Arity hierarchies for 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

Benedikt Pago
Department of Computer Science and Technology, University of Cambridge, UK

Abstract

We investigate the expressive power of generalized quantifiers closed under partial polymorphism conditions motivated by the study of constraint satisfaction problems. We answer a number of questions arising from the work of Dawar and Hella (CSL 2024) where such quantifiers were introduced. For quantifiers closed under partial near-unanimity polymorphisms, we establish hierarchy results clarifying the interplay between the arity of the polymorphisms and of the quantifiers: The expressive power of $(\ell+1)$-ary quantifiers closed under ℓ-ary partial near-unanimity polymorphisms is strictly between the class of all quantifiers of arity $\ell-1$ and ℓ. We also establish an infinite hierarchy based on the arity of quantifiers with a fixed arity of partial near-unanimity polymorphisms. Finally, we prove inexpressiveness results for quantifiers with a partial Maltsev polymorphism. The separation results are proved using novel algebraic constructions in the style of Cai-Fürer-Immerman and the quantifier pebble games of Dawar and Hella (2024).

2012 ACM Subject Classification Theory of computation → Finite Model Theory

Keywords and phrases finite model theory, constraint satisfaction problems, generalized quantifiers

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

Acknowledgements We thank Victor Lagerkvist for pointers to the literature on the use of partial polymorphisms in CSP algorithms.

1 Introduction

The study of generalized quantifiers, also known as Lindström quantifiers [13], has played a major role in the development of finite model theory. This theory is concerned with the expressive power of logics over finite structures. When any particular class $\mathcal{K}$ of structures is not definable in a logic L, we can adjoin to L a Lindström quantifier $\mathcal{Q}_{\mathcal{K}}$, which serves as an oracle for querying membership in $\mathcal{K}$. Extending L with this quantifier yields, in a precise sense, the minimal extension of L that can also express $\mathcal{K}$. Thus, the investigation of the structure and interdefinability of quantifiers is the core of finite model theory.

Results about inexpressibility, commonly referred to as *lower bound* results for logic, are often established by means of exhibiting winning strategies for model-comparison games. Some of the most striking such results in finite model theory have come from model-comparison games developed for Lindström quantifiers. In particular, Hella's bijection game [15] was introduced to establish the arity hierarchy for generalized quantifiers. This game has since been widely used to establish many inexpressibility results in descriptive complexity and beyond. Hella's proof that there is a strict hierarchy of expressive power obtained by classifying quantifiers according to their arity has as a consequence that meaningful logics in descriptive complexity (such as those closed under first-order reductions) should include quantifiers of arbitrary arities. Including all quantifiers clearly leads to logics that are far

too rich. This leads us to study extensions of first-order logic with restricted families of quantifiers, but of unbounded arity.

A natural restriction is to require *closure conditions* on the classes of structures $\mathcal{K}$ whose Lindström quantifiers $\mathcal{Q}_\mathcal{K}$ we include in the logic. By default, such classes $\mathcal{K}$ are closed under isomorphism because logics never distinguish between isomorphic structures. Interestingly, by imposing stricter closure conditions, one can often obtain nicely behaved logics that admit a useful game characterization while still being rich in expressive power (i.e. stronger than first-order logic with counting quantifiers). One such example is first-order logic augmented by all linear-algebraic quantifiers [9]. Its expressive power is characterized by the invertible map game, which was used by Lichter [21] in a highly sophisticated lower bound proof. The classes $\mathcal{K}$ defining linear-algebraic quantifiers are closed under an equivalence relation coarser than isomorphism, which is based on similarity of adjacency matrices over all finite fields [8]. More generally we can ask: *For which closure conditions on classes $\mathcal{K}$ do the Lindström quantifiers $\mathcal{Q}_\mathcal{K}$ give rise to logics that are both powerful and analyzable via a natural model-comparison game?*

The present paper should be seen as a contribution to this research programme. The closure condition we are concerned with here is closure under *partial polymorphisms*, which was first studied by Dawar and Hella in 2024 [10]. An ℓ-ary polymorphism from a structure $\mathbf{A}$ to $\mathbf{B}$ is a homomorphism from the ℓ-th power of $\mathbf{A}$ to $\mathbf{B}$. The interest in the notion comes from the theory of constraint satisfaction problems (CSP). For a fixed finite structure $\mathbf{B}$, called the template, $\mathrm{CSP}(\mathbf{B})$ denotes the class of structures $\mathbf{A}$ that map homomorphically to $\mathbf{B}$. The algebraic approach to CSP relates the computational complexity of $\mathrm{CSP}(\mathbf{B})$ to the algebra of polymorphisms of $\mathbf{B}$ (that is, polymorphisms from $\mathbf{B}$ to itself) and more particularly to the equational theory of that algebra (see e.g. [3]). The success of this programme culminated in the celebrated dichotomy theorem for finite-domain CSPs of Bulatov and Zhuk [4, 23]. Natural examples of quantifiers $\mathcal{Q}_\mathcal{K}$ closed under partial polymorphisms are thus CSP quantifiers, that is, $\mathcal{K} = \mathrm{CSP}(\mathbf{B})$ for a template structure $\mathbf{B}$ that also has the respective (partial) polymorphism. Such CSP quantifiers were first introduced by Hella [16], where a strict hierarchy of expressive power in terms of the size of $\mathbf{B}$ was demonstrated. This size hierarchy is quite different from the arity hierarchy that we establish here. Moreover, it should be stressed that families of quantifiers closed under partial polymorphisms are a more general concept and contain more than just CSP quantifiers.

Our main results are about the interaction of partial polymorphism conditions with restrictions on the arity of quantifiers. The conditions require that the class of structures defining a quantifier is closed under partial functions satisfying certain equational conditions. The conditions we are particularly interested in (already introduced in prior work [10]) are *partial near-unanimity* polymorphisms and *partial Maltsev* polymorphisms. Their total counterparts play a key role in the classification of tractable CSP, since templates with these polymorphisms give rise to CSP with relatively simple polynomial time algorithms [2, 5]. The corresponding *partial* polymorphisms, which we use here, are significant for another reason: For NP-complete CSP, they essentially govern the fine-grained exponential complexity under the strong exponential-time hypothesis, and they have applications in parameterized complexity as well (see e.g. [17, 18, 19]).

Previous work [10] developed games for logics with quantifiers closed under partial polymorphisms and used these to establish that the problem of solving systems of equations over the Boolean field is not definable in the extension of infinitary logic with *all* quantifiers closed under partial near-unanimity polymorphisms.

In the present paper we answer a number of natural questions arising from this previous

work concerning arity hierarchies within the class of quantifiers closed under partial near-unanimity polymorphisms, and also establish the first expressivity bounds for quantifiers closed under partial Maltsev polymorphisms. To be specific, we establish in Section 3 an arity reduction showing that any r-ary quantifier Q which is closed under a partial near-unanimity polymorphism of arity $\ell \leq r$ can be defined using quantifiers of arity at most $\lceil \frac{\ell-1}{\ell} \cdot r \rceil$. On the other hand, our main results in Section 4 show that the arity of such quantifiers does not collapse, and they exhibit in fact a strict infinite hierarchy of expressive power. First, we prove that quantifiers of arity $\ell + 1$ with a partial near-unanimity polymorphism of arity ℓ are, in terms of expressive power, strictly between *all* quantifiers of arity $\ell - 1$ and all quantifiers of arity ℓ (Corollary 17). Secondly, we establish in Theorem 18 that for every fixed arity of polymorphisms, we can always gain in expressive power by increasing the arity of quantifiers. The results can be seen as a refinement of the arity hierarchy of Hella [15], and the separation results require novel constructions in the style of Cai-Fürer-Immerman [6] (called *CFI-like* henceforth). In Section 5 we move on to study a more powerful type of quantifiers, which cannot be fooled by CFI-like constructions, namely those with a partial Maltsev polymorphism. Also for these, we obtain a strict arity hierarchy (Theorem 36). Moreover, we establish a first lower bound: When both the arity of the quantifiers as well as the number of variables are limited to k, Maltsev-closed quantifiers are strictly weaker than general ones (Theorem 37). This proof requires playing the Maltsev-closed quantifier pebble game on another original algebraic construction. It remains an intriguing open problem to prove lower bounds for first-order logic with Maltsev-closed quantifiers of arity k, but an unbounded number of variables. We believe that finding a suitable hard example would also inspire progress on other open problems, as it would have to be radically different from the CFI-like constructions typically used in many known lower bound results.

2 Preliminaries

We denote structures by boldface letters $\mathbf{A}, \mathbf{B}, \ldots$, and their universes by the corresponding Roman letters $A, B, \ldots$. All structures we consider in the paper are finite and relational. Thus, we assume without further mention that the universe A of each structure $\mathbf{A}$ is finite and the vocabulary of $\mathbf{A}$ is a finite set of relation symbols. Let $\mathbf{A}$ and $\mathbf{B}$ be τ-structures for a vocabulary τ. A function $f\colon C \to D$ is a *partial isomorphism* from $\mathbf{A}$ to $\mathbf{B}$ if $C \subseteq A$, $D \subseteq B$ and f is an isomorphism from the substructure of $\mathbf{A}$ induced by C to the substructure of $\mathbf{B}$ induced by D.

Let $\mathbf{A}$ and $\mathbf{B}$ be τ-structures. We write $\mathbf{A} \leq \mathbf{B}$ if $A = B$ and $R^{\mathbf{A}} \subseteq R^{\mathbf{B}}$ for all $R \in \tau$. The *union* $\mathbf{A} \cup \mathbf{B}$ of $\mathbf{A}$ and $\mathbf{B}$ is the τ-structure with universe $A \cup B$ and relations $R^{\mathbf{A}\cup\mathbf{B}} := R^{\mathbf{A}} \cup R^{\mathbf{B}}$ for each $R \in \tau$.

We denote first-order logic by FO, and the infinitary logic with k variables by $\mathcal{L}^k$. Furthermore, we write $\mathcal{L} := \bigcup_{k\in\mathbb{N}} \mathcal{L}^k$. We refer to [12] and [20] for precise definitions of these logics. Given a logic L and two τ-structures $\mathbf{A}$ and $\mathbf{B}$, we write $\mathbf{A} \equiv_L \mathbf{B}$ if $\mathbf{A}$ and $\mathbf{B}$ satisfy exactly the same L-sentences of vocabulary τ. For any positive integer n, we denote the set $\{1, \ldots, n\}$ by $[n]$.

Generalized quantifiers

Let $\tau = \{R_1, \ldots, R_m\}$ be a vocabulary and let $\mathcal{K}$ be a class of τ-structures that is closed under isomorphisms. The extension $L(Q_\mathcal{K})$ of a logic L by the *generalized quantifier* $Q_\mathcal{K}$ (also called *Lindström quantifier*) is obtained by adding the following formula formation rule to the syntax of L:

For any vocabulary σ, if $\Psi = (\psi_1, \dots, \psi_m)$ is a tuple of σ-formulas, and $\bar{y}_1, \dots, \bar{y}_m$ are tuples of variables such that $|\bar{y}_i| = \mathrm{ar}(R_i)$ for $i \in [m]$, then $Q_{\mathcal{K}}\bar{y}_1, \dots, \bar{y}_m\Psi$ is a σ-formula. A variable is free in this formula if, for some $i \in [m]$, it is free in ψ_i, but not in $\bar{y}_i$.

The sequence $\Psi = (\psi_1, \dots, \psi_m)$ of σ-formulas defines an *interpretation* of a τ-structure in any σ-structure $\mathbf{A}$ with an assignment α that interprets all free variables of $Q_{\mathcal{K}}\bar{y}_1, \dots, \bar{y}_m\Psi$: we define $\Psi(\mathbf{A}, \alpha) \coloneqq (A, \psi_1^{\mathbf{A},\alpha}, \dots, \psi_m^{\mathbf{A},\alpha})$, where $\psi_i^{\mathbf{A},\alpha} \coloneqq \{\bar{b} \in A^{\mathrm{ar}(R_i)} \mid (\mathbf{A}, \alpha[\bar{b}/\bar{y}_i]) \models \psi_i\}$ for $i \in [m]$. The semantics of the formula $Q_{\mathcal{K}}\bar{y}_1, \dots, \bar{y}_m\Psi$ can now be defined as follows:

Let $\mathbf{A}$ be a σ-structure and α an assignment on $\mathbf{A}$ that interprets all free variables of $Q_{\mathcal{K}}\bar{y}_1, \dots, \bar{y}_m\Psi$. Then $(\mathbf{A}, \alpha) \models Q_{\mathcal{K}}\bar{y}_1, \dots, \bar{y}_m\Psi$ if, and only if, $\Psi(\mathbf{A}, \alpha) \in \mathcal{K}$.

The extension $L(\mathcal{Q})$ of a logic L by a collection $\mathcal{Q}$ of quantifiers is defined in the natural way by adding the corresponding formula formation and semantic rules for each quantifier $Q_{\mathcal{K}}$ in $\mathcal{Q}$. The *arity* of a quantifier $Q_{\mathcal{K}}$ is the maximum arity of relations in the vocabulary of the class $\mathcal{K}$. For each $r \geq 1$, we denote the collection of all quantifiers of arity at most r by $\mathcal{Q}_r$.

Bijective pebble games

We apply the bijective pebble game of Hella [15] in proving our arity hierarchy results in Section 4. Let $r \leq k$ be positive integers, and $\mathbf{A}$ and $\mathbf{B}$ be structures.

▶ **Definition 1.** *The k, r-*bijection game $\mathrm{BP}_r^k(\mathbf{A}, \mathbf{B})$ *is played between two players,* Duplicator *and* Spoiler*. Positions of the game are pairs (α, β), where α and β are assignments on $\mathbf{A}$ and $\mathbf{B}$, respectively, such that $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta) \subseteq X \coloneqq \{x_1, \dots, x_k\}$. The initial position is $(\emptyset, \emptyset)$. Let (α, β) be the position after some round of the game. The next round consists of the following steps:*

(1) Duplicator chooses a bijection $f \colon A \to B$. If $|A| \neq |B|$, then Spoiler wins the play.
(2) Spoiler chooses an r-tuple $\bar{y}$ of distinct variables in X and an r-tuple $\bar{a} \in A^r$.
(3) The next position is (α', β'), where $\alpha' \coloneqq \alpha[\bar{a}/\bar{y}]$ and $\beta' \coloneqq \beta[f(\bar{a})/\bar{y}]$. If $\alpha' \mapsto \beta'$ is not a partial isomorphism $\mathbf{A} \to \mathbf{B}$, then Spoiler wins the play.

Duplicator wins the play if Spoiler does not win it in a finite number of rounds.

The k, r-bijection game characterizes equivalence with respect to the infinitary k-variable logic with all r-ary quantifiers.

▶ **Proposition 2** ([15])**.** *Let $k \geq r$ be positive integers. Then Duplicator has a winning strategy in $\mathrm{BP}_r^k(\mathbf{A}, \mathbf{B})$ if, and only if, $\mathbf{A} \equiv_{\mathcal{L}^k(\mathcal{Q}_r)} \mathbf{B}$.*

We say that a sentence ψ of some logic separates two disjoint classes $\mathcal{K}$ and $\mathcal{K}'$ of structures if every structure in $\mathcal{K}$ satisfies ψ, but no structure in $\mathcal{K}'$ satisfies ψ.

▶ **Corollary 3.** *Fix a positive integer r. Let $\mathcal{K}$ and $\mathcal{K}'$ be two disjoint classes of structures. If for every $k \geq r$, there exist $\mathbf{A} \in \mathcal{K}$ and $\mathbf{B} \in \mathcal{K}'$ such that Duplicator has a winning strategy in $\mathrm{BP}_r^k(\mathbf{A}, \mathbf{B})$, then no sentence in $\mathcal{L}(\mathcal{Q}_r)$ separates $\mathcal{K}$ from $\mathcal{K}'$.*

Partial polymorphisms

We go briefly through the basic notions related to families of partial polymorphisms. For a more detailed treatment and motivation, we refer to [10], where closure of generalized quantifiers with respect to such families was first considered.

Let $p\colon A^\ell \to A$ be a partial function. Then for any $r \geq 1$, p induces a partial function $\hat{p}\colon (A^r)^\ell \to A^r$ defined by applying p "column wise": if $\bar{a}_i = (a_{i1}, \ldots, a_{ir}) \in A^r$ for $i \in [\ell]$, then $\hat{p}(\bar{a}_1, \ldots, \bar{a}_\ell) \coloneqq (p(a_{11}, \ldots, a_{\ell 1}), \ldots, p(a_{1r}, \ldots, a_{\ell r}))$. Naturally $\hat{p}(\bar{a}_1, \ldots, \bar{a}_\ell)$ is undefined if, and only if, $p(a_{1j}, \ldots, a_{\ell j})$ is undefined for at least one $j \in [r]$. Sometimes we also view the tuples $\bar{a}_1, \ldots, \bar{a}_\ell$ as the rows of an $\ell \times r$-matrix M and write $\hat{p}(M)$ for $\hat{p}(\bar{a}_1, \ldots, \bar{a}_\ell)$. Given a relation $R \subseteq A^r$, we write $p(R) \coloneqq \{\hat{p}(\bar{a}_1, \ldots, \bar{a}_\ell) \mid \bar{a}_1, \ldots, \bar{a}_\ell \in R\}$. Furthermore, if $\mathbf{A}$ is a τ-structure we denote by $p(\mathbf{A})$ the τ-structure with the same universe and relations $p(R^{\mathbf{A}})$, $R \in \tau$.

We say that a partial function $p\colon A^\ell \to A$ is a *partial polymorphism* of $\mathbf{A}$ if $p(\mathbf{A}) \leq \mathbf{A}$. A *family $\mathcal{P}$ of ℓ-ary partial functions* consists of a partial function $p_A\colon A^\ell \to A$ for every finite set A. The family *respects bijections* if $f(p_A(a_1, \ldots, a_\ell)) \simeq p_B(f(a_1), \ldots, f(a_\ell))$ whenever $f\colon A \to B$ is a bijection. Here $s \simeq t$ means that either $s = t$, or both s and t are undefined.

Let $\mathcal{P}$ be a family of partial functions that respects bijections, and let $Q_{\mathcal{K}}$ be a quantifier for a class $\mathcal{K}$ of τ-structures. We say that $Q_{\mathcal{K}}$ is *$\mathcal{P}$-closed* if for every $\mathbf{B} \in \mathcal{K}$ and every $\mathbf{A} \leq \mathbf{B} \cup p_B(\mathbf{B})$, we also have $\mathbf{A} \in \mathcal{K}$. We denote the collection of all $\mathcal{P}$-closed quantifiers by $\mathcal{Q}^{\mathcal{P}}$. It follows immediately from the definition that all $\mathcal{P}$-closed quantifiers $Q_{\mathcal{K}}$ are *downwards monotone*: If $\mathbf{B} \in \mathcal{K}$ and $\mathbf{A} \leq \mathbf{B}$, then $\mathbf{A} \in \mathcal{K}$.

The most interesting families of partial functions are based on equations on the components of the input tuple. Two examples of such families were studied in [10]:

▶ **Example 4.**

(a) The *Maltsev* family $\mathcal{M}$ consists of functions $m_A\colon A^3 \to A$ such that $m_A(a, a, b) = m_A(b, a, a) = b$ and $m_A(a, b, c)$ is undefined if $a \neq b$ and $b \neq c$.
(b) The family $\mathcal{N}^\ell$ of ℓ-ary *partial near-unanimity functions* consists of functions $n^\ell_A\colon A^\ell \to A$ defined as follows: $n^\ell_A(a_1, \ldots, a_\ell) = a$ if at least $\ell - 1$ of the inputs a_i are equal to a; else, $n^\ell_A(a_1, \ldots, a_\ell)$ is undefined.

