Quantifiers Closed Under Partial Polymorphisms

Anuj Dawar
Department of Computer Science and Technology, University of Cambridge, UK

Lauri Hella
Faculty of Information Technology and Communication Sciences, Tampere University, Finland

Abstract

We study Lindström quantifiers that satisfy certain closure properties which are motivated by the study of polymorphisms in the context of constraint satisfaction problems (CSP). When the algebra of polymorphisms of a finite structure $\mathfrak{B}$ satisfies certain equations, this gives rise to a natural closure condition on the class of structures that map homomorphically to $\mathfrak{B}$. The collection of quantifiers that satisfy closure conditions arising from a fixed set of equations are rather more general than those arising as CSP. For any such conditions $\mathcal{P}$, we define a pebble game that delimits the distinguishing power of the infinitary logic with all quantifiers that are $\mathcal{P}$-closed. We use the pebble game to show that the problem of deciding whether a system of linear equations is solvable in $\mathbb{Z}/2\mathbb{Z}$ is not expressible in the infinitary logic with all quantifiers closed under a near-unanimity condition.

2012 ACM Subject Classification Theory of computation → Finite Model Theory

Keywords and phrases generalized quantifiers, constraint satisfaction problems, pebble games, finite variable logics, descriptive complexity theory

Digital Object Identifier 10.4230/LIPIcs.CSL.2024.23

Funding *Anuj Dawar*: funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee: grant number EP/X028259/1.

1 Introduction

Generalized quantifiers, also known as Lindström quantifiers, have played a significant role in the development of finite model theory. The subject of finite model theory is the expressive power of logics in the finite, and Lindström quantifiers provide a very general and abstract method of constructing logics. We can associate with any isomorphism-closed class of structures $\mathcal{K}$, a quantifier $Q_{\mathcal{K}}$ so that the extension $L(Q_{\mathcal{K}})$ of a logic L with the quantifier $Q_{\mathcal{K}}$ is the *minimal* extension of L that can express the class $\mathcal{K}$, subject to certain natural closure conditions. For this reason, comparing the expressive power of logics with Lindström quantifiers is closely related to comparing the descriptive complexity of the underlying classes of structures.

Another reason for the significance of Lindström quantifiers is that we have powerful methods for proving inexpressibility in logics with such quantifiers. In particular, games, based on Hella's bijection games [17], are the basis of the most common inexpressivity results that have been obtained in finite model theory. The k, n-bijection game was introduced by Hella to characterize equivalence in the logic $L^k_{\infty\omega}(\mathbf{Q}_n)$, which is the extension of the infinitary logic with k variables by means of all n-ary Lindström quantifiers. A quantifier $Q_{\mathcal{K}}$ is n-ary if the class $\mathcal{K}$ is defined over a vocabulary σ in which all relation symbols have arity n or less. In particular, the $k, 1$-bijection game, often called the k-pebble bijection game, characterizes equivalence in $L^k_{\infty\omega}(\mathbf{Q}_1)$ which has the same expressive power as $C^k_{\infty\omega}$, the k-variable infinitary logic with counting. Hella uses the k, n-bijection game to show that, for each n, there is an $(n+1)$-ary quantifier that is not definable in $L^k_{\infty\omega}(\mathbf{Q}_n)$ for any k.

The $k, 1$-bijection game has been widely used to establish inexpressibility results for $C^k_{\infty\omega}$. The k, n-bijection game for $n > 1$ has received relatively less attention. One reason is that, while equivalence in $C^k_{\infty\omega}$ is a polynomial-time decidable relation, which is in fact a relation much studied on graphs in the form of the Weisfeiler-Leman algorithm, in contrast the relation induced by the k, n-bijection game for $n > 1$ reduces to isomorphism on graphs and is intractable in general. Nonetheless, there is some interest in studying, for example, the non-trivial equivalence induced by $L^k_{\infty\omega}(\mathbf{Q}_2)$ on structures with a ternary relation. Grochow and Levet [16] investigate this relation on finite groups.

A second reason why the logics $L^\omega_{\infty\omega}(\mathbf{Q}_n)$ have attracted less interest is that in finite model theory we are often interested in logics that are closed under vectorized first-order interpretations. This is especially so in descriptive complexity as the complexity classes we are trying to characterize usually have these closure properties. While $L^\omega_{\infty\omega}(\mathbf{Q}_1)$ is closed under first-order interpretations, this is not the case for $L^\omega_{\infty\omega}(\mathbf{Q}_n)$ for $n > 1$. Indeed, the closure of $L^\omega_{\infty\omega}(\mathbf{Q}_2)$ under interpretations already includes $\mathbf{Q}_n$ for all n and so can express all properties of finite structures. So, it seems that beyond $L^\omega_{\infty\omega}(\mathbf{Q}_1)$, interesting logics from the point of view of complexity necessarily include quantifiers of all arities.

One way of getting meaningful logics that include quantifiers of unbounded arity is to consider quantifiers restricted to stronger closure conditions than just closure under isomorphisms. In recent work, novel game-based methods have established new inexpresibilty results for such logics, i.e. logics with a wide class of quantifiers of unbounded arity, but satisfying further restrictions. An important example is the class of linear-algebraic quantifiers, introduced in [8] which is the closure under interpretations of binary quantifiers invariant under invertible linear maps over finite fields. Equivalence in the resulting logic is characterized by the invertible map games introduced in [10]. These games are used in a highly sophisticated way by Lichter [21] to demonstrate a polynomial-time property that is not definable in fixed-point logic with rank introduced in [9, 15]. The result is extended to the infinitary logic with all linear-algebraic quantifiers in [7].

Another example is the recent result of Hella [18] showing a hierarchy theorem for quantifiers based on *constraint satisfaction problems* (CSP), using a novel game. Recall that for a fixed relational structure $\mathfrak{B}$, $\mathsf{CSP}(\mathfrak{B})$ denotes the class of structures that map homomorphically to $\mathfrak{B}$. Hella establishes that, for each $n > 1$, there is a structure $\mathfrak{B}$ with $n+1$ elements that is not definable in $L^\omega_{\infty\omega}(\mathbf{Q}_1, \mathbf{CSP}_n)$, where $\mathbf{CSP}_n$ denotes the collection of all quantifiers of the form $Q_{\mathsf{CSP}(\mathfrak{B}')}$ where $\mathfrak{B}'$ has at most n elements. Note that $\mathbf{CSP}_n$ includes quantifiers of all arities.

The interest in CSP quantifiers is inspired by the great progress that has been made in classifying constraint satisfaction problems in recent years, resulting in the dichotomy theorem of Bulatov and Zhuk [5, 24] stating that, for any structure $\mathfrak{B}$, $\mathsf{CSP}(\mathfrak{B})$ is either polynomial time computable, or NP-complete. The so-called algebraic approach to the classification of CSP has shown that the dividing line between these alternatives is completely determined by the algebra of polymorphisms of the structure $\mathfrak{B}$. More specifically, it is completely determined by the equational theory of this algebra. As we make explicit in Section 3 below, equations satisfied by the polymorphisms of $\mathfrak{B}$ naturally give rise to certain closure properties for the class of structures $\mathsf{CSP}(\mathfrak{B})$, which we describe by *partial polymorphisms.*

The notion of partial polymorphism, as well as that of polymorphism, goes back to Geiger and Bodnarchuk et al [14, 4], who proved a one-to-one correspondence between sets of relations that are closed under definability by conjunctions of atomic formulas (i.e., positive primitive formulas without existential quantification) and sets of partial functions that contain all projections and are closed under composition and restriction. Later Romov [22] formulated this correspondence as a Galois connection: a relation R is definable in a structure $\mathfrak{B}$ by a conjunction of atomic formulas if, and only if, every partial function that is a partial polymorphism of $\mathfrak{B}$ is also a partial polymorphism of R.

Partial polymorphism offer a more fine-grained tool for comparing the complexity of CSPs and related problems than the method based on total polymorphisms. Schnoor and Schnoor [23] showed that the above mentioned Galois connection for partial polymorphisms can be used for analysing the complexity of enumerating the solutions of $\mathsf{CSP}(\mathfrak{B})$. As an application, they proved that, for a Boolean $\mathsf{CSP}(\mathfrak{B})$, there exists an efficient enumeration algorithm if, and only if, $\mathsf{CSP}(\mathfrak{B})$ itself is polynomial time computable. Furthermore, Jonsson et al. [20] proved that, for two structures $\mathfrak{B}$ and $\mathfrak{C}$ with the same domain, if all partial polymorphisms of $\mathfrak{B}$ are also partial polymorphisms of $\mathfrak{C}$, then $\mathsf{CSP}(\mathfrak{C})$ can be reduced to $\mathsf{CSP}(\mathfrak{B})$ by a polynomial time reduction that increases the size of input structures by at most a constant. As a corollary, they proved a tight result on the relative time complexity of the corresponding CSPs: if $\mathsf{CSP}(\mathfrak{B})$ can be solved in time $2^{(c+\varepsilon)n}$ for every $\varepsilon > 0$, then so can $\mathsf{CSP}(\mathfrak{C})$.

A central aim of the present paper is to initiate the study of quantifiers closed under partial polymorphisms. We present a Spoiler-Duplicator pebble game, based on bijection games, which exactly characterises the expressive power of such quantifiers. More precisely, there is such a game for any suitable family $\mathcal{P}$ of partial polymorphisms. The exact definition of a quantifier being closed under the family $\mathcal{P}$ is given in Section 3; here we just remark that this notion is *not* based on the Galois connection between relations and partial polymorphisms. The definition the game and the proof of the characterization are given in Section 4.

As a case study, we consider the partial polymorphisms described by a *near-unanimity* condition. It is known since the seminal work of Feder and Vardi [13] that if a structure $\mathfrak{B}$ admits a near-unanimity polymorphism, then $\mathsf{CSP}(\mathfrak{B})$ has *bounded width*, i.e. it (or more precisely, its complement) is definable in Datalog. On the other hand, the problem of determining the solvability of a system of equations over the two-element field $\mathbb{Z}/2\mathbb{Z}$ is the classic example of a tractable CSP that is not of bounded width. Indeed, it is not even definable in $C^{\omega}_{\infty\omega}$ [2]. We show that the collection of quantifiers that are closed under near-unanimity partial polymorphisms is much richer than the classes $\mathsf{CSP}(\mathfrak{B})$ where $\mathfrak{B}$ has a near-unanimity polymorphism. The collection not only includes quantifiers which are not CSP, but it also includes CSP quantifiers which are not of bounded width, including intractable ones such as hypergraph colourability. Still, we are able to show that the problem of solving systems of equations over $\mathbb{Z}/2\mathbb{Z}$ is not definable in the extension of $C^{\omega}_{\infty\omega}$ with *all* quantifiers closed under near-unanimity partial polymorphisms. This sheds new light on the inter-definability of constraint satisfaction problems. For instance, while it follows from the arity hierarchy of [17] that the extension of $C^{\omega}_{\infty\omega}$ with a quantifier for graph 3-colourability still cannot define solvability of systems of equations over $\mathbb{Z}/2\mathbb{Z}$, our result shows this also for the extension of $C^{\omega}_{\infty\omega}$ with all hypergraph colourability quantifiers.

2 Preliminaries

We assume basic familiarity with logic, and in particular the logics commonly used in finite model theory (see [11], for example). We write $L^k_{\infty\omega}$ to denote the infinitary logic (that is, the closure of first-order logic with infinitary conjunctions and disjunctions) with k variables and $L^{\omega}_{\infty\omega}$ for $\bigcup_{k\in\omega} L^k_{\infty\omega}$. We are mainly interested in the extensions of these logics with generalized quantifiers, which we introduce in more detail in Section 2.1 below.

We use Fraktur letters $\mathfrak{A}, \mathfrak{B}, \ldots$ to denote structures and the corresponding Roman letters $A, B, \ldots$ to denote their universes. Unless otherwise mentioned, all structures are assumed to be finite. We use function notation, e.g. $f : A \to B$ to denote possibly *partial* functions. If $f : A \to B$ is a function and $\vec{a} \in A^m$ a tuple, we write $f(\vec{a})$ for the tuple in B^m obtained by applying f to $\vec{a}$ componentwise. This extends to functions of arity greater than 1. Thus, if $f : A^n \to B$ is a function of arity n and $\vec{a}_1, \ldots, \vec{a}_m \in A^n$ is a sequence of n-tuples, then $f(\vec{a}_1, \ldots, \vec{a}_m) = (f(\vec{a}_1), \ldots, f(\vec{a}_m))$. It is sometimes convenient to think of the sequence $\vec{a}_1, \ldots, \vec{a}_m \in A^n$ as an $m \times n$ matrix M with $M_{ij} = (\vec{a}_i)_j$ and we may write $f(M)$ for $f(\vec{a}_1, \ldots, a_m)$. On the other hand if N is a $n \times m$ matrix with entries in A, we write $\hat{f}(N)$ to denote $f(N^T)$. That is, for $\vec{a}_1, \ldots, \vec{a}_n \in A^m$ $\hat{f}(\vec{a}_1, \ldots, \vec{a}_n)$ denotes $(f(\vec{b}_1), \ldots, f(\vec{b}_m))$, where $\vec{b}_i = (N^T)_i$ is the tuple of ith components of $\vec{a}_1, \ldots, \vec{a}_n$. For a matrix M, we write M_i to denote the vector formed by the ith row of M.

