infty + a = \infty$ for all $a \in \mathbb{Q} \cup \{\infty\}$, and
- $0 \cdot \infty = \infty \cdot 0 = 0$ and $a \cdot \infty = \infty \cdot a = \infty$ for $a > 0$.

Let C be a set and let $k \in \mathbb{N}$. A *weighted relation of arity k over C* is a function $R \colon C^k \to \mathbb{Q} \cup \{\infty\}$. We write $\mathscr{R}_C^{(k)}$ for the set of all weighted relations of arity k, and define

$$\mathscr{R}_C := \bigcup_{k \in \mathbb{N}} \mathscr{R}_C^{(k)}.$$

A weighted relation is called *finite-valued* if it takes values only in $\mathbb{Q}$.

Example 2.1. *The* weighted equality relation $R_=$ *is the binary weighted relation defined over C by $R_=(x,y) = 0$ if $x = y$ and $R_=(x,y) = \infty$ otherwise. The* empty relation $R_\emptyset$ *is the unary weighted relation defined over C by $R_\emptyset(x) = \infty$ for all $x \in C$.*

A weighted relation $R \in \mathscr{R}_C^{(k)}$ that only takes values from $\{0, \infty\}$ will be identified with the following relation in the usual sense

$$\{a \in C^k \mid R(a) = 0\}.$$

For $R \in \mathscr{R}_C^{(k)}$ the *feasibility relation of R* is defined as

$$\mathrm{Feas}(R) := \{a \in C^k \mid R(a) < \infty\}.$$

A *(relational) signature* τ is a set of *relation symbols*, each equipped with an arity from $\mathbb{N}$. A *valued τ-structure* Γ consists of a set C, which is also called the *domain* of Γ, and a weighted relation $R^\Gamma \in \mathscr{R}_C^{(k)}$ for each relation symbol $R \in \tau$ of arity k. A *τ-structure* in the usual sense may then be identified with a valued τ-structure where all weighted relations only take values from $\{0, \infty\}$.

Example 2.2. *Let* $\tau = \{<\}$ *be a relational signature with a single binary relation symbol* $<$. *Let* $\Gamma_<$ *be the valued* τ*-structure with domain* $\{0,1\}$ *and where* $<(x,y) = 0$ *if* $x < y$, *and* $<(x,y) = 1$ *otherwise.*

Example 2.3. *Let* $\tau = \{E, N\}$ *be a relational signature with two binary relation symbols* E *and* N. *Let* Γ_{LCC} *be the valued* τ*-structure with domain* $\mathbb{N}$ *and where* $E(x,y) = 0$ *if* $x = y$ *and* $E(x,y) = 1$ *otherwise, and where* $N(x,y) = 0$ *if* $x \neq y$ *and* $N(x,y) = 1$ *otherwise.*

An *atomic* τ*-expression* is an expression of the form $R(x_1, \dots, x_k)$ for $R \in \tau$ and (not necessarily distinct) variable symbols $x_1, \dots, x_k$. A τ*-expression* is an expression ϕ of the form $\sum_{i \leq m} \phi_i$ where $m \in \mathbb{N}$ and ϕ_i for $i \in \{1, \dots, m\}$ is an atomic τ-expression. Note that the same atomic τ-expression might appear several times in the sum. We write $\phi(x_1, \dots, x_n)$ for a τ-expression where all the variables are from the set $\{x_1, \dots, x_n\}$. If Γ is a valued τ-structure, then a τ-expression $\phi(x_1, \dots, x_n)$ defines over Γ a member of $\mathscr{R}_C^{(n)}$, which we denote by ϕ^Γ. If ϕ is the empty sum then ϕ^Γ is constant 0.

2.2. **Valued Constraint Satisfaction.** In this section we assume that Γ is a fixed valued τ-structure for a *finite* signature τ. The weighted relations of Γ are also called *cost functions*. The *valued constraint satisfaction problem for* Γ, denoted by VCSP(Γ), is the computational problem to decide for a given τ-expression $\phi(x_1, \dots, x_n)$ and a given $u \in \mathbb{Q}$ whether there exists $a \in C^n$ such that $\phi^\Gamma(a) \leq u$. We refer to $\phi(x_1, \dots, x_n)$ as an *instance* of VCSP(Γ), and to u as the *threshold*. Tuples $a \in C^n$ such that $\phi^\Gamma(a) \leq u$ are called a *solution for* (ϕ, u). The *value* of ϕ (with respect to Γ) is defined to be

$$\inf_{a \in C^n} \phi^\Gamma(a).$$

In some contexts, it will be beneficial to consider only a given τ-expression ϕ to be the input of VCSP(Γ) (rather than ϕ and the threshold u) and a tuple $a \in C^n$ will then be called a *solution for* ϕ if the value of ϕ equals $\phi^\Gamma(a)$. Note that in general there might not be any solution. If there exists a tuple $a \in C^n$ such that $\phi^\Gamma(a) < \infty$ then ϕ is called *satisfiable*.

Note that our setting also captures classical CSPs, which can be viewed as the VCSPs for valued structures Γ that only contain cost functions that take value 0 or ∞. In this case, we will sometimes write CSP(Γ) for VCSP(Γ). Below we give a few examples of known optimisation problems that can be formulated as valued constraint satisfaction problems.

Example 2.4. *The problem* VCSP($\Gamma_<$) *for the valued structure* $\Gamma_<$ *from Example 2.2 models the* directed max-cut *problem: given a finite directed graph* (V, E) *(we do allow loops and multiple edges), partition the vertices* V *into two classes* A *and* B *such that the number of edges from* A *to* B *is maximal. Maximising the number of edges from* A *to* B *amounts to minimising the number* e *of edges within* A, *within* B, *and from* B *to* A. *So when we associate* A *to the preimage of* 0 *and* B *to the preimage of* 1, *computing the number* e *corresponds to finding the evaluation map* $s \colon V \to \{0,1\}$ *that minimises the value* $\sum_{(x,y) \in E} <(s(x), s(y))$, *which can be formulated as an instance of* VCSP($\Gamma_<$). *Conversely, every instance of* VCSP($\Gamma_<$) *corresponds to a directed max-cut instance. It is known that* VCSP($\Gamma_<$) *is NP-complete [23] (even if we do not allow loops and multiple edges in the input). We mention that this problem can be viewed as a resilience problem in database theory as explained in Section 8.*

Example 2.5. *Consider the valued structure* $\Gamma_\geq$ *with domain* $\{0,1\}$ *and the binary weighted relation* $\geq$ *defined by* $\geq(x,y) = 0$ *if* $x \geq y$ *and* $\geq(x,y) = 1$ *otherwise. Similarly to the previous example,* VCSP($\Gamma_\geq$) *models the directed min-cut problem, i.e., given a finite directed graph* (V, E), *partition* *the vertices V into two classes A and B such that the number of edges from A to B is minimal. The min-cut problem is solvable in polynomial time; see, e.g., [24].*

Example 2.6. *The problem of* least correlation clustering with partial information *[43, Example 5] is equal to* $\mathrm{VCSP}(\Gamma_{LCC})$ *where* Γ_{LCC} *is the valued structure from Example 2.3. It is a variant of the min-correlation clustering problem [1], where we have precisely one constraint between any two variables. The problem is NP-complete in both settings [23, 43].*

3. Oligomorphicity

Many facts about VCSPs for valued structures with a finite domain can be generalised to a large class of valued structures over an infinite domain, defined in terms of automorphisms. Automorphisms of valued structures are defined as follows.

Definition 3.1. *Let $k \in \mathbb{N}$, let $R \in \mathscr{R}_C^{(k)}$, and let α be a permutation of C. Then α* preserves *R if for all $a \in C^k$ we have $R(\alpha(a)) = R(a)$. If Γ is a valued structure with domain C, then an* automorphism *of Γ is a permutation of C that preserves all weighted relations of R.*

The set of all automorphisms of Γ is denoted by $\mathrm{Aut}(\Gamma)$, and forms a group with respect to composition. Let $k \in \mathbb{N}$. An *orbit of k-tuples* of a permutation group G is a set of the form $\{\alpha(a) \mid \alpha \in G\}$ for some $a \in C^k$. A permutation group G on a countable set is called *oligomorphic* if for every $k \in \mathbb{N}$ there are finitely many orbits of k-tuples in G. From now on, whenever we write that a structure has an oligomorphic automorphism group, we also imply that its domain is countable. Clearly, every valued structure with a finite domain has an oligomorphic automorphism group. A countable structure has an oligomorphic automorphism group if and only if it is ω-*categorical*, i.e., if all countable models of its first-order theory are isomorphic [25].

Example 3.2. *The automorphism group of Γ_{LCC} from Examples 2.3 and 2.6 is the full symmetric group and hence oligomorpohic.*

Lemma 3.3. *Let Γ be a valued structure with a countable domain C and an oligomorphic automorphism group. Then for every instance $\phi(x_1, \dots, x_n)$ of $\mathrm{VCSP}(\Gamma)$ there exists $a \in C^n$ such that the value of ϕ equals $\phi^\Gamma(a)$.*

Proof. The statement follows from the assumption that there are only finitely many orbits of n-tuples of $\mathrm{Aut}(\Gamma)$, because it implies that there are only finitely many possible values from $\mathbb{Q} \cup \{\infty\}$ for $\phi^\Gamma(a)$. $\square$

A first-order sentence is called *universal* if it is of the form $\forall x_1, \dots, x_l.\ \psi$ where ψ is quantifier-free. Every quantifier-free formula is equivalent to a formula in conjunctive normal form, so we assume in the following that quantifier-free formulas are of this form.

Recall that a τ-structure $\mathfrak{A}$ *embeds* into a τ structure $\mathfrak{B}$ if there is an injective map from A to B that preserves all relations of $\mathfrak{A}$ and their complements; the corresponding map is called an *embedding*. The *age* of a τ-structure is the class of all finite τ-structures that embed into it. A structure $\mathfrak{B}$ with a finite relational signature τ is called

- *finitely bounded* if there exists a universal τ-sentence ϕ such that a finite structure $\mathfrak{A}$ is in the age of $\mathfrak{B}$ if and only if $\mathfrak{A} \models \phi$.
- *homogeneous* if every isomorphism between finite substructures of $\mathfrak{B}$ can be extended to an automorphism of $\mathfrak{B}$.

If $\tau' \subseteq \tau$, then a τ'-structure $\mathfrak{B}'$ is called the *reduct* of $\mathfrak{B}$ if $\mathfrak{B}$ and $\mathfrak{B}'$ have the same domain and $R^{\mathfrak{B}'} = R^{\mathfrak{B}}$ for every $R \in \tau'$.

Note that for every structure $\mathfrak{B}$ with a finite relational signature, for every n there are only finitely many non-isomorphic substructures of $\mathfrak{B}$ of size n. Therefore, all countable homogeneous structures with a finite relational signature and all of their reducts have finitely many orbits of k-tuples for all $k \in \mathbb{N}$, and hence an oligomorphic automorphism group.

Theorem 3.4. *Let Γ be a countable valued structure with finite signature such that there exists a finitely bounded homogeneous structure $\mathfrak{B}$ with $\mathrm{Aut}(\mathfrak{B}) \subseteq \mathrm{Aut}(\Gamma)$. Then $\mathrm{VCSP}(\Gamma)$ is in NP.*

Proof. Let (ϕ, u) be an instance of $\mathrm{VCSP}(\Gamma)$ with n variables. Since $\mathrm{Aut}(\mathfrak{B}) \subseteq \mathrm{Aut}(\Gamma)$, every orbit of n-tuples of $\mathrm{Aut}(\Gamma)$ is determined by the substructure induced by $\mathfrak{B}$ on the elements of some tuple from the orbit. Note that two tuples $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$ lie in the same orbit of $\mathrm{Aut}(\mathfrak{B})$ if and only if the map that maps a_i to b_i for $i \in \{1, \dots, n\}$ is an isomorphism between the substructures induced by $\mathfrak{B}$ on $\{a_1, \dots, a_n\}$ and on $\{b_1, \dots, b_n\}$. Whether a given finite structure $\mathfrak{A}$ is in the age of a fixed finitely bounded structure $\mathfrak{B}$ can be decided in polynomial time: if ϕ is the universal τ-sentence which describes the age of $\mathfrak{B}$, it suffices to exhaustively check all possible instantiations of the variables of ϕ with elements of A and verify whether ϕ is true in $\mathfrak{A}$ under the instantiation. Hence, we may non-deterministically generate a structure $\mathfrak{A}$ with domain $\{1, \dots, n\}$ from the age of $\mathfrak{B}$ and then verify in polynomial time whether the value $\phi^\Gamma(b_1, \dots, b_n)$ is at most u for any tuple $(b_1, \dots, b_n) \in B^n$ such that $i \mapsto b_i$ is an embedding of $\mathfrak{A}$ into $\mathfrak{B}$. □

4. Expressive Power

One of the fundamental concepts in the theory of constraint satisfaction is the concept of *primitive positive definitions*, which is the fragment of first-order logic where only equality, existential quantification, and conjunction are allowed (in other words, negation, universal quantification, and disjunction are forbidden). The motivation for this concept is that relations with such a definition can be added to the structure without changing the complexity of the respective CSP. The natural generalisation to *valued* constraint satisfaction is the following notion of expressibility.

Definition 4.1. *Let Γ be a valued τ-structure. We say that $R \in \mathscr{R}_C^{(k)}$ can be* expressed *by Γ if there exists a τ-expression $\phi(x_1, \dots, x_k, y_1, \dots, y_n)$ such that for all $a \in C^k$ we have*

$$R(a) = \inf_{b \in C^n} \phi^\Gamma(a, b).$$

Note that $\inf_{b \in C^n} \phi^\Gamma(a, b)$ might be irrational or $-\infty$. If this is the case in Definition 4.1, then ϕ does not witness that R can be expressed in Γ since weighted relations must have weights from $\mathbb{Q} \cup \{\infty\}$. If C has an oligomorphic permutation group, however, then Lemma 3.3 guarantees existence. We will further see in Lemma 4.7 that if Γ has an oligomorphic automorphism group, then the addition of weighted relations that are expressible by Γ does not change the computational complexity of $\mathrm{VCSP}(\Gamma)$.

Another way to derive new relations from existing ones that preserves the computational complexity of the original VCSP is introduced in the following definition.

Definition 4.2. *Let $R, R' \in \mathscr{R}_C$. We say that R' can be obtained from R by*

- non-negative scaling *if there exists $r \in \mathbb{Q}_{\geq 0}$ such that $R = rR'$;*
- shifting *if there exists $s \in \mathbb{Q}$ such that $R = R' + s$.*

In the literature about the complexity of finite-domain VCSPs we find another operator on sets of weighted relations that preserves the complexity of the VCSP: the operator Opt (see, e.g., [22,35]).

Definition 4.3. *Let $R \in \mathscr{R}_C^{(k)}$. The relation containing all minimal-value tuples of R is defined as*

$$\mathrm{Opt}(R) := \{a \in \mathrm{Feas}(R) \mid R(a) \leq R(b) \text{ for every } b \in C^k\}.$$

Definition 4.4 (weighted relational clone)**.** *A* weighted relational clone *(over C) is a subset of $\mathscr{R}_C$ that contains $R_=$ and $R_\emptyset$ (from Example 2.1), and is closed under expressibility, shifting, and non-negative scaling,* Feas*, and* Opt*. For a valued structure Γ with domain C, we write $\langle\Gamma\rangle$ for the smallest relational clone that contains the weighted relations of Γ.*

The following example shows that neither the operator Opt nor the operator Feas is redundant in the definition above.

Example 4.5. *Consider the domain $C = \{0, 1, 2\}$ and the unary weighted relation R on C defined by $R(0) = 0$, $R(1) = 1$ and $R(2) = \infty$. Then the relation* $\mathrm{Feas}(R)$ *cannot be obtained from R by expressing, shifting, non-negative scaling and use of* Opt*. Similarly, the relation* $\mathrm{Opt}(R)$ *cannot be obtained from R by expressing, shifting, non-negative scaling and use of* Feas*.*

Remark 4.6. *Note that for every valued structure Γ and $R \in \langle\Gamma\rangle$, every automorphism of Γ is an automorphism of R.*

The motivation of Definition 4.4 for valued CSPs stems from the following lemma, which shows that adding relations in $\langle\Gamma\rangle$ does not change the complexity of the VCSP up to polynomial-time reductions. For valued structures over finite domains this is proved in [16], except for the operator Opt, for which a proof can be found in [22, Theorem 5.13]. Only parts of the proof can be generalised to valued structures over infinite domains in the general case, that is, when oligomorphic automorphism groups are not required; see, e.g., Schneider and Viola [41] and Viola [43, Lemma 7.1.4]. Note, however, that in these works the definition of VCSPs was changed: instead of asking whether a solution can be found of value at most u, they ask whether there exists a solution of value strictly less than u, to circumvent problems about infima that are not realised. Moreover, in [41] the authors restrict themselves to finite-valued weighted relations and hence do not consider the operator Opt. It is visible from Example 4.5 that the operator Opt cannot be simulated by the other ones already on finite domains, which is why it was introduced in [22]. The same is true for the operator Feas, which was included implicitly in [22] by allowing to scale weighted relations by 0 and defining $0 \cdot \infty = \infty$. In our approach, we work under the assumption that the valued structure has an oligomorphic automorphism group, which implies that infima in expressions are realized and the values of VCSPs of such structures can be attained. Therefore, we obtain a polynomial-time reduction for each of the operators in Definition 4.4 as in the finite domain case.

Lemma 4.7. *Let Γ be a valued structure with an oligomorphic automorphism group and a finite signature. Suppose that Δ is a valued structure with a finite signature over the same domain C such that every cost function of Δ is from $\langle\Gamma\rangle$. Then there is a polynomial-time reduction from* $\mathrm{VCSP}(\Delta)$ *to* $\mathrm{VCSP}(\Gamma)$*.*

Proof. Let τ be the signature of Γ. It suffices to prove the statement for expansions of Γ to signatures $\tau \cup \{R\}$ that extend τ with a single relation R, $R^\Delta \in \langle\Gamma\rangle$.

If $R^\Delta = R_\emptyset$, then an instance ϕ of $\mathrm{VCSP}(\Delta)$ with threshold $u \in \mathbb{Q}$ is unsatisfiable if and only if ϕ contains the symbol R or if it does not contain R and is unsatisfiable viewed as an instance of $\mathrm{VCSP}(\Gamma)$. In the former case, choose a k-ary relation symbol $S \in \tau$ and note that S^Γ attains only finitely many values, by the oligomorphicity of $\mathrm{Aut}(\Gamma)$. Let $u' \in \mathbb{Q}$ be smaller than all of them. Then $S(x_1,\dots,x_k)$ is an instance of $\mathrm{VCSP}(\Gamma)$ that never meets the threshold u', so this provides a correct reduction. In the latter case, for every $a \in C^n$ we have that $\phi^\Delta(a) = \phi^\Gamma(a)$; this provides a polynomial-time reduction.

Now suppose that $R^\Delta = R_=$. Let $\psi(x_{i_1},\dots,x_{i_k})$ be obtained from an instance $\phi(x_1,\dots,x_n)$ of $\mathrm{VCSP}(\Delta)$ by identifying all variables x_i and x_j such that ϕ contains the summand $R(x_i,x_j)$. Then ϕ is satisfiable if and only if the instance ψ is satisfiable, and $\inf_{a\in C^n}\phi^\Delta(a) = \inf_{b\in C^k}\psi^\Gamma(b)$; Again, this provides a polynomial-time reduction.

Next, consider the case that for some τ-expression $\delta(y_1,\dots,y_l,z_1,\dots,z_k)$ we have

$$R^\Delta(y_1,\dots,y_l) = \inf_{a\in C^k}\delta^\Gamma(y_1,\dots,y_l,a_1,\dots,a_k).$$

Let $\phi(x_1,\dots,x_n)$ be an instance of $\mathrm{VCSP}(\Delta)$. We replace each summand $R(y_1,\dots,y_l)$ in ϕ by $\delta(y_1,\dots,y_l,z_1,\dots,z_k)$ where $z_1,\dots,z_k$ are new variables (different for each summand). Let $\theta(x_1,\dots,x_n,w_1,\dots,w_t)$ be the resulting τ-expression after doing this for all summands that involve R. For any $a \in C^n$ we have that

$$\phi(a_1,\dots,a_n) = \inf_{b\in C^t}\theta(a_1,\dots,a_n,b)$$

and hence $\inf_{a\in C^n}\phi = \inf_{c\in C^{n+t}}\theta$; here we used the assumption that $\mathrm{Aut}(\Gamma)$ is oligomorphic. Since we replace each summand by an expression whose size is constant (since Γ is fixed and finite) the expression θ can be computed in polynomial time, which shows the statement.

Suppose that $R^\Delta = rS^\Gamma + s$ where $r \in \mathbb{Q}_{\geq 0}$, $s \in \mathbb{Q}$. Let $p \in \mathbb{Z}_{\geq 0}$ and $q \in \mathbb{Z}_{>0}$ be coprime integers such that $p/q = r$. Let (ϕ,u) be an instance of $\mathrm{VCSP}(\Delta)$ where $\phi(x_1,\dots,x_n) = \sum_{i=1}^{\ell}\phi_i + \sum_{j=1}^{k}\psi_j$, the summands ϕ_i contain only symbols from τ, and each ψ_j involves the symbol R. Let ψ'_j be the expression obtained from ψ_j by replacing R with S. For $i \in \{1,\dots,\ell\}$ replace ϕ_i with q copies of itself and for $j \in \{1,\dots,k\}$, replace ψ_j with p copies of ψ'_j; let $\phi'(x_1,\dots,x_n)$ be the resulting τ-expression. Define $u' := q(u-ks)$. Then for every $a \in C^n$ the following are equivalent:

$$\phi(a_1,\dots,a_n) = \sum_{i=1}^{\ell}\phi_i + \sum_{j=1}^{k}\left(\frac{p}{q}\psi'_j + s\right) \leq u$$

$$\phi'(a_1,\dots,a_n) = q\sum_{i=1}^{\ell}\phi_i + p\sum_{j=1}^{k}\psi'_j \leq qu - qks = u'$$

Since (ϕ',u') can be computed from (ϕ,u) in polynomial time, this provides the desired reduction.