Note that the definition of $n^\ell_A(a_1, \ldots, a_\ell)$ is actually independent of the set A. Thus, we omit the subscript A and write simply n^ℓ instead of n^ℓ_A.

It is straightforward to show that n^ℓ is a partial polymorphism of any structure with at most $\ell - 1$-ary relations:

▶ **Lemma 5** ([10, Lemma 9]). *Let $\mathbf{A}$ be a τ-structure such that $\mathrm{ar}(R) < \ell$ for all $R \in \tau$. Then n^ℓ is a partial polymorphism of $\mathbf{A}$.*

CSP quantifiers

Let $\mathbf{A}$ and $\mathbf{B}$ be τ-structures. A function $h\colon A \to B$ is a *homomorphism* from $\mathbf{A}$ to $\mathbf{B}$ if for all $R \in \tau$ and $\bar{a} \in A^{\mathrm{ar}(R)}$, $\bar{a} \in R^{\mathbf{A}}$ implies that $h(\bar{a}) \in R^{\mathbf{B}}$. We write $\mathbf{A} \to \mathbf{B}$ if there exists a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. For any fixed τ-structure $\mathbf{B}$ we denote the class $\{\mathbf{A} \mid \mathbf{A} \to \mathbf{B}\}$ by $\mathrm{CSP}(\mathbf{B})$. The class $\mathrm{CSP}(\mathbf{B})$ is usually regarded as the (non-uniform) *constraint satisfaction problem* (CSP): Given an input structure $\mathbf{A}$, decide whether there exists a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. Here $\mathbf{B}$ is the *template* of the CSP. It is a consequence of the results of Bulatov and Zhuk [4, 23] that, for any finite template $\mathbf{B}$, the corresponding CSP is either decidable in polynomial time or NP-complete, and which of these is the case depends only on the polymorphisms of $\mathbf{B}$. A generalized quantifier $Q_{\mathcal{K}}$ is a *CSP quantifier* if the defining class $\mathcal{K}$ is of the form $\mathrm{CSP}(\mathbf{B})$ for some template $\mathbf{B}$. For CSP quantifiers, being closed under partial polymorphisms is a natural notion: If $\mathbf{B}$ is a structure which has a (total) polymorphism that satisfies an identity, this yields a natural *partial* polymorphism

of the class of structures $\mathrm{CSP}(\mathbf{B})$. Thus, quantifiers closed under this partial polymorphism are a generalization of CSP quantifiers based on a structure with the corresponding total polymorphism. The following collections of CSP quantifiers are of particular interest to us.

- For $r \geq 1$, $\mathcal{Q}_r^{\mathrm{CSP}} \subseteq \mathcal{Q}_r$ is the collection of all CSP quantifiers of arity at most r.
- For $\ell \geq 3$, $\mathcal{Q}^{\mathcal{N}_{\mathrm{CSP}}^\ell} \subseteq \mathcal{Q}^{\mathcal{N}^\ell}$ is the collection of all $\mathcal{N}^\ell$-closed CSP quantifiers.
- For $\ell \geq 3$ and $r \geq 1$, $\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell} := \mathcal{Q}^{\mathcal{N}_{\mathrm{CSP}}^\ell} \cap \mathcal{Q}_r$.
- $\mathcal{Q}^{\mathcal{M}}$ is the collection of all quantifiers closed under the partial Maltsev polymorphism, and $\mathcal{Q}_r^{\mathcal{M}} := \mathcal{Q}^{\mathcal{M}} \cap \mathcal{Q}_r$, for every $r \geq 1$.

Dawar and Hella [10, Proposition 11] proved that if p_B is a partial polymorphism of $\mathbf{B}$, then $Q_{\mathrm{CSP}(\mathbf{B})}$ is $\mathcal{P}$-closed, assuming that $\mathcal{P}$ is projective[1]. In particular, this holds for the family $\mathcal{N}_\ell$ for any $\ell \geq 3$. We conclude this section by proving a converse of this result. A relational structure is a *core* if every homomorphism from it to itself is an automorphism. If a structure is not a core itself, it has a unique (up to isomorphism) substructure that is (see [14, Proposition 3.3]).

▶ **Lemma 6.** *Let $\mathcal{P}$ be a family of partial functions. Assume that $\mathbf{B}$ is a core such that the corresponding CSP quantifier $Q_{\mathrm{CSP}(\mathbf{B})}$ is $\mathcal{P}$-closed. Then p_B is a partial polymorphism of $\mathbf{B}$.*

Proof. We trivially have $\mathbf{B} \in \mathrm{CSP}(\mathbf{B})$, and since $Q_{\mathrm{CSP}(\mathbf{B})}$ is $\mathcal{P}$-closed, also $\mathbf{B} \cup p_B(\mathbf{B}) \in \mathrm{CSP}(\mathbf{B})$. So there is a homomorphism $h\colon \mathbf{B} \cup p_B(\mathbf{B}) \to \mathbf{B}$. Now h maps $R^{\mathbf{B}}$ to itself for every relation R in the vocabulary τ of $\mathbf{B}$, and because $\mathbf{B}$ is a core, h must be a permutation of B. But then we must have $p_B(\mathbf{B}) \leq \mathbf{B}$ because otherwise, for some $R \in \tau$, h would map a tuple in $R^{p_B(\mathbf{B})} \setminus R^{\mathbf{B}}$ to a tuple in $R^{\mathbf{B}}$, which is not possible because h induces a bijection between the tuples in $R^{\mathbf{B}}$. ◀

▶ **Corollary 7.** *Let $\mathbf{B}$ be a structure. The CSP quantifier $Q_{\mathrm{CSP}(\mathbf{B})}$ is in $\mathcal{Q}^{\mathcal{N}_{\mathrm{CSP}}^\ell}$ if, and only if, n^ℓ is a partial polymorphism of the core of $\mathbf{B}$.*

Proof. Note that $\mathrm{CSP}(\mathbf{B}) = \mathrm{CSP}(\mathbf{B}')$ for the core $\mathbf{B}'$ of $\mathbf{B}$. Thus, the implication from left to right follows from Lemma 6, and the implication from right to left follows from [10, Proposition 11]. ◀

3 Arity reductions

Our aim is to prove arity hierarchies for $\mathcal{N}^\ell$-closed quantifiers. Dawar and Hella [10] introduced a model-comparison game for $\mathcal{L}^k(\mathcal{Q}^{\mathcal{N}^\ell})$, and used it to prove that for any $\ell \geq 3$ there is an ℓ-ary CSP quantifier in $\mathcal{Q}^{\mathcal{N}^{\ell+1}}$ that is not definable in $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell})$. However, they left open the question whether for every fixed ℓ, there is an arity hierarchy within the class of $\mathcal{N}^\ell$-closed quantifiers. This question turns out to be nontrivial: In this section, we show that any $\mathcal{N}^\ell$-closed quantifier of arity at least ℓ can be defined by a quantifier with strictly smaller arity. Moreover, the arity of some natural $\mathcal{N}^\ell$-closed quantifiers can even be reduced all the way down to 2 in this way.

[1] See [10, Definition 6] for the definition of projective.

Example: Hypergraph colouring

For $m \geq r \geq 2$ let H^r_m be the *r-uniform hypergraph of size* m, i.e., $H^r_m := ([m], R^r_m)$, where R^r_m is the set of all tuples $(a_1, \ldots, a_r) \in [m]^r$ such that $a_i \neq a_j$ for all $i \neq j$. Note that H^2_m is just the complete graph K_m on m vertices. In [10] the authors proved that the CSP quantifier corresponding to the class $\mathrm{CSP}(H^r_m)$ is $\mathcal{N}^\ell$-closed for any $\ell \geq 3$.

Thus, one might think that the quantifiers $Q_{\mathrm{CSP}(H^r_m)}$, for $m \geq r \geq 2$, would form a natural arity hierarchy within $\mathcal{Q}^{\mathcal{N}^\ell}$. However, this is not the case: Each $\mathrm{CSP}(H^r_m)$ can be defined already by the binary quantifier $Q_{\mathrm{CSP}(K_m)}$:

▶ **Proposition 8.** *There is an* FO*-definable reduction from* $\mathrm{CSP}(H^r_m)$ *to* $\mathrm{CSP}(K_m)$.

Proof sketch. Given $\mathbf{A}$ in the vocabulary $\{R\}$ with $\mathrm{ar}(R) = r$, let $\mathcal{G}(\mathbf{A})$ be the graph (possibly with self-loops) on vertex set A with edges $\{(a_i, a_j) \mid (a_1, \ldots, a_r) \in R^{\mathbf{A}}, i \neq j\}$. Clearly $\mathcal{G}(\mathbf{A})$ is FO-definable in $\mathbf{A}$. It can be verified that $\mathbf{A} \to H^r_m$ if, and only if, $\mathcal{G}(\mathbf{A}) \to K_m$. ◀

Arity reductions for $\mathcal{N}^\ell$-closed quantifiers

We now show a general arity reduction result for N^ℓ-closed quantifiers. We start with the easier case of CSP quantifiers. Let $r \geq 1$ and let $R \subseteq B^r$ be an r-ary relation. For $i, j \in [r]$, $i \leq j$, let $p_{\setminus[i,j]}$ be the projection function $B^r \to B^{r-(j-i+1)}$ that maps $(b_1, \ldots, b_r)$ to $(b_1, \ldots, b_{i-1}, b_{j+1}, \ldots, b_r)$. Furthermore, let $R_{\setminus[i,j]}$ be the $r-(j-i+1)$-ary relation that is the projection of R to $[r] \setminus \{i, i+1, ..., j\}$, i.e., $R_{\setminus[i,j]} = \{p_{\setminus[i,j]}(\bar{b}) \mid \bar{b} \in R\}$.

▶ **Lemma 9.** *Let* $\ell \geq 3$ *and* $r \geq \ell$*, and let* $(i_1, j_1), (i_2, j_2), ..., (i_\ell, j_\ell)$ *be a division of* $[r]$ *into* ℓ *intervals. Let* B *be a finite set and* $R \subseteq B^r$*. Assume that* n^ℓ *is a partial polymorphism of* (B, R)*. Then* $\bar{b} \in R$ *if, and only if,* $p_{\setminus[i_n,j_n]}(\bar{b}) \in R_{\setminus[i_n,j_n]}$ *for every* $n \in [\ell]$.

Proof. Let $\bar{b} \in B^r$. If $\bar{b} \in R$, then $p_{\setminus[i_n,j_n]}(\bar{b}) \in R_{\setminus[i_n,j_n]}$ by definition. For the other direction, assume that $p_{\setminus[i_n,j_n]}(\bar{b}) \in R_{\setminus[i_n,j_n]}$ for every $n \in [\ell]$. This means that there exist tuples $\bar{c}_1, ..., \bar{c}_\ell \in R$ such that $p_{\setminus[i_n,j_n]}(\bar{c}_n) = p_{\setminus[i_n,j_n]}(\bar{b})$ for each $n \in [\ell]$. Clearly then $\widehat{n^\ell}(\bar{c}_1, \ldots, \bar{c}_\ell) = \bar{b}$, and since n^ℓ is a partial polymorphism of (B, R), it follows that $\bar{b} \in R$. ◀

We use the following notation in the rest of this section. Let $\mathbf{C}$ be a τ-structure, where $\mathrm{ar}(R) \leq r$ for all $R \in \tau$, and let $(i_1, j_1), (i_2, j_2), ..., (i_\ell, j_\ell)$ be a division of $[r]$ into intervals. For each $R \in \tau$ and $n \in [\ell]$, let

$$R^{\mathbf{C}}_n := \begin{cases} R^{\mathbf{C}}_{\setminus[i_n,j_n]} & \text{if } j_n \leq \mathrm{ar}(R), \\ R^{\mathbf{C}}_{\setminus[i_n,\mathrm{ar}(R)]} & \text{if } i_n \leq \mathrm{ar}(R) < j_n, \\ \emptyset & \text{if } \mathrm{ar}(R) < i_n. \end{cases}$$

Then we denote by $\mathbf{C}^*$ the structure $(C, (R^{\mathbf{C}}_n)_{R \in \tau, n \in [\ell]})$. Note that $\mathbf{C}^*$ is FO-definable in $\mathbf{C}$.

▶ **Lemma 10.** *Let* ℓ, r *and* $(i_1, j_1), (i_2, j_2), ..., (i_\ell, j_\ell)$ *be as in Lemma 9, and let* $\mathbf{A}$ *and* $\mathbf{B}$ *be* τ*-structures, where* $\mathrm{ar}(R) \leq r$ *for all* $R \in \tau$*. Assume that* n^ℓ *is a partial polymorphism of* $\mathbf{B}$*. Then* $\mathbf{A} \to \mathbf{B}$ *if, and only if,* $\mathbf{A}^* \to \mathbf{B}^*$.

Proof. Let $h : \mathbf{A} \to \mathbf{B}$ be a homomorphism, and let $\bar{a} \in R^{\mathbf{A}}_n$ for some $R \in \tau$ and $n \in [\ell]$. Then there is a tuple $\bar{c} \in R^{\mathbf{A}}$ such that $\bar{a} = p_{\setminus[i_n,j_n]}(\bar{c})$. Since $h(\bar{c}) \in R^{\mathbf{B}}$, also $h(\bar{a}) = p_{\setminus[i_n,j_n]}(h(\bar{c})) \in R^{\mathbf{B}}_n$. Hence, h is a homomorphism from $\mathbf{A}^*$ to $\mathbf{B}^*$.

Conversely, assume that h is a homomorphism from $\mathbf{A}^*$ to $\mathbf{B}^*$. Let $\bar{a} \in R^{\mathbf{A}}$ for some $R \in \tau$. Then for every $n \in [\ell]$, $p_{\setminus[i_n,j_n]}(h(\bar{a})) = h(p_{\setminus[i_n,j_n]}(\bar{a})) \in R^{\mathbf{B}}_n$. By Lemma 9, we have $h(\bar{a}) \in R^{\mathbf{B}}$. ◀

▶ **Corollary 11.** *Let $r \geq \ell \geq 3$. Every quantifier in $\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell}$ is* FO*-definable by a quantifier in $\mathcal{Q}^{\mathrm{CSP}}_{\lceil \frac{\ell-1}{\ell} \cdot r \rceil}$.*

Proof. Let $\mathbf{B}$ be a τ-structure such that $Q_{\mathrm{CSP}(\mathbf{B})}$ is a quantifier in $\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell}$. W.l.o.g. we can assume that $\mathbf{B}$ is a core, and hence by Corollary 7, n^ℓ is a partial polymorphism of $\mathbf{B}$. Consider now a division $(i_1, j_1), (i_2, j_2), ..., (i_\ell, j_\ell)$ of $[r]$ into intervals such that the size of of each $[i_n, j_n]$ is at least $\lfloor \frac{r}{\ell} \rfloor$. Then the arity of each relation $R_n^{\mathbf{B}}$ of $\mathbf{B}^*$ is at most $r - \lfloor \frac{r}{\ell} \rfloor = \lceil \frac{\ell-1}{\ell} \cdot r \rceil$. Thus, $Q_{\mathrm{CSP}(\mathbf{B}^*)} \in \mathcal{Q}^{\mathrm{CSP}}_{\lceil \frac{\ell-1}{\ell} \cdot r \rceil}$.

By Lemma 10, for any τ-structure $\mathbf{A}$ we have $\mathbf{A} \in \mathrm{CSP}(\mathbf{B})$ if, and only if, $\mathbf{A}^* \in \mathrm{CSP}(\mathbf{B}^*)$. Thus, the mapping $\mathbf{A} \mapsto \mathbf{A}^*$ is an FO-definable reduction of $\mathrm{CSP}(\mathbf{B})$ to $\mathrm{CSP}(\mathbf{B}^*)$. ◀

It should be noted that n^ℓ is usually not a partial polymorphism of the structure $\mathbf{B}^*$ in Lemma 10. Otherwise we could iterate the arity reduction trick to obtain even lower arity quantifiers that define $\mathrm{CSP}(\mathbf{B})$.

If we allow ourselves to use the infinitary logic $\mathcal{L}$ rather than just FO in the reductions, then the arity reduction works for *any* $\mathcal{N}^\ell$-closed quantifier, not just CSPs. The general reduction consists in first closing the given structure $\mathbf{A}$ under n^ℓ (which is an $\mathcal{L}$-definable fixed-point induction), and then applying to it the FO-definable operator * defined before Lemma 10. We prove in Appendix A that this yields the following arity reduction:

▶ **Lemma 12.** *Let $r \geq \ell \geq 3$. Let $\mathcal{K}$ be an $\mathcal{N}^\ell$-closed class of τ-structures, where $\mathrm{ar}(R) \leq r$ for all $R \in \tau$. Then there exists a vocabulary σ with at most $\lceil \frac{\ell-1}{\ell} \cdot r \rceil$-ary relations, a class $\mathcal{K}'$ of σ-structures, and an $\mathcal{L}$-definable reduction f that maps τ-structures to σ-structures in such a way that $\mathbf{A} \in \mathcal{K}$ if, and only if, $f(\mathbf{A}) \in \mathcal{K}'$.*

▶ **Corollary 13.** *Let $r \geq \ell \geq 3$. Every quantifier in $\mathcal{Q}_r^{\mathcal{N}^\ell}$ is definable in $\mathcal{L}(\mathcal{Q}_{\lceil \frac{\ell-1}{\ell} \cdot r \rceil})$.*

Here and in Theorem 15 below, the expression $\lceil \frac{\ell-1}{\ell} \cdot r \rceil$ may equivalently be read as $r - i$, for the i such that $i \cdot \ell \leq r < (i+1) \cdot \ell$. Note that the arity reduction result in Corollary 11 for CSP quantifiers is not a special case of Corollary 13, as the latter does not provide an FO-definable reduction. Moreover, in Corollary 11, the defining quantifier is again a CSP quantifier.

Summarizing our arity collapse results, we obtain the following picture. First, we note that if the arity ℓ of the partial near-unanimity function n^ℓ is greater than the arity of the relations, then being $\mathcal{N}^\ell$-closed is not a restriction:

▶ **Theorem 14.** *Let $\ell > r \geq 1$, $\ell \geq 3$. Then $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell}) \equiv \mathcal{L}(\mathcal{Q}_r)$ and $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell}) \equiv \mathcal{L}(\mathcal{Q}_r^{CSP})$.*

Proof. The first claim follows from [10, Lemma 9] and the argument for [10, Corollary 10]. The second claim follows from Corollary 7 and Lemma 5. ◀

Hence, only for $r \geq \ell$, the situation is interesting. In this case, we have:

▶ **Theorem 15.** *Let $r \geq \ell \geq 3$. Then we have $\mathcal{L}(\mathcal{Q}_{\ell-1}^{\mathrm{CSP}}) \leq \mathcal{L}(\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell}) \leq \mathcal{L}(\mathcal{Q}^{\mathrm{CSP}}_{\lceil \frac{\ell-1}{\ell} \cdot r \rceil})$ and $\mathcal{L}(\mathcal{Q}_{\ell-1}) \leq \mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell}) \leq \mathcal{L}(\mathcal{Q}_{\lceil \frac{\ell-1}{\ell} \cdot r \rceil})$.*

In particular, in the case $r = \ell$ we have $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^r}) \equiv \mathcal{L}(\mathcal{Q}_{r-1}^{\mathrm{CSP}})$ and $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^r}) \equiv \mathcal{L}(\mathcal{Q}_{r-1})$.

