# Deciding FO-rewritability of Ontology-Mediated Queries in Linear Temporal Logic

**Vladislav Ryzhikov** ✉
Department of Computer Science, Birkbeck, University of London, UK

**Yury Savateev** ✉
Department of Computer Science, Birkbeck, University of London, UK
HSE University, Moscow, Russia

**Michael Zakharyaschev** ✉
Department of Computer Science, Birkbeck, University of London, UK

## Abstract

Our concern is the problem of determining the data complexity of answering an ontology-mediated query (OMQ) given in linear temporal logic *LTL* over $(\mathbb{Z}, <)$ and deciding whether it is rewritable to an $\mathsf{FO}(<)$-query, possibly with extra predicates. First, we observe that, in line with the circuit complexity and FO-definability of regular languages, OMQ answering in $\mathrm{AC}^0$, $\mathrm{ACC}^0$ and $\mathrm{NC}^1$ coincides with $\mathsf{FO}(<, \equiv)$-rewritability using unary predicates $x \equiv 0 \,(\mathrm{mod}\ n)$, $\mathsf{FO}(<, \mathsf{MOD})$-rewritability, and $\mathsf{FO}(\mathsf{RPR})$-rewritability using relational primitive recursion, respectively. We then show that deciding $\mathsf{FO}(<)$-, $\mathsf{FO}(<, \equiv)$- and $\mathsf{FO}(<, \mathsf{MOD})$-rewritability of *LTL* OMQs is ExpSpace-complete, and that these problems become PSpace-complete for OMQs with a linear Horn ontology and an atomic query, and also a positive query in the cases of $\mathsf{FO}(<)$- and $\mathsf{FO}(<, \equiv)$-rewritability. Further, we consider $\mathsf{FO}(<)$-rewritability of OMQs with a binary-clause ontology and identify OMQ classes, for which deciding it is PSpace-, $\Pi_2^p$- and coNP-complete.

**2012 ACM Subject Classification** Theory of computation → Logic

**Keywords and phrases** FO-rewritability, linear temporal logic, ontology-mediated query, regular language, finite automaton

**Digital Object Identifier** 10.4230/LIPIcs.CVIT.2016.23

**Acknowledgements** This work was supported by the UK EPSRC grant EP/S032282.

## 1 Introduction

**Motivation.** The problem we consider in this paper originates in the area of ontology-based data access (OBDA) to temporal data. The aim of the OBDA paradigm [44, 61] and systems such as Mastro or Ontop[1] is to facilitate management and integration of possibly incomplete and heterogeneous data by providing the user with a view of the data through the lens of a description logic (DL) ontology. Thus, the user can think of the data as a 'virtual knowledge graph' [62], $\mathcal{A}$, whose labels—unary and binary predicates supplied by an ontology, $\mathcal{O}$—are the only thing to know when formulating queries, $\varkappa$. Ontology-mediated queries (OMQs) $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ are supposed to be answered over $\mathcal{A}$ under the open world semantics (taking account of all models of $\mathcal{O}$ and $\mathcal{A}$), which can be prohibitively complex. So the key to practical OBDA is ensuring first-order rewritability of $\boldsymbol{q}$ (aka boundedness in the datalog literature [1]), which reduces open-world reasoning to evaluating an FO-formula over $\mathcal{A}$. The W3C standard ontology language *OWL 2 QL* for OBDA is based on the *DL-Lite* family of DL [3, 19], which uniformly guarantees FO-rewritability of all OMQs with a conjunctive query.

---

[1] `https://www.obdasystems.com`, `https://ontopic.biz`


42nd Conference on Very Important Topics (CVIT 2016).
Editors: John Q. Open and Joan R. Access; Article No. 23; pp. 23:1–23:47
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Other ontology languages with this feature include various dialects of tgds; see, e.g., [8,18,22]. However, by design such languages are rather inexpressive.

