Cohomology in Constraint Satisfaction and Structure Isomorphism

Adam Ó Conghaile
Computer Laboratory, Cambridge University, United Kingdom
The Alan Turing Institute, United Kingdom

Abstract

CSP and SI are among the most well-studied computational problems in Computer Science. While neither problem is thought to be in PTIME, much work is done on PTIME approximations to both problems. Two such historically important approximations are the k-consistency algorithm for CSP and the k-Weisfeiler-Leman algorithm for SI, both of which are based on propagating local partial solutions. The limitations of these algorithms are well-known – k-consistency can solve precisely those CSPs of bounded width and k-Weisfeiler-Leman can only distinguish structures which differ on properties definable in C^k. In this paper, we introduce a novel sheaf-theoretic approach to CSP and SI and their approximations. We show that both problems can be viewed as deciding the existence of global sections of presheaves, $\mathcal{H}_k(A, B)$ and $\mathcal{I}_k(A, B)$ and that the success of the k-consistency and k-Weisfeiler-Leman algorithms correspond to the existence of certain efficiently computable subpresheaves of these. Furthermore, building on work of Abramsky and others in quantum foundations, we show how to use Čech cohomology in $\mathcal{H}_k(A, B)$ and $\mathcal{I}_k(A, B)$ to detect obstructions to the existence of the desired global sections and derive new efficient cohomological algorithms extending k-consistency and k-Weisfeiler-Leman. We show that cohomological k-consistency can solve systems of equations over all finite rings and that cohomological Weisfeiler-Leman can distinguish positive and negative instances of the Cai-Fürer-Immerman property over several important classes of structures.

2012 ACM Subject Classification Theory of computation → Finite Model Theory

Keywords and phrases constraint satisfaction problems, finite model theory, descriptive complexity, rank logic, Weisfeiler-Leman algorithm, cohomology

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

Acknowledgements I thank Anuj Dawar and Samson Abramsky for useful discussions in writing this paper and am especially gratefully to Samson Abramsky for permission to use results from [1].

1 Introduction

Constraint satisfaction problems (CSP) and structure isomorphism (SI) are two of the most well-studied problems in complexity theory. Mathematically speaking, an instance of one of these problems takes a pair of structures (A, B) as input and asks whether there is a homomorphism $A \to B$ for CSP or an isomorphism $A \cong B$ for SI. These problems are not in general thought to be tractable. Indeed the general case of CSP is NP-Complete and restricting our structures to graphs the best known algorithm for SI is Babai's quasi-polynomial time algorithm.[7] As a result, it is common in complexity and finite model theory to study approximations of the relations $\to$ and $\cong$.

The k-consistency and k-Weisfeiler-Leman[1] algorithms efficiently determine two such approximations to $\to$ and $\cong$ which we call $\to_k$ and $\equiv_k$. These relations have many char-

[1] The algorithm we call "k-Weisfeiler-Leman" is more commonly called "$(k-1)$-Weisfeiler-Leman" in the literature, see for example [12]. We prefer "k-Weisfeiler-Leman" to emphasise its relationship to k-variable logic and sets of k-local isomorphisms.

42nd Conference on Very Important Topics (CVIT 2016).
Editors: John Q. Open and Joan R. Access; Article No. 23; pp. 23:1–23:30
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

acterisations in logic and finite model theory, for example in [17] and [12]. One that is particularly useful is that of the existence of winning strategies for Duplicator in certain Spoiler-Duplicator games with k pebbles[25] [23]. For both of these games Duplicator's winning strategies can be represented as non-empty sets $S \subset \mathbf{Hom}_k(A, B)$ of k-local partial homomorphisms which satisfy some extension properties and connections between these games have been studied before. For example, a joint comonadic semantics is given by the pebbling comonad of Abramsky, Dawar and Wang[4].

The limitations of these approximations are well-known. In particular, it is known that k-consistency only solves CSPs of *bounded width* and k-Weisfeiler-Leman can only distinguish structures which differ on properties expressible in the infinitary counting logic $\mathcal{C}^k$. Feder and Vardi[17] showed that CSP encoding linear equations over the finite fields do not have bounded width, while Cai, Fürer, and Immerman[12] demonstrated an efficiently decidable graph property which is not expressible in $\mathcal{C}^k$ for any k.

In the present paper, we introduce a novel approach to the CSP and SI problems based on presheaves of k-local partial homomorphisms and isomorphisms, showing that the problems can be reframed as deciding whether certain presheaves admit global sections. We show that the classic k-consistency and k-Weisfeiler-Leman algorithms can be derived by computing greatest fixpoints of presheaf operators which remove some efficiently computable obstacles to global sections. Furthermore, we show how invariants from sheaf cohomology can be used to find further obstacles to combining local homomorphisms and isomorphisms into global ones. We use these to construct new efficient extensions to the k-consistency and k-Weisfeiler-Leman algorithms computing relations $\rightarrow_k^{\mathbb{Z}}$ and $\equiv_k^{\mathbb{Z}}$ which refine $\rightarrow_k$ and $\equiv_k$.

The application of presheaves has been particularly successful in computer science in recent decades with applications in semantics[27, 18], information theory[28] and quantum contextuality[3, 5, 2]. This work owes draws in particular on the application of sheaf theory to quantum contextuality, pioneered by Abramsky and Brandenburger[3] and developed by Abramsky and others for example in [5] and [2].

Using this work, we prove that these new cohomological algorithms are strictly stronger than k-consistency and k-Weisfeiler-Leman. In particular, we show that cohomological k-consistency decides solvability of linear equations with k variables per equation over all finite rings and that there is a fixed k such that $\equiv_k^{\mathbb{Z}}$ distinguishes structures which differ on Cai, Fürer and Immerman's property.

It is also interesting to compare $\rightarrow_k^{\mathbb{Z}}$ and $\equiv_k^{\mathbb{Z}}$ with other well-studied refinements of $\rightarrow_k$ and $\equiv_k$ such as the algorithms of Bulatov[11] and Zhuk[30] which decide all tractable CSPs, the algorithms of Živný et al.[10, 13] for Promise CSPs and the invertible-map equivalence of Dawar and Holm[16] which bounds the expressive power of rank logic. The latter was recently used by Lichter[26] to demonstrate a property which is decidable in PTIME but not expressible in rank logic. In our paper, we show that $\equiv_k^{\mathbb{Z}}$, for some fixed k , can distinguish structures which differ on this property. Comparing $\rightarrow_k^{\mathbb{Z}}$ to the Bulatov-Zhuk algorithm and algorithms for PCSPs remains a direction for future work.

The rest of the paper proceeds as follows. Section 2 establishes some background and notation. Section 3 introduces the presheaf formulation of CSP and SI and new formulations of k-consistency and k-Weisfeiler-Leman in this framework. Section 4 demonstrates how to apply aspects of sheaf cohomology to CSP and SI and defines new algorithms along these lines. Section 5 surveys the strength of these new cohomological algorithms. Section 6 concludes with some open questions and directions for future work. Major proofs and additional background are left to the appendices.

2 Background and definitions

In this section, we record some definitions and background which are necessary for our work.

2.1 Relational structures & finite model theory

Throughout this paper we use the word *structure* to mean a relational structure over some finite relational signature σ. A structure A consists of an underlying set (which will also call A) and for each relational symbol R of arity r in σ a subset $R^A \subset A^r$ or tuples related by R. A *homomorphism* of structures A, B over a common signature is a function between the underlying sets $f : A \to B$ which preserves related tuples. An *isomorphism* of structures is a bijection between the underlying sets which both preserves and reflects related tuples.

In the paper, we make reference to several important logics from finite model theory and descriptive complexity theory. The logics we make reference to in this paper are as follows.

- Fixed-point logic with counting (written FPC) is first-order logic extended with operators for inflationary fixed-points and and counting, for example see [19].
- For any natural number k, C^k is infinitary first-order logic extended with counting quantifiers with at most k variables. This logic bounds the expressive power of FPC in the sense that, for each k' there exists k such that any FPC formula in k' variables is equivalent to one in C^k. We write C^ω for the union of these logics.
- Rank logic is first-order logic extended with operators for inflationary fixed-points and computing ranks of matrices over finite fields, see [29].
- Linear algebraic logic is first-order infinitary logic extended with quantifiers for computing *all* linear algebraic functions over finite fields, see [14]. This logic bounds rank logic in the sense described above.

At different points in the history of descriptive complexity theory, both FPC and rank logic were considered as candidates for "capturing `PTIME`" and thus refuting a well-known conjecture of Gurevich[21]. Each has since been proven not to capture `PTIME`, for FPC see Cai, Fürer and Immerman[12], for rank logic see Lichter[26]. Infinitary logics such as C^ω and linear algebraic logic are capable of expressing properties which are not decidable in `PTIME` but have been shown not to contain any logic which does not capture `PTIME`. For C^ω, see Cai, Fürer and Immerman [12] and for linear algebraic logic, see Dawar, Grädel, and Lichter[15].

2.2 Constraint satisfaction problems & Structure Isomorphism

Assuming a fixed relational signature σ, we write CSP for the set of all pairs of σ-structures (A, B) such that there is a homomorphism witnessing $A \to B$. We use $CSP(B)$ to denote the set of relational structures A such that $(A, B) \in CSP$. We also use CSP and $CSP(B)$ to denote the decision problem on these sets. For general B, $CSP(B)$ is well-known to be `NP-complete`. However for certain structures B the problem is in `PTIME`. Indeed, the Bulatov-Zhuk Dichotomy Theorem (formerly the Feder-Vardi Dichotomy Conjecture) states that for any B $CSP(B)$ is either `NP`-complete or it is `PTIME`. Working out efficient algorithms which decide $CSP(B)$ for larger and larger classes of B was an active area of research which culminated in Bulatov and Zhuk's exhaustive classes of algorithms [11], [30].

Similarly, we write SI for the set of all pairs of σ-structures (A, B) such that there is an isomorphism witnessing $A \cong B$. The decision problem for this set is also thought not to be in `PTIME` however there are no general hardness results known for this. The best known algorithm (in the case where σ is the signature of graphs) is Babai's[7] which is quasi-polynomial.

The two approximations to CSP and SI which we consider here are the k-consistency and k-Weisfeiler-Leman algorithms. If a pair (A, B) is accepted by k-consistency (resp. k-Weisfeiler-Leman) we write $A \rightarrow_k B$ (resp. $A \equiv_k B$). These relations each have several characterisations in terms of logic, algorithms and games. We use the formulation in terms of positional Duplicator winning strategies for the games of Kolaitis and Vardi[25] and Hella[22] which are respectively downwards-closed sets S of partial homomorphisms or isomorphisms of domain size at most k such that any $s \in S$ of size less than k satisfies the forth property $\mathbf{Forth}(S, s)$ or the bijective forth property $\mathbf{BijForth}(S, s)$. Where $\mathbf{Forth}(S, s)$ holds if $\forall a \in A,\ \exists b \in B$ s.t. $s \cup \{(a, b)\} \in S$ and $\mathbf{BijForth}(S, s)$ holds if $\exists b_s : A \rightarrow B$ a bijection s.t. $\forall a \in A\ s \cup \{(a, b_s(a))\} \in S$. For more details, see Appendix B.

2.3 Presheaves & cohomology

Given two categories $\mathbf{C}$ and $\mathbf{S}$, an $\mathbf{S}$-valued presheaf over $\mathbf{C}$ is a contravariant functor $\mathcal{F} : \mathbf{C}^{op} \rightarrow \mathbf{S}$. We will assume that $\mathbf{C}$ is the posetal category on some subset of the powerset $P(X)$ of some set X, which we will call the underlying space of $\mathbf{C}$. For this reason, when $U' \subset U$ in $\mathbf{C}$ we write $(\cdot)_{|_{U'}}$ for the map $F(U' \subset U)$. We also restrict $\mathbf{S}$ to being either the category $\mathbf{Set}$ of sets or the category $\mathbf{AbGrp}$ of abelian groups. We call $\mathbf{AbGrp}$-valued presheaves, abelian presheaves and $\mathbf{Set}$-valued presheaves are just called presheaves or presheaves of sets where there is ambiguity.

For any $\mathbf{C}$ and $\mathbf{S}$ as above the category of presheaves $\mathbf{PrSh}(\mathbf{C}, \mathbf{S})$ has presheaves $\mathcal{F} : \mathbf{C}^{op} \rightarrow$ as objects and natural transformations as morphisms. If $\mathbf{S}$ has a terminal object 1 (as both $\mathbf{Set}$ and $\mathbf{AbGp}$ do) then the presheaf $\mathbb{I} \in \mathbf{PrSh}(\mathbf{C}, \mathbf{S})$ which sends all elements of $\mathbf{C}$ to 1 is a terminal object in $\mathbf{PrSh}(\mathbf{C}, \mathbf{S})$. For any $\mathcal{F} \in \mathbf{PrSh}(\mathbf{C}, \mathbf{S})$, a *global section* of $\mathcal{F}$ is a natural transformation $S : \mathbb{I} \implies \mathcal{F}$.

3 Presheaves of local homomorphisms and isomorphisms

Some important efficient algorithms for CSP and SI involve working with sets of k-local homomorphisms between the two structures in a given instance. These sets of partial homomorphisms of domain size $\leq k$ are useful for constructing efficient algorithms because computing the sets $\mathbf{Hom}_k(A, B)$ and $\mathbf{Isom}_k(A, B)$ can be done in polynomial time in $|A| \cdot |B|$. In this section, we see that these sets can naturally be given the structure of sheaves, that the CSP and SI problems can be seen as the search for global sections of these sheaves and that the k-consistency and k-Weisfeiler-Leman algorithms can both be seen as determining the existence of certain special subpresheaves. The framework of considering sheaves of local homomorphisms and isomorphisms is novel in this work and essential for the main cohomological algorithms later. The results in Section 3.3 are from a technical report of Samson Abramsky[1] and we are thank him for his permission to include them here.