Proof. The inclusions $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell}) \leq \mathcal{L}(\mathcal{Q}^{\mathrm{CSP}}_{\lceil \frac{\ell-1}{\ell} \cdot r \rceil})$ and $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell}) \leq \mathcal{L}(\mathcal{Q}_{\lceil \frac{\ell-1}{\ell} \cdot r \rceil})$ follow from Corollaries 11 and 13. The inclusions $\mathcal{L}(\mathcal{Q}_{\ell-1}^{\mathrm{CSP}}) \leq \mathcal{L}(\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell})$ and $\mathcal{L}(\mathcal{Q}_{\ell-1}) \leq \mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell})$ follow from Theorem 14, since clearly $\mathcal{L}(\mathcal{Q}_{\ell-1}^{\mathcal{N}_{\mathrm{CSP}}^\ell}) \leq \mathcal{L}(\mathcal{Q}_r^{\mathcal{N}_{\mathrm{CSP}}^\ell})$ and $\mathcal{L}(\mathcal{Q}_{\ell-1}^{\mathcal{N}^\ell}) \leq \mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell})$. ◀

4 Arity hierarchies for partial near-unanimity quantifiers

Theorem 15 shows that restricting quantifiers of arity r to those that are also $\mathcal{N}^r$-closed has the same effect on expressive power as reducing the arity to $r-1$, and does not really give rise to a new interesting logic. Thus, the natural question arises whether there are *any* combinations of polymorphism arities ℓ and quantifier arities r for which the $\mathcal{N}^\ell$-closed quantifiers yield a logic that is strictly different from any $\mathcal{L}(\mathcal{Q}_{r'})$. In other words, we would like to know if the inclusions in Theorem 15 are strict if $r \neq \ell$. The following theorem establishes this already for the case $r = \ell + 1$:

▶ **Theorem 16.** *Fix $\ell \geq 3$. Then $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{\ell+1}) \not\leq \mathcal{L}(\mathcal{Q}_{\ell-1})$.*

▶ **Corollary 17.** *For every $\ell \geq 3$, $\mathcal{L}(\mathcal{Q}_{\ell-1}) \leq \mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell}_{\ell+1}) \leq \mathcal{L}(\mathcal{Q}_\ell)$.*

Proof. The first inclusion follows with Theorem 15, and the fact that it is strict is a consequence of Theorem 16, which shows that already the fragment $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{\ell+1})$ of $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell}_{\ell+1})$ is not contained in $\mathcal{L}(\mathcal{Q}_{\ell-1})$. The second inclusion is a consequence of Theorem 15, and $\mathcal{L}(\mathcal{Q}_\ell) \not\leq \mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell}_{\ell+1})$ was shown in [10, Theorem 30]. ◀

A second natural question that we answer in this section is whether the polymorphism-closed quantifiers exhibit an arity hierarchy with respect to the arity of the quantifiers (keeping the arity of the polymorphism fixed): For the class of all generalized quantifiers, Hella [15] proved that $\mathcal{L}(\mathcal{Q}_r) \leq \mathcal{L}(\mathcal{Q}_{r+1})$ for all $r \in \mathbb{N}$. The following result shows that for every fixed polymorphism-arity ℓ, one can always gain strictly more expressive power by sufficiently increasing the arity of the quantifiers.

▶ **Theorem 18.** *Let $r \geq \ell \geq 3$. Then $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{q(r,\ell)}) \not\leq \mathcal{L}(\mathcal{Q}_r)$, where $q(r,\ell) = (2r+1)(\lfloor \frac{2r+1}{\ell-1} \rfloor + 2)$.*

▶ **Corollary 19.** *Let $r \geq \ell \geq 3$. Then there is an $r' > r$ such that $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{r}) \leq \mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{r'})$.*

Note that the separation $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{\ell+1}) \not\leq \mathcal{L}(\mathcal{Q}_{\ell-1})$ in Theorem 16 is tighter than the one that is implied by Theorem 18 for $r = \ell + 1$. So neither of the two results implies the other.

The proofs of Theorems 16 and 18 are structured similarly but employ different constructions. In both cases we construct a CSP template structure whose associated CSP quantifier can be used to distinguish certain CFI-like structures, which we view as the CSP instances. These instances are such that they cannot be distinguished in the logic $\mathcal{L}(\mathcal{Q}_r)$, for the respective value of r that we have to beat in the theorems. The novelty in both constructions is that we enforce the template structure to be closed under the ℓ-ary partial near-unanimity function, so that the corresponding CSP quantifier is indeed in $\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}$.

The rest of the section is divided into three parts: First, construction of the respective templates needed for the proofs of the two theorems, then construction of the instances, and finally, we provide winning strategies for Duplicator in the bijective pebble game to prove indistinguishability of the instances in the respective logic $\mathcal{L}(\mathcal{Q}_r)$.

4.1 Construction of polymorphism-closed templates

The templates we construct have as relations the solution spaces of certain systems of linear equations. To make them closed under the partial near-unanimity function n^ℓ, we ensure that no collection of tuples in a relation satisfies the conditions for applicability of n^ℓ, except in cases where n^ℓ outputs a tuple that is already in the collection. Then n^ℓ is a partial polymorphism of the template.

Template structure for the proof Theorem 16

We start by developing the ideas to forbid the applicability of n^ℓ. We say that an $n \times m$-matrix over a ring or field has the *n-near-unanimity* property if in every column, all entries except possibly one (the *minority entry*), are equal. Throughout this section, arithmetic operations are always meant in the ring $\mathbb{Z}_3$. Omitted/shortened proofs can be found in Appendix B.

▶ **Lemma 20.** *Let $A \in \mathbb{Z}_3^{3\times 3}$ be a matrix satisfying:*

1. *A has the 3-near-unanimity property.*
2. *$\sum_{j\in[3]} a_{1j} = \sum_{j\in[3]} a_{2j} = \sum_{j\in[3]} a_{3j}$, that is, the row-sums are all equal.*
3. *$\widehat{n^3}(A)$ is a triple that is not a row of A.*

Then $A \cdot (0,1,2)^T$ is a vector in $\mathbb{Z}_3^3$ with all three entries distinct.

Proof. Assumption 3) entails that no two rows of A are identical. Since no two rows of A are equal, the row A_{2-} differs from A_{1-} in at least one position. They cannot differ in all positions because then, by the first property, A_{3-} must be identical to either A_{1-} or A_{2-}, which cannot be the case. If A_{1-} and A_{2-} differ in only one position, then $\widehat{n^3}(A) \in \{A_{1-}, A_{2-}\}$, so this also contradicts assumption 3). Thus, A_{1-} and A_{2-} differ in exactly two positions. Assume first that $a_{11} \neq a_{21}$ and $a_{12} \neq a_{22}$, but $a_{13} = a_{23}$. Let $v_1 = A_{1-} \cdot (0,1,2)^T$ and $v_2 = A_{2-} \cdot (0,1,2)^T$. Then $v_2 - v_1 = a_{22} - a_{12} \neq 0$, so $v_1 \neq v_2$. The same argument works in case that $a_{11} \neq a_{21}$ and $a_{13} \neq a_{23}$. In case that $a_{11} = a_{21}$, $a_{12} \neq a_{22}$, and $a_{13} \neq a_{23}$, we have $v_2 - v_1 = a_{22} - a_{12} + 2(a_{23} - a_{13})$. By condition 2) and because $a_{11} = a_{21}$, we have $a_{12} + a_{13} = a_{22} + a_{23}$. So

$$v_2 - v_1 = a_{22} - a_{12} + 2(a_{23} - a_{13}) = a_{22} + a_{23} + a_{23} - (a_{12} + a_{13} + a_{13}) = a_{23} - a_{13} \neq 0.$$

Therefore, in all cases, $v_1 \neq v_2$. The same reasoning can be applied to any pair of rows, and thus also shows that $v_3 = A_{3-} \cdot (0,1,2)^T \neq v_1$, and $v_3 \neq v_2$. ◀

The lemma also generalizes to square matrices of larger dimensions because one can always find a (3×3)-submatrix with the properties required by Lemma 20 in them.

▶ **Lemma 21.** *Let $\ell \geq 3$ and let $A \in \mathbb{Z}_3^{\ell\times\ell}$ be a matrix satisfying:*

1. *A has the ℓ-near-unanimity property.*
2. *The row-sums of A are all equal.*
3. *$\widehat{n^\ell}(A)$ is not a row of A.*

Then $A \cdot (0,1,2,0,0,...,0)^T$ is a vector in $\mathbb{Z}_3^\ell$ that has at least three distinct entries.

We also need to consider the case where $\widehat{n^\ell}(A)$ is in fact a row of A. The following can be shown with a combinatorial argument.

▶ **Lemma 22.** *Let $\ell \geq 3$, and let $A \in \mathbb{Z}_3^{\ell\times\ell}$ be a matrix that satisfies the first two conditions from Lemma 21. Then $\widehat{n^\ell}(A)$ is a row of A if and only if two rows of A are identical.*

For each $\ell \geq 3$ as in Theorem 16, we now define the template structure $\mathbf{B}_\ell$ that is used to show the separation $\mathcal{L}(\mathcal{Q}_{\ell+1}^{\mathcal{N}_{\mathrm{CSP}}^\ell}) \not\leq \mathcal{L}(\mathcal{Q}_{\ell-1})$.

▶ **Definition 23.** *Let $\ell \geq 3$ be fixed. Let $\bar{u}$ denote the vector $(0,1,2,0,\ldots,0) \in \mathbb{Z}_3^\ell$. Let $\mathbf{B}_\ell = (\mathbb{Z}_3, R_0, R_1, R_2)$ with each R_j defined as follows. For each $j \in \{0,1,2\}$, let $R_j^{\mathbf{B}_\ell} := \{(a_1,...,a_\ell, \bar{u}\cdot\bar{a}) \mid \bar{a} \in \mathbb{Z}_3^\ell, \sum_{i\in[\ell]} a_i = j\}$.*

▶ **Lemma 24.** *For each $\ell \geq 3$, n^ℓ is a partial polymorphism of the structure $\mathbf{B}_\ell$.*

Proof. Let $j \in \mathbb{Z}_3$ and $\ell \geq 3$. Let $\bar{b}_1, \ldots, \bar{b}_\ell \in R_j^{\mathbf{B}_\ell}$. We show that either, $\widehat{n^\ell}(\bar{b}_1, \ldots, \bar{b}_\ell) \in \{\bar{b}_1, \ldots, \bar{b}_\ell\}$, or $\widehat{n^\ell}(\bar{b}_1, \ldots, \bar{b}_\ell)$ is undefined. Let $A \in \mathbb{Z}_3^{\ell \times \ell}$ be the matrix whose rows are the projections of the tuples $\bar{b}_1, \ldots, \bar{b}_\ell$ to their first ℓ components. If A does not have the ℓ-near-unanimity property, then $\widehat{n^\ell}(A)$ is undefined, and hence $\widehat{n^\ell}(\bar{b}_1, \ldots, \bar{b}_\ell)$ is undefined. So suppose A has the ℓ-near-unanimity property. The row-sums of A are all equal to j by definition of $R_j^{\mathbf{B}_\ell}$. If $\widehat{n^\ell}(A)$ is not a row of A, then A satisfies the assumptions of Lemma 21. Hence $A \cdot (0, 1, 2, 0, 0, \ldots, 0)^T$ contains three distinct entries. This means that $n^\ell(\bar{b}_{1(\ell+1)}, \ldots, \bar{b}_{\ell(\ell+1)})$ is undefined. If $\widehat{n^\ell}(A)$ is a row $\bar{a}$ of A, then by Lemma 22, there are two identical rows in A, and these must be equal to $\bar{a}$. Therefore, the value $\bar{a} \cdot \bar{u}$ also appears at least twice as the $(\ell+1)$st component of the tuples $\bar{b}_1, \ldots, \bar{b}_\ell$. Hence, if $n^\ell(\bar{b}_{1(\ell+1)}, \ldots, \bar{b}_{\ell(\ell+1)})$ is defined, it yields the correct value $\bar{a} \cdot \bar{u}$, and thus $\widehat{n^\ell}(\bar{b}_1, \ldots, \bar{b}_\ell) \in \{\bar{b}_1, \ldots, \bar{b}_\ell\}$. ◀

Template structure for the proof Theorem 18

In Theorem 18, the arity of the quantifiers can be much larger than the arity of the polymorphisms. Thus, the matrices of interest are no longer square, but rectangular, and we cannot directly use Lemma 21. Instead, we use the following. It can be shown with similar reasoning as in the proof of Lemma 20.

▶ **Lemma 25.** *Let $\ell \geq 3$, $r \geq \ell$, $q = \lfloor \frac{r}{\ell-1} \rfloor$, and $p = r - q(\ell-1)$. There is a collection $\mathcal{U} \subseteq \mathbb{Z}_3^r$ of size $|\mathcal{U}| = (\ell-1)q(q-1) + 2pq$ such that for every matrix $A \in \mathbb{Z}_3^{\ell \times r}$ with the ℓ-ary near-unanimity property, one of the following two is the case:*

1. *If $\widehat{n^\ell}(A)$ is not a row of A, there is a vector $\bar{u} \in \mathcal{U}$ such that $A \cdot \bar{u}$ contains three distinct entries.*
2. *If $\widehat{n^\ell}(A)$ is a row $\bar{a}$ of A, then either there exists a $\bar{u} \in \mathcal{U}$ such that $A \cdot \bar{u}$ contains three distinct entries, or for all $\bar{u} \in \mathcal{U}$, $\widehat{n^\ell}(A \cdot \bar{u}) = \bar{a} \cdot \bar{u}$.*

Proof sketch. Split $[r]$ into sets $K_1 \uplus \cdots \uplus K_{\ell-1} = [r]$ such that $|K_1| = \cdots = |K_p| = q+1$ and $|K_{p+1}| = \cdots = |K_{\ell-1}| = q$. Define $\mathcal{U}$ as follows. For every 2-element set of columns $\{j_1, j_2\}$ such that $\{j_1, j_2\} \subseteq K_i$ for some $i \in [\ell-1]$, include in $\mathcal{U}$ two vectors $\bar{u}^{11}, \bar{u}^{12} \in \mathbb{Z}_3^r$ such that $\bar{u}^{11}_{j_1} = 1, \bar{u}^{11}_{j_2} = 1$, and $\bar{u}^{12}_{j_1} = 1, \bar{u}^{12}_{j_2} = 2$. In all other components, the two vectors are zero. Note that $|\mathcal{U}| = 2\big(p\binom{q+1}{2} + (\ell-1-p)\binom{q}{2}\big) = 2\big(p\frac{(q+1)q}{2} + (\ell-1-p)\frac{q(q-1)}{2}\big) = (\ell-1)q(q-1)+2pq$.

Assume first that $\widehat{n^\ell}(A)$ is not a row of A. Then A must have ℓ non-constant columns whose minority entry is in a different row, respectively. So there is an $i \in [\ell-1]$ such that K_i contains at least two of these columns, call them j_1, j_2. By definition of $\mathcal{U}$, the vectors $\bar{u}^{11}, \bar{u}^{12}$ for j_1, j_2 are in $\mathcal{U}$. Then it can be verified that $A \cdot \bar{u}^{11}$ or $A \cdot \bar{u}^{12}$ contains three distinct entries.

Now assume that $\widehat{n^\ell}(A) = \bar{a}$ is a row of A. It can be checked that if $\widehat{n^\ell}(A \cdot \bar{u})$ is defined for all $\bar{u} \in \mathcal{U}$, then for every $\bar{u} \in \mathcal{U}$, there must be at least two rows of $A \cdot \bar{u}$ whose entry is $\bar{a} \cdot \bar{u}$. Then it follows that $\widehat{n^\ell}(A \cdot \bar{u}) = \bar{a} \cdot \bar{u}$ for all $\bar{u} \in \mathcal{U}$, as desired. ◀

The template structure used for Theorem 18 is similar to the previous one, except that it has more extra entries in the relations to forbid the applicability of n^ℓ (using the set of vectors $\mathcal{U}$ from the above lemma). For $r \geq \ell$, the instances of this template are supposed to be indistinguishable in $\mathcal{L}(\mathcal{Q}_r)$. But the construction using the vectors in $\mathcal{U}$ effectively makes it possible to fix two entries of the original relation with just one variable. To ensure $\mathcal{L}(\mathcal{Q}_r)$-indistinguishability of the instances, we therefore need to base the template on $(2r+1)$-ary

relations, augmented by the products with all vectors in $\mathcal{U}$. For every $\ell \geq 3$ and $r \geq \ell$, we define the template structure $\mathbf{B}^*_{r,\ell}$ over $\mathbb{Z}_3$ as follows.

▶ **Definition 26.** *Let $\ell \geq 3$ and $r \geq \ell$. Let $\mathcal{U}$ be the set of vectors from Lemma 25, for $(2r+1)$ instead of r. Note that then $m \coloneqq |\mathcal{U}| = (\ell-1)q(q-1) + 2pq$, where $q = \lfloor \frac{2r+1}{\ell-1} \rfloor$ and $p = 2r+1 - \lfloor \frac{2r+1}{\ell-1} \rfloor(\ell-1) < \ell - 1$. Thus, $m < (\ell-1)q(q-1) + 2(\ell-1)q = (\ell-1)q(q+1) \leq (2r+1)(\lfloor \frac{2r+1}{\ell-1} \rfloor + 1)$. In $\mathbf{B}^*_{r,\ell}$, we define the relations R_j, for $j \in \mathbb{Z}_3$ as: $R_j^{\mathbf{B}^*_{r,\ell}} \coloneqq \{(a_1, \ldots, a_{2r+1}, \bar{a} \cdot \bar{u}_1, \ldots, \bar{a} \cdot \bar{u}_m) \mid \sum_{i \in [2r+1]} a_i = j\}$, where $\bar{u}_1, ..., \bar{u}_m$ is an enumeration of the set $\mathcal{U}$.*

Note that for fixed ℓ, the arity of the template structure is quadratic in r.

▶ **Lemma 27.** *For all $\ell \geq 3$ and $r \geq \ell$, n^ℓ is a partial polymorphism of the structure $\mathbf{B}^*_{r,\ell}$.*

Proof. Analogous to the proof of Lemma 24, now using Lemma 25. ◀

4.2 Constructing instances

We aim to prove $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{\ell+1}) \not\leq \mathcal{L}(\mathcal{Q}_{\ell-1})$ by exhibiting two classes $\mathcal{K}_\ell, \tilde{\mathcal{K}}_\ell$ of structures that cannot be distinguished by any sentence in $\mathcal{L}(\mathcal{Q}_{\ell-1})$ but are separated by a sentence in $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{\ell+1})$. Likewise, we prove $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{q(r,\ell)}) \not\leq \mathcal{L}(\mathcal{Q}_r)$ by constructing classes $\mathcal{K}^*_{r,\ell}, \tilde{\mathcal{K}}^*_{r,\ell}$ of structures that cannot be distinguished by any sentence in $\mathcal{L}(\mathcal{Q}_r)$ but are separated by a sentence in $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{q(r,\ell)})$. Both classes of structures are obtained with a CFI-like construction over an ℓ-regular (or $(2r+1)$-regular, respectively) base graph. With each vertex v in the base graph, we associate a substructure, also called *vertex gadget*. In the first construction, this gadget is denoted $\mathbf{A}(v,s)$. In the second construction it is $\mathbf{A}^*(v,s)$. In both cases, $s \in \mathbb{Z}_3$ is a "charge" we associate with v. We only present $\mathbf{A}(v,s)$ here. The gadget $\mathbf{A}^*(v,s)$ is analogous, but such that it matches the template relations from Definition 26 instead of Definition 23. For $v \in V(G)$, let $E(v) \subseteq E(G)$ denote the set of edges incident to v.