Now suppose that $R^\Delta = \mathrm{Feas}(S^\Gamma)$ for some $S \in \tau$. Let (ϕ,u) be an instance of $\mathrm{VCSP}(\Delta)$, i.e., $\phi(x_1,\dots,x_n) = \sum_{i=1}^{\ell}\phi_i + \sum_{j=1}^{k}\psi_j$ where ψ_j, $j \in \{1,\dots,k\}$ are all the atomic expressions in ϕ that involve R. If $R^\Delta = R_\emptyset$, then the statement follows from the reduction for $R_\emptyset$. Therefore, suppose that this not the case and let w be the maximum finite weight assigned by S. Note that there are only finitely many values that the ℓ atoms ϕ_i may take and therefore only finitely many values that $\sum_{i=1}^{\ell}\phi_i$ may take. Let v be the smallest of these values such that $v > u$ and let $d = v - u$; if v does not exist, let $d = 1$. To simplify the notation, set $t = \lceil (kw)/d \rceil + 1$. Let ψ'_j be the τ-expression resulting from ψ_j by replacing the symbol R by the symbol S. Let ϕ' be the τ-expression obtained from ϕ by replacing each atom ϕ_i with t copies of it and replacing every atom ψ_j by ψ'_j. Let $(\phi', tu + kw)$ be the resulting instance of $\mathrm{VCSP}(\Gamma)$; note that it can be computed in polynomial time.

We claim that for every $a \in C^n$, the following are equivalent:

$$\phi(a_1,\dots,a_n) = \sum_{i=1}^{\ell} \phi_i + \sum_{j=1}^{k} \psi_j \leq u \tag{1}$$

$$\phi'(a_1,\dots,a_n) = t \cdot \sum_{i=1}^{\ell} \phi_i + \sum_{j=1}^{k} \psi'_j \leq tu + kw \tag{2}$$

If (1) holds, then by the definition of Feas we must have $\psi_j = 0$ for every $j \in \{1,\dots,k\}$. Thus $\sum_{i=1}^{\ell} \phi_i \leq u$ and $\sum_{j=1}^{k} \psi'_j \leq kw$, which implies (2). Conversely, if (2) holds, then ψ'_j is finite for every $j \in \{1,\dots,k\}$ and hence $\psi_j = 0$. Moreover, (2) implies

$$\sum_{i=1}^{\ell} \phi_i \leq u + \frac{kw}{t}.$$

Note that if v exists, then $u + (kw)/t < v$. Therefore (regardless of the existence of v), this implies $\sum_{i=1}^{\ell} \phi_i \leq u$, which together with what we have observed previously shows (1).

Finally, we consider the case that $R^\Delta = \mathrm{Opt}(S^\Gamma)$ for some relation symbol $S \in \tau$. Since τ is finite and $\mathrm{Aut}(\Gamma)$ is oligomorphic, we may assume without loss of generality that the minimum weight of all weighted relations in Δ equals 0; otherwise, we subtract the smallest weight assigned to a tuple by some weighted relation in Δ. This transformation does not affect the computational complexity of the VCSP (up to polynomial-time reductions). We may also assume that S^Γ takes finite positive values, because otherwise $\mathrm{Opt}(S^\Gamma) = S^\Gamma$ and the statement is trivial. Let m be the smallest positive weight assigned by S^Γ and let M be the largest finite weight assigned by any weighted relation of Γ (again we use that τ is finite and that $\mathrm{Aut}(\Gamma)$ is oligomorphic). Let (ϕ, u), where $\phi(x_1,\dots,x_n) = \sum_{i=1}^{k} \phi_i$, be an instance of $\mathrm{VCSP}(\Delta)$. For $i \in \{1,\dots,k\}$, if ϕ_i involves the symbol R, then replace it by $k \cdot \lceil M/m \rceil + 1$ copies and replace R by S. Let ϕ' be the resulting τ-expression. We claim that $a \in C^n$ is a solution to the instance $(\phi', \min(kM, u))$ of $\mathrm{VCSP}(\Gamma)$ if and only if it is the solution to (ϕ, u).

If $a \in C^n$ is such that $\phi(a) \leq u$ then for every $i \in \{1,\dots,k\}$ such that ϕ_i involves R we have $\phi_i(a) = 0$. In particular, the minimal value attained by S^Γ equals 0 by our assumption, and hence $\phi'(a) = \phi(a) \leq u$ and hence $\phi'(a) \leq \min(kM, u)$. Now suppose that $\phi(a) > u$. Then $\phi'(a) > u \geq \min(kM, u)$ or there exists an $i \in \{1,\dots,k\}$ such that $\phi_i(a) = \infty$. If ϕ_i does not involve the symbol R, then $\phi'(a) = \infty$ as well. If ϕ_i involves the symbol R, then $\phi'(a) \geq (k \cdot \lceil M/m \rceil + 1)m > kM$. In any case, $\phi'(a) > \min(kM, u)$. Since ϕ' can be computed from ϕ in polynomial time, this concludes the proof. □

The next example illustrates the use of Lemma 4.7 for obtaining hardness results.

Example 4.8. *We revisit the countably infinite valued structure Γ_{LCC} from Example 2.3. Recall that $\mathrm{VCSP}(\Gamma_{LCC})$ is the least correlation clustering problem with partial information and that $\mathrm{Aut}(\Gamma_{LCC})$ is oligomorphic. Let Γ_{EC} be the relational structure with the same domain as Γ_{LCC} and the relation $R := \{(x,y,z) \mid (x = y \wedge y \neq z) \vee (x \neq y \wedge y = z)\}$ (attaining values 0 and ∞). Note that*

$$R(x,y,z) = \mathrm{Opt}(N(x,z) + N(x,z) + E(x,y) + E(y,z)).$$

This provides an alternative proof of NP-hardness of the least correlation clustering with partial information via Lemma 4.7, because $\mathrm{CSP}(\Gamma_{EC})$ is known to be NP-hard [6].

5. Hardness from pp-Constructions

A universal-algebraic theory of VCSPs for finite valued structures has been developed in [34], following the classical approach to CSPs which is based on the concepts of cores, addition of constants, and primitive positive interpretations. Subsequently, an important conceptual insight has been made for classical CSPs which states that every structure that can be interpreted in the expansion of the core of the structure by constants can also be obtained by taking a pp-power if we then consider structures up to homomorphic equivalence [3]. We are not aware of any published reference that adapts this perspective to the algebraic theory of VCSPs, so we develop (parts of) this approach here. As in [3], we immediately step from valued structures with a finite domain to the more general case of valued structures with an oligomorphic automorphism group.

Definition 5.1 (pp-power)**.** *Let* Γ *be a valued structure with domain* C *and let* $d \in \mathbb{N}$*. Then a (d-th)* pp-power *of* Γ *is a valued structure* Δ *with domain* C^d *such that for every weighted relation* R *of* Δ *of arity* k *there exists a weighted relation* S *of arity* kd *in* $\langle \Gamma \rangle$ *such that*

$$R((a_1^1,\dots,a_d^1),\dots,(a_1^k,\dots,a_d^k)) = S(a_1^1,\dots,a_d^1,\dots,a_1^k,\dots,a_d^k).$$

The name 'pp-power' comes from 'primitive positive power', since for relational structures expressibility is captured by primitive positive formulas. The following proposition shows that the VCSP of a pp-power reduces to the VCSP of the original structure.

Proposition 5.2. *Let* Γ *and* Δ *be valued structures such that* $\mathrm{Aut}(\Gamma)$ *is oligomorphic and* Δ *is a pp-power of* Γ*. Then* $\mathrm{Aut}(\Delta)$ *is oligomorphic and there is a polynomial-time reduction from* $\mathrm{VCSP}(\Delta)$ *to* $\mathrm{VCSP}(\Gamma)$*.*

Proof. Let d be the dimension of the pp-power and let τ be the signature of Γ. By Remark 4.6, $\mathrm{Aut}(\Gamma) \subseteq \mathrm{Aut}(\Delta)$ and thus $\mathrm{Aut}(\Delta)$ is oligomorphic. By Lemma 4.7, we may suppose that for every weighted relation R of arity k of Δ the weighted relation $S \in \langle \Gamma \rangle$ of arity dk from the definition of pp-powers equals S^Γ for some $S \in \tau$. Let (ϕ, u) be an instance of $\mathrm{VCSP}(\Delta)$. For each variable x of ϕ we introduce d new variables $x_1,\dots,x_d$. For each summand $R(y^1,\dots,y^k)$ we introduce a summand $S(y_1^1,\dots,y_d^1,\dots,y_1^k,\dots,y_d^k)$; let ψ be the resulting τ-expression. It is now straightforward to verify that (ϕ,u) has a solution with respect to Δ if and only if (ψ,u) has a solution with respect to Γ. □

Note that, in particular, if $\mathrm{VCSP}(\Gamma)$, parametrized by the threshold u, is fixed-parameter tractable, then so is $\mathrm{VCSP}(\Delta)$.

If C and D are sets, then we equip the space C^D of functions from D to C with the topology of pointwise convergence, where C is taken to be discrete. In this topology, a basis of open sets is given by

$$\mathscr{S}_{a,b} := \{f \in C^D \mid f(a) = b\}$$

for $a \in D^k$ and $b \in C^k$ for some $k \in \mathbb{N}$. For any topological space T, we denote by $B(T)$ the Borel σ-algebra on T, i.e., the smallest subset of the powerset $\mathcal{P}(T)$ which contains all open sets and is closed under countable intersection and complement. We write $[0,1]$ for the set $\{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$.

Definition 5.3 (fractional map)**.** *Let* C *and* D *be sets. A* fractional map *from* D *to* C *is a probability distribution*

$$(C^D, B(C^D), \omega \colon B(C^D) \to [0,1]),$$

that is,

- *ω is countably additive: if $A_1, A_2, \dots \in B(C^D)$ are disjoint, then*
$$\omega(\bigcup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \omega(A_i).$$
- $\omega(C^D) = 1$.

If $f \in C^D$, we often write $\omega(f)$ instead of $\omega(\{f\})$. Note that $\{f\} \in B(C^D)$ for every f. The set $[0,1]$ carries the topology inherited from the standard topology on $\mathbb{R}$. We also view $\mathbb{R} \cup \{\infty\}$ as a topological space with a basis of open sets given by all open intervals (a,b) for $a,b \in \mathbb{R}$, $a < b$ and additionally all sets of the form $\{x \in \mathbb{R} \mid x > a\} \cup \{\infty\}$.

A *(real-valued) random variable* is a *measurable function* $X : T \to \mathbb{R} \cup \{\infty\}$, i.e., pre-images of elements of $B(\mathbb{R} \cup \{\infty\})$ under X are in $B(T)$. If X is a real-valued random variable, then the *expected value of X (with respect to a probability distribution ω)* is denoted by $E_\omega[X]$ and is defined via the Lebesgue integral
$$E_\omega[X] := \int_T X d\omega.$$

Recall that the Lebesgue integral $\int_T X d\omega$ need not exist, in which case $E_\omega[X]$ is undefined; otherwise, the integral equals a real number, ∞, or $-\infty$. For the convenience of the reader we recall the definition and some properties of the Lebesgue integral, specialised to our setting, in Appendix A. Also recall that the expected value is

- *linear*, i.e., for every $a, b \in \mathbb{R}$ and random variables X, Y such that $E_\omega[X]$ and $E_\omega[Y]$ exist and $aE_\omega[X] + bE_\omega[Y]$ is defined we have
$$E_\omega[aX + bY] = aE_\omega[X] + bE_\omega[Y];$$
- *monotone*, i.e., if X, Y are random variables such that $E_\omega[X]$ and $E_\omega[Y]$ exist and $X(f) \leq Y(f)$ for all $f \in T$, then $E_\omega[X] \leq E_\omega[Y]$.

Let C and D be sets. In the rest of the paper, we will work exclusively on a topological space C^D of maps $f : D \to C$ and the special case $\mathscr{O}_C^{(\ell)}$ for some $\ell \in \mathbb{N}$ (i.e., $D = C^\ell$). Note that if C and D are infinite, then these spaces are uncountable and hence there are probability distributions ω such that $\omega(A) = 0$ for every 1-element set A. Therefore, in these cases, $E_\omega[X]$ for a random variable X might not be expressible as a sum.

Definition 5.4 (fractional homomorphism)**.** *Let Γ and Δ be valued τ-structures with domains C and D, respectively. A* fractional homomorphism *from Δ to Γ is a fractional map from D to C such that for every $R \in \tau$ of arity k and every tuple $a \in D^k$ it holds for the random variable $X : C^D \to \mathbb{R} \cup \{\infty\}$ given by*
$$f \mapsto R^\Gamma(f(a))$$
that $E_\omega[X]$ exists and that
$$E_\omega[X] \leq R^\Delta(a).$$

The following lemma shows that if $\mathrm{Aut}(\Gamma)$ is oligomorphic, then the expected value from Definition 5.4 always exists.

Lemma 5.5. *Let C and D be sets, $a \in D^k$, $R \in \mathscr{R}_C^{(k)}$. Let $X : C^D \to \mathbb{R} \cup \{\infty\}$ be the random variable given by*
$$f \mapsto R(f(a)).$$
If $\mathrm{Aut}(C; R)$ is oligomorphic, then $E_\omega[X]$ exists and $E_\omega[X] > -\infty$.

Proof. It is enough to show that $\int_{C^D} X^- d\omega \neq \infty$. Since $\mathrm{Aut}(C;R)$ is oligomorphic, there are only finitely many orbits of k-tuples in $\mathrm{Aut}(C;R)$. Let $O_1,\dots,O_m$ be all orbits of k-tuples of $\mathrm{Aut}(C;R)$ on which R is negative. For every $i \in \{1,\dots,m\}$, let $b_i \in O_i$. Then we obtain (see (21) in Appendix A for a detailed derivation of the first equality)

$$\begin{aligned}\int_{C^D} X^- d\omega &= \sum_{b\in C^k, R(b)<0} -R(b)\omega(\mathscr{S}_{a,b}) \\ &= -\sum_{i=1}^m R(b_i) \sum_{b\in O_i} \omega(\mathscr{S}_{a,b}) \\ &= -\sum_{i=1}^m R(b_i)\, \omega\left(\bigcup_{b\in O_i} \mathscr{S}_{a,b}\right) \\ &\leq -\sum_{i=1}^m R(b_i) < \infty. \end{aligned}$$ □

Lemma 5.6. *Let Γ_1, Γ_2, Γ_3 be countable valued τ-structures such that there exists a fractional homomorphism ω_1 from Γ_1 to Γ_2 and a fractional homomorphism ω_2 from Γ_2 to Γ_3. Then there exists a fractional homomorphism $\omega_3 := \omega_2 \circ \omega_1$ from Γ_1 to Γ_3.*

Proof. Let C_1, C_2, C_3 be the domains of Γ_1, Γ_2, and Γ_3, respectively. If $a \in C_1^k$ and $c \in C_3^k$, for some $k \in \mathbb{N}$, then define

$$\omega_3(\mathscr{S}_{a,c}) := \sum_{b\in C_2^k} \omega_1(\mathscr{S}_{a,b})\omega_2(\mathscr{S}_{b,c}).$$

Note that on sets of this form, i.e., on basic open sets in $C_3^{C_1}$, ω_3 is countably additive. Since our basis of open sets is closed under intersection, this definition extends uniquely to all of $B(C_3^{C_1})$ by Dynkin's π-λ theorem. □

The following was shown for valued structures over finite domains in [12, Proposition 8.4].

Proposition 5.7. *Let Γ and Δ be valued τ-structures with domains C and D and with a fractional homomorphism ω from Δ to Γ. Then the value of every* VCSP *instance ϕ with respect to Γ is at most the value of ϕ with respect to Δ.*

Proof. Let $\phi(x_1,\dots,x_n) = \sum_{i=1}^m R_i(x_{j_1^i},\dots,x_{j_{k_i}^i})$ be a τ-expression, where $j_1^i,\dots,j_{k_i}^i \in \{1,\dots,n\}$ for every $i \in \{1,\dots m\}$. To simplify the notation in the proof, if $v = (v_1,\dots,v_t)$ is a t-tuple of elements of some domain and $i_1,\dots,i_s \in \{1,\dots,t\}$, we will write $v_{i_1,\dots,i_s}$ for the tuple $(v_{i_1},\dots,v_{i_s})$.

Let $\varepsilon > 0$. From the definition of infimum, there exists $a^* \in D^n$ such that

$$\phi^\Delta(a^*) \leq \inf_{a\in D^n} \phi^\Delta(a) + \varepsilon/2 \tag{3}$$

and $f^* \in C^D$ such that

$$\phi^\Gamma(f^*(a^*)) \leq \inf_{f\in C^D} \phi^\Gamma(f(a^*)) + \varepsilon/2. \tag{4}$$

Note that $E_\omega[f \mapsto R_i^\Gamma(f(a^*)_{j_1^i,\dots,j_{k_i}^i})]$ exists for every $i \in \{1,\dots,m\}$ by the definition of a fractional homomorphism. Suppose first that $\sum_{i=1}^m E_\omega[f \mapsto R_i^\Gamma(f(a^*)_{j_1^i,\dots,j_{k_i}^i})]$ is defined. Then

$$\begin{aligned}
\inf_{b\in C^n} \phi^\Gamma(b) &\leq \phi^\Gamma(f^*(a^*)) && \text{(definition of infimum)}\\
&\leq \inf_{f\in C^D} \phi^\Gamma(f(a^*)) + \varepsilon/2 && \text{(by (4))}\\
&\leq E_\omega[f \mapsto \phi^\Gamma(f(a^*))] + \varepsilon/2 && \text{(by the monotonicity of } E_\omega)\\
&= \sum_{i=1}^m E_\omega[f \mapsto R_i^\Gamma(f(a^*)_{j_1^i,\dots,j_{k_i}^i})] + \varepsilon/2 && \text{(by the linearity of } E_\omega)\\
&\leq \sum_{i=1}^m R_i^\Delta(a^*_{j_1^i,\dots,j_{k_i}^i}) + \varepsilon/2 && \text{(since } \omega \text{ is a fractional homomorphism)}\\
&= \phi^\Delta(a^*) + \varepsilon/2\\
&\leq \inf_{a\in D^n} \phi^\Delta(a) + \varepsilon && \text{(by (3)).}
\end{aligned}$$

Since $\varepsilon > 0$ was chosen arbitrarily, it follows that the value of ϕ with respect to Γ is at most the value of ϕ with respect to Δ.

Suppose now that $\sum_{i=1}^m E_\omega[f \mapsto R_i^\Gamma(f(a^*)_{j_1^i,\dots,j_{k_i}^i})]$ is not defined. Then there exists $i \in \{1,\dots,m\}$ such that $E_\omega[f \mapsto R_i^\Gamma(f(a^*)_{j_1^i,\dots,j_{k_i}^i})] = \infty$. By the definition of a fractional homomorphism, this implies that $R_i^\Delta(a^*_{j_1^i,\dots,j_{k_i}^i}) = \infty$ and hence $\sum_{i=1}^m R_i^\Delta(a^*_{j_1^i,\dots,j_{k_i}^i}) = \infty$. Therefore, we obtain as above that

$$\inf_{b\in C^n} \phi^\Gamma(b) \leq \inf_{a\in D^n} \phi^\Delta(a),$$

which is what we wanted to prove. □

Remark 5.8. *For finite domains, the converse of Proposition 5.7 is true as well [12, Proposition 8.4].*

We say that two valued τ-structures Γ and Δ are *fractionally homomorphically equivalent* if there exists a fractional homomorphisms from Γ to Δ and from Δ to Γ. Clearly, fractional homomorphic equivalence is indeed an equivalence relation on valued structures of the same signature.

Corollary 5.9. *Let Γ and Δ be valued τ-structures with oligomorphic automorphism groups that are fractionally homomorphically equivalent. Then* VCSP(Γ) *and* VCSP(Δ) *are polynomial-time equivalent.*

Proof. In fact, the two problems VCSP(Γ) and VCSP(Δ) coincide. By Proposition 5.7, for every instance ϕ, the values of ϕ with respect to Γ and Δ are equal. By Lemma 3.3, the value is attained in both structures and hence every instance ϕ with a threshold u has a solution with respect to Γ if and only if it has a solution with respect to Δ. □

Remark 5.10. *Note that if Γ and Δ are classical relational τ-structures that are homomorphically equivalent in the classical sense, then they are fractionally homomorphically equivalent when we view them as valued structures: if h_1 is the homomorphism from Γ to Δ and h_2 is the homomorphism from Δ to Γ, then this is witnessed by the fractional homomorphisms ω_1 and ω_2 such that $\omega_1(h_1) = \omega_2(h_2) = 1$.*

Definition 5.11 (pp-construction)**.** *Let* Γ *and* Δ *be valued structures. We say that* Δ *has a* pp-construction *in* Γ *if* Δ *is fractionally homomorphically equivalent to a structure* Δ' *which is a pp-power of* Γ*.*

Corollary 5.12. *Let* Γ *and* Δ *be valued structures with finite signatures and oligomorphic automorphism groups such that* Δ *has a pp-construction in* Γ*. Then there is a polynomial-time reduction from* $\mathrm{VCSP}(\Delta)$ *to* $\mathrm{VCSP}(\Gamma)$*.*

Proof. Combine Proposition 5.2 and Corollary 5.9. □

Let OIT be the following relation

$$\mathrm{OIT} = \{(0,0,1),(0,1,0),(1,0,0)\}.$$

It is well-known (see, e.g., [5]) that $\mathrm{CSP}(\{0,1\};\mathrm{OIT})$ is NP-complete.