Theory and practice of OBDA have revived the interest to the problem of deciding whether an OMQ given in some expressive language is FO-rewritable, which was thoroughly investigated in the 1980–90s for datalog queries; see, e.g., [2,24,42,53,55]. The data complexity and rewritability of OMQs in various DLs and disjunctive datalog have become an active research area in the past decade [15, 27, 31, 41], lying at the crossroads of logic, database theory, knowledge representation, circuit and descriptive complexity, and CSP.

There have been numerous attempts to extend ontology and query languages with constructors capable of representing events over temporal data; see [6, 40] for surveys and [16,59,60] for more recent developments. However, so far the focus has been on the uniform complexity of reasoning with arbitrary ontologies and queries in a given language rather than on understanding the data complexity and FO-rewritability of individual temporal OMQs. On the other hand, the non-uniform analysis of OMQs in DLs or datalog mentioned above is not applicable to standard temporal logics interpreted over linearly-ordered structures.

In this paper, we take a first step towards understanding the problem of FO-rewritability of OMQs over temporal data by focusing on the temporal dimension and considering OMQs given in linear temporal logic *LTL* interpreted over $(\mathbb{Z}, <)$.

► **Example 1.** Let $\mathcal{O}$ be an *LTL* ontology with the following axioms (describing a system's behaviour and) containing the temporal operators $\Box_F/\Box_P$ (always in the future/past), $\Diamond_F/\Diamond_P$ (sometime in the future/past) and $\bigcirc_F/\bigcirc_P$ (the next/previous minute):

$$\Box_P\Box_F\big(\mathit{Malfunction} \to \Diamond_F \mathit{Fixed}\big), \tag{1}$$

$$\Box_P\Box_F\big(\mathit{Fixed} \to \bigcirc_F \mathit{InOperation}\big), \tag{2}$$

$$\Box_P\Box_F\big(\mathit{Malfunction} \land \bigcirc_P \mathit{Malfunction} \land \bigcirc_P^2 \mathit{Malfunction} \to \neg\bigcirc_F \mathit{InOperation}\big). \tag{3}$$

We query temporal data, say

$$\mathcal{A} = \{\mathit{Malfunction}(2), \mathit{Malfunction}(5), \mathit{Malfunction}(6), \mathit{Fixed}(6), \mathit{Malfunction}(7)\}$$

by means of *LTL*-formulas such as

$$\varkappa = \Diamond_P\Diamond_F\big(\mathit{Malfunction} \land \bigvee_{1\le i\le 5} \bigcirc_F^i(\mathit{Fixed} \land \bigvee_{1\le j\le 5} \neg\bigcirc_F^j \mathit{InOperation})\big)$$

asking whether there was a malfunction that was fixed in $\le 5$m but within the next 5m the equipment went out of operation again. The certain answer to the OMQ $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ over $\mathcal{A}$ is yes because $\varkappa$ is true in all models of $\mathcal{O}$ and $\mathcal{A}$. It is readily seen that the certain answer to $\boldsymbol{q}$ over any given data instance $\mathcal{A}'$ in the signature $\{\mathit{Malfunction}, \mathit{Fixed}\}$ can be computed by evaluating over $\mathcal{A}'$ the following $\mathsf{FO}(<)$-sentence, called an $\mathsf{FO}(<)$-rewriting of $\boldsymbol{q}$:

$$\exists x\, \big[\mathit{Malfunction}(x) \land \bigvee_{1\le i\le 5} \big(\mathit{Fixed}(x+i) \land \bigvee_{1\le j\le 5} \bigwedge_{0\le k\le 2} \mathit{Malfunction}(x+i+j-k)\big)\big].$$

**Problem and related work.** The problem we are interested in can be formulated in complexity-theoretic terms: given an *LTL* OMQ $\boldsymbol{q}$, determine the data complexity of answering $\boldsymbol{q}$ over any data instance $\mathcal{A}$ in a given signature $\Xi$. For simplicity's sake, let us assume that $\boldsymbol{q}$ is Boolean (with a yes/no answer). Then the data instances $\mathcal{A}$ over which the answer to $\boldsymbol{q}$ is yes form a language $\boldsymbol{L}(\boldsymbol{q})$ over the alphabet $2^{\Xi}$. In fact, using the automata-theoretic view of *LTL* [58], one can show that $\boldsymbol{L}(\boldsymbol{q})$ is regular, and so can be decided in $\mathrm{NC}^1$ [9, 11].