▶ **Definition 28** (Vertex gadgets)**.**
Fix $\ell \geq 3$ and let $G = (V,E)$ be a connected ℓ-regular graph with an order $\leq^G$ on its vertices. Let $\bar{u} \in \mathbb{Z}_3^\ell$ be the vector from Definition 23, $v \in V$ and $\bar{e}(v) = (e_1, \ldots, e_\ell)$ the set $E(v)$, written as a tuple in the order induced by $\leq^G$. Then for each $s \in \mathbb{Z}_3$, we let $\mathbf{A}(v,s) \coloneqq (A_v, R_0^{\mathbf{A}(v,s)}, R_1^{\mathbf{A}(v,s)}, R_2^{\mathbf{A}(v,s)})$, where

- *$A_v \coloneqq B_v \cup C_v$ for $B_v \coloneqq E(v) \times \mathbb{Z}_3$ and $C_v \coloneqq \{v\} \times \mathbb{Z}_3$,*
- *$R_j^{\mathbf{A}(v,s)} \coloneqq \{((e_1,a_1), \ldots, (e_\ell,a_\ell), (v, \bar{a} \cdot \bar{u})) \mid \sum_{i \in [\ell]} a_i = j - s\}$ for each $j \in \{0,1,2\}$,*

We can now define our CFI structures that separate the desired logics. Starting with a suitable base graph, we simply replace every vertex v with a gadget of the form $\mathbf{A}(v,s)$ or $\mathbf{A}^*(v,s)$. We present in the following only the construction that uses the gadgets $\mathbf{A}(v,s)$ (for Theorem 16). The other construction for the proof of Theorem 18 and its relevant properties are completely analogous. Putting together the structures $\mathbf{A}(v,s)$ we have defined for each vertex of G, we obtain our CFI structures:

▶ **Definition 29.** *Let $G = (V, E, \leq^G)$ be an ℓ-regular connected graph with a linear order on its vertices.*

(a) For each $U \subseteq V$, we define $\mathbf{A}(G,U) = (A_G, R_0^{\mathbf{A}(G,U)}, R_1^{\mathbf{A}(G,U)}, R_2^{\mathbf{A}(G,U)})$, where

- *$A_G \coloneqq \bigcup_{v \in V} A_v$,*

- $R_j^{\mathbf{A}(G,U)} \coloneqq \bigcup_{v \in V \setminus U} R_j^{\mathbf{A}(v,0)} \cup \bigcup_{v \in U} R_j^{\mathbf{A}(v,1)}$ *for each* $j \in \{0,1,2\}$.

(b) Furthermore, we define $\mathbf{A}(G) \coloneqq \mathbf{A}(G,\emptyset)$ *and* $\tilde{\mathbf{A}}(G) \coloneqq \mathbf{A}(G,\{\tilde{v}\})$*, where* $\tilde{v} \in V$ *is the smallest vertex of* G *with respect to the order* $\leq^G$*.*

We write $\mathbf{A}^*(G)$ and $\tilde{\mathbf{A}}^*(G)$ to refer to the analogous structures obtained by using the vertex gadgets $\mathbf{A}^*(v,s)$ instead of $\mathbf{A}(v,s)$.

To define the classes of structures that separate the logics, we apply the above construction to base graphs similar to the ones used by Hella in [15]. For a number $k \in \mathbb{N}$, we call a graph *k-connected* if it is connected and remains so after the removal of any k_1 vertices and k_2 edges, for any $k_1 + k_2 < k$. Arbitrarily large graphs having this and other useful properties exist, for example constructions based on higher-dimensional grids.

▶ **Lemma 30.** *For every* $d \geq 2$ *and for infinitely many* $n \in \mathbb{N}$*, there exists a* d*-regular bipartite graph that is* d*-connected and satisfies that each part of the bipartition contains exactly* n *vertices.*

For fixed $d, n \in \mathbb{N}$, we now define a graph that we denote $G_{d,n}$. This will be used as the base graph for our construction detailed above. Let H denote a fixed graph with the properties from Lemma 30 for values $(d-1), n$. Now $G_{d,n}$ consists of $n+1$ disjoint copies of H, denoted $H_1, \ldots, H_{n+1}$. We also add connections between the H_i such that for each $i \in [n+1]$, every vertex in H_i has precisely one edge leading into a different copy H_j. Formally, name the vertices in the two parts of the bipartition of H_i as follows: $V(H_i) = \{v_{ij} \mid j \in [n+1] \setminus \{i\}\} \uplus \{w_{ij} \mid j \in [n+1] \setminus \{i\}\}$. Then let $E(G_{d,n}) \coloneqq \bigcup_{i \in [n+1]} E(H_i) \cup \{v_{ij}w_{ji} \mid i,j \in [n+1], i \neq j\}$. Note that $G_{d,n}$ is again bipartite with the v-vertices and the w-vertices forming the two parts of the bipartition. Moreover, $G_{d,n}$ is d-regular because H is $(d-1)$-regular. For values $n \in \mathbb{N}$ for which Lemma 30 does not guarantee the existence of a graph, $G_{d,n}$ is simply undefined.

For the proofs of Theorems 16 and 18, we use the following classes of structures, which will be shown to be inseparable in the respective logics of interest. For Theorem 16, for each $\ell \geq 3$, the two classes are $K_\ell \coloneqq \{\mathbf{A}(G_{\ell,n}) \mid n \in \mathbb{N}\}$ and $\tilde{\mathcal{K}}_\ell \coloneqq \{\tilde{\mathbf{A}}(G_{\ell,n}) \mid n \in \mathbb{N}\}$. For Theorem 18, for $r \geq \ell \geq 3$, the classes are $K^*_{r,\ell} \coloneqq \{\mathbf{A}^*(G_{2r+1,n}) \mid n \in \mathbb{N}\}$ and $\tilde{\mathcal{K}}^*_{r,\ell} \coloneqq \{\tilde{\mathbf{A}}^*(G_{2r+1,n}) \mid n \in \mathbb{N}\}$.

4.3 (In-)distinguishability of the instances

We first show that there is a sentence in $\mathcal{L}(\mathcal{Q}^{\mathcal{N}^\ell_{\mathrm{CSP}}}_{\ell+1})$ that separates $\mathcal{K}_\ell$ from $\tilde{\mathcal{K}}_\ell$.

▶ **Lemma 31.** *For each* $n \in \mathbb{N}$*, there exists a homomorphism* $h \colon \mathbf{A}(G_{\ell,n}) \to \mathbf{B}_\ell$*.*

Proof. Let $G_{\ell,n} = (V,E)$. A homomorphism $h \colon \mathbf{A}(G_{\ell,n}) \to \mathbf{B}_\ell$ can be defined by letting $h(e,a) = a$ for each $e \in E, a \in \mathbb{Z}_3$, and likewise, $h(v,a) = a$ for every $v \in V$. By definition of the relations R_j in $\mathbf{A}(G_{\ell,n})$ and $\mathbf{B}_\ell$, this is a homomorphism. ◀

▶ **Lemma 32.** *For each* $n \in \mathbb{N}$*, there exists no homomorphism from* $\tilde{\mathbf{A}}(G_{\ell,n})$ *to* $\mathbf{B}_\ell$*.*

Proof. Again, let $G_{\ell,n} = (V,E)$. Consider the following system of linear equations $\mathcal{E}$ over $\mathbb{Z}_3$: For each $e \in E$, it has one variable x_e, and for each $v \in V \setminus \{\tilde{v}\}$, it has the equation $\sum_{e \in E(v)} x_e = 0$. The equation associated with $\tilde{v}$ reads $\sum_{e \in E(\tilde{v})} x_e = 1$. Now suppose for a contradiction that there exists a homomorphism $h \colon \tilde{\mathbf{A}}(G_{\ell,n}) \to \mathbf{B}_\ell$. Then h defines a solution λ for $\mathcal{E}$ via $\lambda(x_e) \coloneqq h(e,0)$.

But such a solution cannot exist: Let $S \uplus T = V(G)$ be the bipartition of $G_{\ell,n}$. We assume w.l.o.g. that $\tilde{v} \in S$. Then we have $\sum_{v \in S} \sum_{e \in E(v)} \lambda(x_e) = 1$ and $\sum_{v \in T} \sum_{e \in E(v)} \lambda(x_e) =$

0. However, this is a contradiction since clearly $\sum_{v\in S}\sum_{e\in E(v)}\lambda(x_e) = \sum_{e\in E}\lambda(x_e) = \sum_{v\in T}\sum_{e\in E(v)}\lambda(x_e)$. ◀

▶ **Corollary 33.** *There is a sentence $\psi \in \mathcal{L}(\mathcal{Q}_{\ell+1}^{\mathcal{N}_{\mathrm{CSP}}^{\ell}})$ such that for each $n \in \mathbb{N}$, $\mathbf{A}(G_{\ell,n})$ and $\tilde{\mathbf{A}}(G_{\ell,n})$ are separated by ψ.*

Proof. By Lemma 24, n^ℓ is a polymorphism of $\mathbf{B}_\ell$. Thus, the CSP quantifier for the $(\ell+1)$-ary template $\mathbf{B}_\ell$ is in $\mathcal{Q}_{\ell+1}^{\mathcal{N}_{\mathrm{CSP}}^{\ell}}$. By Lemmas 31 and 32, this quantifier distinguishes $\mathbf{A}(G_{\ell,n})$ and $\tilde{\mathbf{A}}(G_{\ell,n})$, for every $n \in \mathbb{N}$. ◀

The preceding assertions can be shown in exactly the same way for the classes $\mathcal{K}^*_{r,\ell}$ and $\tilde{\mathcal{K}}^*_{r,\ell}$, where the polymorphism-closedness of the template $\mathbf{B}^*_{r,\ell}$ is due to Lemma 27.

To finish the proof of the separation $\mathcal{L}(\mathcal{Q}_{\ell+1}^{\mathcal{N}_{\mathrm{CSP}}^{\ell}}) \not\leq \mathcal{L}(\mathcal{Q}_{\ell-1})$ (and thus of Theorem 16), we show that no sentence in $\mathcal{L}(\mathcal{Q}_{\ell-1})$ can distinguish between the classes $\mathcal{K}_\ell$ and $\tilde{\mathcal{K}}_\ell$. To achieve this, we show that for every $k \geq \ell - 1$, there exists $n \in \mathbb{N}$ such that Duplicator has a winning strategy in the game $\mathrm{BP}^k_{\ell-1}(\mathbf{A}(G_{\ell,n}), \tilde{\mathbf{A}}(G_{\ell,n}))$. Then it follows with Corollary 3 that $\mathcal{K}_\ell$ and $\tilde{\mathcal{K}}_\ell$ cannot be separated by a sentence of $\mathcal{L}(\mathcal{Q}_{\ell-1})$.

The indistinguishability of $\mathcal{K}^*_{r,\ell}$ and $\tilde{\mathcal{K}}^*_{r,\ell}$ in $\mathcal{L}(\mathcal{Q}_r)$ is shown with a similar argument. So from now on, let $k \geq \ell - 1$ be a fixed number of pebbles. We choose $n \in \mathbb{N}$ so that $G_{\ell,n}$ is defined, and n is sufficiently larger than k. In the following, write $G \coloneqq G_{\ell,n}$ and let V and E be the vertex and edge set of G. Recall that G is the disjoint union of identical graphs $H_1, \dots H_{n+1}$ with certain connections between them.

▶ **Lemma 34.** *Duplicator has a wining strategy in the game $\mathrm{BP}^k_{\ell-1}(\mathbf{A}(G), \tilde{\mathbf{A}}(G))$.*

The lemma is proved by showing that Duplicator can maintain an invariant which prevents Spoiler from winning. Let (α, β) be a position in the game $\mathrm{BP}^k_{\ell-1}(\mathbf{A}(G), \tilde{\mathbf{A}}(G))$ with $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta) \subseteq \{x_1, \dots, x_k\}$. For the sake of simplicity, we denote the tuples of *pebbled elements* in the image of α and β by $\bar{\alpha}$ and $\bar{\beta}$, respectively. Furthermore, we call $(\bar{\alpha}, \bar{\beta})$ the corresponding *pebble position* of the game.

A vertex $u \in V(G)$ is *safe* in a pebble position $(\bar{\alpha}, \bar{\beta})$ if $u = v_{ij}$ or $u = w_{ij}$ and A_v is pebble-free (i.e., no component of $\bar{\alpha}$ or $\bar{\beta}$ is in A_v) for every $v \in V(H_i) \cup V(H_j)$.

Let $u \in V$. We call a bijection $f\colon A_G \to A_G$ *good bar* u if for every $v \in V \setminus \{u\}$, $f|_{A_v}\colon A_v \to A_v$ is a partial isomorphism $\mathbf{A}(G) \to \tilde{\mathbf{A}}(G)$ (thus, f is an isomorphism between $\mathbf{A}(G)$ and $\tilde{\mathbf{A}}(G)$ except at A_u).

To prove Lemma 34, we show that Duplicator can maintain the following **invariant** for the pebble positions $(\bar{\alpha}, \bar{\beta})$ reached by the moves of the players: There exists a bijection $f\colon A_G \to A_G$ with $f(\bar{\alpha}) = \bar{\beta}$ that is good bar u, for some $u \in V$ which is *safe*.

▶ **Lemma 35.** *Let $(\bar{\alpha}, \bar{\beta})$ be the current pebble position in $\mathrm{BP}^k_{\ell-1}(\mathbf{A}(G), \tilde{\mathbf{A}}(G))$. Assume the invariant holds in the beginning of the round. Then Duplicator has a strategy to ensure the invariant also in the beginning of the following round.*

A full proof of the lemma can be found in Appendix B. The idea is as follows. Duplicator begins the round by playing the bijection $f\colon \mathbf{A}(G) \to \tilde{\mathbf{A}}(G)$ sending $\bar{\alpha}$ to $\bar{\beta}$ which exists because the invariant holds at the beginning of the round. Let u be the vertex such that f is good bar u. Now Spoiler modifies the position of up to $\ell - 1$ pebbles in both structures, leading to a new pebble position $(\bar{\alpha}', \bar{\beta}')$. Duplicator has to find a new bijection f' that satisfies the invariant again. This can be done using a pebble-free path P from u to another vertex u' which is safe with respect to $(\bar{\alpha}', \bar{\beta}')$: Indeed, a standard property of CFI-like constructions is that inconsistencies in bijections can always be "shifted" along a path in

the base graph. More precisely, for every path P, and each constant $c \in \mathbb{Z}_3$, there exists a bijection $g_{P,c}\colon \mathbf{A}(G) \to \tilde{\mathbf{A}}(G)$ that induces a local isomorphism everywhere except at the vertex gadgets belonging to the two endpoints v_1, v_2 of P: For each $j \in \{1,2\}$, $g_{P,c}$ induces an isomorphism from $\mathbf{A}(v_j, s_j)$ to $\mathbf{A}(v_j, s_j + c)$, where $s_j \in \mathbb{Z}_3$ is the charge associated with v_j in $\mathbf{A}(G)$. In other words, for P a pebble-free path from u to some safe u', and for the appropriate choice of $c \in \mathbb{Z}_3$, $f' \coloneqq g_{P,c} \circ f$ will be good bar u'. It remains to argue why such a pebble-free path to a safe vertex always exists. This can be done exactly like in [15]: Because the number of copies of H_i in G is sufficiently larger than k, there always exists some pebble-free copy H_i and safe $u' \in V(H_i)$. The pebble-avoiding path P from u to u' exists because each copy H_i is sufficiently connected. The only difference to [15] is that now, if a pebble has been placed on $C_u = \{u\} \times \mathbb{Z}_3$, one has to take extra care: If the path P starts in u via the edge e_2 or $e_3 \in \bar{e}(u)$, then there is a problem because $\bar{u}_i \neq 0$ for $i \in \{2,3\}$ (see Definition 23 for the vector $\bar{u}$). As a consequence of this, $g_{P,c}$ cannot be defined so that it fixes the pebbled elements in C_u. The workaround in this case is to use two paths P_1, P_2 from u to the same safe vertex u', one via e_2 and one via e_3, and to suitably distribute the value $c \in \mathbb{Z}_3$ by which f has to be corrected at u across the two paths.

The **proof of Theorem 16** is now straightforward: Corollary 33 shows that for every fixed $\ell \geq 3$, our constructed classes $\mathcal{K}_\ell$ and $\tilde{\mathcal{K}}_\ell$ are separated by a sentence of $\mathcal{L}(\mathcal{Q}_{\ell+1}^{\mathcal{N}_{\mathrm{CSP}}^\ell})$. Lemma 34 together with Corollary 3 shows that the classes are inseparable in $\mathcal{L}(\mathcal{Q}_{\ell-1})$. Thus, $\mathcal{L}(\mathcal{Q}_{\ell+1}^{\mathcal{N}_{\mathrm{CSP}}^\ell}) \not\leq \mathcal{L}(\mathcal{Q}_{\ell-1})$. The **proof of Theorem 18** is analogous. We show that for every pebble number $k \geq r$, there is an $n \in \mathbb{N}$ such that Duplicator has a winning strategy in $\mathrm{BP}_r^k(\mathbf{A}^*(G_{2r+1,n}), \tilde{\mathbf{A}}^*(G_{2r+1,n}))$. The invariant that Duplicator maintains in this game is the same as above. From the existence of the winning strategy it follows that $\mathcal{K}_{r,\ell}^*$ and $\tilde{\mathcal{K}}_{r,\ell}^*$ cannot be separated by a sentence in $\mathcal{L}(\mathcal{Q}_r)$. Separability of the classes with a CSP quantifier is shown similarly as in Corollary 33.

5 Limitations of partial-Maltsev-closed quantifiers

Write $\mathcal{Q}_r^{\mathcal{M}}$ for the class of all r-ary quantifiers that are closed under the partial Maltsev operation. Note that the Maltsev operation is a fixed operation of arity three, and thus the arity of the polymorphism is not a variable that plays a role in out study. Instead, we establish a hierarchy based on the arity of the quantifiers and also construct an example that separates the logic with r-ary Maltsev-closed quantifiers from the logic with all r-ary generalized quantifiers, at least in the context of a fixed number of variables. The interest in this lies in the novel nature of the construction and also the novelty in the use of a Spoiler-Duplicator game specific to Maltsev-closed quantifiers. In the present section we give a brief summary of the results. Full details of the construction and proofs can be found in Appendix C.

We obtain the arity hierarchy for Maltsev-closed quantifiers by noting that the quantifiers used to establish the arity hierarchy in [15] are Maltsev-closed. To be specific, define for each $r \geq 3$ the structure $\mathbf{B}_r = (\{0,1\}, R_0, R_1)$ where $R_i = \{(a_1, \ldots, a_r) \mid \sum_{1 \leq j \leq r} a_j = i \pmod 2\}$. Then, it follows from the proof of [15, Theorem 8.6] that $\mathrm{CSP}(\mathbf{B}_{r+1})$ is not definable in $\mathcal{L}(\mathcal{Q}_r^{\mathcal{M}})$. Since $\mathbf{B}_r$ admits a Maltsev polymorphism, the following is immediate.

▶ **Theorem 36.** *For every $r \geq 3$, $\mathcal{L}(\mathcal{Q}_r^{\mathcal{M}}) \lneq \mathcal{L}(\mathcal{Q}_{r+1}^{\mathcal{M}})$.*

Finally, we show that if we limit our logic to exactly k variables, then there is a quantifier of arity k that cannot be defined by using Maltsev-closed quantifiers of arity k, giving us the

following theorem.