Corollary 5.13. *Let* Γ *be a valued structure with a finite signature and oligomorphic automorphism group such that* $(\{0,1\};\mathrm{OIT})$ *has a pp-construction in* Γ*. Then* $\mathrm{VCSP}(\Gamma)$ *is NP-hard.*

Proof. Follows from the NP-hardness of $\mathrm{CSP}(\{0,1\};\mathrm{OIT})$ via Corollary 5.12. □

Lemma 5.14. *The relation of pp-constructibility on the class of countable valued structures is transitive.*

Proof. Clearly, a pp-power of a pp-power is again a pp-power, and fractional homomorphic equivalence is transitive by Lemma 5.6. We are therefore left to prove that if Γ and Δ are valued structures such that Δ is a d-dimensional pp-power of Γ, and if Γ' is fractionally homomorphically equivalent to Γ via fractional homomorphisms $\omega_1 \colon \Gamma \to \Gamma'$ and $\omega_2 \colon \Gamma' \to \Gamma$, then Δ also has a pp-construction in Γ'.

Let C and C' be the domains of Γ and Γ', respectively. Take the τ-expressions that define the weighted relations of Δ over Γ, and interpret them over Γ' instead of Γ; let Δ' be the resulting valued structure. Note that Δ' is a d-dimensional pp-power of Γ'. For a map $f \colon \Gamma \to \Gamma'$, let $\tilde{f} \colon \Delta \to \Delta'$ be given by $(x_1,\dots,x_d) \mapsto (f(x_1),\dots,f(x_d))$. Then for all $S \in B((C')^C)$ we define

$$\tilde{\omega}_1(\{\tilde{f} \mid f \in S\}) := \omega_1(S)$$

and

$$\tilde{\omega}_1(\tilde{S}) := \tilde{\omega}_1(\tilde{S} \cap \{\tilde{f} \mid f \in (C')^C\})$$

for all $\tilde{S} \in B\big(((C')^d)^{C^d}\big)$. Note that $\tilde{\omega}_1$ is a fractional homomorphism from Δ to Δ'. Analogously we obtain from ω_2 a fractional homomorphism $\tilde{\omega}_2$ from Δ' to Δ. Therefore, Δ is fractionally homomorphically equivalent to Δ', which is a pp-power of Γ'. In other words, Δ has a pp-construction in Γ'. □

6. Fractional Polymorphisms

In this section we introduce *fractional polymorphisms* of valued structures; they are an important tool for formulating tractability results and complexity classifications of VCSPs. For valued structures with a finite domain, our definition specialises to the established notion of a fractional polymorphism which has been used to study the complexity of VCSPs for valued structures over finite domains (see, e.g. [42]). Our approach is different from the one of Viola and Schneider [41,43] in that we work with arbitrary probability spaces instead of distributions with finite support. As we will see in Section 7, fractional polymorphisms can be used to give sufficient conditions for tractability of VCSPs of certain valued structures with oligomorphic automorphism groups. This justifies the more general notion of a fractional polymorphism, as it might provide a tractability proof for more problems. We do not know if there are examples in our setting where it is necessary to use the more general notion; see Question 9.2.

Let $\mathscr{O}_C^{(\ell)}$ be the set of all operations $f\colon C^\ell \to C$ on a set C of arity ℓ. We equip $\mathscr{O}_C^{(\ell)}$ with the topology of pointwise convergence, where C is taken to be discrete. That is, the basic open sets are of the form

$$\mathscr{S}_{a^1,\dots,a^\ell,b} := \{f \in \mathscr{O}_C^{(\ell)} \mid f(a^1,\dots,a^\ell) = b\} \tag{5}$$

where $a^1,\dots,a^\ell,b \in C^m$, for some $m \in \mathbb{N}$, and f is applied componentwise. Let

$$\mathscr{O}_C := \bigcup_{\ell \in \mathbb{N}} \mathscr{O}_C^{(\ell)}.$$

Definition 6.1 (fractional operation)**.** *Let $\ell \in \mathbb{N}$. A* fractional operation on C of arity ℓ *is a probability distribution*

$$\big(\mathscr{O}_C^{(\ell)}, B(\mathscr{O}_C^{(\ell)}), \omega\colon B(\mathscr{O}_C^{(\ell)}) \to [0,1]\big).$$

The set of all fractional operations on C of arity ℓ is denoted by $\mathscr{F}_C^{(\ell)}$, and $\mathscr{F}_C := \bigcup_{\ell\in\mathbb{N}} \mathscr{F}_C^{(\ell)}$.

If the reference to C is clear, we occasionally omit the subscript C. We often use ω for both the entire fractional operation and for the map $\omega\colon B(\mathscr{O}_C^{(\ell)}) \to [0,1]$.

Definition 6.2. *A fractional operation $\omega \in \mathscr{F}_C^{(\ell)}$* improves *a k-ary weighted relation $R \in \mathscr{R}_C^{(k)}$ if for all $a^1,\dots,a^\ell \in C^k$*

$$E := E_\omega[f \mapsto R(f(a^1,\dots,a^\ell))]$$

exists and

$$E \leq \frac{1}{\ell}\sum_{j=1}^{\ell} R(a^j). \tag{6}$$

Note that (6) has the interpretation that the expected value of $R(f(a^1,\dots,a^\ell))$ is at most the average of the values $R(a^1),\dots,R(a^\ell)$. Also note that if R is a classical relation improved by a fractional operation ω and $\omega(f) > 0$ for $f \in \mathscr{O}^{(\ell)}$, then f must preserve R in the usual sense. It follows from Lemma 5.5 that if $\mathrm{Aut}(C;R)$ is oligomorphic, then $E_\omega[f \mapsto R(f(a^1,\dots,a^\ell))]$ always exists and is greater than $-\infty$.

Definition 6.3 (fractional polymorphism)**.** *If ω improves every weighted relation in Γ, then ω is called a* fractional polymorphism of Γ*; the set of all fractional polymorphisms of Γ is denoted by* $\mathrm{fPol}(\Gamma)$.

Remark 6.4. *A fractional polymorphism of arity ℓ of a valued structure Γ might also be viewed as a fractional homomorphism from a specific ℓ-th pp-power of Γ to Γ, which we denote by Γ^ℓ: if C is the domain and τ the signature of Γ, then the domain of Γ^ℓ is C^ℓ, and for every $R \in \tau$ of arity k we have*

$$R^{\Gamma^\ell}((a_1^1,\dots,a_\ell^1),\dots,(a_1^k,\dots,a_\ell^k)) := \frac{1}{\ell}\sum_{i=1}^{\ell} R^\Gamma(a_i^1,\dots,a_i^k).$$

Example 6.5. *Let $\pi_i^\ell \in \mathscr{O}_C^{(\ell)}$ be the i-th projection of arity ℓ, which is given by $\pi_i^\ell(x_1,\dots,x_\ell) = x_i$. The fractional operation Id_ℓ of arity ℓ such that $\mathrm{Id}_\ell(\pi_i^\ell) = \frac{1}{\ell}$ for every $i \in \{1,\dots,\ell\}$ is a fractional polymorphism of every valued structure with domain C.*

Example 6.6. *Let Γ be a valued structure and $\alpha \in \mathrm{Aut}(\Gamma)$. The fractional operation $\omega \in \mathscr{F}_C^{(1)}$ defined by $\omega(\alpha) = 1$ is a fractional polymorphism of Γ.*

Let $\mathscr{C} \subseteq \mathscr{F}_C$. We write $\mathscr{C}^{(\ell)}$ for $\mathscr{C} \cap \mathscr{F}_C^{(\ell)}$ and $\mathrm{Imp}(\mathscr{C})$ for the set of weighted relations that are improved by every fractional operation in $\mathscr{C}$.

Lemma 6.7. *Let $R \in \mathscr{R}_C^{(k)}$ and let Γ be a valued structure with domain C and an automorphism $\alpha \in \mathrm{Aut}(\Gamma)$ which does not preserve R. Then $R \notin \mathrm{Imp}(\mathrm{fPol}(\Gamma)^{(1)})$.*

Proof. Since α does not preserve R, there exists $a \in C^k$ such that $R(a) \neq R(\alpha(a))$. If $R(\alpha(a)) > R(a)$, then let $\omega \in \mathscr{F}_C^{(1)}$ be the fractional operation defined by $\omega(\alpha) = 1$. Then ω improves every weighted relation of Γ and does not improve R. If $R(\alpha(a)) < R(a)$, then the fractional polymorphism ω of Γ given by $\omega(\alpha^{-1}) = 1$ does not improve R. □

Parts of the arguments in the proof of the following lemma can be found in the proof of [43, Lemma 7.2.1]; however, note that the author works with a more restrictive notion of fractional operation, so we cannot cite this result.

Lemma 6.8. *For every valued τ-structure Γ over a countable domain C we have*

$$\langle \Gamma \rangle \subseteq \mathrm{Imp}(\mathrm{fPol}(\Gamma)).$$

Proof. Let $\omega \in \mathrm{fPol}(\Gamma)^{(\ell)}$. By definition, ω improves every weighted relation R of Γ. It is clear that ω also preserves $\phi_\emptyset$. To see that ω preserves $\phi_=$, let $a^1, \dots, a^\ell \in C^2$. Note that either $a_1^i = a_2^i$ for every $i \in \{1, \dots, \ell\}$, in which case $f(a_1^1, \dots, a_1^\ell) = f(a_2^1, \dots, a_2^\ell)$ for every $f \in \mathscr{O}_C^{(\ell)}$, and hence

$$E_\omega[f \mapsto \phi_=(f(a^1, \dots, a^\ell))] = 0 = \frac{1}{\ell} \sum_{j=1}^{\ell} \phi_=(a^j).$$

or $a_1^i \neq a_2^i$ for some $i \in \{1, \dots, \ell\}$, in which case $\frac{1}{\ell} \sum_{j=1}^{\ell} \phi_=(a^j) = \infty$ and the inequality in (6) is again satisfied.

The statement is also clear for weighted relations obtained from weighted relations in Γ by non-negative scaling and addition of constants, since these operations preserve the inequality in (6) by the linearity of expectation.

Let $\phi(x_1, \dots, x_k, y_1, \dots, y_n)$ be a τ-expression. We need to show that the fractional operation ω improves the k-ary weighted relation R defined for every $a \in C^k$ by $R(a) = \inf_{b \in C^n} \phi^\Gamma(a, b)$. Since ϕ is a τ-expression, there are $R_i \in \tau$ such that

$$\phi(x_1, \dots, x_k, y_1, \dots, y_n) = \sum_{i=1}^{m} R_i(x_{p_1^i}, \dots, x_{p_{k_i}^i}, y_{q_1^i}, \dots, y_{q_{n_i}^i})$$

for some $k_i, n_i \in \mathbb{N}$, $p_1^i, \dots, p_{k_i}^i \in \{1, \dots, k\}$ and $q_1^i, \dots, q_{n_i}^i \in \{1, \dots, n\}$.

In this paragraph, if $v = (v_1, \dots, v_t) \in C^t$ and $i_1, \dots, i_s \in \{1, \dots, t\}$, we will write $v_{i_1, \dots, i_s}$ for the tuple $(v_{i_1}, \dots, v_{i_s})$ for short. Let $a^1, \dots, a^\ell \in C^k$. Let $\varepsilon > 0$ be a rational number. From the definition of an infimum, for every $j \in \{1, \dots, \ell\}$, there is $b^j \in C^n$ such that

$$R(a^j) \leq \phi(a^j, b^j) < R(a^j) + \varepsilon.$$

Moreover, for every $f \in \mathscr{O}_C^{(\ell)}$,

$$R(f(a^1, \dots, a^\ell)) \leq \phi(f(a^1, \dots, a^\ell), f(b^1, \dots, b^\ell)).$$

By linearity and monotonicity of expectation, we obtain

$$\begin{aligned}E_\omega[f \mapsto R(f(a^1,\dots,a^\ell))] &\le E_\omega[f \mapsto \phi(f(a^1,\dots,a^\ell), f(b^1,\dots,b^\ell))] \\ &= E_\omega[f \mapsto \sum_{i=1}^m R_i((f(a^1,\dots,a^\ell))_{p^i_1,\dots,p^i_{k_i}}, (f(b^1,\dots,b^\ell))_{q^i_1,\dots,q^i_{n_i}})] \\ &= \sum_{i=1}^m E_\omega[f \mapsto R_i((f(a^1,\dots,a^\ell))_{p^i_1,\dots,p^i_{k_i}}, (f(b^1,\dots,b^\ell))_{q^i_1,\dots,q^i_{n_i}})].\end{aligned}$$

Since ω improves R_i for every $i \in \{1,\dots,m\}$, the last row of the inequality above is at most

$$\begin{aligned}\sum_{i=1}^m \frac{1}{\ell}\sum_{j=1}^\ell R_i(a^j_{p^i_1,\dots,p^i_{k_i}}, b^j_{q^i_1,\dots,q^i_{n_i}}) &= \frac{1}{\ell}\sum_{j=1}^\ell\sum_{i=1}^m R_i(a^j_{p^i_1,\dots,p^i_{k_i}}, b^j_{q^i_1,\dots,q^i_{n_i}}) \\ &= \frac{1}{\ell}\sum_{j=1}^\ell \phi(a^j,b^j) < \frac{1}{\ell}\sum_{j=1}^\ell R(a^j) + \varepsilon.\end{aligned}$$

Since ε was arbitrary, it follows that ω improves R.

Finally, we prove that $\mathrm{Imp}(\mathrm{fPol}(\Gamma))$ is closed under Feas and Opt. Let $R \in \tau$ be of arity k and define $S = \mathrm{Feas}(R)$ and $T = \mathrm{Opt}(R)$. We aim to show that $S, T \in \mathrm{Imp}(\mathrm{fPol}(\Gamma))$. Let $s^1,\dots,s^\ell \in C^k$. If $S(s^i) = \infty$ for some $i \in \{1,\dots,\ell\}$, then $\frac{1}{\ell}\sum_{j=1}^\ell S(s^j) = \infty$ and hence ω satisfies (6) (with R replaced by S) for the tuples $s^1,\dots,s^\ell$. So suppose that $S(s^i) = 0$ for all $i \in \{1,\dots,\ell\}$, i.e., $R(s^i)$ is finite for all i. Since ω improves R it holds that

$$E_\omega[f \mapsto R(f(s^1,\dots,s^\ell))] \le \frac{1}{\ell}\sum_{j=1}^\ell R(s^j) \tag{7}$$

and hence the expected value on the left-hand side is finite as well. By (21) in Appendix A,

$$E_\omega[f \mapsto R(f(s^1,\dots,s^\ell))] = \sum_{t \in C^k} R(t)\omega(\mathscr{S}_{s^1,\dots,s^\ell,t}), \tag{8}$$

which implies that $R(t)$ is finite and $S(t) = 0$ unless $\omega(\mathscr{S}_{s^1,\dots,s^\ell,t}) = 0$. Consequently (again by (21)),

$$E_\omega[f \mapsto S(f(s^1,\dots,s^\ell))] = \sum_{t \in C^k} S(t)\omega(\mathscr{S}_{s^1,\dots,s^\ell,t}) = 0 = \frac{1}{\ell}\sum_{j=1}^\ell S(s^j).$$

It follows that ω improves S.

Moving to the weighted relation T, we may again assume without loss of generality that $T(s^i) = 0$ for every $i \in \{1,\dots,\ell\}$ as we did for S. This means that $c := R(s^1) = \dots = R(s^\ell) \le R(b)$ for every $b \in C^k$. Therefore, the right-hand side in (7) is equal to c and by combining it with (8) we get

$$\sum_{t \in C^k} R(t)\omega(\mathscr{S}_{s^1,\dots,s^\ell,t}) \le c.$$

Together with the assumption that $R(t) \ge c$ for all $t \in C^k$ and ω being a probability distribution we obtain that $R(t) = c$ and $T(t) = 0$ unless $\omega(\mathscr{S}_{s^1,\dots,s^\ell,t}) = 0$, and hence

$$E_\omega[f \mapsto T(f(s^1,\dots,s^\ell))] = \sum_{t \in C^k} T(t)\omega(\mathscr{S}_{s^1,\dots,s^\ell,t}) = 0 = \frac{1}{\ell}\sum_{j=1}^\ell T(s^j).$$

This concludes the proof that ω improves T. $\square$

Example 6.9. *Let $<$ be the binary relation on $\{0,1\}$ and $\Gamma_<$ the valued structure from Example 2.2. By definition,* $\mathrm{Opt}(<) \in \langle \Gamma_< \rangle$. *Denote the minimum operation on $\{0,1\}$ by* min *and let ω be a binary fractional operation defined by $\omega(\min) = 1$. Note that $\omega \in \mathrm{fPol}(\{0,1\}; \mathrm{Opt}(<))$. However,*

$$< \left(\min \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) \right) = \,<\!(0,0) = 1,$$

while $(1/2) \cdot {<}(0,1) + (1/2) \cdot {<}(0,0) = 1/2$. This shows that ω does not improve $<$ and hence $< \notin \langle (\{0,1\}; \mathrm{Opt}(<)) \rangle$ by Lemma 6.8.

7. Polynomial-time Tractability via Canonical Fractional Polymorphisms

In this section we make use of a tractability result for finite-domain VCSPs of Kolmogorov, Krokhin, and Rolinek [32], which itself was building on earlier work of Kolmogorov, Thapper, and Živný [33, 42].

Definition 7.1. *An operation $f \colon C^\ell \to C$ for $\ell \geq 2$ is called* cyclic *if*

$$f(x_1, \ldots, x_\ell) = f(x_2, \ldots, x_\ell, x_1)$$

for all $x_1, \ldots, x_\ell \in C$. Let $\mathrm{Cyc}_C^{(\ell)} \subseteq \mathscr{O}_C^{(\ell)}$ be the set of all operations on C of arity ℓ that are cyclic.

If G is a permutation group on a set C, then $\overline{G}$ denotes the closure of G in the space of functions from $C \to C$ with respect to the topology of pointwise convergence. Note that $\overline{G}$ might contain some operations that are not surjective, but if $G = \mathrm{Aut}(\mathfrak{B})$ for some structure $\mathfrak{B}$, then all operations in $\overline{G}$ are still embeddings of $\mathfrak{B}$ into $\mathfrak{B}$ that preserve all first-order formulas.

Definition 7.2. *Let G be a permutation group on the set C. An operation $f \colon C^\ell \to C$ is called* pseudo cyclic with respect to G *if there are $e_1, e_2 \in \overline{G}$ such that for all $x_1, \ldots, x_\ell \in C$*

$$e_1(f(x_1, \ldots, x_\ell)) = e_2(f(x_2, \ldots, x_\ell, x_1)).$$

Let $\mathrm{PC}_G^{(\ell)} \subseteq \mathscr{O}_C^{(\ell)}$ be the set of all operations on C of arity ℓ that are pseudo cyclic with respect to G.

Note that $\mathrm{PC}_G^{(\ell)} \in B(\mathscr{O}_C^{(\ell)})$. Indeed, the complement can be written as a countable union of sets of the form $\mathscr{S}_{a^1,\ldots,a^\ell,b}$ where for all $f \in \mathscr{O}_C^{(\ell)}$ the tuples $f(a^1, \ldots, a^\ell)$ and $f(a^2, \ldots, a^\ell, a^1)$ lie in different orbits with respect to G.

Definition 7.3. *Let G be a permutation group with domain C. An operation $f \colon C^\ell \to C$ for $\ell \geq 2$ is called* canonical with respect to G *if for all $k \in \mathbb{N}$ and $a^1, \ldots, a^\ell \in C^k$ the orbit of the k-tuple $f(a^1, \ldots, a^\ell)$ only depends on the orbits of $a^1, \ldots, a^\ell$ with respect to G. Let $\mathrm{Can}_G^{(\ell)} \subseteq \mathscr{O}_C^{(\ell)}$ be the set of all operations on C of arity ℓ that are canonical with respect to G.*

Note that $\mathrm{Can}_G^{(\ell)} \in B(\mathscr{O}_C^{(\ell)})$, since the complement is a countable union of sets of the form $\mathscr{S}_{a^1,\ldots,a^\ell,b} \cap \mathscr{S}_{c^1,\ldots,c^\ell,d}$ where for all $i \in \{1, \ldots, \ell\}$ the tuples a^i and c^i lie in the same orbit with respect to G, but b and d do not.

Remark 7.4. *Note that if h is an operation over C of arity ℓ which is canonical with respect to G, then h induces for every $k \in \mathbb{N}$ an operation h^* of arity ℓ on the orbits of k-tuples of G. Note that if h is pseudo cyclic with respect to G, then h^* is cyclic.*

Definition 7.5. *A fractional operation ω is called* pseudo cyclic with respect to G *if for every $A \in B(\mathscr{O}_C^{(\ell)})$ we have $\omega(A) = \omega(A \cap \mathrm{PC}_G^{(\ell)})$. Canonicity with respect to G and cyclicity for fractional operations are defined analogously.*

We refer to Section 8.3 for examples of concrete fractional polymorphisms of valued structures Γ that are cyclic and canonical with respect to $\mathrm{Aut}(\Gamma)$. If the reference to a specific permutation group G is clear, then we omit for cyclicity and canonicity the specification 'with respect to G'.

We will prove below that canonical pseudo cyclic fractional polymorphisms imply polynomial-time tractability of the corresponding VCSP. We prove this result by reducing to tractable VCSPs over finite domains. Motivated by Theorem 3.4 and the infinite-domain tractability conjecture from [10], we state these results for valued structures related to finitely bounded homogeneous structures.

Definition 7.6 (Γ_m^*)**.** *Let Γ be a valued structure with signature τ such that $\mathrm{Aut}(\Gamma)$ contains the automorphism group of a homogeneous structure $\mathfrak{B}$ with a finite relational signature. Let m be at least as large as the maximal arity of the relations of $\mathfrak{B}$. Let Γ_m^* be the following valued structure.*

- *The domain of Γ_m^* is the set of orbits of m-tuples of $\mathrm{Aut}(\Gamma)$.*
- *For every $R \in \tau$ of arity $k \leq m$ the signature of Γ_m^* contains a unary relation symbol R^*, which denotes in Γ_m^* the unary weighted relation that returns on the orbit of an m-tuple $t = (t_1, \dots, t_m)$ the value of $R^\Gamma(t_1, \dots, t_k)$ (this is well-defined, because the value is the same for all representatives t of the orbit).*
- *For every $p \in \{1, \dots, m\}$ and $i, j \colon \{1, \dots, p\} \to \{1, \dots, m\}$ there exists a binary relation $C_{i,j}$ which returns 0 for two orbits of m-tuples O_1 and O_2 if for every $s \in O_1$ and $t \in O_2$ we have that $(s_{i(1)}, \dots, s_{i(p)})$ and $(t_{j(1)}, \dots, t_{j(p)})$ lie in the same orbit of p-tuples of $\mathrm{Aut}(\Gamma)$, and returns ∞ otherwise.*

Note that $\mathrm{Aut}(\mathfrak{B})$ and hence $\mathrm{Aut}(\Gamma)$ has finitely many orbits of k-tuples for every $k \in \mathbb{N}$ and therefore Γ_m^* has a finite domain. The following generalises a known reduction for CSPs from [8].