For a pair of structures $\mathfrak{A}$ and $\mathfrak{B}$, a *partial isomorphism* from $\mathfrak{A}$ to $\mathfrak{B}$ is a partial function $f : A \to B$ which is an isomorphism between the substructure of $\mathfrak{A}$ induced by the domain of f and the substructure of $\mathfrak{B}$ induced by the image of f. We write $\mathrm{PI}(\mathfrak{A}, \mathfrak{B})$ to denote the collection of all partial isomorphisms from $\mathfrak{A}$ to $\mathfrak{B}$.

We write $\mathbb{N}$ or ω to denote the natural numbers, and $\mathbb{Z}$ to denote the ring of integers. For any $n \in \mathbb{N}$, we write $[n]$ to denote the set $\{1, \ldots, n\}$. When mentioned without further qualification, a graph $G = (V, E)$ is simple and undirected. That is, it is a structure with universe V and one binary relation E that is irreflexive and symmetric. The *girth* of a graph G is the length of the shortest cycle in G.

A *hypergraph* is a pair (H, E) such that E is a set of subsets of H. (H, E) is *n-uniform* if $|e| = n$ for all $e \in E$. As usual, we treat an n-uniform hypergraph (H, E) as the corresponding relational structure (H, R), where $R := \{(v_1, \ldots, v_n) \in H^n \mid \{v_1, \ldots, v_n\} \in E\}$.

2.1 Generalized quantifiers

Let σ, τ be relational vocabularies with $\tau = \{R_1, \ldots, R_m\}$, and $\mathrm{ar}(R_i) = r_i$ for each $i \in [m]$. An interpretation $\mathcal{I}$ of τ in σ with parameters $\vec{z}$ is a tuple of σ-formulas $(\psi_1, \ldots, \psi_m)$ along with tuples $\vec{y}_1, \ldots, \vec{y}_m$ of variables with $|\vec{y}_i| = r_i$ for $i \in [m]$, such that the free variables of ψ_i are among $\vec{y}_i\vec{z}$. Such an interpretation defines a mapping that takes a σ-structure $\mathfrak{A}$, along with an interpretation α of the parameters $\vec{z}$ in $\mathfrak{A}$ to a τ-structure $\mathfrak{B} := \mathcal{I}(\mathfrak{A}, \alpha)$ as follows. The universe of $\mathfrak{B}$ is A, and the relations $R_i \in \tau$ are interpreted in $\mathfrak{B}$ by $R_i^{\mathfrak{B}} = \{\vec{b} \in A^{r_i} \mid (\mathfrak{A}, \alpha[\vec{b}/\vec{y}_i]) \models \psi_i\}$.

Let L be a logic and $\mathcal{K}$ a class of τ-structures. The extension $L(Q_{\mathcal{K}})$ of L by the *generalized quantifier* for the class $\mathcal{K}$ is obtained by extending the syntax of L by the following formula formation rule:

> For $\mathcal{I} = (\psi_1, \ldots, \psi_m)$ an interpretation of τ in σ with parameters $\vec{z}$, $\psi(\vec{z}) = Q_{\mathcal{K}}\vec{y}_1, \ldots, \vec{y}_m\mathcal{I}$ is a formula over the signature σ, with free variables $\vec{z}$. The semantics of the formula is given by $(\mathfrak{A}, \alpha) \models \psi(\vec{z})$, if, and only if, $\mathcal{I}(\mathfrak{A}, \alpha) \in \mathcal{K}$.

The extension $L(\mathbf{Q})$ of L by a collection $\mathbf{Q}$ of generalized quantifiers is defined by adding the rules above to L for each $Q_{\mathcal{K}} \in \mathbf{Q}$ separately.

The *type* of the quantifier $Q_{\mathcal{K}}$ is $(r_1, \ldots, r_m)$, and the *arity* of $Q_{\mathcal{K}}$ is $\max\{r_1, \ldots, r_m\}$. For the sake of simplicity, we assume in the sequel that the type of $Q_{\mathcal{K}}$ is *uniform*, i.e., $r_i = r_j$ for all $i, j \in [m]$. This is no loss of generality, since any quantifier $Q_{\mathcal{K}}$ is definably equivalent with another quantifier $Q_{\mathcal{K}'}$ of uniform type with the same arity. Furthermore, we restrict the syntactic rule of $Q_{\mathcal{K}}$ by requiring that $\vec{y}_i = \vec{y}_j$ for all $i, j \in [m]$. Then we can denote the formula obtained by applying the rule simply by $\varphi = Q_{\mathcal{K}}\vec{y}\,(\psi_1, \ldots, \psi_m)$.

Let $Q = Q_{\mathcal{K}}$ and $Q' = Q_{\mathcal{K}'}$ be generalized quantifiers. We say that Q is *definable* in $L(Q')$ if the defining class $\mathcal{K}$ is definable in $L(Q')$, i.e., there is a sentence φ of $L(Q')$ such that $\mathcal{K} = \{\mathfrak{A} \mid \mathfrak{A} \models \varphi\}$.

We write $\mathbf{Q}_n$ to denote the collection of all quantifiers of arity at most n. Hella [17] shows that for any n, there is a quantifier of arity $n+1$ that is not definable in $L^{\omega}_{\infty\omega}(\mathbf{Q}_n)$. The logic $L^{\omega}_{\infty\omega}(\mathbf{Q}_1)$ is equivalent to $C^{\omega}_{\infty\omega}$, the infinitary logic with counting. The notion of interpretation we have defined is fairly restricted in that it does not allow for *relativization* or *vectorizations* (see, e.g. [11, Def. 12.3.6]. The relativizations and vectorizations of a quantifer Q can always be seen as a *collection* of simple quantifiers of unbounded arity.

2.2 CSP and polymorphisms

Given relational structures $\mathfrak{A}$ and $\mathfrak{B}$ over the same vocabulary τ, a *homomorphism* $h : \mathfrak{A} \to \mathfrak{B}$ is a function that takes elements of A to elements of B and such that for every $R \in \tau$ of arity r and any $\vec{a} \in A^r$, $\vec{a} \in R^{\mathfrak{A}}$ implies $h(\vec{a}) \in R^{\mathfrak{B}}$. For a fixed structure $\mathfrak{B}$, we write $\mathsf{CSP}(\mathfrak{B})$ to denote the collection of structures $\mathfrak{A}$ for which there is some homomorphism $h : \mathfrak{A} \to \mathfrak{B}$. By the celebrated theorem of Bulatov and Zhuk, every class $\mathsf{CSP}(\mathfrak{B})$ is either decidable in polynomial time or NP-complete.

Given a τ-structure $\mathfrak{B}$ and $m \in \mathbb{N}$, we define a τ-structure $\mathfrak{B}^m$. Its universe is B^m and if R in τ is a relation of arity r, and $\vec{a}_1, \ldots, \vec{a}_r \in B^m$, then $(\vec{a}_1, \ldots, \vec{a}_r) \in R^{\mathfrak{B}^m}$ if, and only if, $(M^T)_j \in R^{\mathfrak{B}}$ for all $j \in [m]$ where M is the $r \times m$ matrix formed by $(\vec{a}_1, \ldots, \vec{a}_r)$. Then, a *polymorphism* of $\mathfrak{B}$ is a homomorphism $p : \mathfrak{B}^m \to \mathfrak{B}$ for some m. The collection of polymorphisms of $\mathfrak{B}$ forms an algebraic *clone* with universe B. It is known that the equational theory of this algebra completely determines the computational complexity of $\mathsf{CSP}(\mathfrak{B})$ (see [3] for an expository account).

A function $m : B^3 \to B$ is a *majority* function if it satisfies the equations $m(a,a,b) = m(a,b,a) = m(b,a,a) = a$ for all $a, b \in B$. More generally, for $\ell \geq 3$, a function $n : B^\ell \to B$ is a *near-unanimity* function of arity ℓ if for any ℓ-tuple $\vec{a}$, we have $n(\vec{a}) = a$ whenever at least $\ell - 1$ components of $\vec{a}$ are a. In particular, a near-unanimity function of arity 3 is a majority function. A function $M : B^3 \to B$ is a *Maltsev* function if it satisfies the identities $M(a,b,b) = M(b,b,a) = a$ for all $a, b \in B$.

For any structure $\mathfrak{B}$ which has a near-unanimity polymorphism, the class $\mathsf{CSP}(\mathfrak{B})$ is decidable in polynomial time, and definable in $L^{\omega}_{\infty\omega}$. If $\mathfrak{B}$ admits a Maltsev polymorphism, then $\mathsf{CSP}(\mathfrak{B})$ is also decidable in polynomial time, but may not be definable in $L^{\omega}_{\infty\omega}$ or $L^{\omega}_{\infty\omega}(\mathbf{Q}_1)$, its extension with all unary quantifiers. The classic example of a CSP with a Maltsev polymorphism that is not definable in $L^{\omega}_{\infty\omega}(\mathbf{Q}_1)$ is solvability of systems of equations over $\mathbb{Z}/2\mathbb{Z}$ with ℓ variables per equation. We can treat this as the class of structures $\mathsf{CSP}(\mathfrak{C}_\ell)$ where $\mathfrak{C}_\ell$ is the structure with universe $\{0,1\}$ and two ℓ-ary relations $R_0 = \{(b_1, \ldots, b_\ell) \mid \sum_i b_i \equiv 0 \pmod 2\}$ and $R_1 = \{(b_1, \ldots, b_\ell) \mid \sum_i b_i \equiv 1 \pmod 2\}$.

If $\mathcal{K} = \mathsf{CSP}(\mathfrak{B})$ for some fixed structure $\mathfrak{B}$, we call $Q_{\mathcal{K}}$ a *CSP quantifier*. Write $\mathbf{CSP}_n$ for the collection of all CSP quantifiers $Q_{\mathcal{K}}$ where $\mathcal{K} = \mathsf{CSP}(\mathfrak{B})$ for a structure with at most n elements. Note that there is no restriction on the number or arity of relations in the signature of $\mathfrak{B}$ and thus $\mathbf{CSP}_n$ contains quantifiers of all arities. Hella [18] defines a pebble game that characterizes equivalence of structures in the logic $L^{\omega}_{\infty\omega}(\mathbf{Q}_1, \mathbf{CSP}_n)$ and shows that there is a structure $\mathfrak{B}$ on $n+1$ elements such that $\mathsf{CSP}(\mathfrak{B})$ is not definable in this logic.

3 Partial polymorphisms

Let τ be a relational vocabulary, and let $\mathfrak{C}$ be a τ-structure with a polymorphism $p\colon \mathfrak{C}^n \to \mathfrak{C}$. This gives rise to a closure condition on the class $\mathsf{CSP}(\mathfrak{C})$. In particular, suppose $\mathfrak{B} \in \mathsf{CSP}(\mathfrak{C})$ by a homomorphism $h : \mathfrak{B} \to \mathfrak{C}$. We can, in a sense, "close" $\mathfrak{B}$ under the polymorphism p

by including in each relation $R^{\mathfrak{B}}$ ($R \in \tau$) any tuple $\vec{a}$ for which $h(\vec{a}) = p(h(\vec{a}_1, \ldots, \vec{a}_n))$ for some $\vec{a}_1, \ldots, \vec{a}_n \in R_i^{\mathfrak{B}}$. The resulting structure $\mathfrak{B}'$ is still in $\mathsf{CSP}(\mathfrak{C})$ as is any structure $\mathfrak{A}$ with the same universe as $\mathfrak{B}$ and for which $R^{\mathfrak{A}} \subseteq R^{\mathfrak{B}'}$ for all $R \in \tau$.

Our aim is to generalize this type of closure property from CSP quantifiers to a larger class of generalized quantifiers. To formally define this, it is useful to introduce some notation. For reasons that will become clear, we use *partial* functions p.

▶ **Definition 1.** *Let $A \neq \emptyset$ be a set, and let p be a partial function $A^n \to A$.*

(a) *If $R \subseteq A^r$, then $p(R) := \{\hat{p}(\vec{a}_1, \ldots, \vec{a}_n) \mid \vec{a}_1, \ldots, \vec{a}_n \in R\}$.*

(b) *If $\mathfrak{A} = (A, R_1^{\mathfrak{A}}, \ldots, R_m^{\mathfrak{A}})$, then we denote the structure $(A, p(R_1^{\mathfrak{A}}), \ldots, p(R_m^{\mathfrak{A}}))$ by $p(\mathfrak{A})$.*

We say that p is a *partial polymorphism* of a τ-structure $\mathfrak{A}$ with domain A if for every $R \in \tau$, the relation $R^{\mathfrak{A}}$ is closed with respect to p, i.e., $p(R^{\mathfrak{A}}) \subseteq R^{\mathfrak{A}}$.

The reason for considering partial functions is that we are usually interested in polymorphisms that satisfy certain equations. The equations specify the polymorphism partially, but not totally. In other words, any polymorphism that extends the given partial function is a polymorphism satisfying the required equations. Thus, we can uniformly specify closure properties on our class of structures for polymorphisms satisfying the equations by only requiring closure for the partial function. This is illustrated in the examples below.

By a *family of partial functions* we mean a class $\mathcal{P}$ that contains a partial function $p_A \colon A^n \to A$ for every finite set A, where n is a fixed positive integer. We give next some important examples of families of partial functions that arise naturally from well-known classes of polymorphisms.