▶ **Theorem 37.** *For every* $k \geq 3$, $\mathcal{L}^k(\mathcal{Q}_k^{\mathcal{M}}) \lneq \mathcal{L}^k(\mathcal{Q}_k)$.

The proof requires a novel construction and substantial technical work to establish. The details are given in Appendix C. It leaves open the interesting question of whether we can show that this quantifier is not definable in $\mathcal{L}(\mathcal{Q}_k^{\mathcal{M}})$, without the limitation on the number of variables. The apparent difficulty in obtaining such inexpressibility results for Maltsev-closed quantifiers lies in the fact that the typical hard examples, which are based on the CFI-construction, cannot be used: Their templates admit a Maltsev polymorphism, and so they are distinguishable by a quantifier in $\mathcal{Q}^{\mathcal{M}}$.

6 Conclusion

Our main results show that the quantifiers closed under ℓ-ary partial near-unanimity polymorphisms introduced at CSL'24 [10] indeed give rise to new and interesting logics whose expressive power, in particular if $r = \ell + 1$, is strictly different from any of the logics $\mathcal{L}(\mathcal{Q}_r)$ that were studied by Hella in 1996 [15].

This establishes the logics $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell})$ as formalisms whose finite-model-theoretic properties deserve further study. One question that comes to mind is what can be said about the equivalence relation induced by $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell})$ on finite structures. Can it be characterized as a *homomorphism indistinguishability relation* or via a *game comonad* [1]? For the logics $\mathcal{L}(\mathcal{Q}_r)$, the answer is affirmative [7], but for example, the invertible-map equivalences (which also stem from certain generalized quantifiers strictly between $\mathcal{Q}_r$ and $\mathcal{Q}_{r+1}$), there provably is no such characterization [22], so it would be interesting to clarify the situation for $\mathcal{L}(\mathcal{Q}_r^{\mathcal{N}^\ell})$. It would also be desirable to tighten the gap between the arity reduction we have in Theorem 15 and the quadratic arity increase in Theorem 18: It seems plausible that already a smaller increase in arity should give strictly more expressive power.

The long-term goal of the study of polymorphism-closed quantifiers is, however, to establish inexpressibility results for quantifiers stronger than $\mathcal{Q}_r^{\mathcal{N}^\ell}$, in particular the partial Maltsev-closed ones. Theorem 37 is a first (and highly non-trivial) step in that direction but ultimately, we would like to exhibit an r-ary class of structures that is undefinable in the full logic $\mathcal{L}(\mathcal{Q}_r^{\mathcal{M}})$, without a restriction on the number of variables.

An even harder open problem is to understand the expressive power of quantifiers closed under partial *weak* near-unanimity polymorphism, for which we do not even know a model-comparison game yet. From a CSP perspective, weak near-unanimity is the most interesting condition as it characterizes precisely the finite-domain CSPs that are solvable in polynomial time (assuming P≠NP).

References

1 S. Abramsky, A. Dawar, and P. Wang. The pebbling comonad in finite model theory. In *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–12. IEEE, 2017.

2 L. Barto and M. Kozik. Constraint satisfaction problems of bounded width. In *2009 50th Annual IEEE symposium on foundations of computer science*, pages 595–603. IEEE, 2009.

3 L. Barto, A. Krokhin, and R. Willard. Polymorphisms, and How to Use Them. In Andrei Krokhin and Stanislav Zivny, editors, *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*, pages 1–44. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2017. URL: `https://drops.dagstuhl.de/entities/document/10.4230/DFU.Vol7.15301.1`, `doi:10.4230/DFU.Vol7.15301.1`.

4 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.

5 A. Bulatov and V. Dalmau. A simple algorithm for Mal'tsev constraints. *SIAM J. Comput.*, 36(1):16–27, 2006. doi:10.1137/050628957.

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

7 A. Ó Conghaile and A. Dawar. Game comonads & generalised quantifiers. In Christel Baier and Jean Goubault-Larrecq, editors, *29th EACSL Annual Conference on Computer Science Logic, CSL 2021, January 25-28, 2021, Ljubljana, Slovenia (Virtual Conference)*, volume 183 of *LIPIcs*, pages 16:1–16:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. URL: https://doi.org/10.4230/LIPIcs.CSL.2021.16, doi:10.4230/LIPICS.CSL.2021.16.

8 A. Dawar. Linear algebraic quantifiers. In K. Meer, A. Rabinovich, E. Ravve, and A. Villaveces, editors, *Model Theory, Computer Science, and Graph Polynomials*, pages 215 – 237. Springer, 2025. Festschrift in Honor of Johann A. Makowsky.

9 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*, 2019. doi:10.4230/LIPIcs.ICALP.2019.112.

10 A. Dawar and L. Hella. Quantifiers Closed Under Partial Polymorphisms. In *32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)*, volume 288 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 23:1–23:19, 2024. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CSL.2024.23, doi:10.4230/LIPIcs.CSL.2024.23.

11 R. Diestel. *Graph Theory*. Springer, 5th edition, 2017.

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

13 H.D. Ebbinghaus. Extended logics: the general framework. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 25 – 76. Springer, 1985.

14 R. Fagin, P. G. Kolaitis, and L. Popa. Data exchange: getting to the core. *ACM Trans. Database Syst.*, 30:174–210, 2005. doi:10.1145/1061318.1061323.

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

16 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.

17 P. Jonsson, V. Lagerkvist, G. Nordh, and B. Zanuttini. Strong partial clones and the time complexity of SAT problems. *J. Comput. Syst. Sci.*, 84:52–78, 2017. URL: https://doi.org/10.1016/j.jcss.2016.07.008, doi:10.1016/J.JCSS.2016.07.008.

18 V. Lagerkvist and M. Wahlström. Sparsification of SAT and CSP problems via tractable extensions. *ACM Trans. Comput. Theory*, 12(2):13:1–13:29, 2020. doi:10.1145/3389411.

19 V. Lagerkvist and M. Wahlström. The (Coarse) Fine-Grained Structure of NP-Hard SAT and CSP Problems. *ACM Trans. Comput. Theory*, 14(1):2:1–2:54, 2022. doi:10.1145/3492336.

20 L. Libkin. *Elements of Finite Model Theory*. Springer, 2004.

21 M. Lichter. Separating rank logic from polynomial time. *J. ACM*, 70(2):14:1–14:53, 2023. doi:10.1145/3572918.

22 M. Lichter, B. Pago, and T. Seppelt. Limitations of Game Comonads for Invertible-Map Equivalence via Homomorphism Indistinguishability. In Aniello Murano and Alexandra Silva, editors, *32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)*, volume 288 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 36:1–36:19, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CSL.2024.36, doi:10.4230/LIPIcs.CSL.2024.36.

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

A Details omitted from Section 3

▶ **Proposition 8.** *There is an* FO*-definable reduction from* $\mathrm{CSP}(H^r_m)$ *to* $\mathrm{CSP}(K_m)$.

Proof. Given $\mathbf{A}$ in the vocabulary $\{R\}$ with $\mathrm{ar}(R) = r$, let $\mathcal{G}(\mathbf{A})$ be the graph (possibly with self-loops) on vertex set A with edges $\{(a_i, a_j) \mid (a_1, \ldots, a_r) \in R^{\mathbf{A}}, i \neq j\}$. Clearly $\mathcal{G}(\mathbf{A})$ is FO-definable in $\mathbf{A}$. Now we claim that $\mathbf{A} \to H^r_m$ if, and only if, $\mathcal{G}(\mathbf{A}) \to K_m$.

If $h : \mathbf{A} \to H^r_m$ is a homomorphism, then h is also a homomorphism from $\mathcal{G}(\mathbf{A})$ to K_m because if (a, b) is an edge in $\mathcal{G}(\mathbf{A})$, then there is a tuple in $R^{\mathbf{A}}$ in which a and b appear together. Thus, h maps a and b to different vertices in $[m]$, as desired.

Conversely, let $h\colon \mathcal{G}(\mathbf{A}) \to K_m$ be a homomorphism, and let $\bar{a} \in R^{\mathbf{A}}$. Then the components of $\bar{a}$ are all distinct because K_m has no self-loops, so otherwise, this homomorphism h would not exist. In $\mathcal{G}(\mathbf{A})$, the components of $\bar{a}$ form a clique, and so h maps them all to distinct vertices in $[m]$, and hence $h(\bar{a}) \in R^r_m$. ◀

▶ **Lemma 12.** *Let* $r \geq \ell \geq 3$. *Let* $\mathcal{K}$ *be an* $\mathcal{N}^\ell$*-closed class of* τ*-structures, where* $\mathrm{ar}(R) \leq r$ *for all* $R \in \tau$. *Then there exists a vocabulary* σ *with at most* $\lceil \frac{\ell-1}{\ell} \cdot r \rceil$*-ary relations, a class* $\mathcal{K}'$ *of* σ*-structures, and an* $\mathcal{L}$*-definable reduction* f *that maps* τ*-structures to* σ*-structures in such a way that* $\mathbf{A} \in \mathcal{K}$ *if, and only if,* $f(\mathbf{A}) \in \mathcal{K}'$.

Proof. We start by defining the reduction f. For any τ-structure $\mathbf{A}$, note that $\mathbf{A} \leq n^\ell(\mathbf{A})$. Thus, iterating the operation n^ℓ results, in a finite number of steps in a structure $\mathbf{A}^+$ with $\mathbf{A} \leq \mathbf{A}^+$ and $n^\ell(\mathbf{A}^+) = \mathbf{A}^+$, i.e. n^ℓ is a partial polymorphism of $\mathbf{A}^+$. It is easily seen that this process can be defined as an inflationary fixed-point and thus $\mathbf{A}^+$ is $\mathcal{L}$-definable in $\mathbf{A}$.

Let $(i_1, j_1), (i_2, j_2), \ldots, (i_\ell, j_\ell)$ be as in the proof of Corollary 11. We define now $f(\mathbf{A}) \coloneqq (\mathbf{A}^+)^*$, where the operator $\mathbf{C} \mapsto \mathbf{C}^*$ is as defined before Lemma 10. Note that f is $\mathcal{L}$-definable as the composition of an $\mathcal{L}$-definable and FO-definable reduction. Furthermore, we define $\mathcal{K}'$ simply as $f(\mathcal{K})$. Thus, the vocabulary σ of $\mathcal{K}'$ is $\{R_n \mid R \in \tau, n \in [\ell]\}$. As in Corollary 11, the arity of all relations in σ is at most $\lceil \frac{\ell-1}{\ell} \cdot r \rceil$.

We still need to show that $\mathbf{A} \in \mathcal{K}$ if, and only if, $f(\mathbf{A}) \in \mathcal{K}'$. The only direction that requires work is to prove that $f(\mathbf{A}) \in \mathcal{K}'$ implies $\mathbf{A} \in \mathcal{K}$. Assume for this that $f(\mathbf{A}) \in \mathcal{K}'$. Then there exists $\mathbf{B} \in \mathcal{K}$ such that $(\mathbf{B}^+)^* = (\mathbf{A}^+)^*$. It follows now from Lemma 9 that $\mathbf{B}^+ = \mathbf{A}^+$: for any $R \in \tau$ and $\bar{a} \in A^r = B^r$ we have the following sequence of equivalences:

$$\begin{aligned}
\bar{a} \in R^{\mathbf{A}^+} &\iff p_{\setminus[i_n,j_n]}(\bar{a}) \in R_n^{\mathbf{A}^+} \text{ for all } n \in [\ell] \text{ such that } i_n \leq \mathrm{ar}(R) \\
&\iff p_{\setminus[i_n,j_n]}(\bar{a}) \in R_n^{\mathbf{B}^+} \text{ for all } n \in [\ell] \text{ such that } i_n \leq \mathrm{ar}(R) \\
&\iff \bar{a} \in R^{\mathbf{B}^+}.
\end{aligned}$$

Since $\mathbf{B} \in \mathcal{K}$ and $\mathcal{K}$ is n^ℓ-closed, we have $\mathbf{B}^+ \in \mathcal{K}$. But then also $\mathbf{A} \in \mathcal{K}$ because $\mathbf{A} \leq \mathbf{A}^+ = \mathbf{B}^+$ and $\mathcal{K}$ is downwards monotone. ◀

B Details omitted from Section 4

We begin with two auxiliary lemmas that do not appear in the main part of the paper but will be needed in some of the omitted proofs.

▶ **Lemma 38.** *Let* $\ell \geq 3$ *and* $r \geq \ell$. *Let* $A \in \mathbb{Z}_3^{\ell \times r}$ *be a matrix that has the* ℓ*-ary near-unanimity property. If* $\widehat{n^\ell}(A)$ *is not a row of* A, *then there are at least* ℓ *columns in* A *that are non-constant and have their minority entry in* ℓ *distinct rows.*

Proof. Suppose for a contradiction that there are at most $\ell - 1$ rows of A which contain the minority values of columns. Then there is at least one row A_{i-} which contains in each column the respective majority value of that column. So $\widehat{n^\ell}(A) = A_{i-}$. ◀

▶ **Lemma 39.** *Let $\ell \geq 3$ and $r \geq \ell$. Let $A \in \mathbb{Z}_3^{\ell \times r}$ be a matrix with the ℓ-ary near-unanimity property. Let A_{-j_1}, A_{-j_2} be two non-constant columns of A that have their respective minority entry in distinct rows. Let $\bar{u}_{11} \in \mathbb{Z}_3^r$ be the vector that has a 1-entry at index j_1 and j_2, and is 0 everywhere else. Let $\bar{u}_{12} \in \mathbb{Z}_3^r$ be the vector that has a 1-entry at index j_1, a 2-entry at index j_2, and is 0 everywhere else. Then the vector $A \cdot \bar{u}_{11}$ or the vector $A \cdot \bar{u}_{12}$ contains three distinct entries.*

Proof. Let $a_{j_1}, a_{j_2} \in \mathbb{Z}_3$ denote the majority values in the respective columns j_1, j_2, and let b_{j_1}, b_{j_2} be the respective uniquely different values in these columns. By the assumption of the lemma, b_{j_1}, b_{j_2} are in distinct rows, say $i_1, i_2 \in [\ell]$. We have $a_{j_1} \neq b_{j_1}$ and $a_{j_2} \neq b_{j_2}$. There are two cases. Suppose first that $b_{j_1} = b_{j_2}$. Let $v = a_{j_1} + a_{j_2}, w = b_{j_1} + a_{j_2}, x = a_{j_1} + b_{j_2}$. Note that these three entries appear in $A \cdot \bar{u}_{11}$ in rows i_1, i_2, and any other row. Now we have $v \neq w$ (since $a_{j_1} \neq b_{j_1}$) and $v \neq x$ (since $a_{j_2} \neq b_{j_2}$). If $w \neq x$, then we are done. If $w = x$, then consider $v' = a_{j_1} + 2a_{j_2}, w' = b_{j_1} + 2a_{j_2}, x' = a_{j_1} + 2b_{j_2}$. These values appear in $A \cdot \bar{u}_{12}$. As before, $v' \neq w'$ and $v' \neq x'$ is easy to see. Since $w = x$, it follows that $w' - x' = (w - x) + a_{j_2} - b_{j_2} = a_{j_2} - b_{j_2} \neq 0$. Thus, $w' \neq x'$, and so, $A \cdot \bar{u}_{12}$ contains three distinct entries. ◀

Now we are ready to present the proofs of the lemmas in Section 4.

▶ **Lemma 21.** *Let $\ell \geq 3$ and let $A \in \mathbb{Z}_3^{\ell \times \ell}$ be a matrix satisfying:*

1. *A has the ℓ-near-unanimity property.*
2. *The row-sums of A are all equal.*
3. *$\widehat{n^\ell}(A)$ is not a row of A.*

Then $A \cdot (0, 1, 2, 0, 0, ..., 0)^T$ is a vector in $\mathbb{Z}_3^\ell$ that has at least three distinct entries.

Proof. By Lemma 38, all the ℓ columns of A are non-constant and have their minority entry in distinct rows. Without loss of generality we can assume that the respective minority entries are on the diagonal (otherwise, reorder the rows). Let $A' \in \mathbb{Z}_3^{3\times 3}$ be the 3×3-submatrix of A in the top left corner. It is easy to see that A' satisfies conditions 1) and 3) from Lemma 20. It also satisfies condition 2) from that lemma because the row sums of A are all equal by assumption 2), and in A, the first three rows are identical in all columns except the first three. Therefore, the first three entries of the vector $A \cdot (0, 1, 2, 0, 0, ..., 0)^T$ are all distinct by Lemma 20. ◀

▶ **Lemma 22.** *Let $\ell \geq 3$, and let $A \in \mathbb{Z}_3^{\ell \times \ell}$ be a matrix that satisfies the first two conditions from Lemma 21. Then $\widehat{n^\ell}(A)$ is a row of A if and only if two rows of A are identical.*

Proof. Since A has the ℓ-near-unanimity property by assumption, $\widehat{n^\ell}(A)$ is defined. If A has at least two rows that are identical, then $\widehat{n^\ell}(A)$ must be that row because for each $j \in [\ell]$, the majority value in the j-th column of A is equal to the j-th entry of the row that appears twice.

Conversely, assume that the polymorphism returns a row $\bar{a}$ of A. Then every entry of $\bar{a}$ is the majority one in its column. Assume for a contradiction that row $\bar{a}$ occurs only once in A. Then for every other row $\bar{a}'$ of A, there exists a column $j \in [r]$ such that $\bar{a}_j \neq \bar{a}'_j$. But this

cannot be the same $j \in [\ell]$ for more than one row $\bar{a}'$ because then, the column j would violate the ℓ-near-unanimity property. Therefore, we can define a submatrix $A' \in \mathbb{Z}_3^{(\ell-1)\times(\ell-1)}$ which arises from A by removing $\bar{a}$ and a column j so that up to reordering of rows, in A', each column contains precisely one value that is different from the others, and this is on the diagonal. This also means that each row of A' is identical to $\bar{a}$ except at precisely one position. Hence, the sum over any row of A' differs from $\sum \bar{a}$. In the removed column j, all entries except possibly one are equal to $\bar{a}_j$. So in each row of A except potentially one, the row sum is not equal to $\sum \bar{a}$ because the rows differ at exactly one position. Hence the second assumption is violated, and this case cannot occur. Thus, there is at least one other row of A that is equal to $\bar{a}$. ◀

▶ **Lemma 25.** *Let $\ell \geq 3$, $r \geq \ell$, $q = \lfloor \frac{r}{\ell-1} \rfloor$, and $p = r - q(\ell-1)$. There is a collection $\mathcal{U} \subseteq \mathbb{Z}_3^r$ of size $|\mathcal{U}| = (\ell-1)q(q-1) + 2pq$ such that for every matrix $A \in \mathbb{Z}_3^{\ell \times r}$ with the ℓ-ary near-unanimity property, one of the following two is the case:*

1. *If $\widehat{n^\ell}(A)$ is not a row of A, there is a vector $\bar{u} \in \mathcal{U}$ such that $A \cdot \bar{u}$ contains three distinct entries.*
2. *If $\widehat{n^\ell}(A)$ is a row $\bar{a}$ of A, then either there exists a $\bar{u} \in \mathcal{U}$ such that $A \cdot \bar{u}$ contains three distinct entries, or for all $\bar{u} \in \mathcal{U}$, $\widehat{n^\ell}(A \cdot \bar{u}) = \bar{a} \cdot \bar{u}$.*

Proof. Split $[r]$ into sets $K_1 \uplus \cdots \uplus K_{\ell-1} = [r]$ such that $|K_1| = \cdots = |K_p| = q+1$ and $|K_{p+1}| = \cdots = |K_{\ell-1}| = q$. Define $\mathcal{U}$ as follows. For every 2-element set of columns $\{j_1, j_2\}$ such that $\{j_1, j_2\} \subseteq K_i$ for some $i \in [\ell-1]$, include in $\mathcal{U}$ the respective vectors $\bar{u}_{11}, \bar{u}_{12}$ given by Lemma 39 for this choice of j_1, j_2. Note that $|\mathcal{U}| = 2\big(p\binom{q+1}{2} + (\ell-1-p)\binom{q}{2}\big) = 2\big(p\frac{(q+1)q}{2} + (\ell-1-p)\frac{q(q-1)}{2}\big) = (\ell-1)q(q-1) + 2pq$.