3.1 Defining presheaves of homomorphisms and isomorphisms

Let A and B be relational structures over the same signature. A *partial homomorphism* is a partial function $s : A \rightharpoonup B$ that preserves related tuples in $\mathbf{dom}(s)$. A *partial isomorphism* is a partial homomorphism $s : A \rightharpoonup B$ which is injective and reflects related tuples from $\mathbf{im}(s)$. A *k-local homomorphism (resp. isomorphism)* is a partial homomorphism (resp. isomorphism) s such that $|\mathbf{dom}(s)| \leq k$. We write $\mathbf{Hom}_k(A, B)$ (resp. $\mathbf{Isom}_k(A, B)$) for the sets of k-local homomorphisms (resp. isomorphisms). We write $\mathbf{Hom}(A, B)$ for the union $\bigcup_{1\le k\le |A|} \mathbf{Hom}_k(A, B)$ and $\mathbf{Isom}(A, B)$ for the union $\bigcup_{1\le k\le |A|} \mathbf{Isom}_k(A, B)$.

It is not hard to see that these sets can be given the structure of presheaves on the underlying space A. Indeed, we define the *presheaf of homomorphisms from A to B* $\mathcal{H}(A, B) : \mathbf{P(A)}^{op} \to \mathbf{Set}$ as $\mathcal{H}(A, B)(U) = \{s \in \mathbf{Hom}(A, B) \mid \mathbf{dom}(s) = U\}$ with restriction maps $\mathcal{H}(A, B)(U' \subset U)$ given by the restriction of partial homomorphisms $(\cdot)_{|_{U'}}$. Similarly, let $\mathcal{I}(A, B)$ be the subpresheaf of $\mathcal{H}(A, B)$ containing only partial isomorphisms. Now, consider the cover of A by subsets of size at most k, written $A^{\le k} \subset P(A)$. We define the *presheaves of k-local homomorphisms and isomorphisms* $\mathcal{H}_k(A, B)$ and $\mathcal{I}_k(A, B)$ as the functors $\mathcal{H}(A, B)$ and $\mathcal{I}(A, B)$ restricted to the subcategory $(\mathbf{A^{\le k}})^{op} \subset \mathbf{P(A)}^{op}$.

We now see how these presheaves and their global sections encode the CSP and SI problems for the instance (A, B).

3.2 CSP and SI as search for global sections

Fix an instance (A, B) for the CSP or SI problem and let $\mathcal{H}$ and $\mathcal{I}$ stand for the presheaves of all partial homomorphisms and isomorphisms between A and B defined in the last section. For either of these sheaves a global section $s : \mathbb{I} \implies \mathcal{S}$ is a collection $\{s_U \in \mathcal{S}(U)\}_{U\in P(A)}$ where naturality implies that for any subsets U and U' of A $(s_U)_{|_{U\cap U'}} = (s_{U'})_{|_{U\cap U'}}$. As the poset $P(A)$ has a maximal element, namely A, any such global section is determined by a choice of $s_A \in \mathcal{S}(A)$. This leads us to the following observation.

▶ **Observation 1.** *Given a pair (A, B) relational structures over the same signature then*

$$(A, B) \in CSP \iff \mathcal{H} \text{ has a global section}$$

and if $|A| = |B|$ then

$$(A, B) \in SI \iff \mathcal{I} \text{ has a global section}$$

This observation reframes the CSP and SI problems in terms of presheaves but algorithmically this not a particularly useful restating as computing the full objects $\mathcal{H}$ and $\mathcal{I}$ requires solving the CSP and SI problems for all subsets of A and B. A much more interesting equivalent condition is that for large enough k, whether or not a particular instance (A, B) is in CSP or SI is determined by the global sections of the presheaves of k-local homomorphisms and isomorphisms.

▶ **Lemma 2.** *For a pair (A, B) relational structures over the same signature, σ, and k at least the arity of sigma then*

$$(A, B) \in CSP \iff \mathcal{H}_k \text{ has a global section}$$

and if $|A| = |B|$ then

$$(A, B) \in SI \iff \mathcal{I}_k \text{ has a global section}$$

Proof. See Appendix A. ◀

This is more interesting than the previous observation as $\mathcal{H}_k$ and $\mathcal{I}_k$ can be computed for any relational structures A and B in $\mathcal{O}(\text{poly}(|A| \cdot |B|))$. Indeed, we can just list all $\mathcal{O}(|A|^k \cdot |B|^k)$ possible k-local functions and check which ones preserve (and reflect) related tuples. This also gives us an interesting starting point for designing efficient algorithms for approximating CSP and SI. In particular, any efficient algorithms which finds obstacles to the existence of global sections in $\mathcal{H}_k$ and $\mathcal{I}_k$ will provide a tractable approximation to CSP and SI. We now see how this approach can be used to capture some classical approximations to these problems.

3.3 Algorithms and games in terms of presheaves

In this section, we consider the approximations $A \to_k B$ and $A \equiv_k B$ to CSP and SI which are computed respectively by the k-consistency and k-Weisfeiler-Leman algorithms and we show that these algorithms can be seen as searching for certain obstructions to global sections in $\mathcal{H}_k(A, B)$ and $\mathcal{I}_k(A, B)$. In particular, we define efficiently computable monotone operators on subpresheaves of $\mathcal{H}_k$ and $\mathcal{I}_k$ and show that they have non-empty greatest fixpoints if and only if (A, B) are accepted by k-consistency and k-Weisfeiler-Leman respectively. Proposition 3 is reproduced with permission from an unpublished technical report of Samson Abramsky and the formulation of the fixpoint operators is inspired by the same report.

3.3.1 Flasque presheaves and k-consistency

Recall that $A \to_k B$ if and only if there is a positional winning strategy for Duplicator in the existential k-pebble game[17] and that a presheaf $\mathcal{F}$ is flasque if all of the restriction maps $\mathcal{F}(U \subset U')$ are surjective. In a recent technical report, Abramsky[1] proves the following characterisation of these strategies in our presheaf setting.

▶ **Proposition 3.** *For A, B relational structures and any k there is a bijection between:*

- *positional strategies in the existential k-pebble game from A to B, and*
- *non-empty flasque subpresheaves $\mathcal{S} \subset \mathcal{H}_k(A, B)$.*

This gives an alternative to the standard k-consistency algorithm which constructs the largest flasque subpresheaf $\overline{\mathcal{H}_k}$ of $\mathcal{H}_k$ and checks if it is empty. This can be computed efficiently as the greatest fixpoint of the presheaf operator $(\cdot)^{\uparrow\downarrow}$ which computes the largest subpresheaf of a presheaf $\mathcal{S} \subset \mathcal{H}_k$ such that every $s \in \mathcal{S}^{\uparrow\downarrow}(C)$ satisfies the forth property $\mathbf{Forth}(\mathcal{S}, s)$. For further details see Appendix B

3.3.2 Greatest fixpoints and k-Weisfeiler-Leman

In a similar way to the k-consistency algorithm, k-Weisfeiler-Leman can be formulated as determining the existence of a positional strategy for Duplicator in the k-pebble bijection game between A and B. This inspires the definition of another efficiently computable presheaf operator $(\cdot)^{\#\downarrow}$ which computes the largest subpresheaf of a presheaf $\mathcal{S} \subset \mathcal{I}_k$ such that for every $s \in \mathcal{S}^{\#\downarrow}(C)$ satisfies the bijective forth property $\mathbf{BijForth}(\mathcal{S}, s)$. We call the greatest fixpoint of this operator $\overline{\overline{\mathcal{S}}}$ and we have that $A \equiv_k B$ if and only if $\overline{\overline{\mathcal{I}_k}}$ is non-empty. For more details, see Appendix B.

To conclude, in this section, we have seen how to reformulate the search for homomorphisms and isomorphisms between relational structures A and B as the search for global sections in the presheaves $\mathcal{H}_k(A, B)$ and $\mathcal{I}_k(A, B)$. We have also seen that common approximations to homomorphism and isomorphism $\to_k$ and $\equiv_k$ can be computed a greatest fixpoints of presheaf operators which remove elements which cannot form part of any global section. In the next section, we look at sheaf-theoretic obstructions to forming a global section which come from cohomology and see how these can be used to define stronger approximations to homomorphism and isomorphism.

4 Cohomology of local homomorphisms and isomorphisms

As we showed in the previous section, an instance of CSP and SI with input (A, B) can be seen as determining the existence of a global section for the presheaf $\mathcal{H}_k(A, B)$ or $\mathcal{I}_k(A, B)$ respectively and that the classic k-consistency and k-Weisfeiler-Leman algorithms can be reformulated as computing greatest fixed points of presheaf operations which successively remove sections which are obstructed from being part of some global section. In this section, we extend these algorithms by considering further efficiently computable obstructions which arise naturally from presheaf cohomology. From this we derive new cohomological algorithms for CSP and SI.

4.1 Cohomology and local vs. global problems

The notion of computing cohomology valued in a **AbGp**-valued presheaf $\mathcal{F}$ on a topological space X has a long history in algebraic geometry and algebraic topology which dates back to Grothendieck's seminal paper on the topic[20]. The cohomology valued in $\mathcal{F}$ consists of a sequence of abelian groups $H^i(X, \mathcal{F})$ where $H^0(X, \mathcal{F})$ is the free $\mathbb{Z}$-module over global sections of $\mathcal{F}$. As seen in the previous section we may be interested in such global sections but their existence may be difficult to determine. This is where the functorial nature of cohomology is extremely useful. Indeed, any short exact sequence of presheaves

$$0 \to \mathcal{F}_L \to \mathcal{F} \to \mathcal{F}_R \to 0$$

lifts to a long exact sequence of cohomology groups

$$0 \to H^0(X, \mathcal{F}_L) \to H^0(X, \mathcal{F}) \to H^0(X, \mathcal{F}_R) \to H^1(X, \mathcal{F}_L) \to \dots$$

This tells us that the global sections of $\mathcal{F}_R$ which are not images of global sections of $\mathcal{F}$ are mapped to non-trivial elements of the group $H^1(X, \mathcal{F}_L)$ by the maps in this sequence. This means that these higher cohomology groups can be seen as a source of obstacles to lifting "local" solutions in $\mathcal{F}_R$ to "global" solutions in $\mathcal{F}$

An important recent example of such an application of cohomology to finite structures can be found in the work of Abramsky et al. [2] in quantum foundations. They show that cohomological obstructions of the type described above can be used to detect contextuality (locally consistent measurements which are globally inconsistent) in quantum systems which were earlier given a presheaf semantics by Abramsky and Brandenburger[3]. In Appendix C, we describe these obstructions in general and show how the presheaves we constructed in the last section admit the same cohomological obstructions. This similarity inspires the definitions and algorithms which follow in the next two sections.

4.2 $\mathbb{Z}$-local sections and $\mathbb{Z}$-extendability

Returning to presheaves of local homomorphisms and isomorphisms let $\mathcal{S}$ be a subpresheaf of $\mathcal{H}_k$. Then we define the presheaf of $\mathbb{Z}$-linear local sections of $\mathcal{S}$ to be the presheaf of formal $\mathbb{Z}$-linear sums of local sections of $\mathcal{S}$. This means that for any $C \in A^{\leq k}$

$$\mathbb{Z}\mathcal{S}(C) := \{ \sum_{s \in \mathcal{S}(C)} \alpha_s s \ \mid \ \alpha_s \in \mathbb{Z}\}$$

This is an abelian presheaf on $A^{\leq k}$ and we call the global sections $\{r_U \in \mathbb{Z}\mathcal{S}(U)\}_{U \in A^{\leq k}}$, $\mathbb{Z}$-linear global sections of $\mathcal{S}$. We say that a local section $s \in \mathcal{S}(C)$ is $\mathbb{Z}$-extendable if there is a $\mathbb{Z}$-linear global section $\{r_U \in \mathbb{Z}\mathcal{S}(U)\}_{U \in A^{\leq k}}$ such that $r_C = s$. We write this condition as $\mathbb{Z}\mathbf{ext}(\mathcal{S}, s)$. As outlined in Appendix C, this condition corresponds to the absence of a cohomological obstruction to $\mathcal{S}$ containing a global section involving s.

Importantly for our purposes, deciding the condition $\mathbb{Z}\mathbf{ext}(\mathcal{S}, s)$ for any $\mathcal{S} \subset \mathcal{H}_k(A, B)$ is computable in polynomial time in the sizes of A and B. This is because the compatibility conditions for a collection $\{r_U \in \mathbb{Z}\mathcal{S}(U)\}_{U \in A^{\leq k}}$ being a global section of $\mathbb{Z}\mathcal{S}$ can be expressed as a system of polynomially many linear equations which by an algorithm of Kannan and Bachem[24] can be solved in polynomial time. This allows us to define the following efficient algorithms for CSP and SI based on removing cohomological obstructions.

4.3 Cohomological algorithms for CSP and SI

We saw in Section 3 that the k-consistency and k-Weisfeiler-Leman algorithms can be recovered as greatest fixpoints of presheaf operators removing local sections which fail the forth and bijective-forth properties respectively. Now that we have from cohomological considerations a new necessary condition $\mathbb{Z}\mathbf{ext}(\mathcal{S}, s)$ for a local section to feature in a global section of $\mathcal{S}$, we can define natural extensions to the k-consistency and k-Weisfeiler-Leman algorithms as follows.

4.3.1 Cohomological k-consistency

To define the cohomological k-consistency algorithm, we first define an operator which removes those local sections which admit a cohomological obstruction. Let $(\cdot)^{\mathbb{Z}\downarrow}$ be the operator which computes for a given presheaf $\mathcal{S} \subset \mathcal{H}_k$ the largest subpresheaf $\mathcal{S}^{\mathbb{Z}\downarrow}$ such that every $s \in \mathcal{S}^{\mathbb{Z}\downarrow}(C)$ satisfies both the forth property $\mathbf{Forth}(\mathcal{S}, s)$ and the $\mathbb{Z}$-extendability property $\mathbb{Z}\mathbf{ext}(\mathcal{S}, s)$. We write $\overline{\mathcal{S}}^{\mathbb{Z}}$ for the greatest fixpoint of this operator starting from $\mathcal{S}$. As both $\mathbf{Forth}(\mathcal{S}, s)$ and $\mathbb{Z}\mathbf{ext}(\mathcal{S}, s)$ are both computable in polynomial time in the size of $\mathcal{S}$ and $\overline{\mathcal{S}}^{\mathbb{Z}}$ has a global section if and only if $\mathcal{S}$ has a global section, this allows us to define the following efficient algorithm for approximating CSP.