▶ **Example 2.**

(a) The *Maltsev family* $\mathcal{M}$ consists of the partial functions $M_A \colon A^3 \to A$ such that $M_A(a, b, b) = M_A(b, b, a) = a$ for all $a, b \in A$, and $M_A(a, b, c)$ is undefined unless $a = b$ or $b = c$. If a structure $\mathfrak{A}$ has a Maltsev polymorphism $p \colon A^3 \to A$, then clearly M_A is a restriction of p, whence it is a partial polymorphism of $\mathfrak{A}$.

(b) The family $\mathcal{MJ}$ of ternary *partial majority functions* consists of the partial functions $m_A \colon A^3 \to A$ such that $m_A(a, a, b) = m_A(a, b, a) = m_A(b, a, a) = a$ for all $a, b \in A$, and $m_A(a, b, c)$ is undefined if a, b and c are all distinct. If $\mathfrak{A}$ has a majority polymorphism, then m_A is a restriction of it, whence it is a partial polymorphism of $\mathfrak{A}$.

(c) More generally, for each $\ell \geq 3$ we define the family $\mathcal{N}_\ell$ of *ℓ-ary partial near-unanimity functions* $n_A^\ell \colon A^\ell \to A$ as follows:
- $n_A^\ell(a_1, \ldots, a_\ell) = a$ if and only if $|\{i \in [n] \mid a_i = a\}| \geq \ell - 1$.

In particular, $\mathcal{MJ} = \mathcal{N}_3$.

We next give a formal definition for the closure property of generalized quantifiers that arises from a family of partial functions. In the definition we use the notation $\mathfrak{A} \leq \mathfrak{B}$ if $\mathfrak{A}$ and $\mathfrak{B}$ are τ-structures such that $A = B$ and $R^{\mathfrak{A}} \subseteq R^{\mathfrak{B}}$ for each $R \in \tau$. Furthermore, we define the union $\mathfrak{A} \cup \mathfrak{B}$ of $\mathfrak{A}$ and $\mathfrak{B}$ to be the τ-structure $\mathfrak{C}$ such that $C = A \cup B$ and $R^{\mathfrak{C}} = R^{\mathfrak{A}} \cup R^{\mathfrak{B}}$ for each $R \in \tau$. Note that we do not assume here that A and B are disjoint. On the contrary, we use the union $\mathfrak{A} \cup \mathfrak{B}$ specifically for structures $\mathfrak{A}$ and $\mathfrak{B}$ that share a common universe $A = B$.

▶ **Definition 3.** *Let $\mathcal{P}$ be a family of n-ary partial functions, and let $Q_{\mathcal{K}}$ be a generalized quantifier of vocabulary τ. We say that $Q_{\mathcal{K}}$ is $\mathcal{P}$-closed if the following holds for all τ-structures $\mathfrak{A}$ and $\mathfrak{B}$ with $A = B$:*
- *if $\mathfrak{B} \in \mathcal{K}$ and $\mathfrak{A} \leq p_A(\mathfrak{B}) \cup \mathfrak{B}$ for some $p_A \in \mathcal{P}$, then $\mathfrak{A} \in \mathcal{K}$.*

We denote the class of all $\mathcal{P}$-closed quantifiers by $\mathbf{Q}_{\mathcal{P}}$.

Note that the condition $\mathfrak{A} \le p_A(\mathfrak{B}) \cup \mathfrak{B}$ holds if and only if for every $R \in \tau$ and every $\vec{a} \in R^{\mathfrak{A}} \setminus R^{\mathfrak{B}}$ there are tuples $\vec{a}_1, \ldots, \vec{a}_n \in R^{\mathfrak{B}}$ such that $\vec{a} = \widehat{p_A}(\vec{a}_1, \ldots, \vec{a}_n)$.

The quantifier $Q_{\mathcal{K}}$ is *downwards monotone* if $\mathfrak{A} \le \mathfrak{B}$ and $\mathfrak{B} \in \mathcal{K}$ implies $\mathfrak{A} \in \mathcal{K}$. It follows directly from Definition 3 that all $\mathcal{P}$-closed quantifiers are downwards monotone.

▶ **Proposition 4.** *If $Q_{\mathcal{K}} \in \mathbf{Q}_{\mathcal{P}}$ for some family $\mathcal{P}$, then $Q_{\mathcal{K}}$ is downwards monotone.*

▶ Remark 5. As far as we know, the notion of $\mathcal{P}$-closed quantifiers (or classes) has not been considered earlier. In particular, as we mentioned in Section 1, a quantifier $Q_{\mathcal{K}}$ being $\mathcal{P}$-closed is not based on the Galois connection between partial polymorphisms and relations: by downwards monotonicity of $Q_{\mathcal{K}}$, the class $\mathcal{K}$ usually contains structures $\mathfrak{A}$ such that p_A is not a partial polymorphism of $\mathfrak{A}$. Note also that the structure $\mathfrak{D}$ with full relations (i.e., $R^{\mathfrak{D}} = D^r$ for each $R \in \tau$ of arity r) is usually not in $\mathcal{K}$ although p_D is a partial polymorphism of $\mathfrak{D}$.

We show next that there are quantifiers that are $\mathcal{P}$-closed for all families $\mathcal{P}$ of partial functions.

▶ **Proposition 6.** *Let $\mathcal{K}_0$ be the class of all $\{R\}$-structures $\mathfrak{A}$ such that $R^{\mathfrak{A}} = \emptyset$. Then $Q_{\mathcal{K}_0} \in \mathbf{Q}_{\mathcal{P}}$ for any family $\mathcal{P}$ of partial functions.*

Proof. If $\mathfrak{B} \in \mathcal{K}_0$, then $R^{\mathfrak{B}} = \emptyset$, whence $p_B(\mathfrak{B}) = \emptyset$. Thus, if $\mathfrak{A} \le p_B(\mathfrak{B}) \cup \mathfrak{B}$, then $R^{\mathfrak{A}} = \emptyset$, and hence $\mathfrak{A} \in \mathcal{K}_0$. ◀

Note that in the case $\mathrm{ar}(R) = 1$, the quantifier $Q_{\mathcal{K}_0}$ of the proposition above is the negation of the existential quantifier: $\mathfrak{A} \models Q_{\mathcal{K}_0} x\, \varphi \iff \mathfrak{A} \models \neg\exists x\, \varphi$. Thus, for any family $\mathcal{P}$, the first-order quantifiers can be defined from a $\mathcal{P}$-closed quantifier using only negation.

Up to now we have not imposed any restrictions on the family $\mathcal{P}$. It is natural to require that the partial functions in $\mathcal{P}$ are uniformly defined, or at least that (A, p_A) and (B, p_B) are isomorphic if $|A| = |B|$. Such requirements are captured by the notions defined below.

▶ **Definition 7.** *Let $\mathcal{P}$ be a family of n-ary partial functions.*

(a) *$\mathcal{P}$* is invariant *if it respects bijections: if $f\colon A \to B$ is a bijection and $a_1, \ldots, a_n \in A$, then $p_B(f(a_1), \ldots, f(a_n)) \simeq f(p_A(a_1, \ldots, a_n))$. Here the symbol $\simeq$ says that either both sides are defined and have the same value, or both sides are undefined.*

(b) *$\mathcal{P}$ is* strongly invariant *if it respects injections: if $f\colon A \to B$ is an injection and $a_1, \ldots, a_n \in A$, then $p_B(f(a_1), \ldots, f(a_n)) \simeq f(p_A(a_1, \ldots, a_n))$.*

(c) *$\mathcal{P}$ is* projective*, if it is strongly invariant and it is preserved by all functions: if $f\colon A \to B$ is a function and $a_1, \ldots, a_n \in A$ are such that $p_A(a_1, \ldots, a_n)$ is defined, then $p_B(f(a_1), \ldots, f(a_n)) = f(p_A(a_1, \ldots, a_n))$.*

It is easy to verify that $\mathcal{P}$ is invariant if, and only if, it is determined by equality types on each cardinality: there are quantifier free formulas in the language of equality $\theta^m_{\mathcal{P}}(\vec{x}, y)$ such that if $|A| = m$, then $p_A(\vec{a}) = b \iff A \models \theta^m_{\mathcal{P}}[\vec{a}/\vec{x}, b/y]$ holds for all $\vec{a} \in A^n$ and $b \in A$. Similarly, $\mathcal{P}$ is strongly invariant if, and only if, the same holds with a single formula $\theta_{\mathcal{P}} = \theta^m_{\mathcal{P}}$ for all $m \in \omega$.

Note that if the family $\mathcal{P}$ is strongly invariant, then for every finite set A, p_A is a *partial choice function*, i.e., $p_A(a_1, \ldots, a_n) \in \{a_1, \ldots, a_n\}$. Indeed, if $b := p_A(a_1, \ldots, a_n) \notin \{a_1, \ldots, a_n\}$ and $B = A \cup \{c\}$, where $c \notin A$, then using the identity function $f = \mathrm{id}_A$ of A in the condition $p_B(f(a_1), \ldots, f(a_n)) = f(p_A(a_1, \ldots, a_n))$, we get $p_B(a_1, \ldots, a_n) = b$. On the other hand, using the injection $f'\colon A \to B$ that agrees with id_A on $A \setminus \{b\}$ but maps b to c, we get the contradiction $p_B(a_1, \ldots, a_n) = c \neq b$.

▶ **Lemma 8.** *Let $\mathcal{P}$ be a family of n-ary partial choice functions. Then $Q_\mathcal{K} \in \mathbf{Q}_\mathcal{P}$ for any unary downwards monotone quantifier $Q_\mathcal{K}$. In particular this holds if $\mathcal{P}$ is strongly invariant.*

Proof. Let τ be the vocabulary of $\mathcal{K}$, and assume that $\mathfrak{B} \in \mathcal{K}$ and $\mathfrak{A} \leq p_A(\mathfrak{B}) \cup \mathfrak{B}$. Then for all $R \in \tau$ and $a \in R^\mathfrak{A} \setminus R^\mathfrak{B}$ there are $a_1, \ldots, a_n \in A$ such that $p_A(a_1, \ldots, a_n) = a$ and $a_i \in R^\mathfrak{B}$ for each $i \in [n]$. Since p_A is a partial choice function, we have $a \in \{a_1, \ldots, a_n\}$, and hence $a \in R^\mathfrak{B}$. Thus we see that $\mathfrak{A} \leq \mathfrak{B}$, and consequently $\mathfrak{A} \in \mathcal{K}$, since $Q_\mathcal{K}$ is downwards monotone. ◀

It is easy to see that the families $\mathcal{M}$ and $\mathcal{N}_\ell$, $\ell \geq 3$, introduced in Example 2, are strongly invariant. Indeed, the defining formulas $\theta_\mathcal{M}$ and $\theta_{\mathcal{N}_\ell}$ are easily obtained from the identities that define these conditions. Thus, all unary downwards monotone quantifiers are $\mathcal{M}$-closed and $\mathcal{N}_\ell$-closed. For the families $\mathcal{N}_\ell$ we can prove a much stronger result:

▶ **Lemma 9.** *Let $\ell \geq 3$, and let $Q_\mathcal{K}$ be a downwards monotone quantifier of arity $r < \ell$. Then $Q_\mathcal{K} \in \mathbf{Q}_{\mathcal{N}_\ell}$.*

Proof. Let τ be the vocabulary of $\mathcal{K}$, and assume that $\mathfrak{B} \in \mathcal{K}$ and $\mathfrak{A} \leq n^\ell_A(\mathfrak{B}) \cup \mathfrak{B}$. Then for all $R \in \tau$ and $\vec{a} = (a_1, \ldots, a_r) \in R^\mathfrak{A} \setminus R^\mathfrak{B}$ there are $\vec{a}_i = (a^1_i, \ldots, a^r_i) \in R^\mathfrak{B}$, $i \in [\ell]$, such that $\widehat{n^\ell_A}(\vec{a}_1, \ldots, \vec{a}_\ell) = \vec{a}$. Thus, for each $j \in [r]$ there is at most one $i \in [\ell]$ such that $a^j_i \neq a_j$, and hence there is at least one $i \in [\ell]$ such that $\vec{a} = \vec{a}_i$. This shows that $\mathfrak{A} \leq \mathfrak{B}$, and since $Q_\mathcal{K}$ is downwards monotone, we conclude that $\mathfrak{A} \in \mathcal{K}$. ◀

Using a technique originally due to Imhof for (upwards) monotone quantifiers (see [19]), we can show that any quantifier $Q_\mathcal{K}$ is definable by a downwards monotone quantifier of the same arity. Indeed, if the vocabulary of $\mathcal{K}$ is $\tau = \{R_1, \ldots, R_m\}$, where $\mathrm{ar}(R_i) = r$ for all $i \in [m]$, we let $\tau' := \{S_1, \ldots, S_m\}$ be a disjoint copy of τ, and $\tau^* := \tau \cup \tau'$. Furthermore, we let $\mathcal{K}^*$ be the class of all τ^*-structures $\mathfrak{A}$ such that $R^\mathfrak{A}_i \cap S^\mathfrak{A}_i = \emptyset$ for all $i \in [m]$, and $(A, R^\mathfrak{A}_1, \ldots, R^\mathfrak{A}_m) \in \mathcal{K}$ or $R^\mathfrak{A}_i \cup S^\mathfrak{A}_i \neq A^r$ for some $i \in [m]$. Then $Q_{\mathcal{K}^*}$ is downwards monotone, and clearly $Q_\mathcal{K}\vec{x}\,(\psi_1, \ldots, \psi_m)$ is equivalent with $Q_{\mathcal{K}^*}\vec{x}\,(\psi_1, \ldots, \psi_m, \neg\psi_1, \ldots, \neg\psi_m)$.