Assume first that $\widehat{n^\ell}(A)$ is not a row of A. By Lemma 38, A has ℓ columns that have a distinguished entry which is in a different row, respectively. So there is an $i \in [\ell-1]$ such that K_i contains at least two of these columns, call them j_1, j_2. So by definition of $\mathcal{U}$, the vectors $\bar{u}_{11}, \bar{u}_{12}$ for j_1, j_2 are in $\mathcal{U}$, and by Lemma 39, one of them satisfies that its product with A contains three distinct entries.

Now assume that $\widehat{n^\ell}(A) = \bar{a}$ is a row of A. Let $\bar{u} \in \mathcal{U}$. Suppose $\widehat{n^\ell}(A \cdot \bar{u})$ is defined, i.e. the vector $A \cdot \bar{u}$ is constant or has at most one entry that is different from the others. We want to show that the value $\bar{a} \cdot \bar{u}$ appears at least twice in the vector $A \cdot \bar{u}$ – then it follows that $\widehat{n^\ell}(A \cdot \bar{u}) = \bar{a} \cdot \bar{u}$. Let j_1, j_2 be the two columns of A where $\bar{u}$ is non-zero. If $\ell \geq 4$, then there are at least two distinct rows i_1, i_2 of A such that $A_{i_1 j_1} = A_{i_2 j_1} = \bar{a}_{j_1}$, and $A_{i_1 j_2} = A_{i_2 j_2} = \bar{a}_{j_2}$ (because $\bar{a}_{j_1}$ and $\bar{a}_{j_2}$ are the respective majority entries in their columns, and all but at most one entry in the column must have this value). This means that $\bar{a} \cdot \bar{u}$ appears in row i_1 and in row i_2 of the vector $A \cdot \bar{u}$, which is at least twice, as desired. In case that $\ell = 3$, it can happen that there are no two distinct rows i_1, i_2 such that $A_{i_1 j_1} = A_{i_2 j_1} = \bar{a}_{j_1}$, and $A_{i_1 j_2} = A_{i_2 j_2} = \bar{a}_{j_2}$. But then w.l.o.g. the situation is this: $A_{1 j_1} \neq \bar{a}_{j_1}$ is the minority value in column j_1, $A_{2 j_2} \neq \bar{a}_{j_2}$ is the minority value in column j_2, and $A_{3 j_1} = \bar{a}_{j_1}$, $A_{3 j_2} = \bar{a}_{j_2}$ are the respective majority entries in the columns. Then by Lemma 39, one of the vectors $A \cdot \bar{u}_{11}$ and $A \cdot \bar{u}_{12}$ for the columns j_1, j_2 has three distinct entries. Since these $\bar{u}_{11}, \bar{u}_{12}$ are in $\mathcal{U}$, there exists a $\bar{u} \in \mathcal{U}$ such that $A \cdot \bar{u}$ has three distinct entries, and this is the other way in which the second statement of the lemma can be satisfied. ◀

▶ **Lemma 30.** *For every $d \geq 2$ and for infinitely many $n \in \mathbb{N}$, there exists a d-regular bipartite graph that is d-connected and satisfies that each part of the bipartition contains exactly n vertices.*

Proof. In case that d is even, a $(d/2)$-dimensional toroidal grid of suitable size satisfies these properties. If d is odd, then we can take two copies of a $(d-1)/2$-dimensional toroidal grid and join every vertex with its copy to achieve degree d at every vertex. ◀

Duplicator winning strategy

We present the details that are needed to formally prove Lemma 34. We first of all consider properties of bijections on the gadgets $\mathbf{A}(v,s)$. Fix a graph $G = (V,E)$ as the base graph. Let $v \in V$ with $\bar{e}(v) = (e_1, \dots, e_\ell)$. We say that a bijection $f\colon A_v \to A_v$ is *edge-preserving* if $\{f(e_i,0), f(e_i,1), f(e_i,2)\} = \{(e_i,0),(e_i,1),(e_i,2)\}$ for all $i \in [\ell]$, and $\{f(v,0), f(v,1), f(v,2)\}$
$= \{(v,0),(v,1),(v,2)\}$. Furthermore, an edge-preserving f is *cyclic* if for each $i \in [\ell]$, there is $c_i \in \mathbb{Z}_3$ such that for all $a \in \mathbb{Z}_3$,

- $f(e_i,a) = (e_i, a + c_i)$,
- $f(v,a) = (v, a + c_s + 2 \cdot c_t)$, where $\bar{u}_s = 1$ and $\bar{u}_t = 2$ are the nonzero components of $\bar{u}$.

The *shift value* of a cyclic bijection f is $c(f) \coloneqq \sum_{i\in[\ell]} c_i$, and its *shift on edge* e_i is c_i. We make the following observation:

▶ **Lemma 40.** *Let $f\colon A_v \to A_v$ be a cyclic bijection. Then for all $s \in \mathbb{Z}_3$*

(a) f is an automorphism of $\mathbf{A}(v,s)$ if and only if $c(f) = 0$.
(b) f is an isomorphism $\mathbf{A}(v,s) \to \mathbf{A}(v,s+1)$ if and only if $c(f) = 1$.
(c) f is an isomorphism $\mathbf{A}(v,s+1) \to \mathbf{A}(v,s)$ if and only if $c(f) = 2$.

In order to show that Duplicator can always keep the mistake in the bijection away from the currently pebbled elements, we need to prove a typical property of CFI-like constructions: Given two vertices $v, w \in V$, and a path P between them in G, we can always find a bijection between $\mathbf{A}(G)$ and $\tilde{\mathbf{A}}(G)$ that defines an isomorphism from $\mathbf{A}(v,s)$ to $\mathbf{A}(v,s+c)$, for any $c \in \mathbb{Z}_3$, and an isomorphism from $\mathbf{A}(w,t)$ to $\mathbf{A}(w,t+c)$ (or, actually, to $\mathbf{A}(w,t-c)$, if P has even length), and an automorphism of every other vertex gadget $\mathbf{A}(v',s')$:

▶ **Lemma 41.** *Let $G = (V,E)$ be a connected regular graph, and let $v, w \in V$ be distinct vertices, and P a path from v to w in G. Let $c \in \mathbb{Z}_3$. There exists a bijection $g_{P,c}\colon A_G \to A_G$ such that:*

1. *$g_{P,c}$ is the identity everywhere except on P, that is, for each $v' \in V$ not on the path P, $g_{P,c}(v',a) = (v'a)$ and $g_{P,c}(e,a) = (e,a)$ for each (v',a) and (e,a) in $A_{v'}$.*
2. *Let $s \in \mathbb{Z}_3$ be the value such that $\mathbf{A}(v,s)$ is the vertex gadget associated with v in $\mathbf{A}(G)$. Then $g_{P,c}|_{A_v}$ is an isomorphism from $\mathbf{A}(v,s)$ to $\mathbf{A}(v,s+c)$.*
3. *Let $t \in \mathbb{Z}_3$ be the value such that $\mathbf{A}(w,t)$ is the vertex gadget associated with w in $\tilde{\mathbf{A}}(G)$. Then $g_{P,c}|_{A_w}$ is an isomorphism from $\mathbf{A}(w,t)$ to $\mathbf{A}(w,t+c)$ if the number of edges in P is odd; else, it is an isomorphism from $\mathbf{A}(w,t)$ to $\mathbf{A}(w,t-c)$.*
4. *For every interior vertex v' on the path P, $g_{P,c}|_{A_{v'}}$ is an automorphism of $\mathbf{A}(v',x)$, where x is the value associated with $\mathbf{A}(v',x)$ in $\mathbf{A}(G)$.*

Proof. We define $g_{P,c}$ individually on the $A_{v'}$, for each $v' \in V$. For each v' that is not on the path P, we define $g_{P,c}|_{A_{v'}}\colon A_{v'} \to A_{v'}$ as the identity map. Thus, it satisfies the first claimed property.

Now consider the first endpoint of P, which is v. Let $e \in E(v)$ be the first edge in P. We define $g_{P,c}|_{A_v}\colon A_v \to A_v$ as the cyclic bijection with shift 0 on each edge except e, where it has shift c. By Lemma 40, this satisfies the second claimed property.

Let v' be the other endpoint of e. If $v' \neq w$, then we define $g_{P,c}|_{A_{v'}} : A_{v'} \to A_{v'}$ as a cyclic bijection with shift value 0. Let $e' \in E(v')$ be the second edge on P. To ensure that $g_{P,c}|_{A_{v'}}$ has shift value 0, we define it such that its shift on e' is $-c$, and it is the identity on all edges in $E(v') \setminus \{e, e'\}$.

In this way, we continue the definition of $g_{P,c}$ along the path. It is easy to see that by Lemma 40, all claimed properties hold. ◀

With this, we can prove the key lemma from which it follows immediately that Duplicator has a winning strategy in the game $\mathrm{BP}^k_{\ell-1}(\mathbf{A}(G), \tilde{\mathbf{A}}(G))$, where $G \coloneqq G_{\ell,n}$ is the graph that we constructed, consisting of identical copies of graphs $H_1, \dots, H_{n+1}$. For the definition of the invariant that Duplicator maintains, we refer back to the text before Lemma 35 in the main part of the paper.

▶ **Lemma 35.** *Let $(\bar{\alpha}, \bar{\beta})$ be the current pebble position in $\mathrm{BP}^k_{\ell-1}(\mathbf{A}(G), \tilde{\mathbf{A}}(G))$. Assume the invariant holds in the beginning of the round. Then Duplicator has a strategy to ensure the invariant also in the beginning of the following round.*

Proof. Duplicator begins the round by playing the bijection $f : \mathbf{A}(G) \to \tilde{\mathbf{A}}(G)$ sending $\bar{\alpha}$ to $\bar{\beta}$, which exists because the invariant holds at the beginning of the round. Let u be the vertex such that f is good bar u. Now Spoiler modifies the position of up to $\ell - 1$ pebbles in both structures, leading to a new pebble position $(\bar{\alpha}', \bar{\beta}')$. Duplicator has to find a new bijection f' that satisfies the invariant again. This can be done using a pebble-free path P from u to another vertex u' which is safe with respect to $(\bar{\alpha}', \bar{\beta}')$. Then essentially f' can be obtained from f by composing it with a bijection as given by Lemma 41 for P. The details are as follows: We make a case distinction according to whether a pebble of $\bar{\alpha}'$ lies on $C_u = \{u\} \times \mathbb{Z}_3$ or not.

Case 1:
Suppose that C_u is pebble-free. Then the argument proceeds exactly like in [15]: There are two subcases. The first is that in $\bar{\alpha}'$, there are still less than $\ell - 1$ pebbles lying on vertices in

$$W \coloneqq \Big(\bigcup_{v \in V(H_i)} A_v \Big) \setminus (C_u \cup \{(e^{ij}, 0), (e^{ij}, 1), (e^{ij}, 2)\}),$$

where $e^{ij} \in E$ denotes the edge connecting $u \in V(H_i)$ with its neighbour in $V(H_j)$. Because H_i is $(\ell - 1)$-connected by definition, there exists a pebble-free path[2] from u to any other vertex of H_i. Moreover, because n is big enough, there is some $i' \in [n+1]$ such that $H_{i'}$ is pebble-free and there exists a *safe* vertex $u' \in H_{i'}$. Hence, we find a pebble-free path P from u to this $u' \in V(H_{i'})$.

The second subcase is that all the $\ell - 1$ pebbles that Spoiler has moved in this round are now on elements of W. But then, H_j and the connecting edge e^{ij} are still pebble-free. Thus, we can again find a pebble-free path P to some safe $u' \in H_{i'}$, namely via e^{ij} through H_j to $H_{i'}$.

In each of the two subcases, we have obtained our desired path P from u to a safe u'. Let $s, t \in \mathbb{Z}_3$ be the values of the gadget of u in $\mathbf{A}(G)$ and $\tilde{\mathbf{A}}(G)$, that is: $\mathbf{A}(u, s)$ is the gadget associated with u in $\mathbf{A}(G)$, and $\mathbf{A}(u, s)$ the gadget associated with u in $\tilde{\mathbf{A}}(G)$. Let $c \in \mathbb{Z}_3$ be such that $f|_{A_u}$ is an isomorphism from $\mathbf{A}(u, s + c)$ to $\mathbf{A}(u, t)$ (in particular, $c \neq 0$ because f

[2] a path P is pebble-free if neither on $\{v\} \times \mathbb{Z}_3$, for a vertex v in P, nor on $\{e\} \times \mathbb{Z}_3$, for an edge e in P, there is a pebble

is not good at u). Let $g_{P,c}\colon \mathbf{A}(G) \to \tilde{\mathbf{A}}(G)$ be a bijection as given by Lemma 41. Then set $f' \coloneqq g_{P,c} \circ f$. This finishes *Case 1.*

Case 2:
In this case, a pebble of $\bar{\alpha}'$ lies on C_u. We can find a safe vertex u' and a path P from u to u' with exactly the same reasoning as in Case 1. If such a path P exists that begins with an edge $e_d \in \bar{e}(u)$ such that $\bar{u}_d = 0$, then we can proceed exactly as before, and define $f' \coloneqq g_{P,c} \circ f$. This is because $g_{P,c}$ is the identity on C_u in this case (see again the definition of *cyclic* bijections for how they act on C_u).

It remains to deal with the case that for every potential path from u to a safe vertex, $\bar{u}_d \neq 0$ for the edge $e_d \in \bar{e}(u)$ that the path begins with. Note that $\bar{u}_d \neq 0$ only for $d \in \{2, 3\}$. In this case we argue as follows. One pebble has been moved to C_u in this round, and another pebble must have been moved to $\{e^{ij}\} \times \mathbb{Z}_3$ or to its other endpoint in H_j: By the choice of the order $\leq$ on G, the edge e^{ij} is in a position $d \notin \{2, 3\}$ in the tuple $\bar{e}(u)$, so $\bar{u}_d = 0$. So if e^{ij} were pebble-free, then contrary to the assumption of this subcase, we can find a path to a safe vertex (with the argument as in the second subcase of Case 1) that "avoids" the non-zero entries of the vector $\bar{u}$. So, at least two pebbles must have been spent on C_u and e^{ij}, which leaves at most $\ell - 3$ pebbles that can lie on vertices in W now. The graph H is $(\ell - 1)$-connected, so there must be two distinct edges $e, e' \in E(u) \setminus \{e^{ij}\}$ via which every other vertex in H_i can be reached from u with a pebble-free path. These two edges must be e_2, e_3 in the tuple $\bar{e}(u)$ because else, we would again have a path that avoids the non-zero entries of $\bar{u}$.

So we can choose two (not necessarily disjoint) pebble-free paths P_1, P_2 from u to a safe vertex u' (using the reasoning from Case 1, subcase 1), one beginning with e_2 and one beginning with e_3.

Now as before, let $c \in \mathbb{Z}_3$ be the constant such that $f|_{A_u}$ is an isomorphism from $\mathbf{A}(u, s+c)$ to $\mathbf{A}(u, t)$. Let $b_1, b_2 \in \mathbb{Z}_3$ be such that $b_1 + b_2 = c$ and $b_1 + 2b_2 = 0$. It is easy to verify that such b_1, b_2 exist for every c. Let g_{P_1,b_1}, g_{P_2,b_2} be the respective bijections obtained with Lemma 41. Set $f' \coloneqq g_{P_1,b_1} \circ g_{P_2,b_2} \circ f$. It remains to check that $g_{P_1,b_1} \circ g_{P_2,b_2}$ is the identity map on C_u, which is required because at least one vertex in C_u is now pebbled. This follows from the fact that $b_1 + 2b_2 = 0$, as can be seen from the definition of cyclic bijections. ◀

C Limitations of partial-Maltsev-closed quantifiers

Write $\mathcal{Q}_r^{\mathcal{M}}$ for the class of all r-ary quantifiers that are closed under the partial Maltsev operation.

We now present and analyze a construction of structures that are distinguishable with r-ary generalized quantifiers but not with r-ary Maltsev-closed ones. To be precise, our lower bound holds for the k-variable fragment of the logic, for $k = r$. For a logic $\mathcal{L}$, let $\mathcal{L}^k$ denote its k-variable fragment.

▶ **Theorem 37.** *For every $k \geq 3$, $\mathcal{L}^k(\mathcal{Q}_k^{\mathcal{M}}) \leq \mathcal{L}^k(\mathcal{Q}_k)$.*

The structures we use in the proof are indistinguishable in $\mathcal{L}^k(\mathcal{Q}_k^{\mathcal{M}})$ but can possibly be distinguished in $\mathcal{L}(\mathcal{Q}_k^{\mathcal{M}})$, i.e. with more variables available.

Constructing the instances

Throughout this section, all arithmetic operations are in the ring $\mathbb{Z}_4$, unless stated otherwise. The structures we construct are over a vocabulary with two relations R_0 and R_1 of arity k,

and one binary relation $\prec$.

As before, we start with a k-regular connected base graph $G = (V, E)$. Again, we first define vertex gadgets $\mathbf{A}^{\mathcal{M}}(v, s)$, for $v \in V$, $s \in \{0, 1\}$ and then put them together to obtain the whole structure.

▶ **Definition 42** (Vertex gadgets, Maltsev counterxample). *Fix $k \geq 3$ and let $G = (V, E)$ be a connected k-regular graph with an order $\leq^G$ on its vertices. Let $\bar{e}(v) = (e_1, \ldots, e_k)$ be the tuple of edges incident to v, in the order induced by $\leq^G$. Then for each $s \in \{0, 1\}$, we let $\mathbf{A}^{\mathcal{M}}(v, s) \coloneqq (A_v^{\mathcal{M}}, R_0^{\mathbf{A}^{\mathcal{M}}(v,s)}, R_1^{\mathbf{A}^{\mathcal{M}}(v,s)})$, where*

- $A_v^{\mathcal{M}} \coloneqq E(v) \times \mathbb{Z}_4$,
- $R_0^{\mathbf{A}^{\mathcal{M}}(v,s)} \coloneqq \{((e_1, a_1), \ldots, (e_k, a_k)) \mid (\sum_{i\in[k]} a_i) - 2s \in \{0, 1\} \pmod 4\}$
- $R_1^{\mathbf{A}^{\mathcal{M}}(v,s)} \coloneqq \{((e_1, a_1), \ldots, (e_k, a_k)) \mid (\sum_{i\in[k]} a_i) - 2s \in \{2, 3\} \pmod 4\}$

We combine the vertex gadgets in the same way as before to obtain our indistinguishable instances:

▶ **Definition 43.** *Let $G = (V, E, \leq^G)$ be an ℓ-regular connected graph with a linear order on its vertices.*

(a) For each $U \subseteq V$, we define
$\mathbf{A}^{\mathcal{M}}(G, U) = (A_G^{\mathcal{M}}, R_0^{\mathbf{A}^{\mathcal{M}}(G,U)}, R_1^{\mathbf{A}^{\mathcal{M}}(G,U)}, R_2^{\mathbf{A}^{\mathcal{M}}(G,U)}, \prec^{\mathbf{A}^{\mathcal{M}}(G,U)})$, *where*

- $A_G^{\mathcal{M}} \coloneqq \bigcup_{v\in V} A_v^{\mathcal{M}}$,
- $R_j^{\mathbf{A}^{\mathcal{M}}(G,U)} \coloneqq \bigcup_{v\in V\setminus U} R_j^{\mathbf{A}^{\mathcal{M}}(v,0)} \cup \bigcup_{v\in U} R_j^{\mathbf{A}^{\mathcal{M}}(v,1)}$ *for each $j \in \{0, 1\}$.*
- $\prec^{\mathbf{A}^{\mathcal{M}}(G,U)}$ *is a preorder with $(e, a) \prec (e', a')$ if and only if $e \leq e'$ in the lexicographic order on E induced by $\leq^G$.*

(b) Furthermore, we define $\mathbf{A}^{\mathcal{M}}(G) \coloneqq \mathbf{A}^{\mathcal{M}}(G, \emptyset)$ and $\tilde{\mathbf{A}}^{\mathcal{M}}(G) \coloneqq \mathbf{A}^{\mathcal{M}}(G, \{\tilde{v}\})$, where $\tilde{v} \in V$ is the smallest vertex of G with respect to the order $\leq^G$.