Theorem 7.7. *Let Γ be a valued structure such that $\mathrm{Aut}(\Gamma)$ equals the automorphism group of a finitely bounded homogeneous structure $\mathfrak{B}$. Let r be the maximal arity of the relations of $\mathfrak{B}$ and the weighted relations in Γ, let v be the maximal number of variables that appear in a single conjunct of the universal sentence ψ that describes the age of $\mathfrak{B}$, and let $m \geq \max(r+1, v, 3)$. Then there is a polynomial-time reduction from $\mathrm{VCSP}(\Gamma)$ to $\mathrm{VCSP}(\Gamma_m^*)$.*

Proof. Let τ be the signature of Γ and τ^* be the signature of Γ_m^*. Let ϕ be an instance of $\mathrm{VCSP}(\Gamma)$ with threshold u and let V be the variables of ϕ. Create a variable $y(\bar{x})$ for every $\bar{x} = (x_1, \dots, x_m) \in V^m$. For every summand $R(x_1, \dots, x_k)$ of ϕ and we create a summand $R^*(y(x_1, \dots, x_k, \dots, x_k))$; this makes sense since $m \geq r$. For every $\bar{x}, \bar{x}' \in V^m$, $p \in \{1, \dots, m\}$, and $i, j \colon \{1, \dots, p\} \to \{1, \dots, m\}$, add the summand $C_{i,j}(y(\bar{x}), y(\bar{x}'))$ if $(x_{i(1)}, \dots, x_{i(p)}) = (x'_{j(1)}, \dots, x'_{j(p)})$; we will refer to these as *compatibility constraints*. Let ϕ^* be the resulting τ^*-expression. Clearly, ϕ^* can be computed from ϕ in polynomial time.

Suppose first that (ϕ, u) has a solution; it will be notationally convenient to view the solution as a function f from the variables of ϕ to the elements of Γ (rather than a tuple). We claim that the map f^* which maps $y(\bar{x})$ to the orbit of $f(\bar{x})$ in $\mathrm{Aut}(\Gamma)$ is a solution for (ϕ^*, u). And indeed, each of the summands involving a symbol $C_{i,j}$ evaluates to 0, and $(\phi^*)^{\Gamma_m^*}$ equals ϕ^Γ.

Now suppose that (ϕ^*, u) has a solution f^*. To construct a solution f to (ϕ, u), we first define an equivalence relation $\sim$ on V. For $x_1, x_2 \in V$, define $x_1 \sim x_2$ if a (equivalently: every) tuple t in $f^*(y(x_1, x_2, \dots, x_2))$ satisfies $t_1 = t_2$. Clearly, $\sim$ is reflexive and symmetric. To verify that $\sim$ is transitive, suppose that $x_1 \sim x_2$ and $x_2 \sim x_3$. In the following we use that $m \geq 3$. Let i be the identity map on $\{1,2\}$, let $j\colon \{1,2\} \to \{2,3\}$ be given by $x \mapsto x+1$, and let $j'\colon \{1,2\} \to \{1,3\}$ be given by $j'(1) = 1$ and $j'(2) = 3$. Then ϕ^* contains the conjuncts

$$C_{i,i}(y(x_1, x_2, x_2, \dots, x_2), y(x_1, x_2, x_3, \dots, x_3)),$$
$$C_{i,j}(y(x_2, x_3, x_3, \dots, x_3), y(x_1, x_2, x_3, \dots, x_3)),$$
$$C_{i,j'}(y(x_1, x_3, x_3, \dots, x_3), y(x_1, x_2, x_3, \dots, x_3)).$$

Let t be a tuple from $f^*(y(x_1, x_2, x_3, \dots, x_3))$. Then it follows from the conjuncts with the relation symbols $C_{i,i}$ and $C_{i,j}$ that $t_1 = t_2$ and $t_2 = t_3$, and therefore $t_1 = t_3$. Thus we obtain from the conjunct with $C_{i,j'}$ that $x_1 \sim x_3$.

Claim 0. For all equivalence classes $[x_1]_\sim, \dots, [x_m]_\sim$, $t \in f^*(y(x_1, \dots, x_m))$, $S \in \sigma$ of arity k, and $j\colon \{1, \dots, k\} \to \{1, \dots, m\}$, whether $\mathfrak{B} \models S(t_{j(1)}, \dots, t_{j(k)})$ does not depend on the choice of the representatives $x_1, \dots, x_m$. It suffices to show this statement if we choose another representative x_i' for $[x_i]_\sim$ for some $i \in \{1, \dots, m\}$, because the general case then follows by induction.

Suppose that for every $t \in f^*(y(x_1, \dots, x_m))$ we have $\mathfrak{B} \models S(t_{j(1)}, \dots, t_{j(k)})$; we have to show that for every $t' \in f^*(y(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_m))$ we have $\mathfrak{B} \models S(t'_{j(1)}, \dots, t'_{j(k)})$. If $i \notin \{j(1), \dots, j(k)\}$, then ϕ^* contains

$$C_{j,j}(y(x_1, \dots, x_m), y(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_m))$$

and hence $\mathfrak{B} \models S(t'_{j(1)}, \dots, t'_{j(k)})$. Now suppose that $i \in \{j(1), \dots, j(k)\}$; for the sake of notation we suppose that $i = j(1)$. By the definition of $\sim$, and since $x_{j(1)} \sim x'_{j(1)}$, every tuple $t'' \in f^*(y(x_{j(1)}, x'_{j(1)}, \dots, x'_{j(1)}))$ satisfies $t''_1 = t''_2$. Let $\tilde{t}$ be a tuple from

$$f^*(y(x_{j(1)}, \dots, x_{j(k)}, x'_{j(1)}, \dots, x'_{j(1)})).$$

(Here we use that $m \geq r+1$.)

- $\mathfrak{B} \models S(\tilde{t}_1, \dots, \tilde{t}_k)$ because of a compatibility constraint between $y(x_1, \dots, x_m)$ and $y(x_{j(1)}, \dots, x_{j(k)}, x'_{j(1)}, \dots, x'_{j(1)})$ in ϕ^*;
- $\tilde{t}_1 = \tilde{t}_{k+1}$ because of a compatibility constraint between $y(x_{j(1)}, \dots, x_{j(k)}, x'_{j(1)}, \dots, x'_{j(1)})$ and $y(x_{j(1)}, x'_{j(1)}, \dots, x'_{j(1)})$ and $x_{j(1)} \sim x'_{j(1)}$ in ϕ^*;
- hence, $\mathfrak{B} \models S(\tilde{t}_{k+1}, \tilde{t}_2, \dots, \tilde{t}_k)$;
- $\mathfrak{B} \models S(t'_{j(1)}, t'_{j(2)}, \dots, t'_{j(k)})$ because of a compatibility constraint between the variables $y(x_{j(1)}, \dots, x_{j(k)},\ x'_{j(1)}, \dots, x'_{j(1)})$ and $y(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_m)$ in ϕ^*: namely, if $j'\colon \{1, \dots, k\} \to \{1, \dots, m\}$ is the map that coincides with the identity map except that $j'(1) := k+1$, then ϕ^* contains
$$C_{j,j'}\big(y(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_m), y(x_{j(1)}, \dots, x_{j(k)}, x'_{j(1)}, \dots, x'_{j(1)})\big).$$

This concludes the proof of Claim 0.

Now we can define a structure $\mathfrak{C}$ in the signature σ of $\mathfrak{B}$ on the equivalence classes of $\sim$. If $S \in \sigma$ has arity k, $j_1, \dots, j_k \in \{1, \dots, m\}$, and $[x_1]_\sim, \dots, [x_m]_\sim$ are equivalence classes of $\sim$ such that the tuples t in $f^*(y(x_1, \dots, x_m))$ satisfy $S^{\mathfrak{B}}(t_{j_1}, \dots, t_{j_k})$ for some representatives $x_1, \dots x_m$ (equivalently, for all representatives, by Claim 0), then add $([x_{j_1}]_\sim, \dots, [x_{j_k}]_\sim)$ to $S^{\mathfrak{C}}$. No other tuples are contained in the relations of $\mathfrak{C}$.

Claim 1. If $[x_1]_\sim, \dots, [x_m]_\sim$ are equivalence classes of $\sim$, and $t \in f^*(y(x_1, \dots, x_m))$, then $[x_i]_\sim \mapsto t_i$, for $i \in \{1, \dots, m\}$, is an isomorphism between a substructure of $\mathfrak{C}$ and a substructure of $\mathfrak{B}$ for any choice of representatives $x_1, \dots, x_m$. First note that $[x_i]_\sim = [x_j]_\sim$ if and only if $t_i = t_j$, so the map is well-defined and bijective. Let $S \in \sigma$ be of arity k and $j \colon \{1, \dots, k\} \to \{1, \dots, m\}$. If $\mathfrak{B} \models S(t_{j(1)}, \dots, t_{j(k)})$, then $\mathfrak{C} \models S([x_{j(1)}]_\sim, \dots, [x_{j(k)}]_\sim)$ by the definition of $\mathfrak{C}$. Conversely, suppose that $\mathfrak{C} \models S([x_{j(1)}]_\sim, \dots [x_{j(k)}]_\sim)$. By Claim 0 and the definition of $\mathfrak{C}$, there is $t' \in f^*(y(x_1, \dots, x_m))$ such that $\mathfrak{B} \models S(t'_{j(1)}, \dots, t'_{j(k)})$. Since $f^*(y(x_1, \dots, x_m))$ is an orbit of $\mathrm{Aut}(\mathfrak{B})$, we have $\mathfrak{B} \models S(t_{j(1)}, \dots, t_{j(k)})$ as well.

Claim 2. $\mathfrak{C}$ embeds into $\mathfrak{B}$. It suffices to verify that $\mathfrak{C}$ satisfies each conjunct of the universal sentence ψ. Let $\psi'(x_1, \dots, x_q)$ be such a conjunct, and let $[c_1]_\sim, \dots, [c_q]_\sim$ be elements of $\mathfrak{C}$. Let t be a tuple from the orbit $f^*(y(c_1, \dots, c_q, \dots, c_q))$ of $\mathrm{Aut}(\Gamma)$; this makes sense since $m \geq v$. Since $t_1, \dots, t_q$ are elements of $\mathfrak{B}$, the tuple $(t_1, \dots, t_q)$ satisfies ψ'. Claim 1 then implies that $([c_1]_\sim, \dots, [c_q]_\sim)$ satisfies ψ'.

Let e be an embedding of $\mathfrak{C}$ to $\mathfrak{B}$. For every $x \in V$, define $f(x) = e([x]_\sim)$. Note that for every summand $R(x_1, \dots, x_k)$ in ϕ and $t \in f^*(y(x_1, \dots, x_k, \dots, x_k))$, we have

$$R^*(f^*(y(x_1, \dots, x_k, \dots, x_k))) = R(t_1, \dots, t_k) = R(e([x_1]_\sim), \dots, e([x_k]_\sim)) = R(f(x_1), \dots, f(x_k)),$$

where the middle equality follows from $t_i \mapsto e([x_i]_\sim)$ being a partial isomorphism of $\mathfrak{B}$ by Claim 1 and 2, which by the homogeneity of $\mathfrak{B}$ extends to an automorphism of $\mathfrak{B}$ and therefore also an automorphism of Γ. Since f^* is a solution to (ϕ^*, u), it follows from the construction of ϕ^* that f is a solution to (ϕ, u). $\square$

Let G be a permutation group that contains the automorphism group of a homogeneous structure with a finite relational signature of maximal arity k. A fractional operation ω over the domain C of Γ of arity ℓ which is canonical with respect to G induces a fractional operation ω^* on the orbits of k-tuples of G, given by

$$\omega^*(A) := \omega\big(\{f \in \mathrm{Can}_G^{(\ell)} \mid f^* \in A\}\big),$$

for every subset A of the set of operations of arity ℓ on the finite domain of Γ_k^* (all such subsets are Borel). Note that $\{f \in \mathrm{Can}_G^{(\ell)} \mid f^* \in A\}$ is a measurable subset of $\mathscr{O}_C^{(\ell)}$. Also note that if ω is pseudo cyclic, then ω^* is cyclic. Statements about the fractional polymorphisms of Γ_m^* lift back to statements about the fractional polymorphisms of Γ via the following useful lemma.

Lemma 7.8. *Let Γ be a valued structure such that $\mathrm{Aut}(\Gamma)$ equals the automorphism group G of a finitely bounded homogeneous structure and let m be as in Theorem 7.7. Let $\nu \in \mathrm{fPol}(\Gamma_m^*)$ be cyclic. Then there exists $\omega \in \mathrm{fPol}(\Gamma)$ which is canonical with respect to G such that $\omega^* = \nu$.*

Proof. Let C be the domain of Γ, let D be the domain of Γ_m^*, and let ℓ be the arity of ν. Suppose that $\nu(f) > 0$ for some operation f. Then there exists a function $g \colon C^\ell \to C$ which is canonical with respect to G such that $g^* = f$ by Lemma 4.9 in [8] (also see Lemma 10.5.12 in [5]). For every such f, choose g such that $g^* = f$ and define $\omega(g) := \nu(f)$ and $\omega(h) := 0$ for all other $h \in \mathscr{O}_C^{(\ell)}$. Since the domain of Γ_m^* is finite, this correctly defines a fractional operation ω of the same arity ℓ

as ν. It also improves every weighted relation R of Γ: if R has arity k, and $a^1,\dots,a^\ell \in C^k$, then

$$\begin{aligned}
E_\omega[g \mapsto R(g(a^1,\dots,a^\ell))] &= \sum_{g \in \mathscr{O}_C^{(\ell)}} \omega(g) R(g(a^1,\dots,a^\ell)) \\
&= \sum_{f \in \mathscr{O}_D^{(\ell)}} \nu(f) R^*(f(a^1,\dots,a^\ell)_1,\dots,f(a^1,\dots,a^\ell)_k,\dots,f(a^1,\dots,a^\ell)_k) \\
&\leq \frac{1}{\ell}\sum_{j=1}^{\ell} R^*(a_1^j,\dots,a_k^j,\dots,a_k^j) \\
&= \frac{1}{\ell}\sum_{j=1}^{\ell} R(a_1^j,\dots,a_k^j). \qquad \square
\end{aligned}$$

Lemma 7.9. *Let G be the automorphism group of a homogeneous structure $\mathfrak{B}$ with a relational signature of maximal arity m. If $\omega \in \mathscr{F}_C^{(\ell)}$ is canonical with respect to G such that ω^* (defined on the orbits of m-tuples of G) is cyclic, then ω is pseudo cyclic with respect to G.*

Proof. We use the fact that if f is canonical with respect to G such that f^* (defined on the orbits of m-tuples) is cyclic, then f is pseudo cyclic (see the proof of Proposition 6.6 in [10]; also see Lemma 10.1.5 in [5]). Let C be the domain of Γ and let $a^1,\dots,a^\ell, b \in C^m$. It suffices to show that $\omega(S_{a^1,\dots,a^\ell,b} \cap \mathrm{PC}_G^{(\ell)}) = \omega(S_{a^1,\dots,a^\ell,b})$. Indeed,

$$\begin{aligned}
\omega(S_{a^1,\dots,a^\ell,b}) &= \omega(S_{a^1,\dots,a^\ell,b} \cap \mathrm{Can}_G^{(\ell)}) && \text{(canonicity of } \omega) \\
&= \omega^*\big(\{f^* \mid f \in S_{a^1,\dots,a^\ell,b} \cap \mathrm{Can}_G^{(\ell)}\}\big) && \text{(definition of } \omega^*) \\
&= \omega^*\big(\{f^* \mid f \in S_{a^1,\dots,a^\ell,b} \cap \mathrm{Can}_G^{(\ell)}\} \cap \mathrm{Cyc}_C^{(\ell)}\,\big) && \text{(by assumption)} \\
&= \omega^*\big(\{f^* \mid f \in S_{a^1,\dots,a^\ell,b} \cap \mathrm{Can}_G^{(\ell)} \cap \mathrm{PC}_G^{(\ell)}\}\big) && \text{(fact mentioned above and Remark 7.4)} \\
&= \omega(S_{a^1,\dots,a^\ell,b} \cap \mathrm{Can}_G^{(\ell)} \cap \mathrm{PC}_G^{(\ell)}) \\
&= \omega(S_{a^1,\dots,a^\ell,b} \cap \mathrm{PC}_G^{(\ell)}). \qquad \square
\end{aligned}$$

7.1. **Fractional Polymorphims on Finite Domains.** For studying canonical operations, we can use known results about operations on finite domains.

Definition 7.10. *Let ω be a fractional operation of arity ℓ on a finite domain C. Then the* support *of ω is the set*

$$\mathrm{Supp}(\omega) = \{f \in \mathscr{O}_C^{(\ell)} \mid \omega(f) > 0\}.$$

If $\mathscr{F}$ is a set of fractional operations, then

$$\mathrm{Supp}(\mathscr{F}) := \bigcup_{\omega \in \mathscr{F}} \mathrm{Supp}(\omega).$$

Note that, a fractional operation ω on a finite domain is determined by the values $\omega(f)$, $f \in \mathrm{Supp}(\omega)$, in contrast to fractional operations on infinite domains. Moreover, a fractional polymorphism ω of a valued structure with a finite domain is cyclic if and only if all operations in its support are cyclic, in accordance to the definitions from [34]. An operation $f\colon C^4 \to C$ is called *Siggers* if $f(a,r,e,a) = f(r,a,r,e)$ for all $a,r,e \in C$.

Lemma 7.11. *Let Γ and Δ be valued structures with finite domains that are fractionally homomorphically equivalent.*

- *If Γ has a cyclic fractional polymorphism, then Δ has a cyclic fractional polymorphism of the same arity.*
- *If $\mathrm{Supp}(\mathrm{fPol}(\Gamma))$ contains a cyclic operation, then $\mathrm{Supp}(\mathrm{fPol}(\Delta))$ contains a cyclic operation of the same arity.*

Proof. Let C be the domain of Γ and let D be the domain of Δ. Let ν_1 be a fractional homomorphism from Γ to Δ, and let ν_2 be a fractional homomorphism from Δ to Γ. Define ν_2' as the fractional homomorphism from Δ^ℓ to Γ^ℓ as follows. If $f\colon D \to C$, then f' denotes the map from D^ℓ to C^ℓ given by $(c_1, \dots, c_\ell) \mapsto (f(c_1), \dots, f(c_\ell))$. Define $\nu_2'(f') := \nu_2(f)$ and $\nu_2'(h) = 0$ for all other $h\colon D^\ell \to C^\ell$; since C and D are finite, this defines a fractional operation.

Suppose that ω is a fractional polymorphism of Γ of arity ℓ. Then $\omega' := \nu_1 \circ \omega \circ \nu_2'$ is a fractional homomorphism from Δ^ℓ to Δ (see Lemma 5.6), and hence a fractional polymorphism of Δ (see Remark 6.4). Note that if ω is cyclic, then ω' is cyclic; this shows that first statement of the lemma.

Next, suppose that there exists $\omega \in \mathrm{fPol}^{(\ell)}(\Gamma)$ such that $\mathrm{Supp}(\omega)$ contains a cyclic operation g of arity ℓ. Since the domain C of Γ is finite, there exists a function $f_1\colon C \to D$ such that $\nu_1(f_1) > 0$ and a function $f_2\colon D \to C$ such that $\nu_2(f_2) > 0$. Note that $f_1 \circ g \circ f_2'\colon D^\ell \to D$ is cyclic since g is cyclic, and that $\omega'(f_1 \circ g \circ f_2') > 0$. $\square$

The following definition is taken from [34].

Definition 7.12 (core)**.** *A valued structure Γ over a finite domain is called a* core *if all operations in $\mathrm{Supp}(\mathrm{fPol}(\Gamma))^{(1)}$ are injective.*

We have been unable to find an explicit reference for the following proposition, but it should be considered to be known; we also present a proof as a guide to the literature.

Proposition 7.13. *Let Γ be a valued structure with a finite domain. Then there exists a core valued structure Δ over a finite domain which is fractionally homomorphically equivalent to Γ.*

Proof. Let C be the domain of Γ. If Γ itself is a core then there is nothing to be shown, so we may assume that there exists a non-injective $f \in \mathrm{Supp}(\mathrm{fPol}^{(1)}(\Gamma))$. Since C is finite, we have that $D := f(C) \neq C$; let Δ be the valued structure with domain D and the same signature as Γ whose weighted relations are obtained from the corresponding weighted relations of Γ by restriction to D. It then follows from Lemma 15 in [34] in combination with Remark 5.8 that Γ and Δ are fractionally homomorphically equivalent. After applying this process finitely many times, we obtain a core valued structure that is fractionally homomorphically equivalent to Γ. $\square$

The following lemma is a variation of Proposition 39 from [34], which is phrased there only for valued structures Γ that are cores and for idempotent cyclic operations.