Using this observation, we get the following corollary to Lemmas 8 and 9.

▶ **Corollary 10.**
(a) *Let $\mathcal{P}$ be as in Lemma 8. Then $L^k_{\infty\omega}(\mathbf{Q}_\mathcal{P} \cup \mathbf{Q}_1) \leq L^k_{\infty\omega}(\mathbf{Q}_\mathcal{P})$.*
(b) *$L^k_{\infty\omega}(\mathbf{Q}_{\mathcal{N}_\ell} \cup \mathbf{Q}_{\ell-1}) \leq L^k_{\infty\omega}(\mathbf{Q}_{\mathcal{N}_\ell})$.*

As explained in the beginning of this section, the definition of $\mathcal{P}$-closed quantifiers was inspired by the closure property of a CSP quantifier $Q_{\mathsf{CSP}(\mathfrak{C})}$ that arises from a polymorphism of $\mathfrak{C}$. Thus, it is natural to look for sufficient conditions on the family $\mathcal{P}$ and the target structure $\mathfrak{C}$ for $Q_{\mathsf{CSP}(\mathfrak{C})}$ to be $\mathcal{P}$-closed. It turns out that the notions of projectivity and partial polymorphism lead to such a condition.

▶ **Proposition 11.** *Let $\mathcal{P}$ be a projective family of n-ary partial functions, and let $\mathfrak{C}$ be a τ-structure. If p_C is a partial polymorphism of $\mathfrak{C}$, then $Q_{\mathsf{CSP}(\mathfrak{C})} \in \mathbf{Q}_\mathcal{P}$.*

Proof. Assume that $\mathfrak{B} \in \mathsf{CSP}(\mathfrak{C})$ and $\mathfrak{A} \leq p_A(\mathfrak{B}) \cup \mathfrak{B}$. Then $A = B$ and there is a homomorphism $h\colon \mathfrak{B} \to \mathfrak{C}$. We show that h is a homomorphism $\mathfrak{A} \to \mathfrak{C}$, and hence $\mathfrak{A} \in \mathsf{CSP}(\mathfrak{C})$. Thus let $R \in \tau$, and let $\vec{a} \in R^\mathfrak{A}$. If $\vec{a} \in R^\mathfrak{B}$, then $h(\vec{a}) \in R^\mathfrak{C}$ by assumption. On the other hand, if $\vec{a} \in R^\mathfrak{A} \setminus R^\mathfrak{B}$, then there exist tuples $\vec{a}_1, \ldots, \vec{a}_n \in R^\mathfrak{B}$ such that $\vec{a} = \widehat{p_A}(\vec{a}_1, \ldots, \vec{a}_n)$. Since h is a homomorphism $\mathfrak{B} \to \mathfrak{C}$, we have $h(\vec{a}_i) \in R^\mathfrak{C}$ for each $i \in [n]$. Since p_C is a partial polymorphism of $\mathfrak{C}$, we have $\widehat{p_C}(h(\vec{a}_1), \ldots, h(\vec{a}_n)) \in R^\mathfrak{C}$. Finally, since $\mathcal{P}$ is projective, we have $h(\vec{a}) = h(\widehat{p_A}(\vec{a}_1, \ldots, \vec{a}_n)) = \widehat{p_C}(h(\vec{a}_1), \ldots, h(\vec{a}_n))$, and hence $h(\vec{a}) \in R^\mathfrak{C}$. ◀

We can now apply Proposition 11 to the families introduced in Example 2.

▶ **Example 12.**

(a) Consider a constraint satisfaction problem $\mathsf{CSP}(\mathfrak{C})$ such that $\mathfrak{C}$ has a Maltsev polymorphism $p\colon \mathfrak{C}^3 \to \mathfrak{C}$. We show that $Q_{\mathsf{CSP}(\mathfrak{C})} \in \mathbf{Q}_{\mathcal{M}}$. As pointed out in Example 2, M_C is a partial polymorphism of $\mathfrak{C}$. Thus, by Proposition 11 it suffices to show that the Maltsev family $\mathcal{M}$ is projective.

Thus, assume that $f\colon A \to B$ is a function, and $M_A(a,b,c)$ is defined. Then $a = b$ and $M_A(a,b,c) = c$, or $b = c$ and $M_A(a,b,c) = a$. In the former case we have $f(a) = f(b)$, whence $M_B(f(a), f(b), f(c)) = f(c) = f(M_A(a,b,c))$. In the latter case we have $f(b) = f(c)$, whence $M_B(f(a), f(b), f(c)) = f(a) = f(M_A(a,b,c))$.

(b) The *n-uniform hypergraph m-colouring* problem is $\mathsf{CSP}(\mathfrak{H}_{n,m})$, where $\mathfrak{H}_{n,m} = ([m], R_{n,m})$ is the complete n-uniform hypergraph with m vertices, i.e.,

- $R_{n,m} := \{(v_1,\ldots,v_n) \in [m]^n \mid v_i \neq v_j \text{ for all } 1 \leq i < j \leq m\}$.

We show that $Q_{\mathsf{CSP}(\mathfrak{H}_{n,m})} \in \mathbf{Q}_{\mathcal{MJ}}$ for all $n \geq 2$ and $m \geq n$. By Proposition 11 it suffices to show that $m_{[m]}$ is a partial polymorphism of $\mathfrak{H}_{n,m}$, and the family $\mathcal{MJ}$ is projective.

To see that $m_{[m]}$ is a partial polymorphism of $\mathfrak{H}_{n,m}$, assume that $\vec{a}_i = (a_i^1, \ldots, a_i^n) \in R_{n,m}$ for $i \in [3]$, and $\vec{a} = (a_1,\ldots,a_n) = \hat{m}_{[m]}(\vec{a}_1, \vec{a}_2, \vec{a}_3)$. By the definition of $m_{[m]}$, for each $j \in [n]$ we have $|\{i \in [3] \mid a_i^j = a_j\}| \geq 2$. Thus for any two distinct $j, k \in [n]$, there is $i \in [3]$ such that $a_j = a_i^j$ and $a_i^k = a_k$, whence $a_j \neq a_k$. Thus we have $\vec{a} \in R_{n,m}$.

To show that $\mathcal{MJ}$ is projective, assume that $f\colon A \to B$ is a function, and $m_A(a,b,c)$ is defined. Then $a = b = m_A(a,b,c)$, $a = c = m_A(a,b,c)$ or $b = c = m_A(a,b,c)$. In the first case we have $f(m_A(a,b,c)) = f(a) = f(b) = m_B(f(a), f(b), f(c))$, as desired. The two other cases are similar.

(c) In the same way we can show that the family $\mathcal{N}_\ell$ of partial near-unanimity polymorphisms is projective for any $\ell \geq 3$. We relax now the notion of hypergraph coloring as follows: Let $\mathfrak{H} = (H, R)$ be an n-uniform hypergraph, and let $k < n$. A *k-weak m-coloring* of $\mathfrak{H}$ is a function $f\colon H \to [m]$ such that for all $(u_1,\ldots,u_n) \in R$ and all $i \in [m]$, $|\{u_1,\ldots,u_n\} \cap f^{-1}[\{i\}]| \leq k$. Thus, instead of requiring that all vertices in a hyperedge $(u_1,\ldots,u_n)$ must have different colors, a k-weak m-coloring allows up to k of them to have the same color. (Note that there are no restrictions on how many at most k-element subsets can be colored with a single color.) Observe now that there exists a k-weak m-coloring of $\mathfrak{H}$ if and only if $\mathfrak{H} \in \mathsf{CSP}(\mathfrak{H}^k_{n,m})$, where $\mathfrak{H}^k_{n,m} = ([m], R^k_{n,m})$ is the structure such that

- $R^k_{n,m} := \{(v_1,\ldots,v_n) \in [m]^n \mid |\{v_i \mid i \in I\}| \geq 2 \text{ for all } I \subseteq [n] \text{ with } |I| = k+1\}$.

Note that $\mathfrak{H}^1_{n,m} = \mathfrak{H}_{n,m}$, whence $m_{[m]} = n^3_{[m]}$ is a partial polymorphism of $\mathfrak{H}^1_{n,m}$. It is straightforward to generalize this to $\ell > 3$: $n^\ell_{[m]}$ is a partial polymorhism of $\mathfrak{H}^{\ell-2}_{n,m}$. Thus by Proposition 11, the CSP quantifier $Q_{\mathsf{CSP}(\mathfrak{H}^{\ell-2}_{n,m})}$ is $\mathbf{Q}_{\mathcal{N}_\ell}$-closed.

▶ Remark 13. As shown in Example 12(b), the partial majority function $m_{[m]}$ is a partial polymorphism of the structure $\mathfrak{H}_{n,m}$. However, there does not exist any polymorphism $p\colon [m]^3 \to [m]$ that extends $m_{[m]}$. This can be verified directly, but it also follows from the fact that $\mathsf{CSP}(\mathfrak{C})$ is of bounded width for any $\mathfrak{C}$ that has a majority polymorphism ([13]), but $\mathsf{CSP}(\mathfrak{H}_{n,m})$ is not of bounded width. The same holds for the partial functions $n^\ell_{[m]}$ and the structures $\mathfrak{H}^k_{n,m}$ in Example 12(c).

4 Pebble game for $\mathcal{P}$-closed quantifiers

In this section we introduce a pebble game that characterizes equivalence of structures with respect to $L^{\omega}_{\infty\omega}(\mathbf{Q}_{\mathcal{P}})$, the extension of the infinitary k-variable logic $L^{\omega}_{\infty\omega}$ by the class of all $\mathcal{P}$-closed quantifiers.

We fix a family $\mathcal{P}$ of n-ary partial functions for the rest of the section. Given two structures $\mathfrak{A}$ and $\mathfrak{B}$ of the same vocabulary, and assignments α and β on $\mathfrak{A}$ and $\mathfrak{B}$, respectively, such that $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta)$, we write $(\mathfrak{A}, \alpha) \equiv^{k}_{\infty\omega,\mathcal{P}} (\mathfrak{B}, \beta)$ if the equivalence

$$(\mathfrak{A}, \alpha) \models \varphi \iff (\mathfrak{B}, \beta) \models \varphi$$

holds for all formulas $\varphi \in L^{k}_{\infty\omega}(\mathbf{Q}_{\mathcal{P}})$ with free variables in $\mathrm{dom}(\alpha)$. If $\alpha = \beta = \emptyset$, we write simply $\mathfrak{A} \equiv^{k}_{\infty\omega,\mathcal{P}} \mathfrak{B}$ instead of $(\mathfrak{A}, \emptyset) \equiv^{k}_{\infty\omega,\mathcal{P}} (\mathfrak{B}, \emptyset)$.

The basic idea of our pebble game for a pair $(\mathfrak{A}, \mathfrak{B})$ of structures is the following. In each round Duplicator gives a bijection $f : A \to B$, just like in the bijection games of [17], but instead of using $\vec{b} = f(\vec{a})$ as answer for Spoiler's move $\vec{a} \in A^r$, she is allowed to give a sequence $\vec{b}_1, \ldots, \vec{b}_n \in B^r$ of alternative answers as long as $\vec{b} = \widehat{p_B}(\vec{b}_1, \ldots, \vec{b}_n)$. In particular, $\vec{b}$ need not be among the list $\vec{b}_1, \ldots, \vec{b}_n$. Spoiler completes the round by choosing one of these alternatives $\vec{b}_i$. Spoiler wins if $\vec{a} \mapsto \vec{b}_i$ is not a partial isomorphism; otherwise the game carries on from the new position. Note that this allows more freedom to Duplicator than in the ordinary k, n-bijection game. She can simulate a winning strategy in that game by simply playing at each move the sequence $\vec{b}_1, \ldots, \vec{b}_n$ where each $\vec{b}_i = \vec{b} = f(\vec{a})$.

Observe now that if Duplicator has a winning strategy for the first round of the game, then $f(\mathfrak{A}) \leq p_B(\mathfrak{B}) \cup \mathfrak{B}$. Indeed, if Spoiler chooses a tuple $\vec{a} \in R^{\mathfrak{A}}$, then Duplicator has to answer by either the tuple $f(\vec{a})$, or a sequence $\vec{b}_1, \ldots, \vec{b}_n \in B^r$ of tuples such that $f(\vec{a}) = \widehat{p_B}(\vec{b}_1, \ldots, \vec{b}_n)$; in the first case she loses if $f(\vec{a}) \notin R^{\mathfrak{B}}$, and in the second case she loses if $\vec{b}_i \notin R^{\mathfrak{B}}$ for some $i \in [n]$. Thus if Duplicator has a winning strategy in the one round game and $\mathfrak{B} \in \mathcal{K}$ for some $\mathcal{P}$-closed quantifier $Q_{\mathcal{K}}$, then $f(\mathfrak{A}) \in \mathcal{K}$, and since f is an isomorphism $\mathfrak{A} \to f(\mathfrak{A})$, also $\mathfrak{A} \in \mathcal{K}$. In other words, if $\mathfrak{B} \models Q_{\mathcal{K}}\vec{y}\,(R_1(\vec{y}), \ldots, R_m(\vec{y}))$, then $\mathfrak{A} \models Q_{\mathcal{K}}\vec{y}\,(R_1(\vec{y}), \ldots, R_m(\vec{y}))$. The reverse implication is obtained by using the move described above with the structures switched.