▶ **Definition 4.** *The* cohomological k-consistency algorithm *accepts an instance* (A, B) *if the greatest fixpoint* $\overline{\mathcal{H}_k(A, B)}^{\mathbb{Z}}$ *is non-empty and otherwise rejects.*
If (A, B) *is accepted by this algorithm we write* $A \to_k^{\mathbb{Z}} B$ *and say that the instance* (A, B) *is cohomologically* k*-consistent.*

We conclude this section by showing that the relation $\to_k^{\mathbb{Z}}$ is transitive.

▶ **Proposition 5.** *For all* k*, given* A, B *and* C *structures over a common finite signature*

$$A \to_k^{\mathbb{Z}} B \to_k^{\mathbb{Z}} C \implies A \to_k^{\mathbb{Z}} C.$$

Proof. See Appendix D. ◀

4.3.2 Cohomological k-Weisfeiler-Leman

We now define cohomological k-equivalence to generalise k-WL-equivalence in the same way as we did for cohomological k-consistency, by removing local sections which are not $\mathbb{Z}$-extendable. As $\mathbb{Z}$-extendability in $S \subset \mathbf{Isom}_k(A, B)$ is not *a priori* symmetric in A and B we need to check that both s is $\mathbb{Z}$-extendable in S and s^{-1} is $\mathbb{Z}$-extendable in $S^{-1} = \{t^{-1} \mid t \in S\}$. We call this s being $\mathbb{Z}$-*bi-extendable* in S and write it as $\mathbb{Z}\mathbf{bext}(\mathcal{S}, s)$. We incorporate this into a new presheaf operator $(\cdot)^{\mathbb{Z}\#}$ as follows. Given a presheaf $\mathcal{S} \subset \mathcal{I}_k$ let $\mathcal{S}^{\mathbb{Z}\#}$ be the largest subpresheaf of $\mathcal{S}$ such that every $s \in \mathcal{S}^{\mathbb{Z}\#}(C)$ satisfies both the bijective forth property $\mathbf{BijForth}(\mathcal{S}, s)$ and the $\mathbb{Z}$-bi-extendability property $\mathbb{Z}\mathbf{bext}(\mathcal{S}, s)$. We write $\overline{\overline{\mathcal{S}}}^{\mathbb{Z}}$ for the greatest fixpoint of this operator starting from $\mathcal{S}$. As both $\mathbf{BijForth}(\mathcal{S}, s)$ and $\mathbb{Z}\mathbf{bext}(\mathcal{S}, s)$ are computable in polynomial time in the size of $\mathcal{S}$ and $\overline{\overline{\mathcal{S}}}^{\mathbb{Z}}$ has a global section if and only if $\mathcal{S}$ has a global section, this allows us to define the following efficient algorithm for approximating SI.

▶ **Definition 6.** *The* cohomological k-Weisfeiler-Leman *accepts an instance* (A, B) *if the greatest fixpoint* $\overline{\overline{\mathcal{I}_k(A, B)}}^{\mathbb{Z}}$ *is non-empty and otherwise rejects.*
If (A, B) *is accepted by this algorithm we write* $A \equiv_k^{\mathbb{Z}} B$ *and say that the instance* (A, B) *is cohomologically* k*-equivalent.*

Finally, we observe that the existence of a non-empty subpresheaf of $\mathcal{I}_k$ satisfying the **BijForth** and $\mathbb{Z}$**bext** properties also satisfies the conditions for witnessing cohomological k-consistency of the pairs (A, B) and (B, A). Formally we have

▶ **Observation 7.** *For any two structures* A *and* B*,* $A \equiv_k^{\mathbb{Z}} B$ *implies that* $A \rightarrow_k^{\mathbb{Z}} B$ *and* $B \rightarrow_k^{\mathbb{Z}} A$*.*

In Section 5, we will demonstrate the power of these new algorithms by showing that both cohomological k-consistency and cohomological k-Weisfeiler-Leman can solve instances of CSP and SI on which the non-cohomological versions fail. Before doing this, we briefly review some other algorithms for CSP and SI which involve solving systems of linear equations and establish a possible connection to be explored in future work.

4.4 Other algorithms for CSP

While the application of $\mathbb{Z}$-linear equations to extend Weisfeiler-Leman is new to this author, the algorithm introduced here is not the first to use solving systems of linear equations to approximate CSP. Some examples of these include the BLP, BLP+AIP[10] and CLAP[13] algorithms studied in the Promise CSP community. One difference here is that for an instance (A, B) the variables in BLP and AIP are indexed by valid assignments to each variable and to each related tuple instead of being indexed by valid k-local homomorphisms as in the algorithm derived above. This means that directly comparing these algorithms as stated is not straightforward and is beyond the scope of this paper. However, it seems likely that these algorithms can also be expressed in terms of appropriate presheaves. For example, let $\mathbf{C}(\mathbf{A})$ be the category whose objects are the elements of A and the related tuples of A and with maps for each projection from a related tuple to an element, and let the **Set**-valued presheaf $\mathcal{H}_C(A, B)$ an $\mathbf{C}(\mathbf{A})$ map any $a \in A$ to the set of all elements in B and any $\mathbf{a} \in R^A$ to the set of all related tuples R^B. Then, in a similar way to above, we can see that global sections of $\mathcal{H}_C$ are homomorphisms from A to B. In future work, we will compare the fixpoints $\overline{\mathcal{H}_C}$ and $\overline{\mathcal{H}_C}^{\mathbb{Z}}$ with solutions to the BLP and AIP systems of equations and we will explore a possible presheaf representation for CLAP.

5 The (unreasonable) effectiveness of cohomology in CSP and SI

In this section, we prove that the new algorithms arising from this cohomological approach to CSP and SI are substantially more powerful than the k-consistency and k-Weisfeiler-Leman algorithms. In particular, we show that cohomological k-consistency resolves CSP over all domains of arity less than or equal to k which admit a ring representation and that for a fixed small k cohomological k-Weisfeiler-Leman can distinguish structures which differ on a very general form of the CFI property, in particular, showing that cohomological k-Weisfeiler-Leman can distinguish a property which Lichter[26] claims not to be expressible in rank logic.

5.1 Cohomological k-consistency solves all affine CSPs

In this section, we demonstrate the power of the cohomological k-consistency algorithm by proving that it can decide the solvability of systems of equations over finite rings.

To express the main theorem of this section in terms of the finite relational structures on which our algorithm is defined, we first need to fix a notion of ring representation of a relational structure. Let A be a relational structure over signature σ with relations given by $\{R^A\}_{R\in\sigma}$. We say that A has a *ring representation* if we can give the set A a ring structure $(A, +, \cdot, 0, 1)$ such that for every relational symbol $R \in \sigma$ the set $R^A \subset A^m$ is an affine subset of the ring $(A, +, \cdot, 0, 1)$, meaning that there exists $b^R_1, \ldots, b^R_m, a^R \in A$ such that

$$R^A = \{\mathbf{x} \in A^m \mid \sum_{i\in[m]} b^R_i \cdot x_i = a^R\}$$

With this necessary background we state the main theorem of this section.

▶ **Theorem 8.** *For any structure B with a ring representation, there is a k such that the cohomological k-consistency algorithm decides* $\mathbf{CSP}(B)$*.*
Alternatively stated, there exists a k such that for all σ-structures A

$$A \to^{\mathbb{Z}}_k B \iff A \to B$$

Proof. See Appendix E. ◀

This theorem is notable because there are relational structures B with ring representations for which there are families of structures A_k such that $A_k \to_k R$ but $A_k \not\to R$, see for example the examples given by Feder and Vardi [17]. Furthermore, there exist pairs (A_k, B_k) where $A_k \equiv_k B_k$, $B_k \to B$ and $A_k \to_k B$ but $A_k \not\to B$, see for example the work of Atserias, Bulatov and Dawar[6]. As the sequence of relations $\equiv_k$ bounds the expressive power of FPC, this effectively proves the solvability of systems of linear equations over $\mathbb{Z}$, which is central to the cohomological k-consistency algorithm is not expressible in FPC, a result which was until now unknown to the author.

5.2 Cohomological k-Weisfeiler-Leman decides the CFI property

The Cai-Fürer-Immerman construction[12] on ordered finite graphs is a very powerful tool for proving expressiveness lower bounds in descriptive complexity theory. While it was originally used to separate the infinitary k variable logic with counting from `PTIME`, it has since been used in adapted forms to prove bounds on invertible maps equivalence[14], computation on Turing machines with atoms[9] and rank logic[26]. In this section, we show that $\equiv^{\mathbb{Z}}_k$ separates a very general form of this

The version we consider in this paper is parameterised by a prime power q and takes any totally ordered graph $(G, <)$ and any map $g : E(G) \to \mathbb{Z}_q$ to a relational structure $\mathbf{CFI}_q(G, g)$. The construction effectively encodes a system of linear equations over $\mathbb{Z}_q$ based on the edges of G and the "twists" introduced by the labels g. The result is the following well-known fact.

▶ **Fact 9.** *For any prime power, q, ordered graph G, and functions $g, hE(G) \to \mathbb{Z}_q$,*

$$\mathbf{CFI}_q(G,g) \cong \mathbf{CFI}_q(G,h) \iff \sum g = \sum h$$

We say that the structure $\mathbf{CFI}_q(G,g)$ has the *CFI property* if $\sum g = 0$. For more details on this construction we refer to Appendix F or the recent paper of Lichter[26] whose presentation we follow.

We now recall the two major separation results based on this construction. The first is a landmark result of descriptive complexity from the early 1990's.

▶ **Theorem 10** (Cai, Fürer, Immerman[12]). *There is a class of ordered (3-regular) graphs $\mathcal{G} = \{G_n\}_{n\in\mathbb{N}}$ such that in the respective class of CFI structures*

$$\mathcal{K} = \{\mathbf{CFI}_2(G,g) \mid G \in \mathcal{G}\}$$

the CFI property is decidable in polynomial-time but cannot be expressed in FPC.

The second is a recent breakthrough due to Moritz Lichter.

▶ **Theorem 11** (Lichter[26]). *There is a class of ordered graphs $\mathcal{G} = \{G_n\}_{n\in\mathbb{N}}$ such that in the respective class of CFI structures*

$$\mathcal{K} = \{\mathbf{CFI}_{2^k}(G,g) \mid G \in \mathcal{G}\}$$

the CFI property is decideable in polynomial-time (indeed, expressible in choiceless polynomial time) but cannot be expressed in rank logic.

Despite this CFI property proving to be inexpressible in both FPC and rank logic, we show that (perhaps surprisingly) there is a fixed k such that cohomological k-consistency algorithm can separate structures which differ on this property in the following general way. The proof of this theorem relies the on showing that $\equiv^{\mathbb{Z}}_k$ behaves well with logical interpretations and the details are left to Appendix F.

▶ **Theorem 12.** *There is a fixed k such that for any q given $\mathbf{CFI}_q(G,g)$ and $\mathbf{CFI}_q(G,h)$ with $\sum g = 0$ we have*

$$\mathbf{CFI}_q(G,g) \equiv^{\mathbb{Z}}_k \mathbf{CFI}_q(G,h) \iff \sum h = 0$$

Proof. See Appendix F. ◀

As a direct consequence of this result, there is some k such that the set of structures with the CFI property in Lichter's class $\mathcal{K}$ from Theorem 11 is closed under $\equiv^{\mathbb{Z}}_k$. This means that, by the conclusion of Theorem 11, the equivalence relation $\equiv^{\mathbb{Z}}_k$ can distinguish structures which disagree on a property that is not expressible in rank logic. Indeed, Dawar, Grädel and Lichter[15] show further that this property is also inexpressible in linear algebraic logic. By the definition of our algorithm for $\equiv^{\mathbb{Z}}_k$ this implies that solvability of systems of $\mathbb{Z}$-linear equations is not definable in linear algebraic logic.

6 Conclusions & future work

In this paper, we have presented novel approach to CSP and SI in terms of presheaves and have used this to derive efficient generalisations of the k-consistency and k-Weisfeiler-Leman algorithms, based on natural considerations of presheaf cohomology. We have shown that the relations, $\to_k^{\mathbb{Z}}$ and $\equiv_k^{\mathbb{Z}}$, computed by these new algorithms are strict refinements of their well-studied classical counterparts $\to_k$ and $\equiv_k$. In particular, we have shown in Theorem 8 that cohomological k-consistency suffices to solve linear equations over all finite rings and in Theorem 12 that cohomological k-Weisfeiler-Leman distinguishes positive and negative instances of the CFI property on the classes of structures studied by Cai, Fürer and Immerman [12] and more recently by Lichter[26]. These results have important consequences for descriptive complexity theory showing, in particular, that the solvability of systems of linear equations over $\mathbb{Z}$ is not expressible in FPC, rank logic or linear algebraic logic. Furthermore, the results of this paper demonstrate the unexpected effectiveness of a cohomological approach to constraint satisfaction and structure isomorphism, analogous to that pioneered by Abramsky and others for the study of quantum contextuality.

The results of this paper suggest several directions for future work to establish the extent and limits of this cohomological approach. We ask the following questions which connect it to important themes in algorithms, logic and finite model theory.

Cohomology and constraint satisfaction: Firstly, Bulatov and Zhuk's recent independent resolutions of the Feder-Vardi conjecture[11][30], show that for all domains B either $\mathbf{CSP}(B)$ is `NP-Complete` or B admits a weak near-unanimity polymorphism and $\mathbf{CSP}(B)$ is tractable. As the cohomological k-consistency algorithm expands the power of the k-consistency algorithm which features as one case of Bulatov and Zhuk's general efficient algorithms, we ask if it is sufficient to decide all tractable **CSP**s.