Lemma 7.14. *Let Γ be a valued structure over a finite domain. Then Γ has a cyclic fractional polymorphism if and only if $\mathrm{Supp}(\mathrm{fPol}(\Gamma))$ contains a cyclic operation.*

Proof. The forward implication is trivial. We prove the reverse implication. Let Δ be a core valued structure over a finite domain that is homomorphically equivalent to Γ, which exists by Proposition 7.13. By Lemma 7.11, $\mathrm{Supp}(\mathrm{fPol}(\Delta))$ contains a cyclic operation. Then $\mathrm{Supp}(\mathrm{fPol}(\Delta))$ contains even an idempotent cyclic operation: If $c \in \mathrm{Supp}(\mathrm{fPol}(\Delta))$ is cyclic, then the operation $c_0\colon x \mapsto c(x, \dots, x)$ is in $\mathrm{Supp}(\mathrm{fPol}(\Delta))$ as well. Since Δ is a finite core, c_0 is bijective and therefore c_0^{-1} (which is just a finite power of c_0) and the idempotent cyclic operation $c_0^{-1} \circ c$

lie in $\mathrm{Supp}(\mathrm{fPol}(\Delta))$. By Proposition 39 in [34], Δ has a cyclic fractional polymorphism and by Lemma 7.11, Γ also has one. □

The following outstanding result classifies the computational complexity of VCSPs for valued structures over finite domains; it does not appear in this form in the literature, but we explain how to derive it from results in [11, 32, 34, 45, 46]. In the proof, if $\mathfrak{C}$ is a finite relational structure (understood also as a valued structure), we denote by $\mathrm{Pol}(\mathfrak{C})$ the set $\mathrm{Supp}(\mathrm{fPol}(\mathfrak{C}))$; this notation is consistent with the literature since the set $\mathrm{Supp}(\mathrm{fPol}(\mathfrak{C}))$ concides with the set of polymorphisms of a relational structure.

Theorem 7.15. *Let Γ be a valued structure with a finite signature and a finite domain. If $(\{0,1\};\mathrm{OIT})$ does not have a pp-construction in Γ, then Γ has a fractional cyclic polymorphism, and $\mathrm{VCSP}(\Gamma)$ is in P, and it is NP-hard otherwise.*

Proof. If $(\{0,1\};\mathrm{OIT})$ has a pp-construction in Γ, then the NP-hardness of $\mathrm{VCSP}(\Gamma)$ follows from Corollary 5.12. So assume that $(\{0,1\};\mathrm{OIT})$ does not have a pp-construction in Γ.

Let $\mathfrak{C}$ be a classical relational structure on the same domain as Γ such that $\mathrm{Pol}(\mathfrak{C}) = \mathrm{Supp}(\mathrm{fPol}(\Gamma))$; it exists since $\mathrm{Supp}(\mathrm{fPol}(\Gamma))$ contains projections by Remark 6.5 and is closed under composition by Lemma 5.6 and Remark 6.4. Note that therefore $\mathrm{fPol}(\Gamma) \subseteq \mathrm{fPol}(\mathfrak{C})$ and since Γ has a finite domain, [22, Theorem 3.3] implies that every relation of $\mathfrak{C}$ lies in $\langle\Gamma\rangle$. Since Γ does not pp-construct $(\{0,1\};\mathrm{OIT})$, neither does $\mathfrak{C}$, and in particular, $\mathfrak{C}$ does not pp-construct $(\{0,1\};\mathrm{OIT})$ in the classical relational setting (see [3, Definition 3.4, Corollary 3.10]). Combining Theorems 1.4 and 1.8 from [3], $\mathrm{Pol}(\mathfrak{C})$ contains a cyclic operation.

Since $\mathrm{Supp}(\mathrm{fPol}(\Gamma))$ contains a cyclic operation, by Lemma 7.14, Γ has a cyclic fractional polymorphism. Then Kolmogorov, Rolínek, and Krokhin [32] prove that in this case $\mathrm{CSP}(\Gamma)$ can be reduced to a finite-domain CSP with a cyclic polymorphism; such CSPs were shown to be in P by Bulatov [11] and, independently, by Zhuk [46]. □

The problem of deciding for a given valued structure Γ with finite domain and finite signature whether Γ satisfies the condition given in the previous theorem can be solved in exponential time [31]. We now state consequences of this result for certain valued structures with an infinite domain.

Proposition 7.16. *Let $\mathfrak{B}$ be a finitely bounded homogeneous structure and let Γ be a valued structure with finite relational signature such that $\mathrm{Aut}(\Gamma) = \mathrm{Aut}(\mathfrak{B})$. Let m be as in Theorem 7.7. Then the following are equivalent.*

(1) $\mathrm{fPol}(\Gamma)$ *contains a fractional operation which is canonical and pseudo cyclic with respect to* $\mathrm{Aut}(\mathfrak{B})$*;*
(2) $\mathrm{fPol}(\Gamma^*_m)$ *contains a cyclic fractional operation;*
(3) $\mathrm{Supp}(\mathrm{fPol}(\Gamma^*_m))$ *contains a cyclic operation.*
(4) $\mathrm{Supp}(\mathrm{fPol}(\Gamma^*_m))$ *contains a Siggers operation.*

Proof. We first prove the implication from (1) to (2). If ω is a fractional polymorphism of Γ, then ω^* is a fractional polymorphism of Γ^*_m: the fractional operation ω^* improves R^* because ω improves R, and ω^* improves $C_{i,j}$ for all i, j because ω is canonical with respect to G. Finally, if ω is pseudo cyclic with respect to G, then ω^* is cyclic.

The implication from (2) to (1) is a consequence of Lemma 7.8 and Lemma 7.9. The equivalence of (2) and (3) follows from Lemma 7.14. The equivalence of (3) and (4) is proved in [5, Theorem 6.9.2]; the proof is based on [2, Theorem 4.1]. □

Note that item (4) in the previous proposition can be decided algorithmically for a given valued structure Γ_m^* (which has a finite domain and finite signature).

Theorem 7.17. *If the conditions from Proposition 7.16 hold, then* $\mathrm{VCSP}(\Gamma)$ *is in P.*

Proof. If Γ_m^* has a cyclic fractional polymorphism of arity $\ell \geq 2$, then the polynomial-time tractability of $\mathrm{VCSP}(\Gamma_m^*)$ follows from Theorem 7.15. For m large enough, we may apply Theorem 7.7 and obtain a polynomial-time reduction from $\mathrm{VCSP}(\Gamma)$ to $\mathrm{VCSP}(\Gamma_m^*)$, which concludes the proof. $\square$

8. Application: Resilience

A research topic that has been studied in database theory is the computational complexity of the so-called *resilience problem* [20, 21, 37]. We formulate it here for the case of conjunctive queries and, more generally, for unions of conjunctive queries. We generally work with Boolean queries, i.e., queries without free variables. Our results, however, can be extended also to the non-Boolean case. A *conjunctive query* is a primitive positive τ-sentence and a *union of conjunctive queries* is a (finite) disjunction of conjunctive queries. Note that every existential positive sentence can be written as a union of conjunctive queries.

Let τ be a finite relational signature and μ a conjunctive query over τ. The input to the *resilience problem for* μ consists of a finite τ-structure $\mathfrak{A}$, called a *database*[1], and the task is to compute the number of tuples that have to be removed from relations of $\mathfrak{A}$ so that $\mathfrak{A}$ does *not* satisfy μ. We call this number the *resilience* of $\mathfrak{A}$ (with respect to μ). As usual, this can be turned into a decision problem where the input also contains a natural number $u \in \mathbb{N}$ and the question is whether the resilience is at most u. Clearly, $\mathfrak{A}$ does not satisfy μ if and only if its resilience equals 0. The computational complexity of this problem depends on μ and various cases that can be solved in polynomial time and that are NP-hard have been described in [20, 21, 37]. A general classification, however, is open.

A natural variation of the problem is that the input database is a *bag database*, meaning that it may contain tuples with *multiplicities*, i.e., the same tuple may have multiple occurrences in the same relation. Formally, a bag database is a valued structure with all weights (which represent multiplicities) taken from $\mathbb{N}$. Resilience on bag databases was introduced by Makhija and Gatterbauer [37] who also present a conjunctive query for which the resilience problem with multiplicities is NP-hard whereas the resilience problem without multiplicities is in P. Note that bag databases are of importance because they represent SQL databases more faithfully than set databases [14]. Bag databases often require different methods than set databases [14, 28]. In this paper, we exclusively consider bag databases. Note that if the resilience problem of a query μ can be solved in polynomial time on bag databases, then also the resilience problem on set databases can be solved in polynomial time.

A natural generalization of the basic resilience problem defined above is obtained by admitting the decoration of databases with a subsignature $\sigma \subseteq \tau$, in this way declaring all tuples in $R^{\mathfrak{A}}$, $R \in \sigma$, to be *exogenous.* This means that we are not allowed to remove such tuples from $\mathfrak{A}$ to make μ false; the tuples in the other relations are then called *endogenous.* For brevity, we also refer to the relations in τ as being exogenous/endogenous. If not specified, then $\sigma = \emptyset$, i.e., all tuples are endogenous. Different variants of exogenous tuples were studied [37]. However, in bag semantics all

[1]To be precise, a finite relational structure is not exactly the same as a database because the latter may not contain elements that do not contain in any relation. This difference, however, is inessential for the problems studied in this paper.

Fixed: a relational signature τ, a subset $\sigma \subseteq \tau$, and a union μ of conjuctive queries over τ.
Input: A bag database $\mathfrak{A}$ in signature τ and $u \in \mathbb{N}$.
$m :=$ minimal number of tuples to be removed from the relations in $\{R^{\mathfrak{A}} \mid R \in \tau \setminus \sigma\}$ so that $\mathfrak{A} \not\models \mu$.
Output: Is $m \leq u$?

FIGURE 1. The resilience problem considered in this paper.

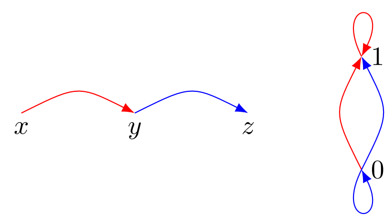

FIGURE 2. The query μ from Example 8.1 (on the left) and the corresponding structure $\mathfrak{B}$ (on the right).

of them are polynomial-time equivalent to a problem of this form, see Remark 8.16. In this paper, we generally admit exogeneous relations. The resilience problem that we study is given in Figure 1.

We next explain how to represent resilience problems as VCSPs using appropriately chosen valued structures with oligomorphic automorphism groups.

Example 8.1. *The following query is taken from Meliou, Gatterbauer, Moore, and Suciu [38]; they show how to solve its resilience problem without multiplicities in polynomial time by a reduction to a max-flow problem. Let μ be the query*

$$\exists x, y, z\big(R(x, y) \wedge S(y, z)\big).$$

Observe that a finite τ-structure satisfies μ if and only if it does not have a homomorphism to the τ-structure $\mathfrak{B}$ with domain $B = \{0, 1\}$ and the relations $R^{\mathfrak{B}} = \{(0, 1), (1, 1)\}$ and $S^{\mathfrak{B}} = \{(0, 0), (0, 1)\}$ (see Figure 2). We turn $\mathfrak{B}$ into the valued structure Γ with domain $\{0, 1\}$ where $R^{\Gamma}(0, 1) = R^{\Gamma}(1, 1) = 0 = S^{\Gamma}(0, 0) = S^{\Gamma}(0, 1)$ and R^{Γ} and S^{Γ} take value 1 otherwise. Then $\mathrm{VCSP}(\Gamma)$ *is precisely the resilience problem for μ (with multiplicities). Our results reprove the result from [37] that even with multiplicities, the problem can be solved in polynomial time (see Theorem 7.17, Proposition 8.15 and Example 8.12).*

Example 8.2. *Let μ be the conjunctive query*

$$\exists x, y, z(R(x, y) \wedge S(x, y, z)).$$

This query is linear *in the sense of Freire, Gatterbauer, Immerman, and Meliou and thus its resilience problem without multiplicities can be solved in polynomial time (Theorem 4.5 in [38]; also see Fact 3.18 in [19]). Our results reprove the result from [37] that this problem remains polynomial-time solvable with multiplicities (see Theorem 7.17, Proposition 8.15 and Example 8.17).*

Remark 8.3. *Note that if the resilience problem (with or without multiplicities) for a union μ of conjunctive queries is in P, then also the computational problem of* finding *tuples to be removed from the input database $\mathfrak{A}$ so that $\mathfrak{A} \not\models \mu$ is in P. To see this, let $u \in \mathbb{N}$ be threshold. If $u = 0$, then no tuple needs to be found and we are done. Otherwise, for every tuple t in a relation $R^{\mathfrak{A}}$, we remove t from $R^{\mathfrak{A}}$ and test the resulting database with the threshold $u - m$, where m is the multiplicity of t. If the modified instance is accepted, then t is a correct tuple to be removed and we may proceed to find a solution of this modified instance.*

8.1. **Connectivity.** We show that when classifying the resilience problem for conjunctive queries, it suffices to consider queries that are connected.

The *canonical database* of a conjunctive query μ with relational signature τ is the τ-structure $\mathfrak{A}$ whose domain are the variables of μ and where $a \in R^{\mathfrak{A}}$ for $R \in \tau$ if and only if μ contains the conjunct $R(a)$. A τ-structure is *connected* if it cannot be written as the disjoint union of two τ-structures with non-empty domains. Conversely, the *canonical query* of a relational τ-structure $\mathfrak{A}$ is the conjunctive query whose variable set is the domain A of $\mathfrak{A}$, and which contains for every $R \in \tau$ and $\bar{a} \in R^{\mathfrak{A}}$ the conjunct $R(\bar{a})$.

Remark 8.4. *All terminology introduced for τ-structures also applies to conjunctive queries with signature τ: by definition, the query has the property if the canonical database has the property.*

In particular, it is clear what it means for a conjunctive query to be connected.

Lemma 8.5. *Let $\mu_1, \dots, \mu_k$ be conjunctive queries such that μ_i does not imply μ_j if $i \neq j$. Then the resilience problem for $\mu := \mu_1 \wedge \cdots \wedge \mu_k$ is NP-hard if the resilience problem for one of the μ_i is NP-hard. Conversely, if the resilience problem is in P (in NP) for each μ_i, then the resilience problem for μ is in P as well (in NP, respectively). The same is true in the setting without multiplicities and/or exogeneous relations.*

Proof. We first present a polynomial-time reduction from the resilience problem of μ_i, for some $i \in \{1, \dots, k\}$, to the resilience problem of μ. Given an instance $\mathfrak{A}$ of the resilience problem for μ_i, let m be the number of tuples in relations of $\mathfrak{A}$. Let $\mathfrak{A}'$ be the disjoint union of $\mathfrak{A}$ with m copies of the canonical database of μ_j for every $j \in \{1, \dots, k\} \setminus \{i\}$. Observe that $\mathfrak{A}'$ can be computed in polynomial time in the size of $\mathfrak{A}$ and that the resilience of $\mathfrak{A}$ with respect to μ_i equals the resilience of $\mathfrak{A}'$ with respect to μ.

Conversely, if the resilience problem is in P for each μ_i, then also the resilience problem for μ is in P: given an instance $\mathfrak{A}$ of the resilience problem for μ, we compute the resilience of $\mathfrak{A}_j$ with respect to μ_i for every $i \in \{1, \dots, k\}$, and the minimum of all the resulting values. The proof for the membership in NP is the same.

The same proof works in the setting without multiplicities. □

When classifying the complexity of the resilience problem for conjunctive queries, by Lemma 8.5 we may restrict our attention to conjunctive queries that are connected. We also formulate an immediate corollary of Lemma 8.5 that, after finitely many applications, establishes the same for unions of conjunctive queries.

Corollary 8.6. *Let $\mu = \mu_1 \wedge \cdots \wedge \mu_k$ be as in Lemma 8.5 and suppose that μ occurs in a union μ' of conjunctive queries. For $i \in \{1, \dots, k\}$, let μ_i' be the union of queries obtained by replacing μ by μ_i in μ'. Then the resilience problem for μ' is NP-hard if the resilience problem for one of the μ_i' is NP-hard. Conversely, if the resilience problem is in P (in NP) for each μ_i', then the resilience problem for μ' is in P as well (in NP, respectively). The same is true in the setting without multiplicities and/or exogeneous relations.*

8.2. **Finite Duals.** If μ is a union of conjunctive queries with signature τ, then a *dual* of μ is a τ-structure $\mathfrak{A}$ with the property that a finite structure $\mathfrak{B}$ has a homomorphism to $\mathfrak{A}$ if and only if $\mathfrak{B}$ does not satisfy μ. The conjunctive query in Example 8.1, for instance, even has a *finite* dual. There is an elegant characterisation of the (unions of) conjunctive queries that have a finite dual. To state it, we need some basic terminology from database theory.

Definition 8.7. *The* incidence graph *of a relational τ-structure $\mathfrak{A}$ is the bipartite undirected multigraph whose first colour class is A, and whose second colour class consists of expressions of the form $R(b)$ where $R \in \tau$ has arity k, $b \in A^k$, and $\mathfrak{A} \models R(b)$. An edge $e_{a,i,R(b)}$ joins $a \in A$ with $R(b)$ if $b_i = a$. A structure is called* acyclic *if its incidence graph is acyclic, i.e., it contains no cycles (if two vertices are linked by two different edges, then they establish a cycle). A structure is called a* tree *if it is acyclic and connected in the sense defined in Section 8.1.*

The following was proved by Nešetřil and Tardif [40]; also see [18, 36].

Theorem 8.8. *A conjunctive query μ has a finite dual if and only if the canonical database of μ is homomorphically equivalent to a tree. A union of conjunctive queries has a finite dual if and only if the canonical database for each of the conjunctive queries is homomorphically equivalent to a tree.*

The theorem shows that in particular Example 8.2 does not have a finite dual, since the query given there is not acyclic and hence cannot be homomorphically equivalent to a tree. To construct valued structures from duals, we introduce the following notation.

Definition 8.9. *Let $\mathfrak{B}$ be a τ-structure and $\sigma \subseteq \tau$. Define $\Gamma(\mathfrak{B}, \sigma)$ to be the valued τ-structure on the same domain as $\mathfrak{B}$ such that*

- *for each $R \in \tau \setminus \sigma$, $R^{\Gamma(\mathfrak{B},\sigma)}(a) := 0$ if $a \in R^{\mathfrak{B}}$ and $R^{\Gamma(\mathfrak{B},\sigma)}(a) := 1$ otherwise, and*
- *for each $R \in \sigma$, $R^{\Gamma(\mathfrak{B},\sigma)}(a) := 0$ if $a \in R^{\mathfrak{B}}$ and $R^{\Gamma(\mathfrak{B},\sigma)}(a) := \infty$ otherwise.*

Note that $\mathrm{Aut}(\mathfrak{B}) = \mathrm{Aut}(\Gamma(\mathfrak{B}, \sigma))$ for any τ-structure $\mathfrak{B}$ and any σ. In the following result we use a correspondence between resilience problems for acyclic conjunctive queries and valued CSPs. The result then follows from the P versus NP-complete dichotomy theorem for valued CSPs over finite domains stated in Theorem 7.15.

Theorem 8.10. *Let μ be a union of acyclic conjunctive queries with relational signature τ and let $\sigma \subseteq \tau$. Then the resilience problem for μ with exogenous relations from σ is in P or NP-complete. Moreover, it is decidable whether the resilience problem for a given union of acyclic conjunctive queries is in P. If μ is a union of queries each of which is homomorphically equivalent to a tree and $\mathfrak{B}$ is the finite dual of μ (which exists by Theorem 8.8), then* $\mathrm{VCSP}(\Gamma(\mathfrak{B}, \sigma))$ *is polynomial-time equivalent to the resilience problem for μ with exogenous relations from σ.*

Proof. By virtue of Corollary 8.6, we may assume for the P versus NP-complete dichotomy that each of the conjunctive queries in μ is connected and thus a tree. The same is true also for the polynomial-time equivalence to a VCSP since replacing a conjunctive query in a union with a homomorphically equivalent one does not affect the complexity of resilience. Define $\Gamma := \Gamma(\mathfrak{B}, \sigma)$. We show that $\mathrm{VCSP}(\Gamma)$ is polynomial-time equivalent to the resilience problem for μ with exogenous relations from σ.

Given a finite bag database $\mathfrak{A}$ with signature τ with exogenous tuples from relations in σ, let ϕ be the τ-expression which contains for every $R \in \tau$ and for every tuple $a \in R^{\mathfrak{A}}$ the summand $R(a)$ with the same number of occurrences as is the multiplicity of a in $R^{\mathfrak{A}}$. Conversely, for every τ-expression ϕ we can create a bag database $\mathfrak{A}$ with signature τ and exogenous relations from σ.

The domain of $\mathfrak{A}$ is the set of variables of ϕ and for every $R \in \tau$ and $a \in R^{\mathfrak{A}}$ with multiplicity equal to the number of occurrences of the summand $R(a)$ in ϕ. In both situations, the resilience of $\mathfrak{A}$ with respect to μ equals the value of ϕ with respect to Γ. This shows the final statement of the theorem. The first statement now follows from Theorem 7.15.

Concerning the decidability of the tractability condition, first note that the finite dual of μ, and hence also Γ, can be effectively computed from μ (e.g., the construction of the dual in [40] is effective). The existence of a fractional cyclic polymorphism for a given valued structure Γ with finite domain and finite signature can be decided (in exponential time in the size of Γ; see [31]). $\square$

Remark 8.11. *We mention that Theorem 8.10 also applies to regular path queries which can be shown to always have a finite dual, see the related [13].*

Theorem 8.10 can be combined with the tractability results for VCSPs from Section 7 that use fractional polymorphisms. To illustrate fractional polymorphisms and how to find them, we revisit a known tractable resilience problem from [19–21, 38] and show that it has a fractional canonical pseudo cyclic polymorphism.

Example 8.12. *We revisit Example 8.1. Consider again the conjunctive query*

$$\exists x, y, z(R(x, y) \wedge S(y, z)).$$

There is a finite dual $\mathfrak{B}$ of μ with domain $\{0, 1\}$ which is finitely bounded homogeneous, as described in Example 8.1. That example also describes a valued structure Γ which is actually $\Gamma(\mathfrak{B}, \emptyset)$. Let ω be the fractional cyclic operation given by $\omega(\min) = \omega(\max) = \frac{1}{2}$. Since $\mathrm{Aut}(\Gamma)$ is trivial, ω is canonical. The fractional operation ω improves both weighted relations R and S (they are submodular*; see, e.g., [35]) and hence is a canonical cyclic fractional polymorphism of Γ.*

Combining Theorem 7.17 and 8.10, Example 8.12 reproves the results from [20] (without multiplicities) and [37] (with multiplicities) that the resilience problem for this query is in P.