By allowing only k variables and repeating rounds indefinitely (unless Spoiler wins at some round), we obtain a game such that Duplicator having a winning strategy implies $\mathfrak{A} \equiv^{k}_{\infty\omega,\mathcal{P}} \mathfrak{B}$. However, in order to prove the converse implication we need to modify the rules explained above. This is because $p_B(\mathfrak{B}) \cup \mathfrak{B}$ is not necessarily closed with respect to the function p_B, and in the argument above it would equally well suffice that $f(\mathfrak{A}) \leq \mathfrak{C}$ for some structure $\mathfrak{C}$ that is obtained by applying p_B repeatedly to $\mathfrak{B}$. In the next definition we formalize the idea of such repeated applications.

▶ **Definition 14.** *Let $p : A^n \to A$ be a partial function, and let $R \subseteq A^r$. We define a sequence $\Gamma^i_p(R)$, $i \in \omega$, of r-ary relations on A by the following recursion:*

- $\Gamma^0_p(R) := R$; $\Gamma^{i+1}_p(R) := p(\Gamma^i_p(R)) \cup \Gamma^i_p(R)$.

Furthermore, we define $\Gamma^{\omega}_p(R) = \bigcup_{i \in \omega} \Gamma^i_p(R)$.

This is generalized to τ-structures in the natural way: for all $i \in \omega \cup \{\omega\}$, $\Gamma^i_p(\mathfrak{A})$ is the τ-structure $\mathfrak{C}$ such that $C = A$ and $R^{\mathfrak{C}} := \Gamma^i_p(R^{\mathfrak{A}})$ for each $R \in \tau$.

Note that since $\Gamma^i_p(R) \subseteq \Gamma^{i+1}_p(R)$ for all $i \in \omega$ (assuming A is finite) there exists $j \leq |A^r|$ such that $\Gamma^{\omega}_p(R) = \Gamma^j_p(R)$. Similarly for any finite structure $\mathfrak{A}$, $\Gamma^{\omega}_p(\mathfrak{A}) = \Gamma^j_p(\mathfrak{A})$ for some $j \leq |A^r|$, where r is the maximum arity of relations in $\mathfrak{A}$.

▶ **Lemma 15.** *Let $\mathcal{P}$ be a family of n-ary partial functions. A quantifier is $\mathcal{P}$-closed if and only if the implication*

$$\mathfrak{B} \in \mathcal{K} \text{ and } \mathfrak{A} \leq \Gamma^{\omega}_{p_A}(\mathfrak{B}) \implies \mathfrak{A} \in \mathcal{K}$$

holds for all structures $\mathfrak{A}$ and $\mathfrak{B}$ with $A = B$.

Proof. Assume first that $Q_{\mathcal{K}}$ is $\mathcal{P}$-closed, $\mathfrak{B} \in \mathcal{K}$ and $\mathfrak{A} \leq \Gamma^{\omega}_{p_A}(\mathfrak{B})$. We show first by induction on i that $\Gamma^{i}_{p_A}(\mathfrak{B}) \in \mathcal{K}$ for all $i \in \omega$. For $i = 0$ this holds by assumption. If $\Gamma^{i}_{p_A}(\mathfrak{B}) \in \mathcal{K}$, then $\Gamma^{i+1}_{p_A}(\mathfrak{B}) = p_A(\mathfrak{C}) \cup \mathfrak{C}$, for $\mathfrak{C} = \Gamma^{i}_{p_A}(\mathfrak{B})$, and hence $\Gamma^{i+1}_{p_A}(\mathfrak{B}) \in \mathcal{K}$ follows from the assumption that $Q_{\mathcal{K}}$ is $\mathcal{P}$-closed.

As noted above, there exists $j \in \omega$ such that $\Gamma^{\omega}_{p_A}(\mathfrak{B}) = \Gamma^{j}_{p_A}(\mathfrak{B})$. Thus we have $\mathfrak{A} \leq \Gamma^{j}_{p_A}(\mathfrak{B}) \leq \Gamma^{j+1}_{p_A}(\mathfrak{B}) = p_A(\Gamma^{j}_{p_A}(\mathfrak{B})) \cup \Gamma^{j}_{p_A}(\mathfrak{B})$. Since $\Gamma^{j}_{p_A}(\mathfrak{B}) \in \mathcal{K}$ and $\mathcal{K}$ is $\mathcal{P}$-closed, it follows that $\mathfrak{A} \in \mathcal{K}$.

Assume then that the implication

$$(*) \quad \mathfrak{B} \in \mathcal{K} \text{ and } \mathfrak{A} \leq \Gamma^{\omega}_{p_A}(\mathfrak{B}) \implies \mathfrak{A} \in \mathcal{K}$$

holds for all $\mathfrak{A}$ and $\mathfrak{B}$ with $A = B$. Assume further that $\mathfrak{B} \in \mathcal{K}$ and $\mathfrak{A} \leq p_A(\mathfrak{B}) \cup \mathfrak{B}$. By definition $p_A(\mathfrak{B}) \cup \mathfrak{B} = \Gamma^{1}_{p_A}(\mathfrak{B})$, and since $\Gamma^{1}_{p_A}(\mathfrak{B}) \leq \Gamma^{\omega}_{p_A}(\mathfrak{B})$, we have $\mathfrak{A} \leq \Gamma^{\omega}_{p_A}(\mathfrak{B})$. Thus $\mathfrak{A} \in \mathcal{K}$ follows from the implication $(*)$. ◀

We are now ready to give the formal definition of our pebble game for $\mathcal{P}$-closed quantifiers. Let k be a positive integer. Assume that $\mathfrak{A}$ and $\mathfrak{B}$ are τ-structures for a relational vocabulary τ. Furthermore, assume that α and β are assignments on $\mathfrak{A}$ and $\mathfrak{B}$, respectively, such that $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta) \subseteq X$, where $X = \{x_1, \ldots, x_k\}$. The *k-pebble $\mathcal{P}$ game* for $(\mathfrak{A}, \alpha)$ and $(\mathfrak{B}, \beta)$ is played between *Spoiler* and *Duplicator*. We denote the game by $\mathrm{PG}^{\mathcal{P}}_k(\mathfrak{A}, \mathfrak{B}, \alpha, \beta)$, and we use the shorthand notation $\mathrm{PG}^{\mathcal{P}}_k(\alpha, \beta)$ whenever $\mathfrak{A}$ and $\mathfrak{B}$ are clear from the context.

▶ **Definition 16.** *The rules of the game $\mathrm{PG}^{\mathcal{P}}_k(\mathfrak{A}, \mathfrak{B}, \alpha, \beta)$ are the following:*

(1) *If $\alpha \mapsto \beta \notin \mathrm{PI}(\mathfrak{A}, \mathfrak{B})$, then the game ends, and Spoiler wins.*

(2) *If (1) does not hold, there are two types of moves that Spoiler can choose to play:*

- **Left $\mathcal{P}$-quantifier move:** *Spoiler starts by choosing $r \in [k]$ and an r-tuple $\vec{y} \in X^r$ of distinct variables. Duplicator responds with a bijection $f \colon B \to A$. Spoiler answers by choosing an r-tuple $\vec{b} \in B^r$. Duplicator answers by choosing $P \subseteq A^r$ such that $f(\vec{b}) \in \Gamma^{\omega}_{p_A}(P)$. Spoiler completes the round by choosing $\vec{a} \in P$. The players continue by playing $\mathrm{PG}^{\mathcal{P}}_k(\alpha', \beta')$, where $\alpha' := \alpha[\vec{a}/\vec{y}]$ and $\beta' := \beta[\vec{b}/\vec{y}]$.*
- **Right $\mathcal{P}$-quantifier move:** *Spoiler starts by choosing $r \in [k]$ and an r-tuple $\vec{y} \in X^r$ of distinct variables. Duplicator chooses next a bijection $f \colon A \to B$. Spoiler answers by choosing an r-tuple $\vec{a} \in A^r$. Duplicator answers by choosing $P \subseteq B^r$ such that $f(\vec{a}) \in \Gamma^{\omega}_{p_B}(P)$. Spoiler completes the round by choosing $\vec{b} \in P$. The players continue by playing $\mathrm{PG}^{\mathcal{P}}_k(\alpha', \beta')$, where $\alpha' := \alpha[\vec{a}/\vec{y}]$ and $\beta' := \beta[\vec{b}/\vec{y}]$.*

(3) *Duplicator wins the game if Spoiler does not win it in a finite number of rounds.*

We now prove that the game $\mathrm{PG}^{\mathcal{P}}_k$ indeed characterizes equivalence of structures with respect to the infinitary k-variable logic with all $\mathcal{P}$-closed quantifiers.

▶ **Theorem 17.** *Let $\mathcal{P}$ be an invariant family of partial functions. Then Duplicator has a winning strategy in $\mathrm{PG}^{\mathcal{P}}_k(\mathfrak{A}, \mathfrak{B}, \alpha, \beta)$ if, and only if, $(\mathfrak{A}, \alpha) \equiv^{k}_{\infty\omega,\mathcal{P}} (\mathfrak{B}, \beta)$.*

Proof.

$\Rightarrow$: We prove by induction on $\varphi \in L^{k}_{\infty\omega}(\mathbf{Q}_{\mathcal{P}})$ that (for any assignments α and β) if Duplicator has a winning strategy in $\mathrm{PG}^{\mathcal{P}}_k(\alpha, \beta)$, then $(\mathfrak{A}, \alpha) \models \varphi \iff (\mathfrak{B}, \beta) \models \varphi$.

- If φ is an atomic formula, the claim follows from the fact that Spoiler always wins the game $\mathrm{PG}_k^{\mathcal{P}}(\alpha,\beta)$ immediately if $\alpha \mapsto \beta \notin \mathrm{PI}(\mathfrak{A},\mathfrak{B})$.
- The cases $\varphi = \neg\psi$, $\varphi = \bigvee \Psi$ and $\varphi = \bigwedge \Psi$ are straightforward.
- By Proposition 6, the negation of the existential quantifier is in $\mathbf{Q}_{\mathcal{P}}$, and hence we do not need to consider the case $\varphi = \exists x_i \psi$ separately.
- Consider then the case $\varphi = Q_{\mathcal{K}}\vec{y}\mathcal{I}$ for some r-ary quantifier $Q_{\mathcal{K}} \in \mathbf{Q}_{\mathcal{P}}$ and interpretation $\mathcal{I} = (\psi_1,\ldots,\psi_\ell)$. We start by assuming that $(\mathfrak{A},\alpha) \models \varphi$. Thus, $\mathcal{I}(\mathfrak{A},\alpha) := (A,R_1,\ldots,R_\ell) \in \mathcal{K}$. Let Spoiler play in the game $\mathrm{PG}_k^{\mathcal{P}}(\alpha,\beta)$ a left $\mathcal{P}$-quantifier move with r and the tuple $\vec{y} \in X^r$, and let $f\colon B \to A$ be the bijection given by the winning strategy of Duplicator. Let $\mathcal{I}(\mathfrak{B},\beta) := (B,R_1',\ldots,R_\ell')$, and for each $i \in [\ell]$, let $S_i := f(R_i')$. We claim that $\mathfrak{D} := (A,S_1,\ldots,S_\ell) \in \mathcal{K}$. Since f is an isomorphism $\mathcal{I}(\mathfrak{B},\beta) \to \mathfrak{D}$, it follows then that $(\mathfrak{B},\beta) \models \varphi$.
  To prove the claim it suffices to show that $\mathfrak{D} \le \Gamma^{\omega}_{p_A}(\mathcal{I}(\mathfrak{A},\alpha))$, since then $\mathfrak{D} \in \mathcal{K}$ by Lemma 15 and the assumption that $Q_{\mathcal{K}}$ is $\mathcal{P}$-closed. To show this, let $i \in [\ell]$ and $\vec{c} \in S_i$. We let Spoiler choose the tuple $\vec{b} = f^{-1}(\vec{c})$ as his answer to the bijection f. Thus, $(\mathfrak{B},\beta[\vec{b}/\vec{y}]) \models \psi_i$. Let $P \subseteq A^r$ be the answer of Duplicator. Then by the rules of the game $\vec{c} \in \Gamma^{\omega}_{p_A}(P)$, and Duplicator has a winning strategy in the game $\mathrm{PG}_k^{\mathcal{P}}(\alpha[\vec{a}/\vec{y}],\beta[\vec{b}/\vec{y}])$ for all $\vec{a} \in P$. Hence by induction hypothesis $(\mathfrak{A},\alpha[\vec{a}/\vec{y}]) \models \psi_i$, i.e., $\vec{a} \in R_i$, holds for all $\vec{a} \in P$. This shows that $S_i \subseteq \Gamma^{\omega}_{p_A}(R_i)$, and since this holds for all $i \in [\ell]$, we see that $\mathfrak{D} \le \Gamma^{\omega}_{p_A}(\mathcal{I}(\mathfrak{A},\alpha))$.
  By using the right $\mathcal{P}$-quantifier move in place of the left quantifier move, we can prove that $(\mathfrak{B},\beta) \models \varphi$ implies $(\mathfrak{A},\alpha) \models \varphi$. Thus, $(\mathfrak{A},\alpha) \models \varphi \iff (\mathfrak{B},\beta) \models \varphi$, as desired.

$\Leftarrow$: Assume then that $(\mathfrak{A},\alpha) \equiv^k_{\infty\omega,\mathcal{P}} (\mathfrak{B},\beta)$. Clearly it suffices to show that Duplicator can play in the first round of the game $\mathrm{PG}_k^{\mathcal{P}}(\alpha,\beta)$ in such a way that $(\mathfrak{A},\alpha') \equiv^k_{\infty\omega,\mathcal{P}} (\mathfrak{B},\beta')$ holds, where α' and β' are the assignments arising from the choices of Spoiler and Duplicator.