▶ **Question 13.** *For all domains B which admit a weak near-unanimity polymorphism, does there exists a k such that for all A*

$$A \to B \iff A \to_k^{\mathbb{Z}} B?$$

Cohomology and structure isomorphism: Secondly, as cohomological k-Weisfeiler-Leman is an efficient algorithm for distinguishing some non-isomorphic relational structures we ask if it distinguishes all non-isomorphic structures. As the best known structure isomorphism algorithm is quasi-polynomial[7], we do not expect a positive answer to this question but expect that negative answers would aid our understanding of the hard cases of structure isomorphism in general.

▶ **Question 14.** *For every signature σ does there exists a k such that for all σ-structures A, B*

$$A \cong B \iff A \equiv_k^{\mathbb{Z}} B?$$

Cohomology and game comonads: Thirdly, as $\to_k$ and $\equiv_k$ have been shown by Abramsky, Dawar, and Wang[4] to be correspond to the coKleisli morphisms and isomorphisms of a comonad $\mathbb{P}_k$, we ask whether a similar account can be given to $\to_k^{\mathbb{Z}}$ and $\equiv_k^{\mathbb{Z}}$. As the coalgebras of the $\mathbb{P}_k$ comonad relate to the combinatorial notion of treewidth, an answer to this question could provide a new notion of "cohomological" treewidth.

▶ **Question 15.** *Does there exist a comonad $\mathbb{C}_k$ for which the notion of morphism and isomorphism in the coKleisli category are $\to_k^{\mathbb{Z}}$ and $\equiv_k^{\mathbb{Z}}$?*

The search for a logic for `PTIME`: Finally, as the algorithms for $\to_k^{\mathbb{Z}}$ and $\equiv_k^{\mathbb{Z}}$ are likely expressible in rank logic extended with a quantifier for solving systems of linear equations over $\mathbb{Z}$ and as $\equiv_k^{\mathbb{Z}}$ distinguishes all the best known family separating rank logic from `PTIME`, we ask if solving systems of equations over $\mathbb{Z}$ is enough to capture all `PTIME` queries.

▶ **Question 16.** *Is there a logic FPC+rk+$\mathbb{Z}$ incorporating solvability of $\mathbb{Z}$-linear equations into rank logic which captures `PTIME`?*

References

1 Samson Abramsky. Notes on cohomological width and presheaf representations. Technical report, University College London, 2022.

2 Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, Cohomology and Paradox. In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*, volume 41 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 211–228, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. URL: http://drops.dagstuhl.de/opus/volltexte/2015/5416, doi:10.4230/LIPIcs.CSL.2015.211.

3 Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, Nov 2011. URL: http://dx.doi.org/10.1088/1367-2630/13/11/113036, doi:10.1088/1367-2630/13/11/113036.

4 Samson Abramsky, Anuj Dawar, and Pengming Wang. The pebbling comonad in finite model theory. In *Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '17. IEEE Press, 2017.

5 Samson Abramsky, Shane Mansfield, and Rui Soares Barbosa. The cohomology of non-locality and contextuality. In B. Jacobs, P. Selinger, and B. Spitters, editors, *Proceedings of 8th International Workshop on Quantum Physics and Logic (QPL 2011)*, volume 95, pages 1–14, 2011. doi:10.4204/EPTCS.95.1.

6 Albert Atserias, Andrei Bulatov, and Anuj Dawar. Affine systems of equations and counting infinitary logic. In *In ICALP'07, volume 4596 of LNCS*, pages 558–570, 2007.

7 László Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 684–697, New York, NY, USA, 2016. Association for Computing Machinery. doi:10.1145/2897518.2897542.

8 L. Barto and M. Kozik. Constraint satisfaction problems of bounded width. pages 595–603, 2009. cited By 99. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-77952395856&doi=10.1109%2fFOCS.2009.32&partnerID=40&md5=8411bfcb009c67c5c9c9765f651f3b45, doi:10.1109/FOCS.2009.32.

9 Mikolaj Bojanczyk, Bartek Klin, Slawomir Lasota, and Szymon Torunczyk. Turing machines with atoms. In *2013 28th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 183–192, 2013. doi:10.1109/LICS.2013.24.

10 Joshua Brakensiek, Venkatesan Guruswami, Marcin Wrochna, and Stanislav Živný. The power of the combined basic linear programming and affine relaxation for promise constraint satisfaction problems. *SIAM Journal on Computing*, 49(6):1232–1248, 2020. arXiv:https://doi.org/10.1137/20M1312745, doi:10.1137/20M1312745.

11 Andrei A. Bulatov. A dichotomy theorem for nonuniform csps. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–330, 2017. doi:10.1109/FOCS.2017.37.

12 Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, December 1992. doi:10.1007/BF01305232.

13 Lorenzo Ciardo and Stanislav Zivný. CLAP: A new algorithm for promise csps. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 1057–1068. SIAM, 2022. doi:10.1137/1.9781611977073.46.

14 Anuj Dawar, Erich Grädel, and Wied Pakusa. Approximations of Isomorphism and Logics with Linear-Algebraic Operators. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 112:1–112:14, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum

fuer Informatik. URL: http://drops.dagstuhl.de/opus/volltexte/2019/10688, doi:10.4230/LIPIcs.ICALP.2019.112.

15 Anuj Dawar, Erich Grädel, and Moritz Lichter. Limitations of the invertible-map equivalences, 2021. arXiv:2109.07218.

16 Anuj Dawar and Bjarki Holm. Pebble games with algebraic rules. In Artur Czumaj, Kurt Mehlhorn, Andrew Pitts, and Roger Wattenhofer, editors, *Automata, Languages, and Programming*, pages 251–262, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

17 Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory. *SIAM Journal on Computing*, 28(1):57–104, 1998. arXiv:https://doi.org/10.1137/S0097539794266766, doi:10.1137/S0097539794266766.

18 Marcelo Fiore and Sam Staton. Comparing operational models of name-passing process calculi. *Information and Computation*, 204(4):524–560, 2006. Seventh Workshop on Coalgebraic Methods in Computer Science 2004. URL: https://www.sciencedirect.com/science/article/pii/S0890540106000058, doi:https://doi.org/10.1016/j.ic.2005.08.004.

19 Jörg Flum and Martin Grohe. On fixed-point logic with counting. *The Journal of Symbolic Logic*, 65(2):777–787, 2000. URL: http://www.jstor.org/stable/2586569.

20 Alexander Grothendieck. Sur quelques points d'algèbre homologique, I. *Tohoku Mathematical Journal*, 9(2):119 – 221, 1957. doi:10.2748/tmj/1178244839.

21 Yuri Gurevich. Logic and the challenge of computer science, 1988.

22 Lauri Hella. Logical hierarchies in ptime. *Information and Computation*, 129(1):1–19, 1996. URL: https://www.sciencedirect.com/science/article/pii/S089054019690070X, doi:https://doi.org/10.1006/inco.1996.0070.

23 Neil Immerman and Eric S. Lander. Describing graphs: A first-order approach to graph canonization. 1990.

24 Ravindran Kannan and Achim Bachem. Polynomial algorithms for computing the smith and hermite normal forms of an integer matrix. *SIAM Journal on Computing*, 8(4):499–9, 11 1979. Copyright - Copyright] © 1979 © Society for Industrial and Applied Mathematics; Last updated - 2021-09-11. URL: https://ezp.lib.cam.ac.uk/login?url=https://www.proquest.com/scholarly-journals/polynomial-algorithms-computing-smith-hermite/docview/918468506/se-2?accountid=9851.

25 P.G. Kolaitis and M.Y. Vardi. On the expressive power of datalog: Tools and a case study. *Journal of Computer and System Sciences*, 51(1):110–134, 1995. URL: https://www.sciencedirect.com/science/article/pii/S0022000085710550, doi:https://doi.org/10.1006/jcss.1995.1055.

26 Moritz Lichter. Separating rank logic from polynomial time. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*, pages 1–13. IEEE, 2021. doi:10.1109/LICS52264.2021.9470598.

27 Cristina Matache, Sean K. Moss, and Sam Staton. Recursion and sequentiality in categories of sheaves. In Naoki Kobayashi, editor, *6th International Conference on Formal Structures for Computation and Deduction, FSCD 2021, July 17-24, 2021, Buenos Aires, Argentina (Virtual Conference)*, volume 195 of *LIPIcs*, pages 25:1–25:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.FSCD.2021.25.

28 Baudot P. and Bennequin D. The homological nature of entropy. *Entropy*, 17(5):3253 – 3318, 2015. Cited by: 27; All Open Access, Gold Open Access, Green Open Access. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84930079211&doi=10.3390%2fe17053253&partnerID=40&md5=07a010ab9781935a2bb2203dd18a971a, doi:10.3390/e17053253.

29 Wied Pakusa. *Linear Equation Systems and the Search for a Logical Characterisation of Polynomial Time*. PhD thesis, RWTH Aachen University, 2016. URL: https://logic.rwth-aachen.de/~pakusa/diss.pdf.

30 Dmitriy Zhuk. A proof of the csp dichotomy conjecture. *J. ACM*, 67(5), aug 2020. doi:10.1145/3402029.

A Proof omitted from Section 3

Proof of Lemma 2. ($\Longrightarrow$) This direction is easy. Suppose that $(A, B) \in CSP$ (resp. $(A, B) \in SI$) then there exists $h : A \to B$ a homomorphism (resp. an isomorphism). Consider the collection of maps $\{h_U\}_{U \in A^{\leq k}}$ defined by $h_U = h_{|_U}$. This forms global section of $\mathcal{H}_k$ (resp. $\mathcal{I}_k$) because firstly $h_U \in \mathcal{H}_k(U)$ (resp. $h_U \in \mathcal{I}_k(U)$) as the restriction of a homomorphism (resp. isomorphism) is a partial homomorphism (resp. isomorphism) and secondly the naturality condition is satisfied as $(h_U)_{|_{U'}} = h_{|_{U'}}$ for any $U' \subset U$.
($\Longleftarrow$) This for this direction we start with a global section $s : \mathbb{I} \implies \mathcal{S}_k$ (for $\mathcal{S}_k = \mathcal{H}_k$ or $\mathcal{I}_k$). In either case, we claim that there is a single function $h : A \to B$ such that $s_U = h_{|_U}$ for all $U \in A^{\leq k}$. Indeed, this is the function h which sends any element $a \in A$ to the element $h(a) := s_{\{a\}}(a) \in B$. This satisfies the required property as for any $U \in A^{\leq k}$ and any $u \in U$, naturality of s along the inclusion $\{u\} \subset U$ ensures that $s_U(u) = s_{\{u\}}(u) = h(u)$ and so $s_U = h_{|_U}$. In the case $\mathcal{S}_k = \mathcal{I}_k$, this map will be injective and so is bijective by the assumption on sizes of A and B. Now we must show that h is an homomorphism, or, in the case of $\mathcal{I}_k$, an isomorphism. Take any related tuple $(a_1, \dots, a_m) \in R^A$ or $(b_1, \dots, b_m) \in R^B$. Let $U = \{a_1, \dots a_m\}$ and $V = \{b_1, \dots, b_m\}$. As k is at least the arity of σ we have that $k \geq m \geq |U|, |V|$. Now, in both cases, $h_{|_U} = s_U \in \mathbf{Hom}_k(A, B)$ is a partial homomorphism. So, $(a_1, \dots, a_m) \in R^A \implies (b_1, \dots, b_m) \in R^B$. Thus h is a homomorphism. In the isomorphism case, as h is bijective $h^{-1}(V)$ is a well-defined member of $A^{\leq k}$ and $h_{|_{h^{-1}(V)}} = s_U \in \mathbf{Isom}_k(A, B)$ is a partial isomorphism. So, $(b_1, \dots, b_m) \in R^B \implies (h^{-1}(b_1), \dots, h^{-1}(b_m)) \in R^A$. Thus h is an isomorphism. ◀

B Algorithms for k-consistency and k-Weisfeiler-Leman

In this appendix, we recall the full definitions of k-consistency and k-Weisfeiler-Leman.

B.1 Classical k-consistency algorithm

We start by recalling some definitions related to the classical k-consistency algorithm on which our algorithm will build.

For A and B finite structures over a common (finite) signature, let $\mathbf{Hom}_k(A, B)$ denote the set of partial homomorphisms from A to B with domain of size less than or equal to k. There is a natural partial order $<$ on this set, defined as follows. For any partial homomorphisms $f, g \in \mathbf{Hom}_k(A, B)$ we say that $f < g$ if $\mathbf{dom}(f) \subset \mathbf{dom}(g)$ and $g_{|_{\mathbf{dom}(f)}} = f$.

We say that any $S \subset \mathbf{Hom}_k(A, B)$ has the *forth* property if for every $f \in S$ with $|\mathbf{dom}(f)| < k$ we have the property $\mathbf{Forth}(S, f)$ which is defined as follows:

$$\forall a \in A,\ \exists b \in B \text{ s.t. } f \cup \{(a, b)\} \in S.$$

Given $S \subset \mathbf{Hom}_k(A, B)$ we define $\overline{S}$ to be the largest subset of S which is downwards-closed and has the forth property. Note that $\emptyset$ satisfies these conditions, so such a set always exists. For a fixed k there is a simple algorithm for computing $\overline{S}$ from S.

This is done by starting with $S_0 = S$ and then entering the following loop with $i = 0$

1. Initialise S_{i+1} as being equal to S_i.
2. For each $s \in S_i$, check if $\mathbf{Forth}(S_i, s)$ holds and if not remove it from S_{i+1} along with all $s' > s$.
3. If none fail this test, halt and output S_i.
4. Otherwise, increment i by one and repeat.

It is easily seen that this runs in polynomial time in $|A||B|$.