| class of OMQs | $\mathsf{FO}(<)$ | $\mathsf{FO}(<,\equiv)$, $\mathrm{AC}^0$ | $\mathsf{FO}(<,\mathsf{MOD})$, $\mathrm{ACC}^0$ |
|---|---|---|---|
| $\mathit{LTL}^{\bigcirc}_{horn}$ OMAQs<br>$\mathit{LTL}_{krom}$ OMPEQs<br>$\mathit{LTL}^{\Box\bigcirc}_{bool}$ OMQs | ExpSpace | ExpSpace | ExpSpace |
| linear $\mathit{LTL}^{\bigcirc}_{horn}$ OMAQ<br>linear $\mathit{LTL}^{\bigcirc}_{horn}$ OMPQs | PSpace | PSpace | PSpace<br>? |
| $\mathit{LTL}^{\bigcirc}_{krom}$ OMAQs<br>$\mathit{LTL}^{\bigcirc}_{core}$ OMPEQs<br>$\mathit{LTL}^{\bigcirc}_{core}$ OMPQs | coNP<br>$\Pi^p_2$<br>PSpace | all in $\mathrm{AC}^0$ [7] | – |

**Table 1** Complexity of deciding FO-rewritability of *LTL* OMQs.

The circuit and descriptive complexity of regular languages was investigated in [10,51], which established an $\mathrm{AC}^0/\mathrm{ACC}^0/\mathrm{NC}^1$ trichotomy, gave algebraic characterisations of languages in these classes (implying that the trichotomy is decidable) and also in terms of extensions of FO. Namely, the languages in $\mathrm{AC}^0$ are definable by $\mathsf{FO}(<,\equiv)$-sentences with unary predicates $x \equiv 0 \,(\mathrm{mod}\ n)$; those in $\mathrm{ACC}^0$ are definable by $\mathsf{FO}(<,\mathsf{MOD})$-sentences with quantifiers $\exists^n x\, \psi(x)$ checking whether the number of positions satisfying $\psi$ is divisible by $n$; and all regular languages are definable in $\mathsf{FO}(\mathsf{RPR})$ with relational primitive recursion [23].

Thus, our problem can be equivalently formulated in logic terms: given an *LTL* OMQ $\boldsymbol{q}$, decide whether $\boldsymbol{L}(\boldsymbol{q})$ is $\mathsf{FO}(<,\equiv)$- or $\mathsf{FO}(<,\mathsf{MOD})$-definable. In the OBDA context, we are also interested in $\mathsf{FO}(<)$-definability (without any extra predicates, quantifiers or recursion), which has been thoroughly investigated in both automata theory and logic; see, e.g., [26] and references therein. In particular, deciding $\mathsf{FO}(<)$-definability of regular languages is known to be PSpace-complete [14, 21, 49]. Note also that, by Kamp's Theorem [35, 45], $\mathsf{FO}(<)$-rewritability reduces answering *LTL* OMQs to model checking *LTL*-formulas.

**Our contribution.** Let $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<,\equiv), \mathsf{FO}(<,\mathsf{MOD})\}$. First, using results of [9, 10], we obtain criteria of $\mathcal{L}$-definability of DFAs in terms of their transition monoids, which are then applied to show that deciding $\mathcal{L}$-definability of the language of a given 2NFA can be done in PSpace. We also establish a matching lower bound for minimal DFAs. These results have been known for $\mathcal{L} = \mathsf{FO}(<)$ and DFAs/NFAs [14, 21, 49]—but otherwise are novel.