Assume first that Spoiler decides to play a left $\mathcal{P}$-quantifier move in the first round of $\mathrm{PG}_k^{\mathcal{P}}(\alpha,\beta)$. Let $\vec{y} \in X^r$ be the tuple of variables he chooses. Since A and B are finite, for each $\vec{a} \in A^r$ there is a formula $\Psi_{\vec{a}} \in L^k_{\infty\omega}(\mathbf{Q}_{\mathcal{P}})$ such that for any τ-structure $\mathfrak{C}$ of size at most $\max\{|A|,|B|\}$, any assignment γ on $\mathfrak{C}$, and any tuple $\vec{c} \in C^r$ we have

- $(\mathfrak{A},\alpha[\vec{a}/\vec{y}]) \equiv^k_{\infty\omega,\mathcal{P}} (\mathfrak{C},\gamma[\vec{c}/\vec{y}])$ if and only if $(\mathfrak{C},\gamma[\vec{c}/\vec{y}]) \models \Psi_{\vec{a}}$.

Let $\vec{c}_1,\ldots,\vec{c}_\ell$ be a fixed enumeration of the set A^r, and let $\mathcal{I}$ be the interpretation $(\Psi_1,\ldots,\Psi_m)$, where $\Psi_j := \Psi_{\vec{c}_j}$ for each $j \in [m]$. We define $\mathcal{K}$ to be the closure of the class $\{\mathfrak{D} \mid \mathfrak{D} \le \Gamma^{\omega}_{p_A}(\mathcal{I}(\mathfrak{A},\alpha))\}$ under isomorphisms. Note that if $\mathfrak{D} \le \Gamma^{\omega}_{p_A}(\mathcal{I}(\mathfrak{A},\alpha))$ and $\mathfrak{E} \le \Gamma^{\omega}_{p_A}(\mathfrak{D})$, then clearly $\mathfrak{E} \le \Gamma^{\omega}_{p_A}(\mathcal{I}(\mathfrak{A},\alpha))$. Hence by Lemma 15, the quantifier $Q_{\mathcal{K}}$ is $\mathcal{P}$-closed. Moreover, since $\mathcal{I}(\mathfrak{A},\alpha) \in \mathcal{K}$, we have $(\mathfrak{A},\alpha) \models Q_{\mathcal{K}}\vec{y}\mathcal{I}$, and consequently by our assumption, $(\mathfrak{B},\beta) \models Q_{\mathcal{K}}\vec{y}\mathcal{I}$. Thus, there is a structure $\mathfrak{D} \le \Gamma^{\omega}_{p_A}(\mathcal{I}(\mathfrak{A},\alpha))$ and an isomorphism $f\colon \mathcal{I}(\mathfrak{B},\beta) \to \mathfrak{D}$. We let Duplicator to use the bijection $f\colon B \to A$ as her answer to the choice $\vec{y}$ of Spoiler.

Let $\vec{b} \in B^r$ be the answer of Spoiler to f, and let $\vec{c} = f(\vec{b})$. Clearly $(\mathfrak{A},\alpha) \models \forall\vec{y} \bigvee_{j\in[\ell]} \Psi_j$, whence there exists $j \in [\ell]$ such that $(\mathfrak{B},\beta[\vec{b}/\vec{y}]) \models \Psi_j$, or in other words, $\vec{b} \in R_j^{\mathcal{I}(\mathfrak{B},\beta)}$. Since f is an isomorphism $\mathcal{I}(\mathfrak{B},\beta) \to \mathfrak{D}$, we have $\vec{c} \in R_j^{\mathfrak{D}}$. We let Duplicator to use $P := R_j^{\mathcal{I}(\mathfrak{A},\alpha)}$ as her answer to the choice $\vec{b}$ of Spoiler; this is a legal move since $\mathfrak{D} \le \Gamma^{\omega}_{p_A}(\mathcal{I}(\mathfrak{A},\alpha))$. Observe now that since $P = R_j^{\mathcal{I}(\mathfrak{A},\alpha)}$, we have $(\mathfrak{A},\alpha[\vec{a}/\vec{y}]) \models \Psi_{\vec{c}_j}$, and consequently $(\mathfrak{A},\alpha[\vec{c}_j/\vec{y}]) \equiv^k_{\infty\omega,\mathcal{P}} (\mathfrak{A},\alpha[\vec{a}/\vec{y}])$, for all $\vec{a} \in P$. On the other hand we also have $(\mathfrak{B},\beta[\vec{b}/\vec{y}]) \models \Psi_{\vec{c}_j}$, and hence $(\mathfrak{A},\alpha[\vec{c}_j/\vec{y}]) \equiv^k_{\infty\omega,\mathcal{P}} (\mathfrak{B},\beta[\vec{b}/\vec{y}])$. Thus the condition $(\mathfrak{A},\alpha') \equiv^k_{\infty\omega,\mathcal{P}} (\mathfrak{B},\beta')$, where $\alpha' = \alpha[\vec{a}/\vec{y}]$ and $\beta' = \beta[\vec{b}/\vec{y}]$, holds after the first round of $\mathrm{PG}_k^{\mathcal{P}}(\alpha,\beta)$ irrespective of the choice $\vec{a} \in P$ of Spoiler in the end of the round.

The case where Spoiler starts with a right $\mathcal{P}$-quantifier move is handled in the same way by switching the roles of $(\mathfrak{A}, \alpha)$ and $(\mathfrak{B}, \beta)$. ◀

5 Playing the game

In this section we use the game $\mathrm{PG}^k_{\mathcal{P}}$ to show inexpressibility of a property of finite structures in the infinitary finite variable logic $L^\omega_{\infty\omega}$ augmented by all $\mathcal{N}_\ell$-closed quantifiers. More precisely, we prove that the Boolean constraint satisfaction problem $\mathsf{CSP}(\mathfrak{C}_\ell)$, where $\mathfrak{C}_\ell$ is the structure with $C = \{0, 1\}$ and two ℓ-ary relations $R_0 = \{(b_1, \ldots, b_\ell) \mid \sum_{i\in[\ell]} b_i \equiv 0 \pmod 2\}$ and $R_1 = \{(b_1, \ldots, b_\ell) \mid \sum_{i\in[\ell]} b_i \equiv 1 \pmod 2\}$, is not definable in $L^\omega_{\infty\omega}(\mathbf{Q}_{\mathcal{N}_\ell})$.

In the proof of the undefinability of $\mathsf{CSP}(\mathfrak{C}_\ell)$ we use a variation of the well-known CFI construction, due to Cai, Fürer and Immerman [6]. Our construction is a minor modification of the one that was used in [17] for producing non-isomorphic structures on which Duplicator wins the k, n-bijection game. We start by explaining the details of the construction.

Let $G = (V, E, \leq^G)$ be a connected ℓ-regular ordered graph. For each vertex $v \in V$, we use the notation $E(v)$ for the set of edges adjacent to v and $\vec{e}(v) = (e_1, \ldots, e_\ell)$ for the tuple that lists $E(v)$ in the order $\leq^G$. The CFI structures we use have in the universe two elements $(e, 1)$ and $(e, 2)$ for each $e \in E$, and two ℓ-ary relations that connect such pairs (e, i) for edges e that are adjacent to some vertex $v \in V$.

▶ **Definition 18.** *Let $G = (V, E, \leq^G)$ be a connected ℓ-regular ordered graph and let $U \subseteq V$. We define a CFI structure $\mathfrak{A}_\ell(G, U) = (A_\ell(G), R_0^{\mathfrak{A}_\ell(G,U)}, R_1^{\mathfrak{A}_\ell(G,U)})$, where* $\mathrm{ar}(R_0) = \mathrm{ar}(R_1) = \ell$*, as follows.*
- $A_\ell(G) := E \times [2]$,
- $R_0^{\mathfrak{A}_\ell(G,U)} := \bigcup_{v\in V\setminus U} R(v) \cup \bigcup_{v\in U} \tilde{R}(v)$ *and* $R_1^{\mathfrak{A}_\ell(G,U)} := \bigcup_{v\in U} R(v) \cup \bigcup_{v\in V\setminus U} \tilde{R}(v)$, *where*
  - $R(v) := \{((e_1, i_1), \ldots, (e_\ell, i_\ell)) \mid (e_1, \ldots, e_\ell) = \vec{e}(v), \sum_{j\in[\ell]} i_j = 0 \pmod 2\}$, *and*
  - $\tilde{R}(v) := \{((e_1, i_1), \ldots, (e_\ell, i_\ell)) \mid (e_1, \ldots, e_\ell) = \vec{e}(v), \sum_{j\in[\ell]} i_j = 1 \pmod 2\}$.

For each $v \in V$, we denote the set $E(v) \times [2]$ by $A(v)$. Furthermore, we define $\mathfrak{A}_\ell(v) := (A(v), R(v), \tilde{R}(v))$ and $\tilde{\mathfrak{A}}_\ell(v) := (A(v), \tilde{R}(v), R(v))$.

By a similar argument as in the CFI structures constructed in [17] and [18] it can be proved that $\mathfrak{A}_\ell(G, U)$ and $\mathfrak{A}_\ell(G, U')$ are isomorphic if and only if $|U|$ and $|U'|$ are of the same parity. We choose $\mathfrak{A}^{\mathrm{ev}}_\ell(G) := \mathfrak{A}_\ell(G, \emptyset)$ and $\mathfrak{A}^{\mathrm{od}}_\ell(G) := \mathfrak{A}_\ell(G, \{v_0\})$ as representatives of these two isomorphism classes, where v_0 is the least element of V with respect to the linear order $\leq^G$. We show first that these structures are separated by $\mathsf{CSP}(\mathfrak{C}_\ell)$.

▶ **Lemma 19.** *$\mathfrak{A}^{\mathrm{ev}}_\ell(G) \in \mathsf{CSP}(\mathfrak{C}_\ell)$, but $\mathfrak{A}^{\mathrm{od}}_\ell(G) \notin \mathsf{CSP}(\mathfrak{C}_\ell)$.*

Proof. Let $h \colon A_\ell(G) \to \{0, 1\}$ be the function such that $h((e, 1)) = 1$ and $h((e, 2)) = 0$ for all $e \in E$. Then for any tuple $((e_1, i_1), \ldots, (e_\ell, i_\ell))$ the parity of $\sum_{j\in[\ell]} h((e_j, i_j))$ is the same as the parity of $\sum_{j\in[\ell]} i_j$. Thus, h is a homomorphism $\mathfrak{A}^{\mathrm{ev}}_\ell(G) \to \mathfrak{C}_\ell$.

To show that $\mathfrak{A}^{\mathrm{od}}_\ell(G) \notin \mathsf{CSP}(\mathfrak{C}_\ell)$, assume towards contradiction that $g \colon A_\ell(G) \to \{0, 1\}$ is a homomorphism $\mathfrak{A}^{\mathrm{od}}_\ell(G) \to \mathfrak{C}_\ell$. Then for every $e \in E$ necessarily $g((e, 1)) \neq g((e, 2))$. Furthermore, for every $v \in V \setminus \{v_0\}$, the number $n_v := |\{e \in E(v) \mid g((e, 2)) = 1\}|$ must be even, while the number n_{v_0} must be odd. Thus, $\sum_{v\in V} n_v$ must be odd. However, this is impossible, since clearly $\sum_{v\in V} n_v = 2|\{e \in E \mid g((e, 2)) = 1\}|$. ◀

Our aim is to prove, for a suitable graph G, that Duplicator has a winning strategy in $\mathrm{PG}^{\mathcal{N}_\ell}_k(\mathfrak{A}^{\mathrm{ev}}_\ell(G), \mathfrak{A}^{\mathrm{od}}_\ell(G), \emptyset, \emptyset)$. For the winning strategy, Duplicator needs a collection of well-behaved bijections. We define such a collection GB in Definition 23 below. One requirement is that the bijections preserve the first component of the elements $(e, i) \in A_\ell(G)$.

▶ **Definition 20.** *A bijection* $f\colon A_\ell(G) \to A_\ell(G)$ *is* edge preserving *if for every* $e \in E$ *and* $i \in [2]$, $f((e,i))$ *is either* $(e,1)$ *or* $(e,2)$.

For any edge preserving f and any $v \in V$ we denote by f_v the restriction of f to the set $E(v) \times [2]$. The *switching number* $\mathrm{swn}(f_v)$ of f_v is $|\{e \in E(v) \mid f_v((e,1)) = (e,2)\}|$. The lemma below follows directly from the definitions of $\mathfrak{A}_\ell(v)$ and $\tilde{\mathfrak{A}}_\ell(v)$.

▶ **Lemma 21.** *Let* $f\colon A_\ell(G) \to A_\ell(G)$ *be an edge preserving bijection, and let* $v \in V$.
(a) *If* $\mathrm{swn}(f_v)$ *is even, then* f_v *is an automorphism of the structures* $\mathfrak{A}_\ell(v)$ *and* $\tilde{\mathfrak{A}}_\ell(v)$.
(b) *If* $\mathrm{swn}(f_v)$ *is odd, then* f_v *is an isomorphism between the structures* $\mathfrak{A}_\ell(v)$ *and* $\tilde{\mathfrak{A}}_\ell(v)$.