Now for a pair of structures A, B we say that the pair (A, B) is k consistent if $\overline{\mathbf{Hom}_k(A, B)} \neq \emptyset$. We denote this by writing $A \rightarrow_k B$ and the algorithm above shows how to decide this relation in polynomial time for fixed k. This relation has many equivalent logical and algorithmic definitions as seen in [17], and [8].

B.2 Classical k-Weisfeiler-Leman algorithm

Immerman and Lander[23] first established that two structures are $\equiv_{k-WL}$-equivalent if and only if they satisfy the same formulas of infinitary k-variable logic with counting quantifiers (written $A \equiv_k B$). Hella[22] showed that this is true if and only if the set of k-local partial isomorphisms $\mathbf{Isom}_k(A, B)$ contains a non-empty subset S which is downward-closed and has the following *bijective forth property* for all $f \in S$ with $|\mathbf{dom}(f)| < k$:

$$\exists b_f : A \to B \text{ a bijection s.t. } \forall a \in A \ f \cup \{(a, b_f(a))\} \in S$$

Whether such a bijection exists can be determined efficiently given A, B, S and f by determining if the bipartite graph with vertices $A \sqcup B$ and edges $\{(a, b) \mid f \cup \{(a, b)\} \in S\}$ has a perfect matching. For $S \subset \mathbf{Isom}_k(A, B)$, let $\overline{\overline{S}}$ be the largest subset of S which is downward-closed and satisfies the bijective forth property. For fixed k this can be computed in polynomial time in the sizes of A and B and so an alternative polynomial time algorithm for determining $\equiv_{k-WL}$ is computing $\overline{\overline{\mathbf{Isom}_k(A, B)}}$ and checking if it is non-empty.

C Cohomological obstructions from quantum contextuality

To understand the cohomological invariants of Abramsky, Barbosa and Mansfield[5] which we need for the main algorithms in this section we first give a brief overview the sheaf-theoretic approach to quantum contextuality introduced by Abramsky and Brandenburger[3] which bears an important resemblance to the set-up in the last section.

A *measurement scenario* is a triple $\mathcal{M} = \langle X, M, O \rangle$ where X and O are finite sets and M is a downward-closed subset of the powerset $P(X)$ which covers X. We interpret such a scenario as a quantum system with a set X of possible measurements, a set M of valid contexts of commuting measurements (which can be done simultaneously) and a set of outcomes O for each measurement. The *sheaf of outcomes* over $\mathcal{M}$ is the presheaf $\mathcal{E} : \mathbf{M}^{op} \to \mathbf{Set}$ defined by $\mathcal{E}(C) = O^C$ with the restriction maps given by normal function restriction. The proof that this is indeed a sheaf is elementary but unimportant for the present work. A *possibilistic empirical model* of $\mathcal{M}$ is any flasque subpresheaf $\mathcal{S}$ of $\mathcal{E}$. For any such model we interpret the set of local sections $\mathcal{S}(C) \subset O^C$ as the set of possible measurement-outcome pairs for the context C. The condition of being flasque is precisely what's required for such a model to satisfy the no-signalling property which is important in quantum mechanical systems. As in the previous section, global sections of these presheaves are important. Indeed Abramsky, Barbosa and Mansfield say that an empirical model $\mathcal{S}$ is *strongly contextual*, written $\mathbf{SC}(\mathcal{S})$ if there is no global section $\{s_C \in \mathcal{S}(C)\}_{C \in M}$ for $\mathcal{S}$. Furthermore, a possible measurement outcome $s \in \mathcal{S}(C')$ is said to be *logically contextual*, written $\mathbf{LC}(\mathcal{S}, s)$ if there is no global section $\{s_C \in \mathcal{S}(C)\}_{C \in M}$ for $\mathcal{S}$ such that $s_{C'} = s$. The whole empirical model $\mathcal{S}$ is said to be *logically contextual*, written $\mathbf{LC}(\mathcal{S})$ if there exists some local section s of $\mathcal{S}$ such that $LC(\mathcal{S}, s)$ holds.

In *The Cohomology of Non-locality and Contextuality*, Abramsky, Barbosa and Mansfield show that contextuality in empirical models, as defined above, can be detected in many cases by considering the cohomology of certain Čech cochain complexes $\check{C}^\bullet(M, \mathcal{F})$ of the cover M valued in abelian presheaves related to $\mathcal{S}$. To do this they first define, for any possibilistic empirical model $\mathcal{S}$, the abelian presheaf $\mathcal{F}_\mathbb{Z} : \mathbf{M}^{op} \to \mathbf{AbGrp}$ which is formed by composing $\mathcal{S}$ with the free $\mathbb{Z}$-module functor $F_\mathbb{Z} : \mathbf{Set} \to \mathbf{AbGrp}$. Local sections $r \in \mathcal{F}_Z(C)$ are simply formal $\mathbb{Z}$-linear combinations of elements of $\mathcal{S}(C)$. For any $U \in M$, they then construct a short exact sequence

$$0 \to \mathcal{F}_{\tilde{U}} \to \mathcal{F}_\mathbb{Z} \to \mathcal{F}_{|U} \to 0$$

which captures the restriction of local sections to the context U. This gives a long exact sequence of cohomology groups. The connection maps in this long exact sequnce allows us to take any $s \in \mathcal{S}(U)$ and send it forward to an element $\delta(s) \in \check{H}^1(M, \mathcal{F}_{\tilde{U}})$. Abramsky et al show that $\delta(s)$ not vanishing is a sufficient condition for $\mathbf{LC}(\mathcal{S}, s)$ and define this condition as $\mathbf{CLC}_\mathbb{Z}(\mathcal{S}, s)$. They also give the following equivalent condition which we use for the rest of the paper. $\mathbf{CLC}_\mathbb{Z}(\mathcal{S}, s)$ holds if and only if there is no global section $\{r_C\}_{C \in M}$ of $\mathcal{F}_\mathbb{Z}$ such that $r_U = s$.

Now we see how this set-up applies equally to the search for global sections in CSP and SI.

C.1 $\mathbb{Z}$-extendability and $\mathbb{Z}$-linear sections

In order to translate the cohomological obstructions from the setting of quantum contextuality to that of constraint satisfaction and structure isomorphism, we first make the following observation.

▶ **Observation 17.** *For any two relational structures A and B and any k, the sheaf of events $\mathcal{E}_\mathcal{M}$ over the measurement scenario $\mathcal{M} = \langle A, A^{\leq k}, B \rangle$ contains both $\mathcal{H}_k(A, B)$ and $\mathcal{I}_k(A, B)$ as subpresheaves.*

Furthermore, as the subpresheaves $\overline{\mathcal{H}_k}$ and $\overline{\overline{\mathcal{I}_k}}$ resulting from the sheaf-theoretic versions of k-consistency and k-Weisfeiler-Leman are flasque, they can be viewed as empirical models for $\mathcal{M}$.

This observation combined with Lemma 2 shows that for k at least as large as the arity of the signature of A and B, strong contextuality of the empirical models $\overline{\mathcal{H}_k}$ and $\overline{\overline{\mathcal{I}_k}}$ is equivalent to the pair (A, B) being rejected by CSP and SI, respectively. Formally this is stated as

▶ **Observation 18.** *For any A and B relational structures and k at least the arity of the largest relation on A then*

$$\mathbf{SC}(\overline{\mathcal{H}_k}(A, B)) \iff A \not\to B$$

and

$$\mathbf{SC}(\overline{\mathcal{I}_k}(A, B)) \iff A \not\cong B$$

Furthermore, the logical contextuality of an individual local section corresponds to the impossibility of extending that section to a full isomorphism or homomorphism.

▶ **Observation 19.** *For any A and B relational structures, $s \in \overline{\mathcal{H}_k}(A, B)(C)$ and $s' \in \overline{\mathcal{I}_k}(A, B)(C)$ then*

$$\mathbf{LC}(\overline{\mathcal{H}_k}(A, B), s) \iff \neg \exists f : A \to B \ s.t. \ f_{|C} = s$$

and

$$\mathbf{LC}(\overline{\mathcal{I}_k}(A,B),s') \iff \neg\exists f : A \to B, \text{ an isomorphism s.t. } f_{|C} = s'$$

As cohomological contextuality gives a sufficient condition for logical contextuality, we now introduce some terminology for cohomological contextuality in subpresheaves $\mathcal{S} \subset \mathcal{H}_k(A,B)$. Firstly, for the abelian presheaf $\mathcal{F} = F_{\mathbb{Z}} \circ \mathcal{S}$, we call any element $r_C \in \mathcal{F}(C)$ a *$\mathbb{Z}$-linear section* of $\mathcal{S}$. Such a $\mathbb{Z}$-linear section can be represented as a formal linear sum

$$r_C = \sum_{s \in \mathcal{S}(C)} \alpha_s s$$

where $\alpha_s \in \mathbb{Z}$ for each $s \in \mathcal{S}(C)$. We say that some $s \in \mathcal{S}(C)$ is $\mathbb{Z}$-extendable in $\mathcal{S}$, write $\mathbb{Z}\mathbf{ext}(\mathcal{S},s)$ if there is a collection $\{r_{C'} \in \mathcal{F}(C')\}_{C' \in M}$ such that $r_C = s$ and for all $C', C'' \in M$ we have

$$(r_{C'})_{|C' \cap C''} = (r_{C''})_{|C' \cap C''}.$$

The following observation is immediate from this definition

▶ **Observation 20.** *For any flasque subpresheaf $\mathcal{S} \subset \mathcal{H}_k(A,B)$ and any $s \in \mathcal{S}(C)$, we have*

$$\mathbb{Z}\mathbf{ext}(\mathcal{S},s) \iff \neg\mathbf{CLC}_{\mathbb{Z}}(\mathcal{S},s)$$

This motivates the definitions of the cohomological algorithms given in the main paper.

D Proofs omitted from Section 4

To aid with the proof of this proposition we observe that the $\mathbb{Z}$-extendability condition subsumes both the forth property and downward closure meaning that we have a slightly simpler condition for the success of the cohomological k-consistency algorithm given as follows.

▶ **Observation 21.** *For any structures A and B $A \to_k^{\mathbb{Z}} B$ if and only if there exists a set $\emptyset \neq S \subset \mathbf{Hom}_k(A,B)$ in which each element $s \in S$ is $\mathbb{Z}$-extendable in S.*

Proof of Proposition 5. Success of the $\to_k^{\mathbb{Z}}$ algorithm for the pairs (A,B) and (B,C) results in two non-empty sets $S^{AB} \subset \mathbf{Hom}_k(A,B)$ and $S^{BC} \subset \mathbf{Hom}_k(B,C)$ in both of which each local section is $\mathbb{Z}$-extendable. By Observation 21, to show that $A \to_k^{\mathbb{Z}} C$, it suffices to show that the set $S^{AC} = \{s \circ t \mid s \in S^{BC}, t \in S^{AB}\}$ has the same property.

To show that every $p_0 = s_0 \circ t_0 \in S^{AC}_{\mathbf{a}_0}$ is $\mathbb{Z}$-extendable in S^{AC} we construct a global $\mathbb{Z}$-linear section extending p_0 from the $\mathbb{Z}$-linear sections $\{r^{t_0}_{\mathbf{a}} := \sum_t z_t t\}_{\mathbf{a} \in A^{\leq k}}$ and $\{r^{s_0}_{\mathbf{b}} := \sum_s w_s s\}_{\mathbf{b} \in B^{\leq k}}$ extending t_0 and s_0 respectively. Define $\{r^{p_0}_{\mathbf{a}}\}_{\mathbf{a} \in A^{\leq k}}$ as

$$r^{p_0}_{\mathbf{a}} = \sum_{t \in S^{AB}_{\mathbf{a}}} \sum_{s \in S^{BC}_{t(\mathbf{a})}} z_t w_s (s \circ t)$$

To show that this is a global $\mathbb{Z}$-linear section extending p_0 we need to show firstly that $r^{p_0}_{\mathbf{a}_0} = p_0$ and secondly that the local sections of r^{p_0} agree on the pairwise intersections of their domains.

To show that $r^{p_0}_{\mathbf{a}_0} = p_0$ we observe that, as r^{t_0} $\mathbb{Z}$-linearly extends t_0, for all $t \in S^{AB}_{\mathbf{a}_0}$ we have

$$z_t = \begin{cases} 1, & \text{for } t = t_0 \\ 0, & \text{otherwise,} \end{cases}$$

and similarly, for all $s \in S^{BC}_{t_0(\mathbf{a}_0)}$

$$w_s = \begin{cases} 1, & \text{for } w = w_0 \\ 0, & \text{otherwise.} \end{cases}$$

From this we have that

$$r^{p_0}_{\mathbf{a}_0} = z_{t_0} w_{s_0} (s_0 \circ t_0) = p_0$$

as required.

Finally, we need to show for any $\mathbf{a}, \mathbf{a}'$ in $A^{\leq k}$ with intersection $\mathbf{a}''$ that

$$r^{p_0}_{\mathbf{a}|_{\mathbf{a}''}} = r^{p_0}_{\mathbf{a}'|_{\mathbf{a}''}}.$$

To do this we show that the left hand side depends only on $\mathbf{a}''$ and not on $\mathbf{a}$. As this argument applies equally to the right hand side, the result follows.

To begin with the left hand side is a dependent sum which loops over $t \in S^{AB}_{\mathbf{a}}$ and $s \in S^{BC}_{t(\mathbf{a})}$ as follows:

$$r^{p_0}_{\mathbf{a}|_{\mathbf{a}''}} = \sum_{t,s} w_s z_t (s \circ t)_{|_{\mathbf{a}''}}$$

To emphasise the dependence on $\mathbf{a}''$ we can group this sum together by pairs t'', s'' with $t'' \in S^{AB}_{\mathbf{a}''}$ and $s'' \in S^{BC}_{t''(\mathbf{a}'')}$. Within each group the the sum loops over $t \in S^{AB}_{\mathbf{a}}$ such that $t_{|_{\mathbf{a}''}} = t''$ and $s \in S^{AB}_{t(\mathbf{a})}$ such that $s_{|_{\mathbf{a}''}} = s''$. We write this as

$$\sum_{t'', s''} \sum_{t_{|_{\mathbf{a}''}} = t''} z_t \sum_{s_{|_{t''(\mathbf{a}'')}} = s''} w_s (s \circ t)_{|_{\mathbf{a}''}}$$

We now show that for each t'', s'' the corresponding part of the sum depends only on t'' and s''. This follows from three observations.