We now aim to show two things. The first is that $\mathbf{A}^{\mathcal{M}}(G)$ and $\mathbf{A}^{\mathcal{M}}(G, \{\tilde{v}\})$ are not isomorphic. Since the vocabulary has relations of arity at most k, this immediately implies that there is a k-ary quantifier Q that distinguishes them. The second is that $\mathbf{A}^{\mathcal{M}}(G)$ and $\mathbf{A}^{\mathcal{M}}(G, \{\tilde{v}\})$ cannot be distinguished in a k-pebble bijective game for Maltsev-closed quantifiers. This is the game defined in [10, Definition 16]. We reproduce the definition for completeness. The following defines a game $\mathcal{M}_k(\mathbf{A}, \mathbf{B}, \alpha, \beta)$ played on structures $\mathbf{A}$ and $\mathbf{B}$ with X a set of k variables and partial assignments α and β of values in $\mathbf{A}$ and $\mathbf{B}$ respectively to the variables in X. Below, for a relation P, $m(P)$ refers to the closure of P under the partial Maltsev operation.

▶ **Definition 44.** *The rules of the game $\mathcal{M}_k(\mathbf{A}, \mathbf{B}, \alpha, \beta)$ are the following:*

1. *If $\alpha \mapsto \beta$ is not a partial isomorphism, 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 $\mathcal{M}$-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 $P = \{\vec{a}_1, \vec{a}_2, \vec{a}_3\} \subseteq A^r$ such that $f(\vec{b}) \in m(P)$. Spoiler completes the round by choosing $\vec{a} \in P$. The players continue by playing $\mathcal{M}_k(\alpha', \beta')$, where $\alpha' \coloneqq \alpha[\vec{a}/\vec{y}]$ and $\beta' \coloneqq \beta[\vec{b}/\vec{y}]$.*

- **Right $\mathcal{M}$-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 $P = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\} \subseteq B^r$ such that $f(\vec{a}) \in m(P)$. Spoiler completes the round by choosing $\vec{b} \in P$. The players continue by playing $\mathcal{M}_k(\alpha', \beta')$, 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.*

By [10, Theorem 17], Duplicator has a winning strategy in $\mathcal{M}_k(\mathbf{A}, \mathbf{B}, \alpha, \beta)$ if, and only if, there is no formula of $\mathcal{L}^k(Q_k^{\mathcal{M}})$ that distinguishes $(\mathbf{A}, \alpha)$ from $(\mathbf{B}, \beta)$.

But first, the Duplicator winning strategy that we are going to define in the game involves playing bijections that do not necessarily permute edge vertices $\{e\} \times \mathbb{Z}_4$ cylically, as before. Now, *reflections* also play a role. We develop the necessary background.

Properties of bijections over $\mathbb{Z}_4$

▶ **Lemma 45.** *Let $\pi\colon \mathbb{Z}_4 \to \mathbb{Z}_4$ be any permutation of $\mathbb{Z}_4$. Then one of the following two cases applies:*

1. *There exists $x \in \mathbb{Z}_4$ such that $\pi(x+1) = \pi(x) + 1$.*
2. *There exists $a \in \mathbb{Z}_4$ such that for every $x \in \mathbb{Z}_4$, $\pi(x) = a - x$.*

Proof. We show that if π does not satisfy (1), then it must satisfy (2). Let $a := \pi(0)$. Because (1) does not apply, we have $\pi(1) \in \{a+2, a+3\}$. Suppose $\pi(1) = a+2$. Then $\pi(2) = a+1$ and $\pi(3) = a+3$. But then $\pi(0) = \pi(3+1) = (a+3)+1 = \pi(3)+1$. So condition (1) is satisfied for $x = 3$. Therefore, we must have $\pi(1) = a+3$. Then $\pi(2) \in \{a+1, a+2\}$, but we cannot have $\pi(2) = a+1$ because then, $\pi(3) = a+2$, so we would be in case (1) again, for $x = 2$. Therefore, $\pi(2) = a+2$ and $\pi(3) = a+1$. That means, π is exactly as described in case (2). ◀

We call permutations of the second kind *reflections*. A permutation $\pi\colon \mathbb{Z}_4 \to \mathbb{Z}_4$ for which there is a constant $a \in \mathbb{Z}_4$ such that for each $x \in \mathbb{Z}_4$, $\pi(x) = x + a$, is called a *rotation* (or *cyclic*). For an edge $e \in E$, a bijection $f\colon \{e\} \times \mathbb{Z}_4 \to \{e\} \times \mathbb{Z}_4$ is said to be a rotation or a reflection if its induced permutation on $\mathbb{Z}_4$ has this property.

In what follows, for $v \in V, j \in \{0,1\}$, we use $R_j(v)$ as shorthand to denote the relation $R_j^{\mathbf{A}^{\mathcal{M}}(v,0)}$.

▶ **Lemma 46.** *Let $v \in V(G)$ and let f be a bijection defined from the set $\bigcup_{e \in E(v)} \{e\} \times \mathbb{Z}_4$ to itself, such that $f(\{e\} \times \mathbb{Z}_4) = \{e\} \times \mathbb{Z}_4$ for each $e \in E(v)$. Suppose f is either a rotation on every $e \in E(v)$ or a reflection on every $e \in E(v)$. For each $e \in E(v)$, let $c_e \in \mathbb{Z}_4$ denote the* shift value *of f on e, i.e such that $f(e,x) = (e, x + c_e)$ or $f(e,x) = (e, -x + c_e)$, respectively, for each $x \in \mathbb{Z}_4$. Let $c := \sum_{e \in E(v)} c_e$.*

1. *If f is a rotation on every $e \in E(v)$, then $f(R_j(v)) = R_j(v)$ for each $j \in \{0,1\}$ if and only if $c = 0$. Moreover, $f(R_j(v)) = R_{1-j}(v)$ for each $j \in \{0,1\}$ if and only if $c = 2$.*
2. *If f is a reflection on every $e \in E(v)$, then $f(R_j(v)) = R_j(v)$ for each $j \in \{0,1\}$ if and only if $c = 1$. Moreover, $f(R_j(v)) = R_{1-j}(v)$ for each $j \in \{0,1\}$ if and only if $c = 3$.*

Proof. Suppose first that f is a rotation on every $e \in E(v)$. If $c \in \{1,3\}$, then we see that f neither preserves, nor exchanges $R_0(v)$ and $R_1(v)$: If, for example, $c = 1$, then for $(e_1, \ldots, e_k) = E(v)$, f maps the tuples $((e_1, a_1), \ldots, (e_k, a_k))$ with $\sum_{i \in [k]} a_i = 0$ to tuples whose sum is 1 (keeping them in $R_0(v)$), but it also maps the tuples whose sum is 1 to tuples

with sum 2 (moving them from $R_0(v)$ to $R_1(v)$). The situation is similar if $c = 3$. Hence, if f maps $R_j(v)$ to $R_j(v)$ or to $R_{1-j}(v)$, then we must have $c \in \{0, 2\}$. Clearly, if $c = 0$, then f preserves $R_j(v)$ for each $j \in \{0, 1\}$. It is also not hard to see that adding 2 to the sum of any tuple moves it from $R_0(v)$ to $R_1(v)$ or vice versa. This shows the first part of the lemma. The second part is easily seen with a similar analysis. ◀

▶ **Lemma 47.** *Let $v \in V(G)$ and let f be as in Lemma 46. If there is an $e \in E(v)$ on which f is not a rotation, and an $e' \in E(v)$ on which f is not a reflection, then f does not map $R_j(v)$ to itself, for both $j \in \{0, 1\}$.*

Proof. Assume for a contradiction that f maps each $R_j(v)$ to itself. Since f is not a rotation on e, there exist $a, b \in \mathbb{Z}_4$, $a \neq b$, and $s, t \in \mathbb{Z}_4$, $s < t$, such that $f(e, s) = (e, s + a)$ and $f(e, t) = (e, t + b)$. W.l.o.g. suppose that $0 \notin \{s, t\}$ (otherwise we can modify the first condition below so that $a_1 \notin \{s, t\}$, and reason analogously). Pick a tuple $\bar{\alpha} := ((e, a_1), (e', a_2), ..., (e_k, a_k)) \in R_0(v)$ with the following properties:

1. $a_1 = 0$.
2. $\sum \bar{\alpha} = \sum_{i=1}^{k} a_i = 0$.
3. a_2 is such that $f(e', a_2 + 1) = (e', f(a_2) + 1)$.

Such a tuple can always be chosen because $k \geq 3$, and the third property can be achieved because f induces a permutation on e' that is not a reflection, so it must satisfy case (1) of Lemma 45.

Write $\bar{\alpha}^s$ and $\bar{\alpha}^t$ for the tuples obtained from $\bar{\alpha}$ by exchanging $(e, 0)$ in the first entry of $\bar{\alpha}$ with (e, s) or (e, t), respectively. Then $\sum \bar{\alpha}^s = s$, and $\sum \bar{\alpha}^t = t$. Define $\bar{\alpha}_{+1}$ as the tuple obtained from $\bar{\alpha}$ by exchanging the second entry (e', a_2) with $(e', a_2 + 1)$. Thus, $\sum \bar{\alpha}_{+1} = 1$. Also, let $\bar{\alpha}^s_{+1}$ and $\bar{\alpha}^t_{+1}$ be the tuples obtained from $\bar{\alpha}_{+1}$ by exchanging the first entry with (e, s) or (e, t), respectively. So $\sum \bar{\alpha}^s_{+1} = s + 1$, $\sum \bar{\alpha}^t_{+1} = t + 1$.

Consider now the tuples $f(\bar{\alpha})$ and $f(\bar{\alpha}_{+1})$. Because $\bar{\alpha}, \bar{\alpha}_{+1} \in R_0(v)$ and we are assuming that f preserves R_0, we have $f(\bar{\alpha}), f(\bar{\alpha}_{+1}) \in R_0(v)$. Moreover, by the third property of $\bar{\alpha}$, we know that $\sum f(\bar{\alpha}_{+1}) = \sum f(\bar{\alpha}) + 1$. Therefore, we must have $\sum f(\bar{\alpha}) = 0$ and $\sum f(\bar{\alpha}_{+1}) = 1$. The entry sums of $\bar{\alpha}, \bar{\alpha}^s, \bar{\alpha}^t$ are pairwise distinct (as we are assuming $0 \notin \{s, t\}$). That means that two of them are in the same relation $R_j(v)$ and one of them is in $R_{1-j}(v)$. Suppose first that $j = 1$. Then, because $s < t$, we must have $s = 2$ and $t = 3$. Then $\sum \bar{\alpha}^s_{+1} = 1+s = 3, \sum \bar{\alpha}^t_{+1} = 1+t = 0$. Now $\delta := \sum f(\bar{\alpha}^t) - \sum f(\bar{\alpha}^s) = t-s+b-a = 1+b-a$, and we have $b - a \neq 0$. So $\delta \neq 1$. Because f is assumed to preserve the relations, and $\bar{\alpha}^s, \bar{\alpha}^t \in R_1(v)$, also $f(\bar{\alpha}^s), f(\bar{\alpha}^t) \in R_1(v)$. Since $\delta \neq 1$, we can only have $\delta = 0$ or $\delta = 3$. The case $\delta = 0$ is ruled out as follows: Consider $\bar{\alpha}^s_{+1}$ and $\bar{\alpha}^t_{+1}$. Also for these tuples, we have $\sum f(\bar{\alpha}^t_{+1}) - \sum f(\bar{\alpha}^s_{+1}) = \delta$. So if $\delta = 0$, then $f(\bar{\alpha}^s_{+1})$ and $f(\bar{\alpha}^t_{+1})$ are in the same relation $R_j(v)$. But this cannot be the case because as already established, $\sum \bar{\alpha}^s_{+1} = 3$ and $\sum \bar{\alpha}^t_{+1} = 0$, so the preimages under f are not in the same relation $R_j(v)$. Therefore, we must have $\delta = 3$. This means that $\sum f(\bar{\alpha}^s) = 3$ and $\sum f(\bar{\alpha}^t) = 2$ as we have $\sum \bar{\alpha}^s = 2$ and $\sum \bar{\alpha}^t = 3$, and f preserves $R_1(v)$. From this, we can conclude $\sum f(\bar{\alpha}^s_{+1}) = 0$ and $\sum f(\bar{\alpha}^t_{+1}) = 3$ using again the third property of $\bar{\alpha}$. But as already mentioned, we have $\sum \bar{\alpha}^s_{+1} = 3$, so f maps $\bar{\alpha}^s_{+1}$ incorrectly. This finishes the case that $R_1(v)$ is the relation which contains two of the tuples $\bar{\alpha}, \bar{\alpha}^s, \bar{\alpha}^t$. If $R_0(v)$ contains two of these tuples, then $\sum \bar{\alpha}^s = 1$. In that case, a similar argument leads to a contradiction. ◀

Properties of the constructed instances over $\mathbb{Z}_4$

First, we need to prove that, as desired, $\mathbf{A}^{\mathcal{M}}(G) \not\cong \tilde{\mathbf{A}}^{\mathcal{M}}(G)$. The following fact is standard (see [11, Corollary 2.13]

▶ **Lemma 48.** *Let $k \geq 3$. Every bipartite k-regular graph contains a perfect matching.*

▶ **Lemma 49.** *Let $k \geq 3$ and let $G = (V, E)$ be a bipartite connected ℓ-regular graph. Let $\mathcal{E}$ be a system of linear equations over $\mathbb{Z}_4$ with variable set $X = \{x_e \mid e \in E(G)\}$ and one equation of the form $\sum_{e \in E(v)} x_e = c_v$ for every vertex $v \in V$, where $c_v \in \mathbb{Z}_4$ is a constant. If there is a constant $a \in \mathbb{Z}_4$ such that $c_v = a$ for all but one vertex $\tilde{v} \in V$, where $c_{\tilde{v}} \neq a$, then $\mathcal{E}$ has no $\mathbb{Z}_4$-solution.*

Proof. Let $S \uplus T = V$ be the bipartition. Because G is k-regular, $|T| = |E|/k = |S|$. Suppose for a contradiction that $\lambda \colon X \to \mathbb{Z}_4$ is a solution for $\mathcal{E}$. We assume w.l.o.g. that $\tilde{v} \in S$. Then we have $\sum_{v \in S} \sum_{e \in E(v)} \lambda(x_e) = (|S| - 1) \cdot a + c_{\tilde{v}}$ and $\sum_{v \in T} \sum_{e \in E(v)} \lambda(x_e) = |T| \cdot a$. Since $|S| = |T|$ and $c_{\tilde{v}} \neq a$, these values are not equal. However, this is a contradiction since clearly $\sum_{v \in S} \sum_{e \in E(v)} \lambda(x_e) = \sum_{e \in E(G_n)} \lambda(x_e) = \sum_{v \in T} \sum_{e \in E(v)} \lambda(x_e)$. ◀

▶ **Lemma 50.** *Let $k \geq 3$ and let G be a G be a bipartite connected k-regular graph. Let $\mathcal{E}$ be a system of linear equations with variables X over $\mathbb{Z}_4$ as in Lemma 49. Assume again that there exists an $a \in \mathbb{Z}_4$ and a $\tilde{v} \in V(G)$ so that $c_{\tilde{v}} \neq a$ and $c_v = a$ for every $v \in V(G) \setminus \{\tilde{v}\}$. Fix an arbitrary $v' \in V(G)$. There is a $\lambda \colon X \to \mathbb{Z}_4$ that satisfies every equation in $\mathcal{E}$ except potentially the equation for v', and $\sum_{e \in E(v')} \lambda(x_e) - c_{v'} \in \{a - c_{\tilde{v}}, c_{\tilde{v}} - a\}$.*

Proof. Suppose first that $a = 0$. Then consider the assignment $\lambda_0 \colon X \to \mathbb{Z}_4$ that is 0 at every edge. It satisfies every equation except the one for $\tilde{v}$, where $\sum_{e \in E(v')} \lambda(x_e) - c_{\tilde{v}} = 0 - c_{\tilde{v}} = a - c_{\tilde{v}}$. This shows the lemma for $v' = \tilde{v}$. If $v' \neq \tilde{v}$, then by adding $c_{\tilde{v}}$ to an incident edge of $\tilde{v}$, and propagating this alternatingly with $-c_{\tilde{v}}$ along a path from $\tilde{v}$ to v', we can define a new assignment λ' that satisfies every equation except the one for v'. Moreover, $\sum_{e \in E(v')} \lambda(x_e) - c_{v'} \in \{-c_{\tilde{v}}, c_{\tilde{v}}\}$ (depending whether the length of the path is odd or even). In case that $a \neq 0$, we can take an assignment as in the case $a = 0$, and add a to $\lambda(x_e)$, for every edge e of a fixed perfect matching of G – a perfect matching exists by Lemma 48. ◀

▶ **Lemma 51.** *Let $G = (V, E)$ be a connected k-regular bipartite graph. Then for any $k \geq 3$, $\mathbf{A}^{\mathcal{M}}(G) \not\cong \tilde{\mathbf{A}}^{\mathcal{M}}(G)$.*

Proof. Suppose for a contradiction that $f \colon A_G^{\mathcal{M}} \to A_G^{\mathcal{M}}$ is a bijection that defines an isomorphism from $\mathbf{A}^{\mathcal{M}}(G)$ to $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$. Summation in the following is always in $\mathbb{Z}_4$, i.e. mod 4. Because of the preorder $\preceq^{\mathbf{A}^{\mathcal{M}}(G)}$, f must map the set $\{e\} \times \mathbb{Z}_4$ to itself, for each $e \in E$. In the following, we also write $f(i)$ instead of $f(e, i)$ if the edge e is clear from the context.

Claim: It is neither the case that f is a rotation on every $e \in E$, nor that f is a reflection on every $e \in E$.
Proof of claim. Suppose first that f is a rotation on every edge. Let, for every $e \in E$, $a_e \in \mathbb{Z}_4$ be the cyclic shift applied to edge e, as described before. By Lemma 46, the sum of the shifts must be 0 at every $v \in V \setminus \{\tilde{v}\}$, and it must be 2 at $\tilde{v}$. For each $e \in E$, let $a_e \in \mathbb{Z}_4$ denote the value of the cyclic shift that f applies to $\{e\} \times \mathbb{Z}_4$. Now the admissible values of the a_e can be seen as a solution to a system of linear equations as in Lemma 49. But by that lemma, such a solution does not exist. This shows that f cannot be a rotation on every edge. If f is a reflection on every edge, then according to Lemma 46, the sum of the shifts must be 1 at every $v \in V \setminus \{\tilde{v}\}$, and it must be 3 at $\tilde{v}$. So again, the shifts a_e, for

each $e \in E(G)$, must be a solution to an equation system as in Lemma 49, which cannot exist.