To investigate $\mathcal{L}$-rewritability of *LTL* OMQs $\boldsymbol{q} = (\mathcal{O}, \varkappa)$, we follow the classification of [7], according to which the axioms of every *LTL* ontology $\mathcal{O}$ are given in the clausal form

$$\Box_P \Box_F \big(C_1 \land \dots \land C_k \;\to\; C_{k+1} \lor \dots \lor C_{k+m}\big), \tag{4}$$

where the $C_i$ are atoms, possibly prefixed by the temporal operators $\bigcirc_F$, $\bigcirc_P$, $\Box_F$, $\Box_P$. Given some $\boldsymbol{o} \in \{\Box, \bigcirc, \Box\bigcirc\}$ and $\boldsymbol{c} \in \{\mathit{bool}, \mathit{horn}, \mathit{krom}, \mathit{core}\}$, we denote by $\mathit{LTL}^{\boldsymbol{o}}_{\boldsymbol{c}}$ the fragment of *LTL* with clauses of the form (4), where the $C_i$ can only use the (future and past) operators indicated in $\boldsymbol{o}$, and $m \le 1$ if $\boldsymbol{c} = \mathit{horn}$; $k+m \le 2$ if $\boldsymbol{c} = \mathit{krom}$; $k+m \le 2$ and $m \le 1$ if $\boldsymbol{c} = \mathit{core}$; and arbitrary $k$, $m$ if $\boldsymbol{c} = \mathit{bool}$. If $\boldsymbol{o}$ is omitted, the $C_i$ are atomic. An $\mathit{LTL}^{\boldsymbol{o}}_{horn}$-ontology $\mathcal{O}$ is linear if, in each of its axioms (4), at most one $C_i$, for $1 \le i \le k$, can occur on the right-hand side of an axiom in $\mathcal{O}$ (is an IDB predicate, in datalog parlance). We distinguish between arbitrary $\mathit{LTL}^{\boldsymbol{o}}_{\boldsymbol{c}}$ OMQs $\boldsymbol{q} = (\mathcal{O}, \varkappa)$, where $\mathcal{O}$ is any $\mathit{LTL}^{\boldsymbol{o}}_{\boldsymbol{c}}$ ontology and $\varkappa$ any *LTL*-formula with $\bigcirc$-, $\Box$- and $\Diamond$-operators; positive OMQs (OMPQs), where $\varkappa$ is $\to,\neg$-free; existential OMPQs (OMPEQs) with $\Box$-free $\varkappa$; and atomic OMQs (OMAQs) with atomic $\varkappa$.

The main result of this paper is the tight complexity bounds on deciding $\mathcal{L}$-rewritability (and so data complexity) of *LTL* OMQs in various classes defined above, which are summarised in Table 1. The ExpSpace upper bound in the first stripe is shown using our $\mathcal{L}$-definability criteria and exponential-size NFAs for *LTL* akin to those in [57]; in the proof of the matching lower bound, an exponential-size automaton is encoded in a polynomial-size ontology. If the ontology in an $LTL^{\bigcirc}_{horn}$ OMAQ is linear, we show that its language (yes-data instances) can be captured by a polynomial-size 2NFA, which allows us to reduce the complexity of deciding $\mathcal{L}$-rewritability to PSpace. However, for linear $LTL^{\bigcirc}_{horn}$ OMPQs (with more expressive queries $\varkappa$), the existence of polynomial-size 2NFAs remains open; instead, we show how the structure of the canonical (minimal) models for $LTL^{\bigcirc}_{horn}$-ontologies can be utilised to yield a PSpace algorithm. In the third stripe of the table, we deal with binary-clause ontologies. The coNP-completeness of deciding FO-rewritability of $LTL^{\bigcirc}_{krom}$ OMAQs is established using unary NFAs and results from [50]. The $\Pi^p_2$-completeness for $LTL^{\bigcirc}_{core}$ OMPEQs (without $\vee$ in ontologies but with $\wedge$, $\vee$, $\Diamond$ in queries) and the PSpace-completeness for $LTL^{\bigcirc}_{core}$ OMPQs (admitting $\Box$ in queries, too) can be explained by the fact that the combined complexity of answering such OMPEQs and OMPQs is, respectively, NP- and $\mathrm{P}^{\mathrm{NP}}[O(\log n)]$-complete (like validity in Carnap's modal logic [32]), rather than tractable as in the previous case.