Given an edge preserving bijection $f\colon A_\ell(G) \to A_\ell(G)$ we denote by $\mathrm{Odd}(f)$ the set of all $v \in V$ such that $\mathrm{swn}(f_v)$ is odd. Observe that $|\mathrm{Odd}(f)|$ is necessarily even.

▶ **Corollary 22.** *An edge preserving bijection* $f\colon A_\ell(G) \to A_\ell(G)$ *is an automorphism of the structures* $\mathfrak{A}^{\mathrm{ev}}_\ell(G)$ *and* $\mathfrak{A}^{\mathrm{od}}_\ell(G)$ *if and only if* $\mathrm{Odd}(f) = \emptyset$.

Proof. If $\mathrm{Odd}(f) = \emptyset$, then by Lemma 21(a) f_v is an automorphism of $\mathfrak{A}_\ell(v)$ and $\tilde{\mathfrak{A}}_\ell(v)$ for all $v \in V$. Clearly this means that f is an automorphism of $\mathfrak{A}^{\mathrm{ev}}_\ell(G)$ and $\mathfrak{A}^{\mathrm{od}}_\ell(G)$. On the other hand, if $v \in \mathrm{Odd}(f)$, then by Lemma 21(b), for any tuple $\vec{a} \in A(v)^\ell$, we have $\vec{a} \in R(v) \iff f(\vec{a}) \in \tilde{R}(v)$. Since $R(v) \cap \tilde{R}(v) = \emptyset$, it follows that f is not an automorphism of $\mathfrak{A}^{\mathrm{ev}}_\ell(G)$ and $\mathfrak{A}^{\mathrm{od}}_\ell(G)$. ◀

▶ **Definition 23.** *Let* $f\colon A_\ell(G) \to A_\ell(G)$ *be edge preserving bijection. Then* f *is a* good bijection *if either* $\mathrm{Odd}(f) = \emptyset$ *or* $\mathrm{Odd}(f) = \{v_0, v\}$ *for some* $v \in V \setminus \{v_0\}$. *We denote the set of all good bijections by* GB.

Note that if $f\colon A_\ell(G) \to A_\ell(G)$ is a good bijection, then there is exactly one vertex $v^* \in V$ such that f_{v^*} is not a partial isomorphism $\mathfrak{A}^{\mathrm{ev}}_\ell(G) \to \mathfrak{A}^{\mathrm{od}}_\ell(G)$. In case $\mathrm{Odd}(f) = \emptyset$, $v^* = v_0$, while in case $\mathrm{Odd}(f) = \{v_0, v\}$ for some $v \neq v_0$, $v^* = v$. We denote this vertex v^* by $\mathrm{tw}(f)$ (the *twist* of f).

Assume now that Duplicator has played a good bijection f in the game $\mathrm{PG}^{\mathcal{N}_\ell}_k$ on the structures $\mathfrak{A}^{\mathrm{ev}}_\ell(G)$ and $\mathfrak{A}^{\mathrm{od}}_\ell(G)$. Then it is sure that Spoiler does not win the game in the next position (α, β) if $(e,1)$ and $(e,2)$ are not in the range of α (and β) for any $e \in E(\mathrm{tw}(f))$. This leads us to the following notion.

▶ **Definition 24.** *Let* f *be a* good *bijection, and let* $F \subseteq E$. *Then* f *is* good for F *if* $E(\mathrm{tw}(f)) \cap F = \emptyset$. *We denote the set of all bijections that are good for* F *by* $\mathrm{GB}(F)$.

▶ **Lemma 25.** *If* $f \in \mathrm{GB}(F)$, *then* $f \restriction (F \times [2])$ *is a partial isomorphism* $\mathfrak{A}^{\mathrm{ev}}_\ell(G) \to \mathfrak{A}^{\mathrm{od}}_\ell(G)$.

Proof. Clearly $f \restriction (F \times [2]) \subseteq \bigcup_{v \in V \setminus \{\mathrm{tw}(f)\}} f_v$. By Lemma 21, f_v is an automorphism of $\mathfrak{A}_\ell(v)$ for any $v \in V \setminus \{\mathrm{tw}(f), v_0\}$, and if $v_0 \neq \mathrm{tw}(f)$, f_{v_0} is an isomorphism $\mathfrak{A}_\ell(v) \to \tilde{\mathfrak{A}}_\ell(v)$. The claim follows from this. ◀

Given a good bijection f with $\mathrm{tw}(f) = u$ and an E-path $P = (u_0, \ldots, u_m)$ from $u = u_0$ to $u' = u_m$, we obtain a new edge preserving bijection f_P by switching f on the edges $e_i := \{u_i, u_{i+1}\}$, $i < m$, of P: $f_P((e_i, j)) = (e_i, 3 - j)$ for $i < m$, and $f_P(c) = f(c)$ for other $c \in A_\ell(G)$. Clearly f_P is also a good bijection, and $\mathrm{tw}(f_P) = u'$.

In order to prove that Duplicator has a winning strategy in $\mathrm{PG}^{\mathcal{N}_\ell}_k(\mathfrak{A}^{\mathrm{ev}}_\ell(G), \mathfrak{A}^{\mathrm{od}}_\ell(G), \emptyset, \emptyset)$ we need to assume that the graph G has a certain largeness property with respect the number k. We formulate this largeness property in terms of a game, $\mathrm{CRG}^\ell_k(G)$, that is a new variation of the *Cops&Robber games* used for similar purposes in [17] and [18].

▶ **Definition 26.** *The game* $\mathrm{CRG}_k^\ell(G)$ *is played between two players, Cop and Robber. The positions of the game are pairs* (F, u)*, where* $F \subseteq E$*,* $|F| \leq k$*, and* $u \in V$*. The rules of the game are the following:*

- *Assume that the position is* (F, u)*.*
- *If* $E(u) \cap F \neq \emptyset$*, the game ends and Cop wins.*
- *Otherwise Cop chooses a set* $F' \subseteq E$ *such that* $|F'| \leq k$*. Then Robber answers by giving mutually disjoint* $E \setminus (F \cap F')$*-paths* $P_i = (u, u_1^i, \ldots, u_{n_i}^i)$*,* $i \in [\ell]$*, from* u *to vertices* $u_i := u_{n_i}^i$*; here mutual disjointness means that* P_i *and* $P_{i'}$ *do not share edges for* $i \neq i'$ *(i.e.,* $u_1^i \neq u_1^{i'}$ *and* $\{u_j^i, u_{j+1}^i\} \neq \{u_{j'}^{i'}, u_{j'+1}^{i'}\}$ *for all* j *and* j'*). Then Cop completes the round by choosing* $i \in [\ell]$*. The next position is* (F', u_i)*.*

The intuition of the game $\mathrm{CRG}_k^\ell(G)$ is that Cop has k pebbles that he plays on edges of G forming a set $F \subseteq E$; these pebbles mark the edges e such that $(e, 1)$ or $(e, 2)$ is in the range of α or β in a position (α, β) of the game $\mathrm{PG}_k^{\mathcal{N}_\ell}$ on $\mathfrak{A}_\ell^{\mathrm{ev}}(G)$ and $\mathfrak{A}_\ell^{\mathrm{od}}(G)$. Robber has one pebble that she plays on the vertices of G; this pebble marks the vertex $\mathrm{tw}(f)$, where f is the good bijection played by Duplicator in the previous round of $\mathrm{PG}_k^{\mathcal{N}_\ell}$.

Cop captures Robber and wins the game if after some round (at least) one of his pebbles is on an edge that is adjacent to the vertex containing Robber's pebble. This corresponds to a position (α, β) in the game $\mathrm{PG}_k^{\mathcal{N}_\ell}$ such that $\alpha \mapsto \beta$ is potentially not a partial isomorphism. Otherwise Lemma 25 guarantees that $\alpha \mapsto \beta$ is a partial isomorphism. Cop can then move any number of his pebbles to new positions on G. While the pebbles Cop decides to move are still on their way to their new positions, Robber is allowed to prepare ℓ mutually disjoint escape routes along edges of G that do not contain any stationary pebbles of Cop. We show in the proof of Theorem 29 that these escape routes generate tuples $\vec{a}_1, \ldots, \vec{a}_\ell$ such that $f(\vec{b}) = \hat{q}(\vec{a}_1, \ldots, \vec{a}_\ell)$, where $q = n_{A_\ell(G)}^\ell$ and $\vec{b}$ is the tuple chosen by Spoiler after Duplicator played f. This gives Duplicator a legal answer $P = \{\vec{a}_1, \ldots, \vec{a}_\ell\}$ to $\vec{b}$. Then Spoiler completes the round by choosing one of the tuples in P. Correspondingly, in the end of the round of $\mathrm{CRG}_k^\ell(G)$ Cop chooses which escape route Robber has to use by blocking all but one of them.

▶ **Definition 27.** *Assume that* $u \in V$ *and* $F \subseteq E$ *is a set of edges such that* $|F| \leq k$*. We say that* u *is* k-safe for F *if Robber has a winning strategy in the game* $\mathrm{CRG}_k^\ell(G)$ *starting from position* (F, u)*.*

We prove next the existence of graphs G such that Robber has a winning strategy in the game $\mathrm{CRG}_k^\ell(G)$.

▶ **Theorem 28.** *For every* $\ell \geq 3$ *and every* $k \geq 1$*, there is an* ℓ*-regular graph* $G = (V, E)$ *such that every vertex* $v \in V$ *is* k*-safe for* $\emptyset$*.*

Proof. Clearly if Robber has a winning strategy in $\mathrm{CRG}_k^\ell(G)$, it also has a winning strategy in $\mathrm{CRG}_{k'}^\ell(G)$ for $k' < k$. Thus, it suffices to prove the theorem for $k \geq \ell$.

By a well-known result of Erdös and Sachs [12] (see also [1] for a more accessible construction), there exist ℓ-regular connected graphs of *girth* g for arbitrarily large g. Choose a positive integer d with $d > \frac{\log 2k}{\log(\ell-1)} + 1$ and let G be an ℓ-regular graph of girth $g > 6d$. We claim that any vertex v in G is k-safe for $\emptyset$.

To prove this, we show inductively that Robber can maintain the following invariant in any position (F, u) reached during the game:

$(*)$ for each edge $e \in F$, neither end point of e is within distance d of u in G.

Note that, from the assumption that $k \geq \ell$ and $d > \frac{\log 2k}{\log(\ell-1)} + 1$, it follows that $d \geq 2$. Thus, the invariant $(*)$ guarantees, in particular, that Cop does not win at any point.

Clearly the invariant $(*)$ is satisfied at the initial position, since F is empty. Suppose now that it is satisfied in some position (F,u) and Cop chooses a set F' in the next move. Let $C \subseteq V$ denote the set of end points of all edges in F'. Since $|F'| \leq k$, we have $|C| \leq 2k$.

Let $N \subseteq V$ denote the collection of vertices which are at distance at most $3d$ from u. By the assumption on the girth of G, the induced subgraph $G[N]$ is a tree. We can consider it as a rooted tree, with root u. Then, u has exactly ℓ children. All vertices in N at distance less than $3d$ from u have exactly $\ell-1$ children (and one parent), and all vertices at distance exactly $3d$ from u are leaves of the tree. This allows us to speak, for instance, of the subtree rooted at a vertex u' meaning the subgraph of G induced by the vertices x in N such that the unique path from u to x in $G[N]$ goes through u'.

Let $u_1,\dots,u_\ell$ be the children of u. For each i, let U_i denote the set of descendants of u_i that are at distance exactly d from u (and so at distance $d-1$ from u_i). Note that the collection $U_1,\dots,U_\ell$ forms a partition of the set of vertices in N that are at distance exactly d from u. Each $x \in U_i$ is the root of a tree of height $2d$. Moreover, since the tree below u_i is $(\ell-1)$-regular, U_i contains exactly $(\ell-1)^{d-1}$ vertices. By the assumption that $d > \frac{\log 2k}{\log(\ell-1)}+1$, it follows that $(\ell-1)^{d-1} > 2k \geq |C|$ and therefore each U_i contains at least one vertex x_i such that the subtree rooted at x_i contains no vertex in C. Let y_i be any descendant of x_i at distance d from x_i and let P_i denote the unique path in $G[N]$ from u to y_i. Robber's move is to play the paths $P_1,\dots,P_\ell$. We now verify that this is a valid move, and that it maintains the required invariant $(*)$.

First, note that the paths $P_1,\dots,P_\ell$ are paths in the tree $G[N]$ all starting at u and the second vertex in path P_i is u_i. It follows that the paths are pairwise edge disjoint. We next argue that no path P_i goes through an edge in $F \cap F'$. Indeed, by the inductive assumption, no endpoint of an edge in F appears within distance d of u and therefore the path from u to x_i does not go through any such vertex. Moreover, by the choice of x_i, no endpoint of an edge in F' appears in the subtree rooted at x_i and therefore the path from x_i to y_i does not go through any such vertex. Together these ensure that the path P_i does not visit any vertex that is an endpoint of an edge in $F \cap F'$.

Finally, to see that the invariant $(*)$ is maintained, note that all vertices that are at distance at most d from y_i are in the subtree of $G[N]$ rooted at x_i. The choice of x_i means this contains no vertex in C. This is exactly the condition that we wished to maintain. ◀

We are now ready to prove that a winning strategy for Robber in $\mathrm{CRG}^\ell_k(G)$ generates a winning strategy for Duplicator in the game $\mathrm{PG}^{\mathcal{N}_\ell}_k$ on the structures $\mathfrak{A}^{\mathrm{ev}}_\ell(G)$ and $\mathfrak{A}^{\mathrm{od}}_\ell(G)$.