From the claim it follows that there is at least one edge e on which f is not a rotation. In fact, this e can be chosen in such a way that one of its endpoints is a vertex $v \neq \tilde{v}$ such that not on all edges in $E \setminus \{e\}$, f is a reflection. To see this, start with an arbitrary e where f is not a rotation. If it does not have an endpoint that satisfies the desired condition, then move from e along a path consisting only of edges where f is a reflection, until an edge e is reached that does satisfy the condition. This is always possible because by the claim, there exist edges on which f is not a reflection. Hence, we can find the following set-up: An edge $e \in E$ on which f is not a rotation, and an endpoint $v \neq \tilde{v}$ of e such that there is $e' \in E(v) \setminus \{e\}$ on which f is not a reflection. Then by Lemma 47, f does not preserve $R_0(v)$ and $R_1(v)$. Since $v \neq \tilde{v}$, any isomorphism has to preserve these relations. Therefore, f cannot be an isomorphism between $\mathbf{A}^{\mathcal{M}}(G)$ and $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$. ◀

Distinguishability of the structures in the logic $\mathcal{L}^k(\mathcal{Q}_k)$ is now an immediate consequence:

▶ **Corollary 52.** *Let $G = (V, E)$ be connected, k-regular and bipartite. Then $\mathbf{A}^{\mathcal{M}}(G)$ and $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$ are distinguished by a sentence in $\mathcal{L}^k(Q_k)$.*

Proof. Let Q be a quantifier of arity k corresponding to the class of structures isomorphic to $\mathbf{A}^{\mathcal{M}}(G)$ for some graph G. Then, trivially, Q is definable in $\mathcal{L}^k(\mathcal{Q}_k)$. ◀

Indistinguishability of the instances in the logic with partial Maltsev quantifiers

To complete the proof of Theorem 37, for every $k \geq 3$, we define for each k, a graph G such that $\mathbf{A}^{\mathcal{M}}(G)$ and $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$ cannot be distinguished in $\mathcal{L}^k(\mathcal{Q}_k^{\mathcal{M}})$. Here it suffices to use one base graph for each $k \in \mathbb{N}$, rather than a family of graphs, because the arity *and* the number of variables are fixed to k.

We call a k-regular graph G *near-k-connected* if the removal of any k edges does not disconnect the graph, except in the case that the k removed edges are the incident edges of a vertex v (in which case v becomes isolated from the rest).

For any $k \geq 3$ let G^k be the graph obtained by removing from the biclique $K_{k+1,k+1}$ a set of edge M that form a perfect matching. Thus, G^k is a k-regular bipartite graph on two sets (say A and B) of $k+1$ vertices each.

▶ **Lemma 53.** *G^k is near-k-connected.*

Proof. Let S be a set of edges of G^k, not all incident on a single vertex, with $|S| \leq k$. We wish to argue that the graph G' obtained by removing S from G^k is connected. Since A has $k+1$ vertices, there is at least one $a \in A$ such that none of the edges of S is incident on A. We show that all vertices in G' are reachable from a. Let b be the unique non-neighbour of a in B and let $A' = A \setminus \{a\}$ and $B' = B \setminus \{b\}$. Then, every vertex in B' is a neighbour of a and hence reachable from a. Moreover, if $u \in A'$ is such that some edge between u and B' is not in S, then u is also reachable from a. Hence, if for all $u \in A'$ there is at least one edge between u and B' that is not in S then all of A' is reachable from a. Also, since all k edges incident on b are to A' and they cannot all be in S, b is also reachable from a in G'.

Thus, we are only left with the case that there is a vertex $u \in A'$ such that all $k-1$ edge between u and B' are in S. Then, since S cannot contain all edges incident on u, b is still a neighbour of u in G'. Moreover, there cannot be more than one vertex in A' satisfying these conditions. Indeed, if $u, u' \in A'$ are such that each has $k-1$ incident edges in S, then S has

at least $2k-2$ edges (since the edges incident on u and u' form disjoint sets) and $2k-2 > k$ (since $k \geq 3$). Now, since S includes $k-1$ edges incident on u and none of these is incident on b, S can contain at most one edge incident on b. Thus, b still has at least two neighbours in A', and the one that is not u is reachable from a. We conclude that b, and therefore also u, is reachable from a. ◀

▶ **Lemma 54.** *For every $k \geq 3$, the Duplicator has a winning strategy in the game $\mathcal{M}_k(\mathbf{A}^{\mathcal{M}}(G^k), \tilde{\mathbf{A}}^{\mathcal{M}}(G^k), \emptyset, \emptyset)$.*

The winning strategy again consists in Duplicator maintaining a certain invariant. In the following, fix $k \geq 3$, and let $G = (V, E)$ be the graph G^k.

A bijection $f\colon A_G^{\mathcal{M}} \to A_G^{\mathcal{M}}$ is called *edge-preserving* if f maps the set $\{e\} \times \mathbb{Z}_4$ to itself, for every $e \in E$. Duplicator only ever plays edge-preserving bijections. Again we have the notion of a bijection being good bar a vertex but it is slightly different than before: In this proof, an edge-preserving bijection f is called *good bar* w, for a $w \in V$ if all of the following hold:

1. f is either a rotation on every $\{e\} \times \mathbb{Z}_4$ or a reflection on every $\{e\} \times \mathbb{Z}_4$, for $e \in E$.
2. For every $v \in V \setminus \{w\}$, $f|_{A_v^{\mathcal{M}}}\colon \mathbf{A}^{\mathcal{M}}(v, s) \to \mathbf{A}^{\mathcal{M}}(v, t)$ is an isomorphism, where $s, t \in \mathbb{Z}_2$ are the charges of the gadgets in $\mathbf{A}^{\mathcal{M}}(G)$ and $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$, respectively.
3. At w, the sum of the shift values of f on the edges in $E(w)$, $c \coloneqq \sum_{e \in E(w)} c_e$, is precisely 2 off from being a local isomorphism: That is, if f is a rotation on every edge, then $c = 0$ if $w = \tilde{v}$, and $c = 2$ if $w \neq \tilde{v}$. If f is a reflection on every edge, then $c = 1$ if $w = \tilde{v}$, and $c = 3$ if $w \neq \tilde{v}$ (cf. Lemma 46).

After every round, let $\bar{\alpha}, \bar{\beta}$ denote the tuples of pebbles on $\mathbf{A}(G)$ and $\tilde{\mathbf{A}}(G)$, respectively. Let $F(\bar{\alpha}) \subseteq E$ denote the set of edges e such that a pebble in $\bar{\alpha}$ lies on an element of $\{e\} \times \mathbb{Z}_4$. Duplicator ensures the following *invariant for the Maltsev game.* There exists a $w \in V$ such that:

1. Every $e \in E(w)$ is not in $F(\bar{\alpha})$.
2. There is an edge-preserving bijection f that is good bar w and satisfies $f(\alpha) = \beta$.

In the beginning of the game, this invariant is fulfilled for $w = \tilde{v}$ with $f\colon A_G^{\mathcal{M}} \to A_G^{\mathcal{M}}$ being the identity map.

▶ **Lemma 55.** *If the invariant is satisfied in a game position (α, β) at the beginning of the round, then Duplicator has a strategy to ensure the invariant also at the beginning of the following round.*

Proof. Spoiler chooses $r \in [k]$ and announces which r-tuple $\bar{y}$ of pebbles is moved this round, and whether he is playing a left or a right CSP quantifier move. Duplicator plays the bijection f that is good bar w which exists by the invariant (it does not matter whether it maps from the left to the right structure or the other way round). Now Spoiler chooses an r-tuple in either $\mathbf{A}^{\mathcal{M}}(G)$ or $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$, depending on which move he has chosen. Suppose w.l.o.g. that Spoiler is playing in $\mathbf{A}^{\mathcal{M}}(G)$, and let $\bar{a}$ be his chosen tuple of vertices. Let $\bar{\alpha}' \coloneqq \alpha[\bar{a}/\bar{y}]$ be the new tuple of pebbled vertices in $\mathbf{A}^{\mathcal{M}}(G)$.

Case 1: There is an edge $e \in E(w) \setminus F(\bar{\alpha}')$.
In this case, Duplicator plays $P \coloneqq \{f(\bar{a})\}$ i.e. effectively replies according to the bijection without using the power of the partial Maltsev move. Since f defines a local isomorphism everywhere except at the tuples in $R_j(w)$, for $j \in \{0, 1\}$, and because all edges incident to

w were pebble-free before Spoiler's move, the pebbles now still define a local isomorphism (since no tuple in $R_j(w)$ is fully pebbled). Thus, Duplicator has not lost. It remains to define a new bijection f' that is good bar w' for some suitable $w' \in V$. Since f is good bar w, it is either a rotation or a reflection on every edge, and it does not map $R_j^{\mathbf{A}^{\mathcal{M}}(G)}(w)$ to $R_j^{\tilde{\mathbf{A}}^{\mathcal{M}}(G)}(w)$, for each $j \in \{0,1\}$. If f is a rotation on every edge, then in order to change it to a local isomorphism from $R_j^{\mathbf{A}^{\mathcal{M}}(G)}(w)$ to $R_j^{\tilde{\mathbf{A}}^{\mathcal{M}}(G)}(w)$ for each $j \in \{0,1\}$, we have to make sure that the sum of shifts on the edges in $E(w)$ is 0 if $w \neq \tilde{v}$, and 2 if $w = \tilde{v}$. If f is a reflection on every edge, then the sum of shifts has to be 1 if $w \neq \tilde{v}$, and 3 if $w = \tilde{v}$ (both according to Lemma 46). In either case, this is achieved by letting $f'(e,i) \coloneqq (e, f(e,a)+b)$ for every $a \in \mathbb{Z}_4$ for a suitable choice of b that corrects the sum of the shifts in $E(w)$ as desired. Let $v \neq w$ be the other endpoint of e. To compensate for the shift by b, we let $f'(e',a) \coloneqq (e', f(e',a)-b)$ for some other edge $e' \in E(v)$, and we proceed like this along a pebble-free escape path to some vertex w' that satisfies condition 1 of the invariant. Such a path exists because G is k-near-connected, and such a vertex w' exists because there are only k pebbles, but $k+1$ vertices on each side of the bipartition of G.

On all edges that are not on this escape path, we define f' like f (see also Lemmas 41 and the proof of Lemma 34). By definition, the new bijection f' is good bar w', and w' satisfies all conditions of the invariant for the new game position.

Case 2: Each $e \in E(w)$ is in $F(\bar{\alpha})$, that is, $E(w) = F(\bar{\alpha})$.
In this case, Spoiler can in principle expose the fact that f is not a local isomorphism at w. Duplicator can prevent this by playing a Maltsev move. By the invariant, the sum of shifts c that f applies to the edges in $E(w)$ is off by 2 from inducing a partial isomorphism. So by Lemma 46, f maps $R_j^{\mathbf{A}(G)}(w)$ to $R_{1-j}^{\tilde{\mathbf{A}}(G)}(w)$ for each $j \in \{0,1\}$. Thus, $\bar{a} \in R_j^{\mathbf{A}(G)}(w)$ and $f(\bar{a}) \in R_{1-j}^{\tilde{\mathbf{A}}(G)}(w)$ for some $j \in \{0,1\}$. Let $\bar{a} = ((e_1,a_1),...,(e_k,a_k))$. Assume first that $\theta \coloneqq \sum_{i\in[k]} f(a_i) \in \{1,3\}$. Then $\sum_{i\in[k]} a_i \in \{0,1\}$ in case that $\theta = 3$, and $\sum_{i\in[k]} a_i \in \{2,3\}$ in case that $\theta = 1$, because f does not preserve $R_j^{\mathbf{A}(G)}(w)$. That means, to turn f into a local isomorphism, the value of θ must be increased by 1 or 2.

Let $f(\bar{a})_{+1,-}$ be a tuple like $f(\bar{a})$, but with $(e_1, f(a_1))$ replaced with $(e_1, f(a_1)+1)$. Let $f(\bar{a})_{+1,+1}$ be like $f(\bar{a})$, but $(e_1, f(a_1))$ replaced with $(e_1, f(a_1)+1)$ and $(e_2, f(a_2))$ replaced with $(e_2, f(a_2)+1)$. Let $f(\bar{a})_{-,+1}$ be like $f(\bar{a})$ but $(e_2, f(a_2))$ replaced with $(e_2, f(a_2)+1)$.

Then Duplicator plays: $P = (f(\bar{a})_{+1,-}, f(\bar{a})_{+1,+1}, f(\bar{a})_{-,+1})$. One can see that applying the Maltsev operation to these tuples yields $\bar{a}$, so this choice of P is admissible. Now Spoiler picks a $\bar{b} = ((e_1,b_1),\dots,(e_k,b_k)) \in P$. In any case, $\sum_{i\in[k]} b_i \in \{1+\theta, 2+\theta\}$. This means that $\bar{b} \in R_j^{\tilde{\mathbf{A}}(G)}(w)$, so the pebbles $\bar{a} \mapsto \bar{b}$ define a local isomorphism, as desired. It remains to show that Duplicator can choose a bijection f' that satisfies the invariant in the next round. In case that $\bar{b} = f(\bar{a})_{+1,+1}$, we just define $f'(e_1,a) \coloneqq (e_1, f(e_1,a)+1)$ for all $a \in \mathbb{Z}_4$, and likewise on the edge e_2. Then we propagate these two shifts along two disjoint and pebble-free paths to an escape vertex w' that satisfies condition 1 of the invariant for next round. This is possible because all the k pebbles lie on $E(w)$ now, so all other edges are pebble-free, and we can find 2 disjoint paths because G is 2-connected.

Because G is bipartite, the escape paths have the same length modulo 2 and so, indeed, the bijection f' thus defined has shift value 2 at w' (cf. Lemma 41). At w, we have corrected the error of 2 in the shift value, which existed by definition of f being good bar w. We furthermore define f' like f on all other edges that are not on any of the two escape paths.

In case that $\bar{b} = f(\bar{a})_{+1,-}$ or $\bar{b} = f(\bar{a})_{-,+1}$, the situation is more difficult. Assume w.l.o.g. that $\bar{b} = f(\bar{a})_{+1,-}$, so the $+1$ has been added to the first entry of the tuple. We have to

define f' so that it satisfies $\bar{b} = f'(\bar{a})$ and so that it defines a local isomorphism at w. This is not possible by adding cyclic shifts to f, but it requires to invert the current bijection f. For each edge $e_i \in E(w)$, let $f_i \colon \mathbb{Z}_4 \to \mathbb{Z}_4$ denote the projection of f to the $\mathbb{Z}_4$-part of $\{e_i\} \times \mathbb{Z}_4$. We now define for each $e_i \in E(w)$ the bijection $f'_i \colon \mathbb{Z}_4 \to \mathbb{Z}_4$ that is the projection of the new f' to the respective edge gadgets. Recall that $\bar{a} = ((e_1, a_1), \ldots, (e_k, a_k))$. Let

$$f'_1(x) \coloneqq -f_1(x) + 2f_1(a_1) + 1.$$

For every $1 < i \leq k$, let

$$f'_i(x) \coloneqq -f_i(x) + 2f_i(a_i).$$

It is easy to see that, by definition, we have for every $i \geq 2$: $f'_i(a_i) = f(a_i)$, and $f'_1(a_1) = f(a_1) + 1$. So if we define $f' \colon A_G^{\mathcal{M}} \to A_G^{\mathcal{M}}$ so that its restriction to the copy of $\mathbb{Z}_4$ in $\{e_i\} \times \mathbb{Z}_4$ is f'_i, for each $i \in [k]$, then, as desired, $f'(\bar{a}) = f(\bar{a})_{+1,-} = \bar{b}$. Moreover, the f' thus defined yields a local isomorphism that maps $R_j^{\mathbf{A}^{\mathcal{M}}(G)}(w)$ to $R_j^{\tilde{\mathbf{A}}^{\mathcal{M}}(G)}$, for each $j \in \{0, 1\}$. To see this, let $c_e \in \mathbb{Z}_4$, for each $e \in E(w)$ denote the additive shift constant in the rotation or reflection induced by f on $\{e\} \times \mathbb{Z}_4$. Likewise, let $c'_e \in \mathbb{Z}_4$ be that shift for f'. Then:

$$\begin{aligned}
\sum_{e \in E(w)} c'_e &= -\sum_{e \in E(w)} c_e + 1 + 2\sum_{i \in [k]} f_i(a_i) \\
&= -\sum_{e \in E(w)} c_e + 1 + 2\theta \\
&= -\sum_{e \in E(w)} c_e + 3.
\end{aligned}$$

The last equality is due to our assumption above that $\theta \in \{1, 3\}$. Now by property 3 of what it means for f to be good bar w, the sum $\sum_{e \in E(w)} c_e$ is precisely 2 off from yielding a local isomorphism. We go through one case as an example: Suppose f is a rotation and $c = \sum_{e \in E(w)} c_e = 0$. Then any local isomorphism from $R_j^{\mathbf{A}}(G)(w)$ to $R_j^{\tilde{\mathbf{A}}}(G)(w)$ has shift value 2 (according to Lemma 46, this is the case where $w = \tilde{v}$, and so, any local isomorphism must exchange $R_j(w)$ and $R_{1-j}(w)$). Then, $-\sum_{e \in E(w)} c_e + 3 = 3$, which means by Lemma 46 that f' does $R_j(w)$ and $R_{1-j}(w)$, as desired. One can check that for all other values of $\sum_{e \in E(w)} c_e$, a similar effect takes place, and f' always induces a local isomorphism. It only remains to extend the definition of f' from $E(w)$ to all edges. Since there exist only k pebbles, all of which lie on $E(w)$ now, we are free to choose f' arbitrarily on every other edge. Indeed, we need to invert f everywhere, so that if f was a rotation on every edge, then f' is a reflection on every edge, and vice versa. We simply choose f' in such a way that its sum of shifts is correct at every vertex except at one arbitrarily chosen w' that is not a neighbour of w. At w', we ensure that the sum of shifts is off by 2 from being a local isomorphism, as required by the invariant. This is possible by Lemma 50: If f' is a rotation on every edge, then its shift values of the edges can be written as a solution to a system of equations where every right hand side is 0 except at w', where it is 2; if f' is a reflection on every edge, the system of equations has 1 in every right hand side except for w', where it is 3.

This finishes the subcase where $\theta \in \{1, 3\}$. The case where $\theta \in \{0, 2\}$ is symmetric. Then, Duplicator chooses $P = (f(\bar{a})_{+3,-}, f(\bar{a})_{+3,+3}, f(\bar{a})_{-,+3})$, and the rest of the above proof goes through similarly, where one has to exchange the role of 1 and 3 everywhere. ◀

Now the **proof of Theorem 37** is a combination of the preceding results: For every $k \in \mathbb{N}$, choose G according to Lemma 53. Then $\mathbf{A}^{\mathcal{M}}(G)$ and $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$ are not isomorphic

(Lemma 51) but indistinguishable by any sentence in $\mathcal{L}^k(\mathcal{Q}_k^{\mathcal{M}})$ (Lemma 54). Finally, by Corollary 52, there is a sentence in the logic $\mathcal{L}^k(\mathcal{Q}_k)$ that does distinguish $\mathbf{A}^{\mathcal{M}}(G)$ from $\tilde{\mathbf{A}}^{\mathcal{M}}(G)$.