It might be of interest to compare the results in Table 1 with the complexity of deciding FO-rewritability (aka boundedness) of datalog queries, which is

- undecidable for linear datalog queries with binary predicates and for ternary linear datalog queries with a single recursive rule [33, 43];
- 2NExpTime-complete for monadic disjunctive datalog queries [17, 27];
- 2ExpTime-complete for monadic datalog queries [12, 24];
- PSpace-complete for linear monadic programs [24, 54];
- NP-complete for linear monadic single rule programs [55].

## 2 Preliminaries: *LTL* OMQs

In our setting, the alphabet of linear temporal logic *LTL* comprises a set of *atomic concepts* $A_i$, $i < \omega$. *Basic temporal concepts*, $C$, are defined by the grammar

$$C \;::=\; A_i \;\mid\; \Box_F C \;\mid\; \Box_P C \;\mid\; \bigcirc_F C \;\mid\; \bigcirc_P C$$

with the *temporal operators* $\Box_F/\Box_P$ (always in the future/past) and $\bigcirc_F/\bigcirc_P$ (at the next/previous moment). A *temporal ontology*, $\mathcal{O}$, is a finite set of *axioms* of the form

$$C_1 \wedge \dots \wedge C_k \;\to\; C_{k+1} \vee \dots \vee C_{k+m}, \tag{5}$$

where $k, m \geq 0$, the $C_i$ are basic temporal concepts, the empty $\wedge$ is $\top$, and the empty $\vee$ is $\bot$. Following the *DL-Lite* convention [3, 5], we classify ontologies by the shape of their axioms and the temporal operators that can occur in them. Suppose $\boldsymbol{c} \in \{horn, krom, core, bool\}$ and $\boldsymbol{o} \in \{\Box, \bigcirc, \Box\bigcirc\}$. The axioms of an $LTL^{\boldsymbol{o}}_{\boldsymbol{c}}$-*ontology* may only contain occurrences of the (future and past) temporal operators in $\boldsymbol{o}$ and satisfy the following restrictions on $k$ and $m$ in (5) indicated by $\boldsymbol{c}$: *horn* requires $m \leq 1$, *krom* requires $k + m \leq 2$, *core* both $k + m \leq 2$ and $m \leq 1$, while *bool* imposes no restrictions. For example, axiom (2) from Example 1 is allowed in all of these fragments, (3) is equivalent to a *Horn* axiom (with $\bot$ on the right-hand side), and (1) can be expressed in *Krom* as explained in Remark 3 below. A basic concept is called an *IDB* (intensional database) *concept* in an ontology $\mathcal{O}$ if its atom occurs on the right-hand side of some axiom in $\mathcal{O}$. The set of IDB atomic concepts in $\mathcal{O}$ is denoted by $idb(\mathcal{O})$. An $LTL^{\boldsymbol{o}}_{horn}$-ontology is called *linear* if each of its axioms $C_1 \wedge \dots \wedge C_k \to C_{k+1}$ contains *at most one* IDB concept $C_i$, for $1 \leq i \leq k$.

A *data instance*—*ABox* in description logic parlance—is a finite set $\mathcal{A}$ of atoms $A_i(\ell)$, for $\ell \in \mathbb{Z}$, together with a finite interval $\mathsf{tem}(\mathcal{A}) = [m, n] \subseteq \mathbb{Z}$, called the *active domain* of $\mathcal{A}$, such that $m \leq \ell \leq n$, for all $A_i(\ell) \in \mathcal{A}$. If $\mathcal{A} = \emptyset$, then $\mathsf{tem}(\mathcal{A})$ may also be $\emptyset$. Otherwise, we assume (without loss of generality) that $m = 0$. If $\mathsf{tem}(\mathcal{A})$ is not specified explicitly, it is assumed to be either empty or $[0, n]$, where $n$ is the maximal timestamp in $\mathcal{A}$. By a *signature*, $\Xi$, we mean any finite set of atomic concepts. An ABox $\mathcal{A}$ is said to be a $\Xi$-*ABox* if $A_i(\ell) \in \mathcal{A}$ implies $A_i \in \Xi$.