The first observation is that in the sum

$$\sum_{t_{|_{\mathbf{a}''}} = t''} z_t \sum_{s_{|_{t''(\mathbf{a}'')}} = s''} w_s (s \circ t)_{|_{\mathbf{a}''}}$$

the formal variables $(s \circ t)_{|_{\mathbf{a}''}}$ are, by definition, all equal to the variable $(s'' \circ t'')$. Thus we need only consider the coefficients, given by the sum

$$\sum_{t_{|_{\mathbf{a}''}} = t''} z_t \sum_{s_{|_{t''(\mathbf{a}'')}} = s''} w_s$$

The second observation is that for each t such that $t_{|_{\mathbf{a}''}} = t''$ the sum

$$\sum_{s_{|_{t''(\mathbf{a}'')}} = s''} w_s$$

is simply the s'' component of $(r^{s_0}_{t(\mathbf{a})})_{|_{t''(\mathbf{a}'')}}$. As r^{s_0} is a global $\mathbb{Z}$-linear section this is equal to the fixed parameter $w_{s''}$. So the sum in question reduces to

$$w_{s''} \cdot \left(\sum_{t_{|_{\mathbf{a}''}} = t''} z_t \right)$$

The final observation, is that the remaining sum is the t'' component of $(r^{t_0}_{\mathbf{a}})_{|_{\mathbf{a}''}}$ which, as r^{t_0} is a global $\mathbb{Z}$-linear section, is equal to $r^{t_0}_{t''}$. This gives the final form of the expression for $(r^{p_0}_{\mathbf{a}})_{|_{\mathbf{a}''}}$ as

$$\sum_{t'',s''} z_{t''} w_{s''} (t'' \circ s'')$$

It is easy to see that the same arguments apply to $r^{p_0}_{\mathbf{a}'}$ and so

$$(r^{p_0}_{\mathbf{a}})_{|_{\mathbf{a}''}} = (r^{p_0}_{\mathbf{a}'})_{|_{\mathbf{a}''}}$$

as required. ◀

E Proof of Theorem 8

To prove this theorem we invoke a result from [2] which considers a similar set-up to that seen in the previous sections and proves a result relating the non-existence of solutions to a system of linear equations over a ring R to the non-triviality of a family of cohomological "obstructions". We will recall their set-up, the relevant result and a characterisation of these cohomological "obstructions" in terms of global $\mathbb{Z}$-linear sections before proving Theorem 8.

E.1 Result from *Contextuality, cohomology & paradox*

In order to state the relevant theorem, we start with some preliminary definitions. Let a *ring-valued measurement scenario* be a triple $\langle X, \mathcal{M}, R\rangle$ where X is a finite set, $\mathcal{M}$ is a downward closed cover of X and R is a ring. An *R-linear equation* on $\langle X, \mathcal{M}, R\rangle$ is a triple $\phi = (V_\phi, a, b)$ where $V_\phi \in \mathcal{M}$, $a : V_\phi \to R$ and $b \in R$. Then for any $s \in R^{V_\phi}$ we say that $s \models \phi$ if

$$\sum_{m \in V_\phi} a(m)s(m) = b$$

in the ring R.

An *empirical model* S on $\langle X, \mathcal{M}, R\rangle$ is a collection of sets $\{S_C\}_{C\in\mathcal{M}}$ where for each C, $S_C \subset R^C$ satisfying the following compatibility condition for all $C, C' \in \mathcal{M}$

$$\{s_{|_{C\cap C'}} \mid s \in S_C\} = \{s'_{|_{C\cap C'}} \mid s' \in S_{C'}\}$$

We make the following observation linking relational structures over signatures $\sigma \subset \sigma_R$ and empirical models which will be useful later.

▶ **Observation 22.** *For any* $\mathbf{CSP}(A, R)$ *and* $S \subset \mathbf{Hom}_k(A, R)$ *which is non-empty, and downward-closed and satisfies the forth property then the local sections of* S *form an empirical model for the measurement scenario* $\langle A, A^{\leq k}, R\rangle$.

For an empirical model S on an R-valued measurement scenario, the *R-linear theory* of S is the set of R-linear equations

$$\mathbb{T}_R(S) = \{\phi \mid \forall s \in S_{V_\phi},\ s \models \phi\}$$

If $\mathbb{T}_R(S)$ is inconsistent (i.e. there is no R-assignment to all the variables in X simultaneously satisfying each of the R-linear equations in the theory), then the empirical model S is said to be "all-vs-nothing for R", written $\mathbf{AvN}_R(S)$.

We can now state the following results that we need for Theorem 8. The first result shows an important implication about the cohomological obstructions in an empirical model which has an inconsistent R-linear theory.

▶ **Theorem 23** (Abramsky, Barbosa, Kishida, Lal, Mansfield [2]). *For any ring R and any R-valued measurement scenario $\langle X, \mathcal{M}, R\rangle$ and any empirical model S we have that*

$$\mathbf{AvN}_R(S) \implies \mathbf{CSC}_{\mathbb{Z}}(S)$$

where $\mathbf{CSC}_{\mathbb{Z}}(S)$ means that for every local section s in S the "cohomological obstruction" of Abramsky, Barbosa and Mansfield $\gamma(s)$ is non-zero.

Next we have a result due to Abramsky, Barbosa and Mansfield which establishes this useful equivalent condition for $\mathbf{CSC}_{\mathbb{Z}}(S)$

▶ **Theorem 24** (Abramsky, Barbosa, Mansfield [5]). *For any empirical model S, $\mathbf{CSC}_{\mathbb{Z}}(S)$ if and only if for every $s \in S_C$ there is no collection $\{r_{C'} \in \mathbb{Z}S_{C'}\}_{C' \in \mathcal{M}}$ such that $r_C = s$ and for all $C_1, C_2 \in \mathcal{M}$*

$$r_{C_1|_{C_1 \cap C_2}} = r_{C_2|_{C_1 \cap C_2}}$$

This condition is precisely what inspired the cohomological k-consistency algorithm and in the next section we show how these two results imply Theorem 8.

E.2 Proof of Theorem 8

We now prove the following equivalent formulation of Theorem 8 which replaces a structure with a ring representation with the underlying ring R "represented as a relational structure". This means simply that the relational symbols (which are affine subsets of R under the ring R) are labelled as $E^m_{\mathbf{a},b}$ for each $\mathbf{a}$ an m-tuple of elements of the ring R and b an element of R such that $(E^m_{\mathbf{a},b})^R = \{(r_1, \dots, r_m) \mid \sum_i a_i \cdot r_i = b\}$.

▶ **Theorem 25.** *For any finite ring R represented as a relational structure over a finite signature σ, there is a k such that the cohomological k-consistency algorithm decides $\mathbf{CSP}(R)$. Alternatively, there exists a k such that for all σ-structures A*

$$A \to^{\mathbb{Z}}_k R \iff A \to R$$

Proof. The direction $A \to R \implies A \to^{\mathbb{Z}}_k R$ is easy and is true for all signatures σ and all $k \leq |A|$. Indeed note that to any homomorphism $f : A \to R$ we can associate the set $S_f = \{f_{|\mathbf{a}}\}_{\mathbf{a} \in A^{\leq k}} \subset \mathbf{Hom}_k(A, R)$. It is not hard to see that S_f is downward closed, has the forth property and that S_f is itself a global section witnessing the $\mathbb{Z}$-extendability of each $f_{|\mathbf{a}} \in S_f$. By Observation 21, this implies that $A \to^{\mathbb{Z}}_k R$.

This leaves the more challenging direction, that there exists a k such that $A \not\to R \implies A \not\to^{\mathbb{Z}}_k R$ for all A. Suppose that the maximum arity of a relation in σ is n. Then as R is a relational model of a finite ring we know that each relation on R is of the form $E^m_{\mathbf{a},b} = \{(r_1, \dots, r_m) \mid \sum_i a_i \cdot r_i = b\}$ where $\mathbf{a}$ is an m-tuple of elements of the ring R and b is an element of R. We show that $k = n$ will suffice to identify all unsatisfiable instances A.

For R and σ as above any instance $\mathbf{CSP}(A, R)$ is specified by a set A of variables where each related tuple $(x_1, \dots, x_m) \in (E^m_{\mathbf{a},b})^A$ specifies an R-linear equation $\sum_i a_i \cdot x_i = b$. Call the collection of such equations $\mathbb{T}^A$. The fact that there is no homomorphism $A \to R$ is exactly the statement that $\mathbb{T}^A$ is unsatisfiable. Taking $k = n$, we have that the R-linear theory $\mathbb{T}_R(\mathbf{Hom}_k(A, R))$ (as defined in the previous section) contains $\mathbb{T}^A$ and so is unsatisfiable. We now show how this is sufficient to prove the theorem.

Consider running the cohomological k-consistency algorithm on the pair (A, R) we get $S_0 = \overline{\mathbf{Hom}_k(A, R)}$. If $S_0 = \emptyset$ we are done. Otherwise, by Observation 22, S_0 can be considered as an empirical model on the measurement scenario $\langle A, A^{\leq k}, R\rangle$. Furthermore, as $S_0 \subset \mathbf{Hom}_k(A,R)$, we have that $\mathbb{T}_R(S_0) \supset \mathbb{T}_R(\mathbf{Hom}_k(A,R))$. This means in particular that $\mathbb{T}_R(S_0)$ is unsatisfiable by the assumption that $A \not\to R$. By Theorems 23 and 24, this means that no local section s of S_0 is $\mathbb{Z}$-extendable in S_0, so $S_1 = \emptyset$. So the cohomological k-consistency algorithm rejects (A,R) and $A \not\to_k^{\mathbb{Z}} R$, as required. ◀

It is notable that in the proof of this theorem, we see that the cohomological k-consistency algorithm decides unsatisfiability of these systems of equations after just one iteration of its loop. A future version of this work will investigate whether multiple iterations are required in over different CSP domains. For now, we retain the iterative nature of the algorithm to guarantee the conclusion in Observation 21.

F The strength of cohomological k-Weisfeiler-Leman

In this appendix, we demonstrate the power of $\equiv_k^{\mathbb{Z}}$ to distinguish structures which disagree on the CFI property, proving Theorem 12. To do this we give an equivalent definition of the cohomological k-consistency algorithm and prove that this behaves well with appropriate logical interpretations.

F.1 Cohomological k-Weisfeiler-Leman Equivalence

The following is an alternative way of computing the $\equiv_k^{\mathbb{Z}}$ relation defined in the main article. Begin by computing $S_0 = \overline{\overline{\mathbf{Isom}_k(A,B)}}$ as in the k-WL equivalence algorithm. If $S_0 = \emptyset$, then reject the pair (A,B) and halt. Otherwise we enter the following loop with $i = 0$:

1. Compute $S_i^{\mathbb{Z}} = \{s \in S_i \mid s \text{ is } \mathbb{Z}\text{-bi-extendable in } S_i\}$
2. Compute $S_{i+1} = \overline{\overline{S_i^{\mathbb{Z}}}}$
3. If $S_{i+1} = \emptyset$, then reject (A,B) and halt
4. If $S_{i+1} = S_i$ then accept (A,B) and halt.
5. Return to Step 1 with $i = i+1$.

If this algorithm accepts a pair (A,B) we say that A and B are cohomologically k-equivalent and we write $A \equiv_k^{\mathbb{Z}} B$.

We now record some simple facts about this equivalence. Firstly, by definition, this generalises k-equivalence and so (k)-WL equivalence, i.e.

$$A \equiv_k^{\mathbb{Z}} B \implies A \equiv_k B \iff A \equiv_{(k-1)-WL} B$$

Secondly, this algorithm determines a maximal set $S \subset \mathbf{Isom}_k(A,B)$ which is downward-closed, has the bijective forth property and for which each $f \in S$ is $\mathbb{Z}$-extendable in S and f^{-1} is $\mathbb{Z}$-extendable in S^{-1}. However, analogously to Observation 21, we note that the existence of any non-empty S satisfying these properties is a witness of $\equiv_k^{\mathbb{Z}}$.

▶ **Observation 26.** *For any two structures A and B, $A \equiv_k^{\mathbb{Z}} B$ if and only if there exists a subset $S \subset \mathbf{Isom}_k(A,B)$ such that both S and S^{-1} are downward-closed, has the bijective forth property and have $\mathbb{Z}$-extendability for each of their elements.*

Finally, we observe that such a set also satisfies the conditions for witnessing cohomological k-consistency of $\mathbf{CSP}(A,B)$ and $\mathbf{CSP}(B,A)$. Formally we have

▶ **Observation 27.** *For any two structures A and B, $A \equiv_k^{\mathbb{Z}} B$ implies that $A \to_k^{\mathbb{Z}} B$ and $B \to_k^{\mathbb{Z}} A$.*

In the next section we establish how this equivalence relation behaves with respect to logical interpretations.

F.2 $\equiv_k^{\mathbb{Z}}$ and interpretations

There are many different notions of logical interpretation in finite model theory. The one we consider is defined as follows. A $\mathcal{C}^l$*-interpretation* Φ (of order n) of signature τ in signature σ is a tuple of $\mathcal{C}^l[\sigma]$ formulas $\langle \phi_R \rangle_{R\in\tau}$. For each relation symbol $R \in \tau$ of arity r, the formula ϕ_R has nr free variables and is written as $\phi_R(\mathbf{x}_1, \dots, \mathbf{x}_r)$, where the $\mathbf{x}_i$ are n-tuples of variables. Such an interpretation defines a map from σ-structures to τ-structures as follows. For any A, $\Phi(A)$ has universe A^n and for each relational symbol $R \in \tau$, the set of related tuples is given by

$$R^{\Phi(A)} := \{(\mathbf{a}_1, \dots, \mathbf{a}_r) \in (A^n)^r \ \mid \ A, \mathbf{a}_1, \dots, \mathbf{a}_r \models \phi_R\}$$

In the next result, we show that the equivalence $\equiv_k^{\mathbb{Z}}$ is preserved by C^l-interpretations in the following way.