▶ **Theorem 29.** *Let G be a connected ℓ-regular ordered graph. If v_0 is k-safe for the empty set, then Duplicator has a winning strategy in the game $\mathrm{PG}^{\mathcal{N}_\ell}_k(\mathfrak{A}^{\mathrm{ev}}_\ell(G),\mathfrak{A}^{\mathrm{od}}_\ell(G),\emptyset,\emptyset)$.*

Proof. We show that Duplicator can maintain the following invariant for all positions (α,β) obtained during the play of the game $\mathrm{PG}^{\mathcal{N}_\ell}_k(\mathfrak{A}^{\mathrm{ev}}_\ell(G),\mathfrak{A}^{\mathrm{od}}_\ell(G),\emptyset,\emptyset)$:

(†) There exists a bijection $f \in \mathrm{GB}(F_\alpha)$ such that $p := \alpha \mapsto \beta \subseteq f$ and $\mathrm{tw}(f)$ is k-safe for F_α, where $F_\alpha := \{e \in E \mid \mathrm{rng}(\alpha) \cap \{e\} \times [2] \neq \emptyset\}$.

Note that if (†) holds, then $p \subseteq f \restriction (F_\alpha \times [2])$ and, by Lemma 25, $f \restriction (F_\alpha \times [2]) \in \mathrm{PI}(\mathfrak{A}^{\mathrm{ev}}_\ell(G),\mathfrak{A}^{\mathrm{od}}_\ell(G))$, whence Spoiler does not win the game in position (α,β). Thus, maintaining the invariant (†) during the play guarantees a win for Duplicator.

Note first that (†) holds in the initial position $(\alpha,\beta) = (\emptyset,\emptyset)$ of the game: if $f_0 \in \mathrm{GB}$ is the bijection with $\mathrm{tw}(f_0) = v_0$, as $\emptyset \mapsto \emptyset = \emptyset \subseteq f_0$ and $\mathrm{tw}(f_0)$ is k-safe for $F_\emptyset = \emptyset$.

Assume then that (†) holds for a position (α,β), and assume that Spoiler plays a left $\mathcal{N}_\ell$-quantifier move by choosing $r \leq k$ and $\vec{y} \in X^r$. Duplicator answers this by giving the bijection f^{-1}. Let $\vec{b} = (b_1,\dots,b_r) \in A_\ell(G)^r$ be the second part of Spoiler's move, and let

F' be the set $\{e \in E \mid \mathrm{rng}(\beta[\vec{b}/\vec{y}]) \cap \{e\} \times [2] \neq \emptyset\}$. Since $\mathrm{tw}(f)$ is k-safe for F_α, there are mutually disjoint $E \setminus (F_\alpha \cap F')$-paths P_i, $i \in [\ell]$, from $\mathrm{tw}(f)$ to some vertices u_i that are k-safe for the set F'. Let f_{P_i}, $i \in [\ell]$, be the good bijections obtained from f as explained before Definition 24. Now Duplicator answers the move $\vec{b}$ of Spoiler by giving the set $P = \{\vec{a}_1, \ldots, \vec{a}_\ell\}$ of r-tuples, where $\vec{a}_i := f_{P_i}^{-1}(\vec{b})$ for each $i \in [\ell]$.

To see that this is a legal move, observe that since the paths P_i are disjoint, for each $j \in [r]$ there is at most one $i \in [\ell]$ such that $f_{P_i}^{-1}(b_j) \neq f^{-1}(b_j)$. Thus we have $\hat{q}(\vec{a}_1, \ldots, \vec{a}_\ell) = f^{-1}(\vec{b})$, and hence $f^{-1}(\vec{b}) \in q(P) \subseteq \Gamma_q^\omega(P)$ for $q = n^\ell_{A_\ell(G)}$, as required. Let Spoiler complete the round of the game by choosing $i \in [\ell]$; thus, the next position is $(\alpha', \beta') := (\alpha[\vec{a}_i/\vec{y}], \beta[\vec{b}/\vec{y}])$. It suffices now to show that (†) holds for the position (α', β') and the bijection $f' := f_{P_i}$.

Note first that $F_{\alpha'} = F'$, since clearly $\mathrm{rng}(\alpha[\vec{a}_i/\vec{y}]) \cap \{e\} \times [2] \neq \emptyset$ if, and only if, $\mathrm{rng}(\beta[\vec{b}/\vec{y}]) \cap \{e\} \times [2] \neq \emptyset$. Thus, $\mathrm{tw}(f') = u_i$ is k-safe for $F_{\alpha'}$. This implies that $f' \in \mathrm{GB}(F_{\alpha'})$, since otherwise by Definition 26, Cop would win the game $\mathrm{CRG}_k(G)$ immediately in position $(F_{\alpha'}, \mathrm{tw}(f'))$. It remains to show that $p' := \alpha' \mapsto \beta'$ is contained in f'. For all components a_i^j of $\vec{a}_i$ we have $p'(a_i^j) = b_j = f'(a_i^j)$ by definition of $\vec{a}_i$. On the other hand, for any element $a \in \mathrm{dom}(p') \setminus \{a_i^1, \ldots, a_i^r\}$ we have $p'(a) = p(a) = f(a)$. Furthermore, since the path P_i does not contain any edges in $F_\alpha \cap F_{\alpha'}$, we have $f' \restriction (F_\alpha \cap F_{\alpha'}) \times [2] = f \restriction (F_\alpha \cap F_{\alpha'}) \times [2]$, and since clearly $a \in (F_\alpha \cap F_{\alpha'}) \times [2]$, we see that $f'(a) = f(a)$. Thus, $p'(a) = f'(a)$.

The case where Spoiler plays a right $\mathcal{N}_\ell$-quantifier move is similar. ◀

Note that the vocabulary of the structures $\mathfrak{A}_\ell^{\mathrm{ev}}(G)$ and $\mathfrak{A}_\ell^{\mathrm{od}}(G)$ consists of two ℓ-ary relation symbols. The presence of at least ℓ-ary relations is actually necessary: Duplicator cannot have a winning strategy in $\mathrm{PG}_{\ell-1}^{\mathcal{N}_\ell}$ on structures containing only relations of arity less than ℓ, since by Corollary 10(b), all properties of such structures are definable in $L_{\infty\omega}^{\ell-1}(\mathbf{Q}_{\mathcal{N}_\ell})$.

From Lemma 19, Theorem 28 and Theorem 29, we immediately obtain the result.

▶ **Theorem 30.** *For any $\ell \geq 3$, $\mathsf{CSP}(\mathfrak{C}_\ell)$ is not definable in $L_{\infty\omega}^\omega(\mathbf{Q}_{\mathcal{N}_\ell})$.*

Note that $\mathsf{CSP}(\mathfrak{C}_\ell)$ corresponds to solving systems of linear equations over $\mathbb{Z}/2\mathbb{Z}$ with all equations containing (at most) ℓ variables. Thus, as a corollary we see that solvability of such systems of equations cannot be expressed in $L_{\infty\omega}^\omega(\mathbf{Q}_{\mathcal{N}_\ell})$ for any ℓ. Furthermore, since systems of linear equations over $\mathbb{Z}/2\mathbb{Z}$ can be solved in polynomial time, we see that the complexity class PTIME is not contained in $L_{\infty\omega}^\omega(\mathbf{Q}_{\mathcal{N}_\ell})$ for any ℓ.

Finally, note that since the class $\mathsf{CSP}(\mathfrak{C}_\ell)$ is downwards monotone, by Lemma 9 the quantifier $Q_{\mathsf{CSP}(\mathfrak{C}_\ell)}$ is $\mathcal{N}_{\ell+1}$-closed. Thus, we get the following hierarchy result for the near-unanimity families $\mathcal{N}_\ell$ with respect to the arity ℓ of the partial functions.

▶ **Theorem 31.** *For every $\ell \geq 3$ there is a quantifier in $\mathbf{Q}_{\mathcal{N}_{\ell+1}}$ which is not definable in $L_{\infty\omega}^\omega(\mathbf{Q}_{\mathcal{N}_\ell})$.*

6 Conclusion

We have introduced new methods, in the form of pebble games, for proving inexpressibility in logics extended with generalized quantifiers. There is special interest in proving inexpressibility in logics with quantifiers of unbounded arity. We introduced a general method of defining such collections inspired by the equational theories of polymorphisms arising in the study of constraint satisfaction problems. Perhaps surprisingly, while the collection of CSP that have near-unanimity polymorphisms is rather limited (as they all have bounded width), the collection of quantifiers with the corresponding closure property is much richer, including even CSP that are intractable. The pebble game gives a general method of proving inexpressibility that works for a wide variety of closure conditions. We were able to deploy it to prove that solvability of systems of equations over $\mathbb{Z}/2\mathbb{Z}$ is not definable using only quantifiers closed under near-unanimity conditions.

It would be interesting to use the pebble games we have defined to show undefinability with other collections of quantifiers closed under partial polymorphisms. Showing some class is not definable with quantifiers closed under partial Maltsev polymorphisms would be especially instructive. It would require using the pebble games with a construction that looks radically different from the CFI-like constructions most often used. This is because CFI constructions encode problems of solvability of equations over finite fields (or more generally finite rings), and all of these problems are Maltsev-closed.

References

1 N. Alon. Tools from higher algebra. In *Handbook of Combinatorics (Vol. 2)*, pages 1749–1783. MIT Press, 1996.

2 A. Atserias, A. Bulatov, and A. Dawar. Affine systems of equations and counting infinitary logic. *Theoretical Computer Science*, 410(18):1666–1683, 2009.

3 L. Barto, A. A. Krokhin, and R. Willard. Polymorphisms, and how to use them. In A. A. Krokhin and S. Zivný, editors, *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*, pages 1–44. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. doi:10.4230/DFU.Vol7.15301.1.

4 V. G. Bodnarchuk, L. A. Kaluzhnin, V. N. Kotov, and B. A. Romov. Galois theory for post algebras. i. *Cybernetics*, 5(3):243–252, 1969. doi:10.1007/BF01070906.

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

6 J-Y. Cai, M. Fürer, and N. Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, 1992.

7 A. Dawar, E. Grädel, and M. Lichter. Limitations of the invertible-map equivalences. *arXiv*, abs/2109.07218, 2021. arXiv:2109.07218.

8 A. Dawar, E. Grädel, and W. Pakusa. Approximations of isomorphism and logics with linear-algebraic operators. In *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019.*, pages 112:1–112:14, 2019. doi:10.4230/LIPIcs.ICALP.2019.112.

9 A. Dawar, M. Grohe, B. Holm, and B. Laubner. Logics with rank operators. In *Proc. 24th IEEE Symp. on Logic in Computer Science*, pages 113–122, 2009.

10 A. Dawar and B. Holm. Pebble games with algebraic rules. *Fundam. Inform.*, 150(3-4):281–316, 2017.

11 H-D. Ebbinghaus and J. Flum. *Finite Model Theory.* Springer, 2nd edition, 1999.

12 P. Erdös and H. Sachs. Reguläre Graphen gegebener Taillenweite mit minimaler Knotenzahl. *Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe*, 12(251-257):22, 1963.

13 T. Feder and M.Y. Vardi. Computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal of Computing*, 28:57–104, 1998.

14 David Geiger. Closed systems of functions and predicates. *Pacific Journal of Mathematics*, 27(1):95–100, 1968.

15 E. Grädel and W. Pakusa. Rank logic is dead, long live rank logic! *J. Symb. Log.*, 84:54–87, 2019. doi:10.1017/jsl.2018.33.

16 J. A. Grochow and M. Levet. On the descriptive complexity of groups without abelian normal subgroups. *arXiv*, abs/2209.13725, 2022. doi:10.48550/arXiv.2209.13725.

17 L. Hella. Logical hierarchies in PTIME. *Information and Computation*, 129:1–19, 1996.

18 L. Hella. The expressive power of CSP-quantifiers. In *31st EACSL Annual Conference on Computer Science Logic, CSL*, pages 25:1–25:19, 2023. doi:10.4230/LIPIcs.CSL.2023.25.

19 L. Hella and H. Imhof. Enhancing fixed point logic with cardinality quantifiers. *J. Log. Comput.*, 8(1):71–86, 1998. doi:10.1093/logcom/8.1.71.

20 Peter Jonsson, Victor Lagerkvist, Gustav Nordh, and Bruno Zanuttini. Strong partial clones and the time complexity of SAT problems. *J. Comput. Syst. Sci.*, 84:52–78, 2017. doi:10.1016/J.JCSS.2016.07.008.

21 M. Lichter. Separating rank logic from polynomial time. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS*. IEEE, 2021. doi:10.1109/LICS52264.2021.9470598.

22 B. A. Romov. The algebras of partial functions and their invariants. *Cybernetics*, 17(2):157–167, 1981. doi:10.1007/BF01069627.

23 Henning Schnoor and Ilka Schnoor. Partial polymorphisms and constraint satisfaction problems. In Nadia Creignou, Phokion G. Kolaitis, and Heribert Vollmer, editors, *Complexity of Constraints – An Overview of Current Research Themes [Result of a Dagstuhl Seminar]*, volume 5250 of *Lecture Notes in Computer Science*, pages 229–254. Springer, 2008. doi:10.1007/978-3-540-92800-3_9.

24 D. Zhuk. A proof of the CSP dichotomy conjecture. *J. ACM*, 67:30:1–30:78, 2020. doi:10.1145/3402029.