We query ABoxes by means of *temporal concepts*, $\varkappa$, which are *LTL*-formulas built from the atoms $A_i$, Booleans $\wedge$, $\vee$, $\neg$, temporal operators $\bigcirc_F$, $\Box_F$, $\Diamond_F$ (eventually) and their past-time counterparts $\bigcirc_P$, $\Box_P$, $\Diamond_P$ (previously). If $\varkappa$ does not contain $\neg$, we call it *positive*; if $\varkappa$ does not contain $\Box_P$ and $\Box_F$ either, we call *positive existential*.

An *interpretation* is a structure $\mathcal{I} = (\mathbb{Z}, A_0^{\mathcal{I}}, A_1^{\mathcal{I}}, \dots)$ with $A_i^{\mathcal{I}} \subseteq \mathbb{Z}$, for every $i < \omega$. The *extension* $\varkappa^{\mathcal{I}}$ of a temporal concept $\varkappa$ in $\mathcal{I}$ is defined inductively as usual in *LTL* under the 'strict semantics' [25, 30]:

$$(\bigcirc_F \varkappa)^{\mathcal{I}} = \{\, n \in \mathbb{Z} \mid n+1 \in \varkappa^{\mathcal{I}} \,\},$$
$$(\Box_F \varkappa)^{\mathcal{I}} = \{\, n \in \mathbb{Z} \mid k \in \varkappa^{\mathcal{I}}, \text{ for all } k > n \,\},$$
$$(\Diamond_F \varkappa)^{\mathcal{I}} = \{\, n \in \mathbb{Z} \mid \text{there is } k > n \text{ with } k \in \varkappa^{\mathcal{I}} \,\},$$

and symmetrically for the past-time operators. We regard $\mathcal{I}, n \models \varkappa$ as synonymous to $n \in \varkappa^{\mathcal{I}}$. We say that an axiom (5) is *true* in $\mathcal{I}$ if $C_1^{\mathcal{I}} \cap \dots \cap C_k^{\mathcal{I}} \subseteq C_{k+1}^{\mathcal{I}} \cup \dots \cup C_{k+m}^{\mathcal{I}}$, that is, if it holds at every moment of time; cf. (4). An interpretation $\mathcal{I}$ is a *model* of $\mathcal{O}$ if all axioms of $\mathcal{O}$ are true in $\mathcal{I}$; it is a *model* of $\mathcal{A}$ if $A_i(\ell) \in \mathcal{A}$ implies $\ell \in A_i^{\mathcal{I}}$.

An $LTL_{\boldsymbol{c}}^{\boldsymbol{o}}$ *ontology-mediated query* (OMQ) is a pair of the form $\boldsymbol{q} = (\mathcal{O}, \varkappa)$, where $\mathcal{O}$ is an $LTL_{\boldsymbol{c}}^{\boldsymbol{o}}$ ontology and $\varkappa$ a temporal concept. If $\varkappa$ is positive, we call $\boldsymbol{q}$ a *positive OMQ* (OMPQ, for short), if $\varkappa$ is positive existential, we call $\boldsymbol{q}$ a *positive existential OMQ* (OMPEQ), and if $\varkappa$ is an atomic concept, we call $\boldsymbol{q}$ *atomic* (OMAQ). The set of atomic concepts occurring in $\boldsymbol{q}$ is denoted by $\mathsf{sig}(\boldsymbol{q})$.