▶ **Proposition 28.** *For any (finite, relational) signatures σ and τ, σ-structures A and B, natural numbers n and k, and any order n $\mathcal{C}^{nk}$-interpretation Φ of τ in σ we have that*

$$A \equiv_{nk}^{\mathbb{Z}} B \implies \Phi(A) \equiv_k^{\mathbb{Z}} \Phi(B)$$

Proof. By Observation 26, it suffices to show that there is a set $S' \subset \mathbf{Isom}_k(\Phi(A), \Phi(B))$ which is downward-closed, satisfies the bijective forth property and in which every map is $\mathbb{Z}$-extendable. As $A \equiv_{nk}^{\mathbb{Z}} B$, there is already a set $S \subset \mathbf{Isom}_{nk}(A, B)$ satisfying these properties. For any $Q \subset A$ we use S_Q to mean the elements of S with domain Q. We now show how to construct a suitable S' from S.

For any $C \subset \Phi(A)$, let $\pi(C)$ be the set of element in A which appear in some tuple of C. As elements of $\Phi(A)$ are n-tuples over A, it is clear that $|\pi(C)| \leq n|C|$. We can now define S'_C as the set of partial isomorphisms in $S_{\pi(C)}$ applied coordinatewise to C, namely,

$$\{(f, \dots, f)_{|C} \ \mid \ f \in S_{\pi(C)}\}$$

This is well defined for all $C \in (\Phi(A))^{\leq k}$ as $|\pi(C)| \leq nk$. That these maps define partial isomorphisms between $\Phi(A)$ and $\Phi(B)$ follows from Hella's Lemma 5.1 in [22] which states that the elements of $\mathbf{Isom}_{nk}(A, B)$ are exactly those which preserve and reflect C^{nk} formulas. As the relations on $\Phi(A)$ and $\Phi(B)$ are defined by C^{nk} formulas they are preserved and reflected by the members of S. We now show that $S' = \bigcup_{C\in\Phi(A)^{\leq k}} S'_C$ satisfies the required properties.

Downward-closure This follows easily from downward-closure of S. Suppose $\mathbf{f} = (f, \dots, f)_{|C} \in S'$ and $\mathbf{g} \leq \mathbf{f}$. Then there is some $C' \subset C$ such that $\mathbf{g} = \mathbf{f}_{|C'}$ and $\mathbf{g} = (f_{|\pi(C')}, \dots, f_{|\pi(C')})_{|C'}$. but $f_{|\pi(C')} \leq f$ and so is an element of S.

Bijective forth property Let $\mathbf{f} \in S'_C$ with $|C| < k$, with $\mathbf{f}$ given as the coordinatewise application of some $f \in S_{\pi(C)}$. To show that S' has the bijective forth property we must show that there is a bijection $b : \Phi(A) \to \Phi(B)$ such that for any $\mathbf{a} \in \Phi(A)$ the function

$\mathbf{f} \cup \{(\mathbf{a}, b(\mathbf{a}))\}$ is in $S'_{C\cup\{\mathbf{a}\}}$. For any such $\mathbf{f}$, we can construct a bijection b whose image on any $\mathbf{a} \in \Phi(A)$ is given as

$$b(\mathbf{a}) = (b^{\epsilon}(a_1), b^{\mathbf{a}_1}(a_2), \ldots, b^{(\mathbf{a}_{n-1})}(a_n))$$

where $\mathbf{a}_i$ is the i-tuple of the first i elements in $\mathbf{a}$ and each $b^{\mathbf{a}_i}$ is a bijection $A \to B$. For any $\mathbf{a} \in \Phi(A)$ we choose the bijections $b^{\mathbf{a}_i}$ using the bijective forth property on S. As $\mathbf{f}$ is a coordinatewise application of some $f \in S_{\pi(C)}$ and as $|C| < k$ implies $|\pi(C)| \leq nk - n < nk$, the bijective forth property for S implies the existence of a b_1 such that $f_1 = f \cup \{a_1, b_1(a_1)\} \in S_{\pi(C)\cup\{a_1\}}$. Let $b^{\epsilon} := b_1$. Now suppose for any $i < n$ we have defined the bijections $b^{\epsilon}, b^{\mathbf{a}_1}, \ldots, b^{\mathbf{a}_i}$ and $f_i = f \cup \{(a_j, b^{\mathbf{a}_{j-1}}(a_j))\}_{1\leq j\leq i} \in S_{\pi(C)\cup\{a_1,\ldots,a_i\}}$. We still have $|\pi(C) \cup \{a_1, \ldots, a_i\}| < nk$ so can use the bijective forth property on S again to find a bijection $b^{\mathbf{a}_i}$ such that $f_{i+1} = f_i \cup \{(a_i, b_{\mathbf{a}_i}(a_i))\} \in S_{\pi(C)\cup\{a_1,\ldots,a_{i+1}\}}$. This inductive procedure defines all the required bijections and furthermore shows that $\mathbf{f} \cup \{(\mathbf{a}, b(\mathbf{a})\}$ is the coordinatewise application of some $f_n \in S_{\pi(C\cup\{\mathbf{a}\})}$. This means in particular that $\mathbf{f} \cup \{(\mathbf{a}, b(\mathbf{a})\}$ is in $S'_{C\cup\{\mathbf{a}\}}$, as required.

$\mathbb{Z}$-extendability Our choice of S' makes $\mathbb{Z}$-extendability rather easy. Indeed, we see that any $\mathbf{f} = (f, \ldots, f) \in S'_C$ is $\mathbb{Z}$-extendable because the $\mathbb{Z}$-linear global section extending $f \in S_{\pi(C)}$ given as $s_C = \sum_{g\in S_C} \alpha_g g$ can be lifted to a $\mathbb{Z}$-linear extension of $\mathbf{f}$ by defining $s'_C = \sum_{g\in S_{\pi(C)}} \alpha_g(g, \ldots, g)$. The properties of being a $\mathbb{Z}$-linear extension follow from those properties on S.

◀

F.3 Deciding the CFI property

Cai, Fürer and Immerman[12] showed that there is a property of relational structures which can be decided in polynomial time but which cannot be expressed in infinitary first-order logic with counting quantifiers for any number of variables. This construction essentially encodes certain systems of linear equations (over $\mathbb{Z}_2$) on top of graphs in such a way that isomorphism of the constructed structures is determined by checking solvability of the systems of equations. In their seminal paper[12], Cai, Fürer and Immerman show that the solvable and unsolvable versions of their construction cannot be distinguished in fixed point logic with counting. Adaptations of this construction, encoding equations over different finite fields were used by Dawar, Grädel and Pakusa to show that adding rank quantifiers over each finite field added distinct expressive power to FPC and a version using equations over the rings $\mathbb{Z}_{2^q}$ was used by Lichter[26] to separate rank logic from `PTIME`.

As cohomological k-consistency was shown in the previous section to simultaneously decide solvability over any finite ring, it is natural to ask whether the related equivalence $\equiv^{\mathbb{Z}}_k$ can decide these CFI properties which are not definable in FPC, rank logic or linear algebraic logic. We show in this section that it can.

Following Lichter[26], we define the general CFI construction $\mathbf{CFI}_q(G, g)$ for q a prime power, $G = (G, <)$ an ordered undirected graph and g a function from the edge set of G to $\mathbb{Z}_q$. The idea is that the construction encodes a system of linear equations over $\mathbb{Z}_q$ into G while the function g "twists" these equations in a certain way. For CFI structures, $\mathbf{CFI}_q(G, g)$ the property $\sum g = 0$ is sometimes called the *CFI property*. The following well-known fact (see [29], for example) shows that this property is closed under isomorphisms and is useful in our later arguments.

▶ **Fact 29.** *For any prime power, q, ordered graph G, and functions g, h from the edges of G to $\mathbb{Z}_q$,*

$$\mathbf{CFI}_q(G, g) \cong \mathbf{CFI}_q(G, h) \iff \sum g = \sum h$$

$\mathbf{CFI}_q(G, g)$ is built in three steps. First, we define a gadget which replaces each vertex of x with elements that form a ring. Secondly, we define relations between gadgets which impose consistency equations between gadgets. Finally, the function g is used to insert the important twists into the consistency equations. We now describe this in detail below, following a presentation by Lichter[26].

Vertex gadgets For any vertex $x \in G$, let $N(x)$ be the neighbourhood of x in G (i.e. those vertices which share edges with x) and let $\mathbb{Z}_q^{N(x)}$ denote the ring of functions from $N(x)$ to the ring $\mathbb{Z}_q$. We will replace each vertex x of the base graph with a gadget whose vertices are the following subset of $\mathbb{Z}_q^{N(x)}$,

$$A_x = \{\mathbf{a} \in \mathbb{Z}_q^{N(x)} \ \mid \ \sum_{y \in N(x)} \mathbf{a}(y) = 0\}$$

The relations on the gadget are for each y in $N(x)$ a symmetric relation

$$I_{x,y} = \{(\mathbf{a}, \mathbf{b}) \ \mid \ \mathbf{a}(y) = \mathbf{b}(y)\}$$

and a directed cycle encoded by the relation

$$C_{x,y} = \{(\mathbf{a}, \mathbf{b}) \ \mid \ \mathbf{a}(y) = \mathbf{b}(y) + 1\}$$

Together these impose the ring structure of $\mathbb{Z}_q^{N(x)}$ onto the vertices of the gadget.

Edge equations Next define a relation between gadgets for each edge $\{x, y\}$ in G and each constant $c \in \mathbb{Z}_q$ of the form

$$E_{\{x,y\},c} = \{(\mathbf{a}, \mathbf{b}) \ \mid \ \mathbf{a} \in A_x, \ \mathbf{b} \in A_y, \ \mathbf{a}(y) + \mathbf{b}(x) = c\}$$

Putting it together with a twist We finally define the structure $\mathbf{CFI}_q(G, g)$ as $\langle A, \prec, R_I, R_C, R_{E,0}, R_{E,1}, \ldots, R_{E,q-1} \rangle$ where the universe is $A = \cup_x A_x$ where $\prec$ is the linear pre-order

$$\prec = \bigcup_{x<y} A_x \times A_y$$

and the edge equations $R_{E,c}$ are interpreted according to the twists in g as

$$R_{E,c} = \bigcup_{e \in E} E_{e,c+g(e)}$$

where the sum in the subscript is over $\mathbb{Z}_q$ For the relations R_I and R_C we deviate slightly from Lichter's construction and interpret these as ternary relations of the following form

$$R_I = \bigcup_{\{x,y\} \in E} I_{x,y} \times A_y$$

$$R_C = \bigcup_{\{x,y\} \in E} C_{x,y} \times A_y$$

We now use recall the two major separation results based on this construction. The first is a landmark result of descriptive complexity from the early 1990's.

▶ **Theorem 30** (Cai, Furer, Immerman[12]). *There is a class of ordered (3-regular) graphs $\mathcal{G} = \{G_n\}_{n \in \mathbb{N}}$ such that in the respective class of CFI structures*

$$\mathcal{K} = \{\mathbf{CFI}_2(G, g) \ \mid \ G \in \mathcal{G}\}$$

the CFI property is decidable in polynomial-time but cannot be expressed in FPC.

The second is a recent breakthrough due to Moritz Lichter.

▶ **Theorem 31** (Lichter[26]). *There is a class of ordered graphs $\mathcal{G} = \{G_n\}_{n\in\mathbb{N}}$ such that in the respective class of CFI structures*

$$\mathcal{K} = \{\mathbf{CFI}_{2^k}(G, g) \mid G \in \mathcal{G}\}$$

the CFI property is decidable in polynomial-time (indeed, expressible in choiceless polynomial time) but cannot be expressed in rank logic.

We now show that in both of these classes there exists a fixed k such that $\equiv^{\mathbb{Z}}_k$ distinguishes structures which differ on the CFI property. This relies on two lemmas. The first shows that this property is equivalent to the solvability of a certain system of equations over $\mathbb{Z}_q$, while the second shows that this system of equations can be interpreted in on the classes above with a uniform bound on the number of variables per equation.

The first lemma is an adaptation of Lemma 4.36 from Wied Pakusa's PhD thesis[29]. We begin by defining for any $\mathbf{CFI}q(G, g)$ a system of linear equations over $\mathbb{Z}_q$. This system, $\mathbf{Eq}_q(G, g)$, is the following collection of equations:

- $X_{\mathbf{a},u}$ for all $u \in G$ and all $\mathbf{a} \in A_u \subset \mathbf{CFI}_q(G, g)$,
- $I_{\mathbf{a},\mathbf{b},v}$ for all $u \in G$ and $\mathbf{a}, \mathbf{b} \in A_u$ such that there exists $v \in N(u)$ and $\mathbf{c} \in A_v$ such that $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in R_I$,
- $C_{\mathbf{a},\mathbf{b},v}$ for all $u \in G$ and $\mathbf{a}, \mathbf{b} \in A_u$ such that there exists $v \in N(u)$ and $\mathbf{c} \in A_v$ $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in R_C$, and
- $E_{\mathbf{a},\mathbf{b},c}$ for all $\mathbf{a} \in A_u, \mathbf{b} \in A_v$ and $(\mathbf{a}, \mathbf{b}) \in R_{E,c}$

where the variables are $w_{\mathbf{a},v}$ for every $u \in G$, $\mathbf{a} \in A_u$ and $v \in N(u)$ and the equations are given as:

$$X_{\mathbf{a},u} : \quad \sum_{v\in N(u)} w_{\mathbf{a},v} = 0$$
$$I_{\mathbf{a},\mathbf{b},v} : \quad w_{\mathbf{a},v} - w_{\mathbf{b},v} = 0$$
$$C_{\mathbf{a},\mathbf{b},v} : \quad w_{\mathbf{a},v} - w_{\mathbf{b},v} = 1$$
$$E_{\mathbf{a},\mathbf{b},c} : \quad w_{\mathbf{a},v} + w_{\mathbf{b},u} = c$$

Then we have the following lemma.