8.3. **Infinite Duals.** Conjunctive queries might not have a finite dual (see Example 8.2), but unions of connected conjunctive queries always have a countably infinite dual. Cherlin, Shelah and Shi [15] showed that in this case we may even find a dual with an oligomorphic automorphism group (see Theorem 8.13 below). This is the key insight to phrase resilience problems as VCSPs for valued structures with oligomorphic automorphism groups. The not necessarily connected case again reduces to the connected case by Corollary 8.6.

In Theorem 8.13 below we state a variant of a theorem of Cherlin, Shelah, and Shi [15] (also see [5, 26, 27]). If $\mathfrak{B}$ is a structure, we write $\mathfrak{B}_{\mathrm{pp}(m)}$ for the expansion of $\mathfrak{B}$ by all relations that can be defined with a connected primitive positive formula (see Remark 8.4) with at most m variables, at least one free variable, and without equality. For a union of conjunctive queries μ over the signature τ, we write $|\mu|$ for the maximum of the number of variables of each conjunctive query in μ, the maximal arity of τ, and 2.

Theorem 8.13. *For every union μ of connected conjunctive queries over a finite relational signature τ there exists a τ-structure $\mathfrak{B}_\mu$ such that the following statements hold:*

(1) $(\mathfrak{B}_\mu)_{pp(|\mu|)}$ *is homogeneous.*
(2) $\mathrm{Age}(\mathfrak{B}_{pp(|\mu|)})$ *is the class of all substructures of structures of the form $\mathfrak{A}_{pp(|\mu|)}$ for a finite structure $\mathfrak{A}$ that satisfies $\neg\mu$.*
(3) *A countable τ-structure $\mathfrak{A}$ satisfies $\neg\mu$ if and only if it embeds into $\mathfrak{B}_\mu$.*
(4) $\mathfrak{B}_\mu$ *is finitely bounded.*

(5) $\mathrm{Aut}(\mathfrak{B}_\mu)$ *is oligomorphic.*
(6) $(\mathfrak{B}_\mu)_{pp(|\mu|)}$ *is finitely bounded.*

Proof. The construction of a structure $\mathfrak{B}_\mu$ with the given properties follows from a proof of Hubička and Nešetřil [26,27] of the theorem of Cherlin, Shelah, and Shi [15], and can be found in [5, Theorem 4.3.8]. Properties (1), (2) and property (3) restricted to finite structures $\mathfrak{A}$ are explicitly stated in [5, Theorem 4.3.8]. Property (3) restricted to finite structures clearly implies property (4). Property (5) holds because homogeneous structures with a finite relational signature have an oligomorphic automorphism group. Property (3) for countable structures now follows from [5, Lemma 4.1.7].

Since we are not aware of a reference for (6) in the literature, we present a proof here. Let σ be the signature of $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$. We claim that the following universal σ-sentence ψ describes the structures in the age of $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$. If ϕ is a σ-sentence, then ϕ' denotes the τ-sentence obtained from ϕ by replacing every occurrence of $R(\bar{x})$, for $R \in \sigma \setminus \tau$, by the primitive positive τ-formula $\eta(\bar{x})$ for which R was introduced in $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$. Then ψ is a conjunction of all σ-sentences $\neg\phi$ such that ϕ is primitive positive, ϕ' has at most $|\mu|$ variables, and ϕ' implies μ. Clearly, there are finitely many conjuncts of this form.

Suppose that $\mathfrak{A} \in \mathrm{Age}(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$. Then $\mathfrak{A}$ satisfies each conjunct $\neg\phi$ of ψ, because otherwise $\mathfrak{B}_\mu$ satisfies ϕ', and thus satisfies μ, contrary to our assumptions.

The interesting direction is that if a finite σ-structure $\mathfrak{A}$ satisfies ψ, then $\mathfrak{A}$ embeds into $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$. Let ϕ be the canonical query of $\mathfrak{A}$. Let $\mathfrak{A}'$ be the canonical database of the τ-formula ϕ'. Suppose for contradiction that $\mathfrak{A}' \models \mu$. Let χ be a minimal subformula of ϕ such that the canonical database of χ models μ. Then χ has at most $|\mu|$ variables and implies μ, and hence $\neg\chi$ is a conjunct of of ψ which is not satisfied by $\mathfrak{A}$, a contradiction to our assumptions. Therefore, $\mathfrak{A}' \models \neg\mu$ and by Property (2), we have that $\mathfrak{A}'_{\mathrm{pp}(|\mu|)}$ has an embedding f into $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$.

We claim that the restriction of f to the elements of $\mathfrak{A}$ is an embedding of $\mathfrak{A}$ into $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$. Clearly, if $\mathfrak{A} \models R(\bar{x})$ for some relation R that has been introduced for a primitive positive formula η, then $\mathfrak{A}'$ satisfies $\eta(\bar{x})$, and hence $\mathfrak{B}_\mu \models \eta(f(\bar{x}))$, which in turn implies that $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)} \models R(f(\bar{x}))$ as desired. Conversely, if $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)} \models R(f(\bar{x}))$, then $\mathfrak{A}'_{\mathrm{pp}(|\mu|)} \models R(\bar{x})$, and hence $\mathfrak{A}' \models \eta(\bar{x})$. This in turn implies that $\mathfrak{A} \models R(\bar{x})$. Since the restriction of f and its inverse preserve the relations from τ trivially, we conclude that $\mathfrak{A}$ embeds into $(\mathfrak{B}_\mu)_{\mathrm{pp}(|\mu|)}$. □

By Properties (1) and (6) of Theorem 8.13, $\mathfrak{B}_\mu$ is always a reduct of a finitely bounded homogeneous structure. For short, we write Γ_μ for $\Gamma(\mathfrak{B}_\mu, \emptyset)$ and $\Gamma_{\mu,\sigma}$ for $\Gamma(\mathfrak{B}_\mu, \sigma)$, see Definition 8.9. For some queries μ, the structure $\mathfrak{B}_\mu$ can be replaced by a simpler structure $\mathfrak{C}_\mu$. This will be convenient for some examples that we consider later, because the structure $\mathfrak{C}_\mu$ is finitely bounded and homogeneous itself and hence admits the application of Theorem 7.17. To define the respective class of queries, we need the following definition. The *Gaifman graph* of a relational structure $\mathfrak{A}$ is the undirected graph with vertex set A where $a, b \in A$ are adjacent if and only if $a \neq b$ and there exists a tuple in a relation of $\mathfrak{A}$ that contains both a and b. The Gaifman graph of a conjunctive query is the Gaifman graph of the canonical database of that query.

Theorem 8.14. *For every union μ of connected conjunctive queries over a finite relational signature τ such that the Gaifman graph of each of the conjunctive queries in μ is complete, there exists a countable τ-structure $\mathfrak{C}_\mu$ such that the following statements hold:*

(1) $\mathfrak{C}_\mu$ *is homogeneous.*
(2) $\mathrm{Age}(\mathfrak{C}_\mu)$ *is the class of all finite structures $\mathfrak{A}$ that satisfy $\neg\mu$.*

Moreover, $\mathfrak{C}_\mu$ is finitely bounded, $\operatorname{Aut}(\mathfrak{C}_\mu)$ is oligomorphic, and a countable τ-structure satisfies $\neg\mu$ if and only if it embeds into $\mathfrak{C}_\mu$.

Proof. Let $\mathfrak{A}_1$ and $\mathfrak{A}_2$ be finite τ-structures that satisfy $\neg\mu$. Since the Gaifman graph of each of the conjunctive queries in μ is complete, the union of the structures $\mathfrak{A}_1$ and $\mathfrak{A}_2$ satisfies $\neg\mu$ as well. By Fraïssé's Theorem (see, e.g., [25]) there is a countable homogeneous τ-structure $\mathfrak{C}_\mu$ such that $\operatorname{Age}(\mathfrak{C}_\mu)$ is the class of all finite structures that satisfy $\neg\mu$; this shows that $\mathfrak{C}_\mu$ is finitely bounded. Homogeneous structures with finite relational signature clearly have an oligomorphic automorphism group. For the final statement, see [5, Lemma 4.1.7]. □

Note that $\mathfrak{C}_\mu$ is homomorphically equivalent to $\mathfrak{B}_\mu$ by [5, Lemma 4.1.7]. Therefore, $\Gamma(\mathfrak{C}_\mu, \sigma)$ is homomorphically equivalent to $\Gamma_{\mu,\sigma}$ for any $\sigma \subseteq \tau$.

The following proposition follows straightforwardly from the definitions and provides a valued constraint satisfaction problem that is polynomial-time equivalent to the resilience problem for μ, similar to Theorem 8.10.

Proposition 8.15. *The resilience problem for a union of connected conjunctive queries μ where the relations from $\sigma \subseteq \tau$ are exogenous is polynomial-time equivalent to $\operatorname{VCSP}(\Gamma(\mathfrak{B}, \sigma))$ for any dual $\mathfrak{B}$ of μ; in particular, to $\operatorname{VCSP}(\Gamma_{\mu,\sigma})$.*

Proof. Let $\mathfrak{B}$ be a dual of μ. For every bag database $\mathfrak{A}$ over signature τ and with exogenous relations from σ, let ϕ be the τ-expression obtained by adding atomic τ-expressions $S(x_1, \dots, x_n)$ according to the multiplicity of the tuples $(x_1, \dots, x_n)$ in $S^{\mathfrak{A}}$ for all $S \in \tau$. Note that ϕ can be computed in polynomial time. Then the resilience of $\mathfrak{A}$ with respect to μ is at most u if and only if (ϕ, u) has a solution over $\Gamma(\mathfrak{B}, \sigma)$.

To prove a polynomial-time reduction in the other direction, let ϕ be a τ-expression. We construct a bag database $\mathfrak{A}$ with signature τ. The domain of $\mathfrak{A}$ are the variables that appear in ϕ and for every $S \in \tau$ and tuple $(x_1, \dots, x_n) \in S^{\mathfrak{A}}$, its multiplicity in $\mathfrak{A}$ is the number of times that $(x_1, \dots, x_n)$ in $S^{\mathfrak{A}}$ occurs as a summand of ϕ. The relations $S^{\mathfrak{A}}$ with $S \in \sigma$ are exogenous in $\mathfrak{A}$, the remaining ones are endogenous. Again, $\mathfrak{A}$ can be computed in polynomial time and the resilience of $\mathfrak{A}$ with respect to μ is at most u if and only if (ϕ, u) has a solution over $\Gamma(\mathfrak{B}, \sigma)$. □

In [37] one may find a seemingly more general notion of exogenous tuples, where in a single relation there might be both endogenous and exogenous tuples. Using Proposition 8.15 and Lemma 4.7, however, we can show that classifying the complexity of resilience problems according to our original definition also entails a classification of this variant.

Remark 8.16. *Let μ be a union of conjunctive queries with the signature τ, let $\sigma \subseteq \tau$, and let $\rho \subseteq \tau \setminus \sigma$. Suppose we would like to model the resilience problem for μ where the relations in σ are exogenous and the relations in ρ might contain both endogenous and exogenous tuples. Let $\mathfrak{B}$ be a dual of μ and Γ be the expansion of $\Gamma(\mathfrak{B}, \sigma)$ where for every relational symbol $R \in \rho$, there is also a relation $(R^x)^\Gamma = R^{\mathfrak{B}}$, i.e., a classical relation that takes values 0 and ∞. The resilience problem for μ with exogenous tuples specified as above is polynomial-time equivalent to $\operatorname{VCSP}(\Gamma)$ by analogous reductions as in Proposition 8.15. Note that $(R^x)^\Gamma = \operatorname{Opt}\left(R^{\Gamma(\mathfrak{B},\sigma)}\right)$ for every $R \in \rho$, and therefore by Lemma 4.7, $\operatorname{VCSP}(\Gamma)$ is polynomial-time equivalent to $\operatorname{VCSP}(\Gamma(\mathfrak{B}, \sigma))$ and thus to the resilience problem for μ where the relations in σ are exogeneous and the relations in $\tau \setminus \sigma$ are purely endogeneous. This justifies the restriction to our setting for exogenous tuples. Moreover, the same argument shows that if resilience of μ with all tuples endogenous is in P, then all variants of resilience of μ with exogenous tuples are in P as well.*

Similarly as in Example 8.12, Proposition 8.15 can be combined with the tractability results for VCSPs from Section 7 that use fractional polymorphisms to prove tractability of resilience problems.

Example 8.17. *We revisit Example 8.2. Consider the conjunctive query* $\exists x, y, z\, (R(x, y) \wedge S(x, y, z))$ *over the signature* $\tau = \{R, S\}$*. Note that the Gaifman graph of* μ *is complete; let* $\mathfrak{C}_\mu$ *be the structure from Theorem 8.14. We construct a binary pseudo cyclic canonical fractional polymorphism of* $\Gamma(\mathfrak{C}_\mu, \emptyset)$*. Let* $\mathfrak{M}$ *be the* τ*-structure with domain* $(C_\mu)^2$ *and where*

- $((b_1^1, b_1^2), (b_2^1, b_2^2)) \in R^{\mathfrak{M}}$ *if and only if* $(b_1^1, b_2^1) \in R^{\mathfrak{C}_\mu}$ *and* $(b_1^2, b_2^2) \in R^{\mathfrak{C}_\mu}$*, and*
- $((b_1^1, b_1^2), (b_2^1, b_2^2), (b_3^1, b_3^2)) \in S^{\mathfrak{M}}$ *if and only if* $(b_1^1, b_2^1, b_3^1) \in S^{\mathfrak{C}_\mu}$ *or* $(b_1^2, b_2^2, b_3^2) \in S^{\mathfrak{C}_\mu}$*.*

Similarly, let $\mathfrak{N}$ *be the* τ*-structure with domain* $(C_\mu)^2$ *and where*

- $((b_1^1, b_1^2), (b_2^1, b_2^2)) \in R^{\mathfrak{N}}$ *if and only if* $(b_1^1, b_2^1) \in R^{\mathfrak{C}_\mu}$ *or* $(b_1^2, b_2^2) \in R^{\mathfrak{C}_\mu}$*, and*
- $((b_1^1, b_1^2), (b_2^1, b_2^2), (b_3^1, b_3^2)) \in S^{\mathfrak{N}}$ *if and only if* $(b_1^1, b_2^1, b_3^1) \in S^{\mathfrak{C}_\mu}$ *and* $(b_1^2, b_2^2, b_3^2) \in S^{\mathfrak{C}_\mu}$*.*

Note that $\mathfrak{M} \not\models \mu$ *and hence there exists an embedding* $f \colon \mathfrak{M} \to \mathfrak{C}_\mu$*. Similarly, there exists an embedding* $g \colon \mathfrak{N} \to \mathfrak{C}_\mu$*. Clearly, both* f *and* g *regarded as operations on the set* C_μ *are pseudo cyclic (but in general not cyclic) and canonical with respect to* $\mathrm{Aut}(\mathfrak{C}_\mu)$ *(see Claim 6 in Proposition 8.21 for a detailed argument of this type). Let* ω *be the fractional operation given by* $\omega(f) = \frac{1}{2}$ *and* $\omega(g) = \frac{1}{2}$*. Then* ω *is a binary fractional polymorphism of* $\Gamma := \Gamma(\mathfrak{C}_\mu, \emptyset)$*: for* $b^1, b^2 \in (C_\mu)^2$ *we have*

$$\sum_{h \in \mathscr{O}^{(2)}} \omega(h) R^\Gamma(h(b^1, b^2)) = \frac{1}{2} R^\Gamma(f(b^1, b^2)) + \frac{1}{2} R^\Gamma(g(b^1, b^2))$$
$$= \frac{1}{2} \sum_{j=1}^{2} R^\Gamma(b^j). \tag{9}$$

so ω *improves* R*, and similarly we see that* ω *improves* S*.*

We proved that the corresponding valued structure has a binary canonical pseudo cyclic fractional polymorphism. By Theorem 7.17 and 8.15, this reproves the results from [20] (without multiplicities) and [37] (with multiplicities) that the resilience problem for this query is in P.

8.4. **The Resilience Tractability Conjecture.**

In this section we present a conjecture which implies, together with Corollary 5.13 and Lemma 8.5, a P versus NP-complete dichotomy for resilience problems for finite unions of conjunctive queries.

Conjecture 8.18. *Let* μ *be a union of connected conjunctive queries over the signature* τ*, and let* $\sigma \subseteq \tau$*. If the structure* $(\{0, 1\}; \mathrm{OIT})$ *has no pp-construction in* $\Gamma := \Gamma_{\mu,\sigma}$*, then* Γ *has a fractional polymorphism of arity* $\ell \geq 2$ *which is canonical and pseudo cyclic with respect to* $\mathrm{Aut}(\Gamma)$ *(and in this case,* $\mathrm{VCSP}(\Gamma)$ *is in P by Theorem 7.17).*

The conjecture is intentionally only formulated for VCSPs that stem from resilience problems, because it is known to be false for the more general situation of VCSPs for valued structures Γ that have the same automorphisms as a reduct of a finitely bounded homogeneous structure [5] (Section 12.9.1; the counterexample is even a CSP). However, see Conjecture 9.3 for a conjecture that could hold for VCSPs in this more general setting.

For the following conjunctive query μ, the NP-hardness of the resilience problem without multiplicites was shown in [20]; to illustrate our condition, we verify that $(\{0, 1\}; \mathrm{OIT})$ has a pp-construction in Γ_μ and thus prove in a different way that the resilience problem (with multiplicities) for μ is NP-hard.

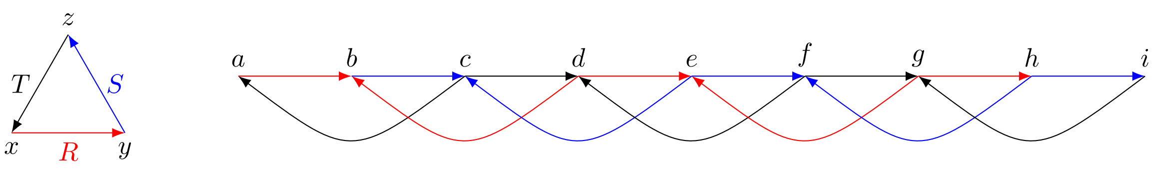

FIGURE 3. Example 8.19, visualisation of μ and ϕ.

Example 8.19 (Triangle query)**.** *Let τ be the signature that consists of three binary relation symbols R, S, and T, and let μ be the conjunctive query*

$$\exists x, y, z\big(R(x,y) \wedge S(y,z) \wedge T(z,x)\big).$$

The resilience problem without multiplicities for μ is NP-complete [20], and hence $\mathrm{VCSP}(\Gamma_\mu)$ is NP-hard (Proposition 8.15). Since the Gaifman graph of μ is NP-complete, the structure $\mathfrak{C}_\mu$ from Theorem 8.14 exists. Let $\Gamma := \Gamma(\mathfrak{C}_\mu, \emptyset)$. We provide a pp-construction of $(\{0,1\}; \mathrm{OIT})$ in Γ, which also proves NP-hardness of $\mathrm{VCSP}(\Gamma)$ and hence the resilience problem of μ with multiplicities by Corollary 5.13. Since Γ is homomorphically equivalent to Γ_μ, this also provides a pp-construction of $(\{0,1\}; \mathrm{OIT})$ in Γ_μ (see Lemma 5.14).

Let C be the domain of Γ. Let $\phi(a,b,c,d,e,f,g,h,i)$ be the τ-expression

$$R(a,b) + S(b,c) + T(c,d) + R(d,e) + S(e,f) + T(f,g) + R(g,h) + S(h,i) \quad (10)$$

$$+ T(i,g) + S(h,f) + R(g,e) + T(f,d) + S(e,c) + R(d,b) + T(c,a). \quad (11)$$

For an illustration of μ and ϕ, see Figure 3. Note that ϕ can be viewed as seven overlapping copies of μ.

In what follows, we say that an atomic τ-expression holds if it evaluates to 0. Note that every atom in (10) except the first and the last ones appears in exactly two copies of μ in ϕ, whereas all other atoms of ϕ occur in only one copy of μ in ϕ. Hence, since there are seven copies of μ in ϕ, in the optimal solution of the instance ϕ of $\mathrm{VCSP}(\Gamma)$ all atoms in (11) hold, and either every atom at even position or every atom at odd position in (10) holds. Let $RT \in \langle\Gamma\rangle$ be given by

$$RT(a,b,f,g) := \mathrm{Opt} \inf_{c,d,e,h,i \in C} \phi.$$

Note that $RT(a,b,f,g)$ holds if and only if

- *$R(a,b)$ holds and $T(f,g)$ does not hold, or*
- *$T(f,g)$ holds and $R(a,b)$ does not hold,*

where the reverse implication uses that $\mathfrak{C}_\mu$ is homogeneous. Similarly, define $RS \in \langle\Gamma\rangle$ by

$$RS(a,b,h,i) := \mathrm{Opt} \inf_{c,d,e,f,g \in C} \phi.$$

Note that $RS(a,b,h,i)$ holds if and only if

- *$R(a,b)$ holds and $S(h,i)$ does not hold, or*
- *$S(h,i)$ holds and $R(a,b)$ does not hold.*

Next, we define the auxiliary relation $\widetilde{RS}(a,b,e,f)$ to be

$$\mathrm{Opt} \inf_{c,d,g,h,i \in C} \phi.$$

Note that $\widetilde{RS}(a,b,e,f)$ holds if and only if

- *both $R(a,b)$ and $S(e,f)$ hold, or*
- *neither $R(a,b)$ and nor $S(e,f)$ holds.*

This allows us to define the relation

$$RR(u,v,x,y) := \inf_{w,z \in C} RS(u,v,w,z) + \widetilde{RS}(x,y,w,z)$$

which holds if and only if

- *$R(u,v)$ holds and $R(x,y)$ does not hold, or*
- *$R(x,y)$ holds and $R(u,v)$ does not hold.*

Define $M \in \langle \Gamma \rangle$ as

$$M := \mathrm{Opt} \inf_{x,y,z \in C} \big(RR(u,v,x,y) + RS(u',v',y,z) + RT(u'',v'',z,x) \\ + R(x,y) + S(y,z) + T(z,x) \big).$$

Note that $R(x,y)$, $S(y,z)$ and $T(z,x)$ cannot hold at the same time and therefore $(u,v,u',v',u'',v'') \in M$ if and only if exactly one of of $R(u,v)$, $R(u',v')$, and $R(u'',v'')$ holds. Let Δ be the pp-power of $(C;M)$ of dimension two with signature $\{\mathrm{OIT}\}$ such that

$$\mathrm{OIT}^{\Delta}((u,v),(u',v'),(u'',v'')) := M(u,v,u',v',u'',v'').$$

Then Δ is homomorphically equivalent to $(\{0,1\};\mathrm{OIT})$, witnessed by the homomorphism from Δ to $(\{0,1\};\mathrm{OIT})$ that maps (u,v) to 1 if $R(u,v)$ and to 0 otherwise, and the homomorphism $(\{0,1\};\mathrm{OIT}) \to \Delta$ that maps 1 to any pair of vertices $(u,v) \in R$ and 0 to any pair of vertices $(u,v) \notin R$. Therefore, Γ pp-constructs $(\{0,1\};\mathrm{OIT})$.