We can treat $\boldsymbol{q}$ as a *Boolean* OMQ, which returns a yes/no answer, or as a *specific* OMQ, which returns timestamps from the ABox in question assigned to the free variable, say $x$, in the standard FO-translation of $\varkappa$. In the latter case, we write $\boldsymbol{q}(x) = (\mathcal{O}, \varkappa(x))$. More precisely, a *certain answer* to a Boolean OMQ $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ over an ABox $\mathcal{A}$ is yes if, for every model $\mathcal{I}$ of $\mathcal{O}$ and $\mathcal{A}$, there is $k \in \mathbb{Z}$ such that $k \in \varkappa^{\mathcal{I}}$, in which case we write $(\mathcal{O}, \mathcal{A}) \models \exists x \varkappa(x)$. If $(\mathcal{O}, \mathcal{A}) \not\models \exists x \varkappa(x)$, the certain answer to $\boldsymbol{q}$ over $\mathcal{A}$ is no. We write $(\mathcal{O}, \mathcal{A}) \models \varkappa(k)$, for $k \in \mathbb{Z}$, if $k \in \varkappa^{\mathcal{I}}$ in all models $\mathcal{I}$ of $\mathcal{O}$ and $\mathcal{A}$. A *certain answer* to a specific OMQ $\boldsymbol{q}(x) = (\mathcal{O}, \varkappa(x))$ over $\mathcal{A}$ is any $k \in \mathsf{tem}(\mathcal{A})$ with $(\mathcal{O}, \mathcal{A}) \models \varkappa(k)$. By the *evaluation* (or *answering*) *problems* for $\boldsymbol{q}$ or $\boldsymbol{q}(x)$ we understand the decision problem '$(\mathcal{O}, \mathcal{A}) \models^? \exists x \varkappa(x)$' or '$(\mathcal{O}, \mathcal{A}) \models^? \varkappa(k)$' with input $\mathcal{A}$ or, respectively, $\mathcal{A}$ and $k \in \mathsf{tem}(\mathcal{A})$. We say that $\boldsymbol{q}$ or $\boldsymbol{q}(x)$ is in a complexity class $\mathcal{C}$ if the corresponding evaluation problem is in $\mathcal{C}$.

▶ **Example 2.** (*i*) Suppose $\mathcal{O}_1 = \{A \to \Box_F B,\ \Box_F B \to C\}$ and $\boldsymbol{q}_1 = (\mathcal{O}_1, C \wedge D)$. The certain answer to $\boldsymbol{q}_1$ over $\mathcal{A}_1 = \{D(0), B(1), A(1)\}$ is yes, and no over $\mathcal{A}_2 = \{D(0), A(1)\}$. The only answer to $\boldsymbol{q}_1(x) = \big(\mathcal{O}_1, (C \wedge D)(x)\big)$ over $\mathcal{A}_1$ is 0.

(*ii*) Let $\mathcal{O}_2 = \{\, \bigcirc_P A \to B,\ \bigcirc_P B \to A,\ A \wedge B \to \bot \,\}$. The certain answer to $\boldsymbol{q}_2 = (\mathcal{O}_2, C)$ over $\mathcal{A}_1 = \{A(0)\}$ is no, and yes over $\mathcal{A}_2 = \{A(0), A(1)\}$. There are no certain answers to $\boldsymbol{q}_2(x) = (\mathcal{O}_1, C(x))$ over $\mathcal{A}_1$, while over $\mathcal{A}_2$ the answers are 0 and 1.

(*iii*) Consider now the ontology

$$\mathcal{O}_3 = \{\bigcirc_P B_k \wedge A_0 \to B_k,\ \bigcirc_P B_{1-k} \wedge A_1 \to B_k \mid k = 0, 1\}.$$

For any word $\boldsymbol{e} = e_1 \dots e_n \in \{0,1\}^n$, let $\mathcal{A}_{\boldsymbol{e}} = \{B_0(0)\} \cup \{A_{e_i}(i) \mid 0 < i \leq n\} \cup \{E(n)\}$. The answer to $\boldsymbol{q}_3 = (\mathcal{O}_3, B_0 \wedge E)$ over the ABox $\mathcal{A}_{\boldsymbol{e}}$ is yes iff the number of 1s in $\boldsymbol{e}$ is even.