▶ **Lemma 32.** *$\mathbf{CFI}_q(G, g)$ a CFI structure, has $\sum g = 0$ if and only if $\mathbf{Eq}_q(G, g)$ is solvable in $\mathbb{Z}_q$*

Proof. Firstly we recall Fact 9 that $\sum g = 0$ if and only if there is an isomorphism $f : \mathbf{CFI}_q(G, g) \to \mathbf{CFI}_q(G, \mathbf{0})$, where $\mathbf{0}$ is the constant 0 function. We now show that there is such an isomorphism if and only if there is a solution to $\mathbf{Eq}_q(G, g)$.

For the forward direction, suppose that we have an isomorphism $f : \mathbf{CFI}_q(G, g) \to \mathbf{CFI}_q(G, \mathbf{0})$. Now as f is a bijection and preserves the pre-order $\prec$, we have that for any $u \in G$, f maps A_u to A_u. This means that for any $\mathbf{a} \in A_u$ $f(\mathbf{a})$ is a function in $\mathbb{Z}_q^{N(u)}$. This means that the assignment $w_{\mathbf{a},v} \mapsto f(\mathbf{a})(v)$ is well-defined for all the variables in $\mathbf{Eq}_q(G, g)$. We now show that this assigment satisfies the system of equations. The X equations in $\mathbf{Eq}_q(G, g)$ become the statement that for all $u \in G$ and $\mathbf{a} \in A_u$

$$\sum_{v\in N(u)} f(\mathbf{a})(v) = 0$$

which follows directly from the fact that $f(\mathbf{a}) \in A_u$. For the I and C equations, we note that as f preserves all relations from $\mathbf{CFI}_q(G, g)$. So for any $\mathbf{a}, \mathbf{b} \in A_u$ and $\mathbf{c} \in A_v$ such that $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is related by R_I or R_C in $\mathbf{CFI}_q(G, g)$ then $(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c}))$ is similarly related in $\mathbf{CFI}_q(G, \mathbf{0})$. The definitions of these relations imply that $f(\mathbf{a})(v) - f(\mathbf{b})(v)$ is 0 or 1 respectively, which implies that our assignment to the variables $w_{\mathbf{a},v}$ and $w_{\mathbf{b},v}$ satisfies the relevant I or C equation. A similar argument applies to the E equations except that the conclusion from $(f(\mathbf{a}), f(\mathbf{b})) \in R_{E,c}$ in $\mathbf{CFI}_q(G, \mathbf{0})$ that the relevant E equation is satisfied follows from the fact that there is no twisting of the $R_{E,c}$ relation in $\mathbf{CFI}_q(G, \mathbf{0})$.

The reverse direction is the observation that any satisfying assignment to the variables $w_{\mathbf{a},v}$ in $\mathbf{Eq}_q(G, g)$ defines an isomorphism from $\mathbf{CFI}_q(G, g)$ to $\mathbf{CFI}_q(G, \mathbf{0})$ where $f(\mathbf{a})(v) = w_{\mathbf{a},v}$. Satisfying the X equation guarantees that for $\mathbf{a} \in A_u$ its image $f(\mathbf{a})$ is also in A_u. Satisfying the I and C equations ensures that the R_I and R_C relations are preserved. So, the additive structure of $\mathbb{Z}_q^{N(u)}$ is preserved in A_u and thus f is bijective. Finally the E equations define the $R_{E,c}$ relation in $\mathbf{CFI}_q(G, \mathbf{0})$ and so satisfying these ensures that f preserves the $R_{E,c}$ relation. ◀

It is not hard to see that the system $\mathbf{Eq}_q(G, g)$ is first order interpretable in $\mathbf{CFI}_q(G, g)$. However, Theorem 8 shows that cohomological k-consistency decides satisfiability of systems of equations over any ring in with up to k variables per equation. Thus to show that cohomological k-equivalence distinguishes positive and negative instances of the CFI property for some fixed k we need to show that an equivalent system of equations can be interpreted which fixes the number of variables per equation. This is the content of the following lemma.

▶ **Lemma 33.** *For any prime power q, there is an interpretation Φ_q from the signature of the CFI structures $\mathbf{CFI}_q(G, g)$ to the signature of the ring $\mathbb{Z}_q$ with relations of arity at most 3 such that*

$$\Phi_q(\mathbf{CFI}_q(G, g)) \to \mathbb{Z}_q \iff \sum g = 0$$

Proof. From Lemma 32, we know that interpreting the system of equations $\mathbf{Eq}_q(G, g)$ would suffice for this purpose. However, the X equations in $\mathbf{Eq}_q(G, g)$ contain a number of variables which grows with the size of the maximum degree of a vertex in G. As this is, in general, unbounded - and in particular is unbounded in Lichter's class - we need to introduce some equivalent equations in a bounded number of variables. To do this we will introduce some slack variables and utilise the ordering on G to turn any such equation in n variables into a series of equations in 3 variables. We now describe the interpretation Φ_q as follows.

Let 3-$\mathbb{Z}_q$ denote the relational structure which contains a relation $T_{\alpha,\beta}$ for each α a tuple of elements of $\mathbb{Z}_q$ size up to 3 and $\beta \in \mathbb{Z}_q$. Each related tuple $(x, y, z) \in T_{\alpha,\beta}$ in a 3-$\mathbb{Z}_q$ structure is an equation

$$\alpha_1 x + \alpha_2 y + \alpha_3 z = \beta$$

To help define the interpretation we introduce some shorthand for some easily interpretable relations on CFI structures A. For $\mathbf{a}, \mathbf{b} \in A$ write $\mathbf{a} \sim \mathbf{b}$ if the two elements belong to the same gadget in A and $\mathbf{a} \frown \mathbf{b}$ if they belong to adjacent gadgets. Both of these relations are easily first-order definable as $\mathbf{a} \sim \mathbf{b}$ if and only if they are incomparable in the $\prec$ relation and $\mathbf{a} \frown \mathbf{b}$ if and only if $(\mathbf{a}, \mathbf{b}) \in R_{E,c}$ for some c. For $\mathbf{a} \frown \mathbf{b}$ in A we will refer to the elements $(\mathbf{a}, \mathbf{a}, \mathbf{b})$ and $(\mathbf{a}, \mathbf{b}, \mathbf{b})$ as $w_{\mathbf{a},\mathbf{b}}$ and $z_{\mathbf{a},\mathbf{b}}$. These will be the variables in the interpreted system of equations. As A comes with a linear pre-order $\prec$ inherited from the order on G, we can also define a local predecessor relation in the neighbourhood of any $\mathbf{a} \in A$. We say that $\mathbf{b}$ is a local predecessor of $\mathbf{b}'$ at $\mathbf{a}$ and write $\mathbf{b} \vdash_{\mathbf{a}} \mathbf{b}'$ if $\mathbf{a} \frown \mathbf{b}$ and $\mathbf{a} \frown \mathbf{b}'$ and there is no $\mathbf{b}''$ with $\mathbf{a} \frown \mathbf{b}''$ such that $\mathbf{b} \prec \mathbf{b}'' \prec \mathbf{b}'$.

Now we define the interpretation on A^3 in three steps, resulting in a system of equations which is solvable if and only if $\mathbf{Eq}_q(G,g)$ is solvable. **Step 1: Reducing variables** We note that in $\mathbf{Eq}_q(G,g)$ there are only variables $w_{\mathbf{a},y}$ for $\mathbf{a} \in A_x$ and $y \in N(x)$, whereas the shorthand above describes variables $w_{\mathbf{a},\mathbf{b}}$ and $z_{\mathbf{a},\mathbf{b}}$ for all $\mathbf{a} \in A_x$ and $\mathbf{b} \in A_y$. To reduce the number of variables we want to interpret, for all $\mathbf{a} \frown \mathbf{b}$ and $\mathbf{b} \sim \mathbf{b}'$, the equations $w_{\mathbf{a},\mathbf{b}} = w_{\mathbf{a},\mathbf{b}'}$ and $z_{\mathbf{a},\mathbf{b}} = z_{\mathbf{a},\mathbf{b}'}$. This is done by add the pairs $(w_{\mathbf{a},\mathbf{b}}, w_{\mathbf{a},\mathbf{b}'})$ and $(z_{\mathbf{a},\mathbf{b}}, z_{\mathbf{a},\mathbf{b}'})$ to the relation $T_{(1,-1),0}$ which can be done as $\frown$ and $\sim$ are definable.

Step 2: Interpreting I, C and E equations Defining these equations in $\Phi(A)$ is straightforward as they all have fewer than 3 variables. In particular we want to add equations

$$w_{\mathbf{a},\mathbf{b}} - w_{\mathbf{a}',\mathbf{b}} = 0$$

for any $(\mathbf{a}, \mathbf{a}', \mathbf{b}) \in R_I$,

$$w_{\mathbf{a},\mathbf{b}} - w_{\mathbf{a}',\mathbf{b}} = 1$$

for any $(\mathbf{a}, \mathbf{a}', \mathbf{b}) \in R_C$, and

$$w_{\mathbf{a},\mathbf{b}} + w_{\mathbf{b},\mathbf{a}} = c$$

for any $(\mathbf{a}, \mathbf{b}) \in R_{E,c}$. These are all easily first-order definable in the $\mathbf{CFI}_q$ signature.

Step 3: Interpreting X equations To interpret the equations for each $u \in G$ and $\mathbf{a} \in A_u$

$$\sum_{v \in N(u)} w_{\mathbf{a},v} = 0$$

in $\Phi(A)$, we first note that the linear order on G restricts to a linear order on $N(u)$ which we can write as $\{v_1, \dots, v_n\}$ where $i < j$ if and only if $v_i < v_j$. To do this it suffices to impose the equations

$$w_{\mathbf{a},\mathbf{b}_1} + \dots + w_{\mathbf{a},\mathbf{b}_n} = 0$$

for each sequence of elements $\mathbf{b}_1 \vdash_{\mathbf{a}} \dots \vdash_{\mathbf{a}} \mathbf{b}_n$ with $\mathbf{b}_i \in A_{v_i}$. To do this in equations with at most three variables we employ the auxiliary z variables in the following way. For any $\mathbf{ab} \in A$ such that $\mathbf{a} \frown \mathbf{b}$, if there is no $\mathbf{b}'$ such that $\mathbf{b}' \vdash_{\mathbf{a}} \mathbf{b}$, then we interpret the equation

$$w_{\mathbf{a},\mathbf{b}} - z_{\mathbf{a},\mathbf{b}} = 0$$

if there is $\mathbf{b}'$ such that $\mathbf{b}' \vdash_{\mathbf{a}} \mathbf{b}$ then interpret for all such $\mathbf{b}'$ the equation

$$z_{\mathbf{a},\mathbf{b}'} + w_{\mathbf{a},\mathbf{b}} - z_{\mathbf{a},\mathbf{b}} = 0$$

and if there is no $\mathbf{b}'$ such that $\mathbf{b} \vdash_{\mathbf{a}} \mathbf{b}'$ then interpret the equation

$$z_{\mathbf{a},\mathbf{b}} = 0$$

In this system of equations the $z_{\mathbf{a},\mathbf{b}}$ variables act as running totals for the sum $\sum w_{\mathbf{a},\mathbf{b}_i}$ and so it is not hard to see that solutions to these equations are precisely solutions to the equations $\sum w_{\mathbf{a},\mathbf{b}_i} = 0$. Furthermore, as the relation $\vdash_{\mathbf{a}}$ is definable in the signature of the $\mathbf{CFI}_q$ structures so too are these equations.

To conclude, we have interpreted in $\Phi(\mathbf{CFI}_q(G,g))$ a system of linear equations with three variables per equation which is solvable over $\mathbb{Z}_q$ if and only if $\mathbf{Eq}_q(G,g)$ is solvable. Thus there is a homomorphism $\Phi(\mathbf{CFI}_q(G,g)) \to \mathbb{Z}_q$ (as 3-$\mathbb{Z}_q$ structures) if and only if $\sum g = 0$. ◀

We can now conclude with the proof of Theorem 12.

Proof of Theorem 12. By Fact 9, the reverse implication is easy as $\sum h = 0$ implies that $\mathbf{CFI}_q(G, g) \cong \mathbf{CFI}_q(G, h)$ and so the structures are cohomologically k-equivalent for any k. The converse follows from the series of lemmas we have just presented. If $\sum h \neq 0$ then by Lemma 33 there is an interpretation Φ_q of order 3 such that $\Phi_q(\mathbf{CFI}_q(G, g)) \to \mathbb{Z}_q$ but $\Phi_q(\mathbf{CFI}_q(G, h)) \not\to \mathbb{Z}_q$. By Theorem 8, This is means that $\Phi_q(\mathbf{CFI}_q(G, g)) \to_3^{\mathbb{Z}} \mathbb{Z}_q$ but $\Phi_q(\mathbf{CFI}_q(G, h)) \not\to_3^{\mathbb{Z}} \mathbb{Z}_q$. So by Observation 7, we must have that $\Phi_q(\mathbf{CFI}_q(G, g)) \not\equiv_3^{\mathbb{Z}} \Phi_q(\mathbf{CFI}_q(G, h))$. Then noting that the number of variables used in the interpretation Φ_q is some constant c not depending on q and assuming without loss of generality that k is greater than $3c$ then Proposition 28 implies that $\mathbf{CFI}_q(G, g) \not\equiv_k^{\mathbb{Z}} \mathbf{CFI}_q(G, h)$, as required. ◀