We mention that another conjecture concerning a P vs. NP-complete complexity dichotomy for resilience problems appears in [37, Conjecture 7.7]. The conjecture has a similar form as Conjecture 8.18 in the sense that it states that a sufficient hardness condition for resilience is also necessary. The relationship between our hardness condition from Corollary 5.13 and the condition from [37] remains to be studied.

8.5. **An example of formerly open complexity.** We use our approach to settle the complexity of the resilience problem for a conjunctive query that was mentioned as an open problem in [21] (Section 8.5):

$$\mu := \exists x,y (S(x) \wedge R(x,y) \wedge R(y,x) \wedge R(y,y)) \tag{12}$$

Let $\tau = \{R,S\}$ be the signature of μ. To study the complexity of resilience of μ, it will be convenient to work with a dual which has different model-theoretic properties than the duals $\mathfrak{B}_\mu$ from Theorem 8.13 and $\mathfrak{C}_\mu$ from Theorem 8.14, namely a dual that is a model-complete core.

Definition 8.20. *A structure $\mathfrak{B}$ with an oligomorphic automorphism group is* model-complete *if every embedding of $\mathfrak{B}$ into $\mathfrak{B}$ preserves all first-order formulas. It is a* core *if every endomorphism is an embedding.*

Note that the definition of cores of valued structures with finite domain (Definition 7.12) and the definition above specialise to the same concept for relational structures over finite domains. A structure with an oligomorphic automorphism group is a model-complete core if and only if for every $n \in \mathbb{N}$ every orbit of n-tuples can be defined with an existential positive formula [5]. Every countable structure $\mathfrak{B}$ is homomorphically equivalent to a model-complete core, which is unique up

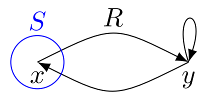

FIGURE 4. Visualisation of the query μ from (12).

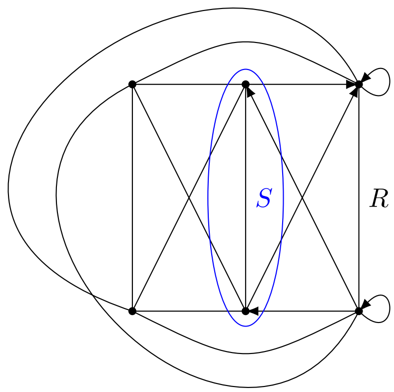

FIGURE 5. Illustration of a finite substructure of $\mathfrak{B}$ that contains representatives for all orbits of pairs of $\mathrm{Aut}(\mathfrak{B})$. Arrows are not drawn on undirected edges.

to isomorphism [4, 5]; we refer to this structure as the model-complete core of $\mathfrak{B}$. The advantage of working with model-complete cores is that the structure is in a sense 'minimal' and therefore easier to work with in concrete examples.[2]

Proposition 8.21. *There is a finitely bounded homogeneous dual $\mathfrak{B}$ of μ such that the valued τ-structure $\Gamma := \Gamma(\mathfrak{B}, \emptyset)$ has a binary fractional polymorphism which is canonical and pseudo cyclic with respect to* $\mathrm{Aut}(\Gamma)$. *Hence,* $\mathrm{VCSP}(\Gamma)$ *and the resilience problem for μ are in P. As a consequence, the polynomial-time tractability result even holds for resilience of μ with exogeneous relations from any $\sigma \subseteq \tau$.*

Proof. Since the Gaifman graph of μ is a complete graph, there exists the structure $\mathfrak{C}_\mu$ as in Theorem 8.14. Let $\mathfrak{B}$ be the model-complete core of $\mathfrak{C}_\mu$. Note that $\mathfrak{B}$ has the property that a countable structure $\mathfrak{A}$ maps homomorphically to $\mathfrak{B}$ if and only if $\mathfrak{A} \models \neg\mu$; in particular, $\mathfrak{B}$ is a dual of μ and $\mathfrak{B} \models \neg\mu$. The structure $\mathfrak{C}_\mu$ is homogeneous, and it is known that the model-complete core of a homogeneous structure is again homogeneous (see Proposition 4.7.7 in [5]), so $\mathfrak{B}$ is homogeneous. Let $\Gamma := \Gamma(\mathfrak{B}, \emptyset)$.

Note that

$$\mathfrak{B} \models \forall x\big(\neg S(x) \vee \neg R(x, x)\big) \tag{13}$$

$$\text{and } \mathfrak{B} \models \forall x, y\big(x = y \vee R(x, y) \vee R(y, x)\big). \tag{14}$$

To see (14), suppose for contradiction that $\mathfrak{B}$ contains distinct elements x, y such that neither (x, y) nor (y, x) is in $R^{\mathfrak{B}}$. Let $\mathfrak{B}'$ be the structure obtained from $\mathfrak{B}$ by adding (x, y) to $R^{\mathfrak{B}}$. Then

[2] The model-complete core of $\mathfrak{B}_\mu$ would be a natural choice for the canonical dual of μ to work with instead of $\mathfrak{B}_\mu$. However, proving that the model-complete core has a finitely bounded homogeneous expansion (so that, for example, Theorem 3.4 applies) requires introducing further model-theoretical notions [39] which we want to avoid in this article.

$\mathfrak{B}' \models \neg\mu$ as well, and hence there is a homomorphism from $\mathfrak{B}'$ to $\mathfrak{B}$ by the properties of $\mathfrak{B}$. This homomorphism is also an endomorphism of $\mathfrak{B}$ which is not an embedding, a contradiction to the assumption that $\mathfrak{B}$ is a model-complete core.

Also observe that

$$\mathfrak{B} \models \forall x, y\big(x = y \vee (R(x,y) \wedge R(y,x)) \vee (S(x) \wedge R(y,y)) \vee (R(x,x) \wedge S(y))\big). \tag{15}$$

Suppose for contradiction that (15) does not hold for some distinct x and y. Then $\neg S(x) \vee \neg R(y,y)$ and $\neg R(x,x) \vee \neg S(y)$, i.e., $\neg S(x) \wedge \neg R(x,x)$, or $\neg S(x) \wedge \neg S(y)$, or $\neg R(y,y) \wedge \neg R(x,x)$, or $\neg R(y,y) \wedge \neg S(y)$. In each of these cases we may add both R-edges between the distinct elements x and y to $\mathfrak{B}$ and obtain a structure not satisfying μ, which leads to a contradiction as above.

For an illustration of a finite substructure of $\mathfrak{B}$ which contains a representative for every orbit of pairs in $\mathrm{Aut}(\mathfrak{B})$, see Figure 5.

Claim 1. For every a finite τ-structure $\mathfrak{A}$ that satisfies $\neg\mu$ and the sentences in (14) and (15), there exists a *strong* homomorphism to $\mathfrak{B}$, i.e., a homomorphism that also preserves the complements of R and S. First observe that $\mathfrak{B}$ embeds the countably infinite complete graph, where R is the edge relation and precisely one element lies in the relation S; this is because this structure maps homomorphically to $\mathfrak{B}$ and unless embedded, it contradicts $\mathfrak{B} \not\models \mu$. In particular, there are infinitely many $x \in B$ such that $\mathfrak{B} \models \neg S(x) \wedge \neg R(x,x)$ and by (15), for every $y \in B$, $x \neq y$, we have $\mathfrak{B} \models R(x,y) \wedge R(y,x)$.

To prove the claim, let $\mathfrak{A}$ be a finite structure that satisfies $\neg\mu$ and the sentences in (14) and (15). For a homomorphism h from $\mathfrak{A}$ to $\mathfrak{B}$, let

$$s(h) := |\{x \in A \mid \mathfrak{A} \models \neg S(x) \wedge \mathfrak{B} \models S(h(x))\}|$$

and

$$r(h) := |\{(x,y) \in A^2 \mid \mathfrak{A} \models \neg R(x,y) \wedge \mathfrak{B} \models R(h(x),h(y))\}|.$$

Let h be a homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$, which exists since $\mathfrak{A} \models \neg\mu$. If $s(h) + r(h) = 0$, then h is a strong homomorphism and there is nothing to prove. Suppose therefore $s(h) + r(h) > 0$. We construct a homomorphism h' such that $r(h') + s(h') < r(h) + s(h)$. Since $r(h) + s(h)$ is finite, by applying this construction finitely many times, we obtain a strong homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$.

If $s(h) > 0$, then there exists $a \in A \setminus S^{\mathfrak{A}}$ such that $h(a) \in S^{\mathfrak{B}}$. By (13), $\mathfrak{B} \not\models R(h(a),h(a))$ and hence $\mathfrak{A} \not\models R(a,a)$. Pick $b \in B \setminus h(A)$ such that $\mathfrak{B} \models \neg S(b) \wedge \neg R(b,b)$ and define

$$h'(x) := \begin{cases} b \text{ if } x = a, \\ h(x) \text{ otherwise.} \end{cases}$$

Observe that h' is a homomorphism, $s(h') < s(h)$ and $r(h') = r(h)$. If $r(h) > 0$, then there exists $(x,y) \in A^2 \setminus R^{\mathfrak{A}}$ such that $(h(x),h(y)) \in R^{\mathfrak{B}}$. If $x = y$, the argument is similar as in the case $s(h) > 0$. Finally, if $x \neq y$, then $\mathfrak{A} \models (S(x) \wedge R(y,y)) \vee (R(x,x) \wedge S(y))$, because $\mathfrak{A}$ satisfies the sentence in (15). Since $\mathfrak{A}$ satisfies the sentence in (14), $\mathfrak{A} \models R(y,x)$. Since h is a homomorphism, we have

$$\mathfrak{B} \models R(h(x),h(y)) \wedge R(h(y),h(x)) \wedge ((S(h(x)) \wedge R(h(y),h(y))) \vee (R(h(x),h(x)) \wedge S(h(y)))),$$

which contradicts $\mathfrak{B} \not\models \mu$.

Claim 2. Every finite τ-structure $\mathfrak{A}$ that satisfies $\neg\mu$ and the sentences in (14) and (15) embeds into $\mathfrak{B}$. In particular, $\mathfrak{B}$ is finitely bounded. Let $\mathfrak{A}$ be such a structure. By Theorem 8.13, there is an embedding e of $\mathfrak{A}$ into $\mathfrak{B}_\mu$. Since $\mathfrak{B}_\mu$ is homogeneous and embeds every finite τ-structure that satisfies $\neg\mu$, there exists a finite substructure $\mathfrak{A}'$ of $\mathfrak{B}_\mu$ satisfying the sentences in (14) and (15) such that $e(\mathfrak{A})$ is a substructure of $\mathfrak{A}'$ and for all distinct $a, b \in A$ there exists $s \in S^{\mathfrak{A}'}$ such that $\mathfrak{B}_\mu \models R(e(a), s) \wedge R(s, e(b))$. By Claim 1, there is a strong homomorphism h from $\mathfrak{A}'$ to $\mathfrak{B}$.

We claim that $h \circ e$ is injective and therefore an embedding of $\mathfrak{A}$ into $\mathfrak{B}$. Suppose there exist distinct $a, b \in A$ such that $h(e(a)) = h(e(b))$. Since $e(\mathfrak{A})$ satisfies the sentence in (14) and h is a strong homomorphism, we obtain that $\mathfrak{B}_\mu \models R(e(a), e(a)) \wedge R(e(b), e(b))$. Let $s \in S^{\mathfrak{A}'}$ be such that $\mathfrak{B}_\mu \models R(e(a), s) \wedge R(s, e(b))$. Hence,

$$\mathfrak{B} \models S(h(s)) \wedge R(h(e(a)), h(s)) \wedge R(h(s), h(e(a))) \wedge R(h(e(a)), h(e(a))),$$

a contradiction to $\mathfrak{B} \not\models \mu$. It follows that $h \circ e$ is an embedding of $\mathfrak{A}$ into $\mathfrak{B}$.

We define two $\{R, S\}$-structures $\mathfrak{M}, \mathfrak{N}$ with domain B^2 as follows. For all $x_1, x_2, y_1, y_2, x, y \in B$ define

$$\mathfrak{M}, \mathfrak{N} \models R\big((x_1, y_1), (x_2, y_2)\big) \qquad \text{if } \mathfrak{B} \models R(x_1, x_2) \wedge R(y_1, y_2), \tag{16}$$

$$\mathfrak{M}, \mathfrak{N} \models S\big((x, y)\big) \qquad \text{if } \mathfrak{B} \models S(x) \wedge S(y) \tag{17}$$

$$\mathfrak{M} \models S\big((x, y)\big) \qquad \text{if } \mathfrak{B} \models S(x) \vee S(y) \tag{18}$$

$$\mathfrak{N} \models R\big((x, y), (x, y)\big) \qquad \text{if } \mathfrak{B} \models R(x, x) \vee R(y, y). \tag{19}$$

Add pairs of distinct elements to $R^{\mathfrak{M}}$ and $R^{\mathfrak{N}}$ such that both $\mathfrak{M}$ and $\mathfrak{N}$ satisfy the sentence in (15) (note that no addition of elements to $S^{\mathfrak{M}}$ and $S^{\mathfrak{N}}$ is needed). Finally, add $((x_1, y_1), (x_2, y_2))$ to $R^{\mathfrak{M}}$ and $((x_2, y_2), (x_1, y_1))$ to $R^{\mathfrak{N}}$ if at least one of the following cases holds:

(A) $\mathfrak{B} \models S(x_1) \wedge R(x_1, x_2) \wedge R(x_2, x_2) \wedge R(y_2, y_2) \wedge R(y_2, y_1) \wedge S(y_1)$,
(B) $\mathfrak{B} \models R(x_1, x_1) \wedge R(x_1, x_2) \wedge S(x_2) \wedge y_1 = y_2 \wedge R(y_1, y_2)$,
(C) $\mathfrak{B} \models S(y_1) \wedge R(y_1, y_2) \wedge R(y_2, y_2) \wedge R(x_2, x_2) \wedge R(x_2, x_1) \wedge S(x_1)$,
(D) $\mathfrak{B} \models R(y_1, y_1) \wedge R(y_1, y_2) \wedge S(y_2) \wedge x_1 = x_2 \wedge R(x_1, x_2)$.

Conditions (A) and (B) are illustrated in Figure 6; conditions (C) and (D) are obtained from (A) and (B) by replacing x by y. Note that for $(x_1, y_1) = (x_2, y_2)$, none of the conditions (A)-(D) is ever satisfied. No other atomic formulas hold on $\mathfrak{M}$ and $\mathfrak{N}$. Note that both $\mathfrak{M}$ and $\mathfrak{N}$ satisfy the property stated for $\mathfrak{B}$ in (13).

Claim 3. $\mathfrak{M}$ and $\mathfrak{N}$ satisfy the sentence in (14). We prove the statement for $\mathfrak{M}$; the proof for $\mathfrak{N}$ is similar. Let $(x_1, y_1), (x_2, y_2) \in B$ be such that $(x_1, y_1) \neq (x_2, y_2)$ and $\mathfrak{M} \models \neg R((x_2, y_2), (x_1, y_1))$. Since $\mathfrak{M}$ satisfies the sentence in (15), we must have either $\mathfrak{M} \models S(x_1, y_1) \wedge R((x_2, y_2), (x_2, y_2))$ or $\mathfrak{M} \models S(x_2, y_2) \wedge R((x_1, y_1), (x_1, y_1))$. Suppose the former is true; the other case is treated analogously. Then $\mathfrak{B} \models R(x_2, x_2) \wedge R(y_2, y_2)$ and $\mathfrak{B} \models S(x_1) \vee S(y_1)$. If $\mathfrak{B} \models S(x_1)$, then $x_1 \neq x_2$ and by (14) we have $\mathfrak{B} \models R(x_1, x_2) \vee R(x_2, x_1)$. By (14) and (15) for (y_1, y_2), we obtain that $\mathfrak{M} \models R((x_1, y_1), (x_2, y_2))$ by (16) or one of the conditions (A)-(D). The argument if $\mathfrak{B} \models S(y_1)$ is similar with x and y switched.

Claim 4. $\mathfrak{M}$ and $\mathfrak{N}$ satisfy $\neg\mu$. Let $x_1, x_2, y_1, y_2 \in B$. Suppose for contradiction that

$$\mathfrak{M} \models S(x_1, y_1) \wedge R((x_1, y_1), (x_2, y_2)) \wedge R((x_2, y_2), (x_1, y_1)) \wedge R((x_2, y_2), (x_2, y_2)).$$

By the definition of $\mathfrak{M}$, we have $\mathfrak{B} \models R(x_2, x_2) \wedge R(y_2, y_2)$ and $\mathfrak{B} \models S(x_1) \vee S(y_1)$. Assume that $\mathfrak{B} \models S(x_1)$; the case $\mathfrak{B} \models S(y_1)$ is analogous.

By the assumption, $\mathfrak{M} \models R((x_1, y_1), (x_2, y_2))$. Then, by the definition of $\mathfrak{M}$, one of the conditions (16), (A)-(D) holds, or

$$\mathfrak{M} \models \neg\big(S(x_1, y_1) \wedge R((x_2, y_2), (x_2, y_2))\big)$$

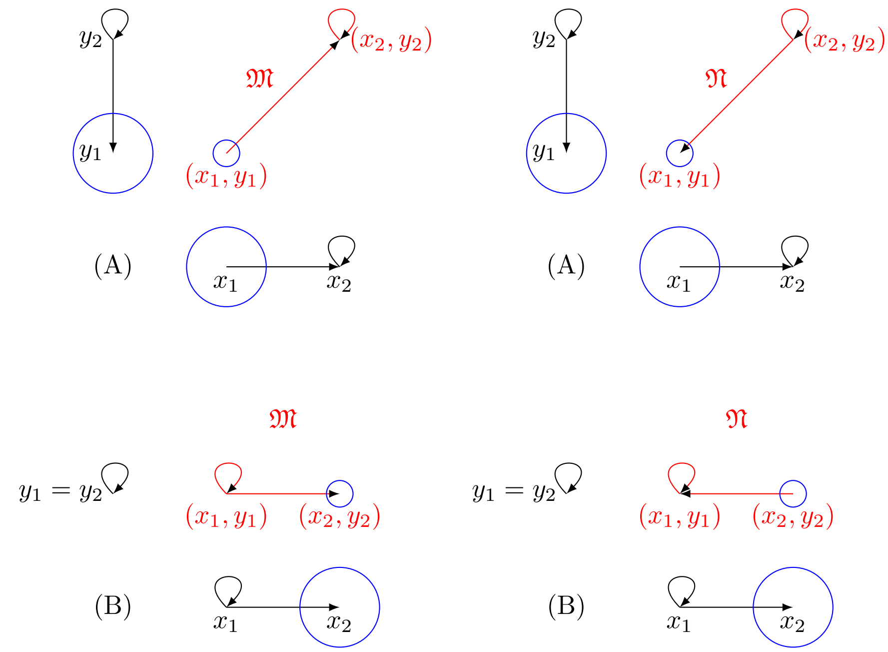

FIGURE 6. An illustration of the conditions (A) and (B) in $\mathfrak{M}$ and $\mathfrak{N}$.

(recall that $((x_1, y_1), (x_2, y_2))$ might have been added to $R^{\mathfrak{M}}$ so that $\mathfrak{M}$ satisfies the sentence in (15)). The last option is false by the assumption and by (13), $\mathfrak{B} \models \neg S(x_2) \wedge \neg S(y_2)$, and hence neither (B) nor (D) holds. Therefore, one of the conditions (16), (A), or (C) holds for $((x_1, y_1), (x_2, y_2))$. Similarly, we obtain that one of the conditions (16) or (B) holds for $((x_2, y_2), (x_1, y_1))$, since $\mathfrak{M} \models R((x_2, y_2), (x_1, y_1))$ (to exclude (D) we use the assumption that $\mathfrak{B} \models S(x_1)$ and hence $x_1 \neq x_2$). This yields six cases and in each of them we must have that $\mathfrak{B} \models R(x_1, x_2) \wedge R(x_2, x_1)$ or $\mathfrak{B} \models S(y_1) \wedge R(y_1, y_2) \wedge R(y_2, y_1)$. Since $\mathfrak{B} \models S(x_1) \wedge R(x_2, x_2) \wedge R(y_2, y_2)$, this contradicts $\mathfrak{B} \models \neg\mu$. Since $(x_1, y_1), (x_2, y_2) \in M$ were chosen arbitrarily, this shows that $\mathfrak{M} \models \neg\mu$. The argument for $\mathfrak{N}$ is similar.

Claim 5. There is an embedding f of $\mathfrak{M}$ into $\mathfrak{B}$ and an embedding g of $\mathfrak{N}$ into $\mathfrak{B}$. We show the claim for $\mathfrak{M}$; the proof for $\mathfrak{N}$ is analogous. By [5, Lemma 4.1.7], it is enough to show that every finite substructure of $\mathfrak{M}$ embeds into $\mathfrak{B}$. By the definition of $\mathfrak{M}$ and Claims 3 and 4, every finite substructure $\mathfrak{M}$ satisfies (14), (15) and $\neg\mu$ and hence, by Claim 2, it embeds into $\mathfrak{B}$.

Let ω be the fractional operation over B defined by $\omega(f) = \frac{1}{2}$ and $\omega(g) = \frac{1}{2}$.

Claim 6. ω is pseudo cyclic and canonical with respect to $\mathrm{Aut}(\mathfrak{B}) = \mathrm{Aut}(\Gamma)$. Note that since $\mathfrak{B}$ is homogeneous in a finite relational signature, two k-tuples of elements of $\mathfrak{B}$ lie in the same orbit if and only if they satisfy the same atomic formulas. Therefore, the canonicity of f and g with respect to $\mathrm{Aut}(\mathfrak{B})$ follows from the definition of $\mathfrak{M}$ and $\mathfrak{N}$: for $(a, b) \in B^2$, whether $\mathfrak{B} \models S(f(a, b))$ only depends on whether $\mathfrak{M} \models S(a, b)$ by Claim 5, which depends only on the atomic formulas that hold on a and on b in $\mathfrak{B}$. An analogous statement is true for atomic formulas of the form $R(x, y)$ and $x = y$. Therefore, f is canonical. The argument for the canonicity of g is analogous.

To see that f and g are pseudo cyclic, we show that f^* and g^* defined on 2-orbits (using the terminology of Remark 7.4) are cyclic. By the definition of f^*, we need to show that for any $a_1, a_2, b_1, b_2 \in B$, the two pairs $(f(a_1, b_1), f(a_2, b_2))$ and $(f(b_1, a_1), f(b_2, a_2))$ satisfy the same atomic formulas. For the formulas of the form $S(x)$ and $R(x, y)$, this can be seen from Claim 5 and the definition of $\mathfrak{M}$ and $\mathfrak{N}$, since each of the conditions (16),(17),(18),(19),(15) and the union of (A), (B), (C), (D) is symmetric with respect to exchanging x and y. For the atomic formulas of the form $x = y$, this follows from the injectivity of f. This shows that f^* is cyclic; the argument for g^* is the same. Hence, the pseudo-cyclicity of f and g is a consequence of Lemma 7.9 for $m = 2$.

Claim 7. ω improves S.
By the definition of $\mathfrak{M}$ and $\mathfrak{N}$ and Claim 5, we have for all $x, y \in B$

$$\omega(f)S^\Gamma(f(x,y)) + \omega(g)S^\Gamma(g(x,y)) = \frac{1}{2}(S^\Gamma(x) + S^\Gamma(y)).$$

Claim 8. ω improves R.
Let $x_1, y_1, x_2, y_2 \in B$. We have to verify that

$$\omega(f)R^\Gamma(f(x_1,y_1), f(x_2,y_2)) + \omega(g)R^\Gamma(g(x_1,y_1), g(x_2,y_2)) \leq \frac{1}{2}(R^\Gamma(x_1,x_2) + R^\Gamma(y_1,y_2)). \quad (20)$$

We distinguish four cases.

- $\mathfrak{M}, \mathfrak{N} \models R((x_1, y_1), (x_2, y_2))$. Then Inequality (20) holds since the left-hand side is zero, and the right-hand side is non-negative (each weighted relation in Γ is non-negative).
- $\mathfrak{M}, \mathfrak{N} \models \neg R((x_1, y_1), (x_2, y_2))$. Since $\mathfrak{M}$ and $\mathfrak{N}$ satisfy the sentences in (14) and (15) and $\mathfrak{B}$ satisfies (14) we must have $\mathfrak{B} \models \neg R(x_1, x_2) \wedge \neg R(y_1, y_2)$, and both sides of the inequality evaluate to 1.
- $\mathfrak{M} \models \neg R((x_1, y_1), (x_2, y_2))$ and $\mathfrak{N} \models R((x_1, y_1), (x_2, y_2))$. By Claim 5, the left-hand side evaluates to $\frac{1}{2}$. By (16), we have $\mathfrak{B} \models \neg R(x_1, x_2)$ or $\mathfrak{B} \models \neg R(y_1, y_2)$. Therefore, the right-hand side of (20) is at least $\frac{1}{2}$ and the inequality holds.
- $\mathfrak{M} \models R((x_1, y_1), (x_2, y_2))$ and $\mathfrak{N} \models \neg R((x_1, y_1), (x_2, y_2))$. Similar to the previous case.

This exhausts all cases and concludes the proof of Claim 8.

It follows that ω is a binary fractional polymorphism of Γ which is canonical and pseudo cyclic with respect to $\mathrm{Aut}(\Gamma)$. Polynomial-time tractability of VCSP(Γ) follows by Theorem 7.17 and 8.15. The final statement follows from Remark 8.16. ☐

9. Conclusion and Future Work

We formulated a general hardness condition for VCSPs of valued structures with an oligomorphic automorphism group and a new polynomial-time tractability result. We use the latter to resolve a resilience problem whose complexity was left open in the literature and conjecture that our conditions exactly capture the hard and easy resilience problems for conjunctive queries (with multiplicities), respectively. In fact, a full classification of resilience problems for conjunctive queries based on our approach seems feasible, but requires further research, as discussed in the following.

We have proved that if Γ is a valued structure with an oligomorphic automorphism group and R is a weighted relation in the smallest weighted relational clone that contains the weighted relations of Γ, then R is preserved by all fractional polymorphisms of Γ (Lemma 6.8). We do not know whether the converse is true. Note that it is known to hold for the special cases of finite-domain valued structures [17, 22] and for classical relational structures with 0-∞ valued relations (CSP setting) having an oligomorphic automorphism group [9].

Question 9.1. *Let Γ be a valued structure with an oligomorphic automorphism group. Is it true that $R \in \langle \Gamma \rangle$ if and only if $R \in \mathrm{Imp}(\mathrm{fPol}(\Gamma))$?*

Note that a positive answer to this question would imply that the computational complexity of VCSPs for valued structures Γ with an oligomorphic automorphism group, and in particular the complexity of resilience problems, is fully determined by the fractional polymorphisms of Γ.

Fractional polymorphisms are probability distributions on operations. In all the examples that arise from resilience problems that we considered so far, it was sufficient to work with fractional polymorphisms ω that are *finitary*, i.e., such that there are finitely many operations $f_1, \dots, f_k \in \mathscr{O}_C$ such that $\sum_{i \in \{1,\dots,k\}} \omega(f_i) = 1$. This motivates the following question.

Question 9.2. *Does our notion of pp-constructability change if we restrict to finitary fractional homomorphisms ω? Is there a valued structure Γ with an oligomorphic automorphism group and a weighted relation R such that R is not improved by all fractional polymorphism of Γ, but* is *improved by all finitary fractional polymorphisms ω? In particular, are these statements true if we restrict to valued τ-structures Γ that arise from resilience problems as described in Proposition 8.15?*

In the following, we formulate a common generalisation of the complexity-theoretic implications of Conjecture 8.18 and the infinite-domain tractability conjecture from [10] that concerns a full complexity classification of VCSPs for valued structures from reducts of finitely bounded homogeneous structures.

Conjecture 9.3. *Let Γ be a valued structure with finite signature such that $\mathrm{Aut}(\Gamma) = \mathrm{Aut}(\mathfrak{B})$ for some reduct $\mathfrak{B}$ of a countable finitely bounded homogeneous structure. If $(\{0,1\}; \mathrm{OIT})$ has no pp-construction in Γ, then $\mathrm{VCSP}(\Gamma)$ is in P (otherwise, we already know that $\mathrm{VCSP}(\Gamma)$ is NP-complete by Theorem 3.4 and Corollary 5.13).*

One might hope to prove this conjecture under the assumption of the infinite-domain tractability conjecture. Recall that also the finite-domain VCSP classification was first proven conditionally on the finite-domain tractability conjecture [32, 34], which was only confirmed later [11, 46].

We also believe that the 'meta-problem' of deciding whether for a given conjunctive query the resilience problem with multiplicities is in P is decidable. This would follow from a positive answer to Conjecture 8.18 because Γ_m^* can be computed and Item 4 of Proposition 7.16 for the finite-domain valued structure Γ_m^* can be decided algorithmically using linear programming [31].

References

[1] N. Bansal, A. Blum, and S. Chawla. Correlation clustering. *Machine Learning*, 56(1-3):89–113, 2004.

[2] L. Barto and M. Kozik. Absorbing subalgebras, cyclic terms and the constraint satisfaction problem. *Logical Methods in Computer Science*, 8/1(07):1–26, 2012.

[3] L. Barto, J. Opršal, and M. Pinsker. The wonderland of reflections. *Israel Journal of Mathematics*, 223(1):363–398, 2018.

[4] M. Bodirsky. Cores of countably categorical structures. *Logical Methods in Computer Science (LMCS)*, 3(1):1–16, 2007.

[5] M. Bodirsky. *Complexity of Infinite-Domain Constraint Satisfaction*. Lecture Notes in Logic (52). Cambridge University Press, 2021.

[6] M. Bodirsky and J. 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).

[7] M. Bodirsky, M. Mamino, and C. Viola. Piecewise linear valued CSPs solvable by linear programming relaxation. *ACM Transactions of Computational Logic*, 23(1):1–35, 2022. Preprint arXiv:1912.09298.

[8] M. Bodirsky and A. Mottet. A dichotomy for first-order reducts of unary structures. *Logical Methods in Computer Science*, 14(2), 2018.

[9] M. Bodirsky and J. Nešetřil. Constraint satisfaction with countable homogeneous templates. In *Proceedings of Computer Science Logic (CSL)*, pages 44–57, Vienna, 2003.

[10] M. Bodirsky, M. Pinsker, and A. Pongrácz. Projective clone homomorphisms. *Journal of Symbolic Logic*, 86(1):148–161, 2021.

[11] A. A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17*, pages 319–330, 2017.

[12] S. Butti. *Symmetries in Constraint Satisfaction: Weisfeiler-Leman Invariance and Promise Problems*. PhD thesis, Universitat Pompeu Fabra, Barcelona, 2022.

[13] D. Calvanese, G. D. Giacomo, M. Lenzerini, and M. Y. Vardi. View-based query processing and constraint satisfaction. In *15th Annual IEEE Symposium on Logic in Computer Science, Santa Barbara, California, USA, June 26-29, 2000*, pages 361–371, 2000.

[14] S. Chaudhuri and M. Y. Vardi. Optimization of real conjunctive queries. In *Proceedings of the Twelfth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, PODS '93, pages 59–70, New York, NY, USA, 1993. Association for Computing Machinery.

[15] G. Cherlin, S. Shelah, and N. Shi. Universal graphs with forbidden subgraphs and algebraic closure. *Advances in Applied Mathematics*, 22:454–491, 1999.

[16] D. A. Cohen, M. C. Cooper, P. Creed, P. G. Jeavons, and S. Živný. An algebraic theory of complexity for discrete optimization. *SIAM J. Comput.*, 42(5):1915–1939, 2013.

[17] D. A. Cohen, M. C. Cooper, and P. Jeavons. An algebraic characterisation of complexity for valued constraints. In *Principles and Practice of Constraint Programming - CP 2006, 12th International Conference, CP 2006, Nantes, France, September 25-29, 2006, Proceedings*, pages 107–121, 2006.

[18] J. Foniok. Homomorphisms and structural properties of relational systems. Doctoral Thesis, Charles University Prague, Faculty of Mathematics and Physics, 2007.

[19] C. Freire, W. Gatterbauer, N. Immerman, and A. Meliou. A characterization of the complexity of resilience and responsibility for self-join-free conjunctive queries. *CoRR*, abs/1507.00674, 2015.

[20] C. Freire, W. Gatterbauer, N. Immerman, and A. Meliou. The complexity of resilience and responsibility for self-join-free conjunctive queries. *Proc. VLDB Endow.*, 9(3):180–191, 2015.

[21] C. Freire, W. Gatterbauer, N. Immerman, and A. Meliou. New results for the complexity of resilience for binary conjunctive queries with self-joins. In *Proceedings of the 39th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2020, Portland, OR, USA, June 14-19, 2020*, pages 271–284, 2020.

[22] P. Fulla and S. Živný. A Galois connection for weighted (relational) clones of infinite size. *TOCT*, 8(3):9:1–9:21, 2016.

[23] M. Garey and D. Johnson. *A guide to NP-completeness*. CSLI Press, Stanford, 1978.

[24] A. V. Goldberg and R. E. Tarjan. A new approach to the maximum-flow problem. *J. Assoc. Comput. Mach.*, 35(4):921–940, 1988.

[25] W. Hodges. *A shorter model theory*. Cambridge University Press, Cambridge, 1997.

[26] J. Hubička and J. Nešetřil. Universal structures with forbidden homomorphisms. In *Logic Without Borders - Essays on Set Theory, Model Theory, Philosophical Logic and Philosophy of Mathematics*, pages 241–264. De Gruyter, 2015. arXiv:0907.4079.

[27] J. Hubička and J. Nešetřil. Homomorphism and embedding universal structures for restricted classes. *Multiple-Valued Logic and Soft Computing*, 27(2-3):229–253, 2016. arXiv:0909.4939.

[28] M. A. Khamis, P. G. Kolaitis, H. Q. Ngo, and D. Suciu. Bag query containment and information theory. *ACM Trans. Database Syst.*, 46(3), 2021.

[29] E. J. Kim, S. Kratsch, M. Pilipczuk, and M. Wahlström. Solving hard cut problems via flow-augmentation. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 149–168, 2021.

[30] E. J. Kim, S. Kratsch, M. Pilipczuk, and M. Wahlström. Directed flow-augmentation. In *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 938–947, 2022.

[31] V. Kolmogorov. Testing the complexity of a valued CSP language. *CoRR*, abs/1803.02289, 2018.

[32] V. Kolmogorov, A. A. Krokhin, and M. Rolínek. The complexity of general-valued CSPs. *SIAM J. Comput.*, 46(3):1087–1110, 2017.

[33] V. Kolmogorov, J. Thapper, and S. Živný. The power of linear programming for general-valued CSPs. *SIAM J. Comput.*, 44(1):1–36, 2015.

[34] M. Kozik and J. Ochremiak. Algebraic properties of valued constraint satisfaction problem. In *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, pages 846–858, 2015.
[35] A. Krokhin and S. Zivny. The Complexity of Valued CSPs. In A. Krokhin and S. Zivny, editors, *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*, pages 233–266. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2017.
[36] B. Larose, C. Loten, and C. Tardif. A characterisation of first-order constraint satisfaction problems. *Logical Methods in Computer Science*, 3(4:6), 2007.
[37] N. Makhija and W. Gatterbauer. A Unified Approach for Resilience and Causal Responsibility with Integer Linear Programming (ILP) and LP Relaxations, 2022.
[38] A. Meliou, W. Gatterbauer, K. F. Moore, and D. Suciu. The complexity of causality and responsibility for query answers and non-answers. *Proc. VLDB Endow.*, 4(1):34–45, 2010.
[39] A. Mottet and M. Pinsker. Cores over Ramsey structures. *Journal of Symbolic Logic*, 86(1):352–361, 2021.
[40] J. Nešetřil and C. Tardif. Duality theorems for finite structures. *Journal of Combinatorial Theory, Series B*, 80:80–97, 2000.
[41] F. M. Schneider and C. Viola. An application of Farkas' lemma to finite-valued constraint satisfaction problems over infinite domains. *CoRR*, abs/2208.04912, 2022.
[42] J. Thapper and S. Živný. The complexity of finite-valued CSPs. In *Proceedings of the Symposium on Theory of Computing Conference (STOC), Palo Alto, CA, USA, June 1-4, 2013*, pages 695–704, 2013.
[43] C. Viola. *Valued Constraint Satisfaction Problems over Infinite Domains*. PhD thesis, TU Dresden, 2020.
[44] C. Viola and S. Zivný. The combined basic LP and affine IP relaxation for promise VCSPs on infinite domains. *ACM Trans. Algorithms*, 17(3):21:1–21:23, 2021.
[45] D. Zhuk. A proof of the CSP dichotomy conjecture. *J. ACM*, 67(5):30:1–30:78, 2020.
[46] D. N. Zhuk. A proof of CSP dichotomy conjecture. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17*, pages 331–342, 2017. https://arxiv.org/abs/1704.01914.

Appendix A. The Lebesgue Integral

It will be convenient to use an additional value $-\infty$ that has the usual properties:

- $-\infty < a$ for every $a \in \mathbb{R} \cup \{\infty\}$,
- $a + (-\infty) = (-\infty) + a = -\infty$ for every $a \in \mathbb{R}$,
- $a \cdot \infty = \infty \cdot a = -\infty$ for $a < 0$,
- $0 \cdot (-\infty) = (-\infty) \cdot 0 = 0$.
- $a \cdot (-\infty) = (-\infty) \cdot a = -\infty$ for $a > 0$ and $a \cdot (-\infty) = (-\infty) \cdot a = \infty$ for $a < 0$.

The sum of ∞ and $-\infty$ is undefined.

Let C and D be sets. We define the Lebesgue integration over the space C^D of all functions from D to C. We usually (but not always) work with the special case $D = C^\ell$, i.e. the space is $\mathscr{O}_C^{(\ell)}$ for some set C and $\ell \in \mathbb{N}$.

To define the Lebesgue integral, we need the definition of a *simple function*: this is a function $Y : C^D \to \mathbb{R}$ given by

$$\sum_{k=1}^{n} a_k 1_{S_k}$$

where $n \in \mathbb{N}$, $S_1, S_2, \dots$ are disjoint elements of $B(C^D)$, $a_k \in \mathbb{R}$, and $1_S : C^D \to \{0,1\}$ denotes the indicator function for $S \subseteq C^D$. If Y is a such a simple function, then the Lebesgue integral is defined as follows:

$$\int_{C^D} Y d\omega := \sum_{k=1}^{n} a_k \omega(S_k).$$

If X and Y are two random variables, then we write $X \leq Y$ if $X(f) \leq Y(f)$ for every $f \in C^D$. We say that X is *non-negative* if $0 \leq X$. If X is a non-negative measurable function, then the Lebesgue integral is defined as

$$\int_{C^D} X d\omega := \sup\left\{\int_{C^D} Y d\omega \mid 0 \leq Y \leq X, Y \text{ simple}\right\}.$$

For an arbitrary measurable function X, we write $X = X^+ - X^-$, where

$$X^+(x) := \begin{cases} X(x) & \text{if } X(x) > 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$X^-(x) := \begin{cases} -X(x) & \text{if } X(x) < 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then both X^+ and X^- are measurable, and both $\int_{C^D} X^- d\omega$ and $\int_{C^D} X^+ d\omega$ take values in $\mathbb{R}_{\geq 0} \cup \{\infty\}$. If both take value ∞, then the integral is undefined (see Remark A.1). Otherwise, define

$$\int_{C^D} X d\omega := \int_{C^D} X^+ d\omega - \int_{C^D} X^- d\omega.$$

In particular, note that for $X \geq 0$ the integral is always defined.

Let ω be a fractional map from D to C, let $R \in \mathscr{R}_C^{(k)}$ be a weighted relation, and let $s \in D^k$. Then $X\colon C^D \to \mathbb{R} \cup \{\infty\}$ given by

$$f \mapsto R(f(s))$$

is a random variable: if (a, b) is a basic open subset of $\mathbb{R} \cup \{\infty\}$, then

$$\begin{aligned} X^{-1}((a,b)) &= \{f \in C^D \mid R(f(s)) \in (a,b)\} \\ &= \bigcup_{t \in C^k, R(t) \in (a,b)} \mathscr{S}_{s,t} \end{aligned}$$

is a union of basic open sets in C^D, hence open. The argument for the other basic open sets in $\mathbb{R} \cup \{\infty\}$ is similar.

If the set C is countable and X is as above, we may express $E_\omega[X]$ as a sum, which is useful in proofs in Sections 5 and 6. If $E_\omega[X]$ exists, then

$$\begin{aligned} E_\omega[X] &= \int_{C^D} X^+ d\omega - \int_{C^D} X^- d\omega \\ &= \sup\left\{\int_{C^D} Y d\omega \mid 0 \leq Y \leq X^+, Y \text{ simple}\right\} - \sup\left\{\int_{C^D} Y d\omega \mid 0 \leq Y \leq X^-, Y \text{ simple}\right\} \\ &= \sum_{t \in C^k, R(t) \geq 0} R(t)\omega(\mathscr{S}_{s,t}) + \sum_{t \in C^k, R(t) < 0} R(t)\omega(\mathscr{S}_{s,t}) \\ &= \sum_{t \in C^k} R(t)\omega(\mathscr{S}_{s,t}). \end{aligned} \tag{21}$$

Remark A.1. *The Lebesgue integral*

$$\int_{\mathscr{O}_C^{(\ell)}} X d\omega = \int_{\mathscr{O}_C^{(\ell)}} X^+ d\omega - \int_{\mathscr{O}_C^{(\ell)}} X^- d\omega$$

need not exist: e.g., consider $C = \mathbb{N}$, $k = \ell = 1$, *and* $R(x) = -2^x$ *if* $x \in \mathbb{N}$ *is even and* $R(x) = 2^x$ *otherwise. Let* $s \in C$ *and define* $X \colon \mathscr{O}_C^{(1)} \to \mathbb{R} \cup \{\infty\}$ *by*

$$f \mapsto R(f(s)).$$

Let ω *be a unary fractional operation such that for every* $t \in C$ *we have* $\omega(\mathscr{S}_{s,t}) = \frac{1}{2^{t+1}}$. *Then*

$$\begin{aligned}\int_{\mathscr{O}_C^{(\ell)}} X^+ d\omega &= \sup\{\int_{\mathscr{O}_C^{(\ell)}} Y d\omega \mid 0 \leq Y \leq X^+, Y \text{ simple}\} \\ &= \sum_{t \in C, R(t) \geq 0} R(t)\omega(\mathscr{S}_{s,t}) \\ &= \sum_{t \in C, R(t) \geq 0} \frac{1}{2} = \infty\end{aligned}$$

and, similarly, $\int_{\mathscr{O}_C^{(\ell)}} X^- d\omega = \infty$.

Institut für Algebra, TU Dresden
Email address: manuel.bodirsky@tu-dresden.de

Institut für Algebra, TU Dresden
Email address: zaneta.semanisinova@tu-dresden.de

Institut für Informatik, Leipzig University
Email address: clu@informatik.uni-leipzig.de