($iv$) Let $\mathcal{O}_4 = \{A \to \bigcirc_F B\}$ and $\boldsymbol{q}_4 = (\mathcal{O}_4, B)$. Then, the answer to $\boldsymbol{q}_4$ over $\mathcal{A} = \{A(0)\}$ is yes; however, there are no certain answers to $\boldsymbol{q}_4(x) = (\mathcal{O}_4, B(x))$ over $\mathcal{A}$.

($v$) Let $\mathcal{O}_5 = \{A \to B \vee \bigcirc_F B\}$. The certain answer to $\boldsymbol{q}_5 = (\mathcal{O}_5, B)$ over $\mathcal{A} = \{A(0), C(1)\}$ is yes; however, there are no certain answers to $\boldsymbol{q}_5(x)$ over $\mathcal{A}$.

▶ Remark 3. As follows from [4, 28], if arbitrary *LTL*-formulas are used as axioms of an ontology $\mathcal{O}$, then one can construct an $LTL^{\Box\bigcirc}_{bool}$ ontology $\mathcal{O}'$ that is a model conservative extension of $\mathcal{O}$. For example, let $\mathcal{O}'$ be the result of replacing (1) in $\mathcal{O}$ from Example 1 by $\mathit{Malfunction} \wedge \Box_F X \to \bot$ and $\top \to X \vee \mathit{Fixed}$, for a fresh concept name $X$. Then the OMQ $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ is equivalent to $\boldsymbol{q}' = (\mathcal{O}', \varkappa)$ in the sense that $\boldsymbol{q}$ and $\boldsymbol{q}'$ have the same certain answers over any $\mathsf{sig}(\boldsymbol{q})$-ABox.

Let $\mathcal{L}$ be a class of FO-formulas that can be interpreted over finite linear orders. A Boolean OMQ $\boldsymbol{q}$ is *$\mathcal{L}$-rewritable over $\Xi$-ABoxes* if there is an $\mathcal{L}$-sentence $\boldsymbol{Q}$ such that, for any $\Xi$-ABox $\mathcal{A}$, the certain answer to $\boldsymbol{q}$ over $\mathcal{A}$ is yes iff $\mathfrak{S}_{\mathcal{A}} \models \boldsymbol{Q}$. Here, $\mathfrak{S}_{\mathcal{A}}$ is a structure with domain $\mathsf{tem}(\mathcal{A})$ ordered by $<$, in which $\mathfrak{S}_{\mathcal{A}} \models A_i(\ell)$ iff $A_i(\ell) \in \mathcal{A}$. A specific OMQ $\boldsymbol{q}(x)$ is *$\mathcal{L}$-rewritable over $\Xi$-ABoxes* if there is an $\mathcal{L}$-formula $\boldsymbol{Q}(x)$ with one free variable $x$ such that, for any $\Xi$-ABox $\mathcal{A}$, $k$ is a certain answer to $\boldsymbol{q}(x)$ over $\mathcal{A}$ iff $\mathfrak{S}_{\mathcal{A}} \models \boldsymbol{Q}(k)$. The sentence $\boldsymbol{Q}$ and the formula $\boldsymbol{Q}(x)$ are called *$\mathcal{L}$-rewritings* of the OMQs $\boldsymbol{q}$ and $\boldsymbol{q}(x)$, respectively.

We require four languages $\mathcal{L}$ for rewriting *LTL* OMQs, which are listed below in order of increasing expressive power:

**FO($<$):** (monadic) first-order formulas with the built-in predicate $<$ for order;

**FO($<, \equiv$):** FO($<$)-formulas with unary (numerical) predicates $x \equiv 0 \,(\mathrm{mod}\ N)$, for $N > 1$;

**FO($<$, MOD):** FO($<$)-formulas with quantifiers $\exists^N x$, for $N > 1$, that are defined by taking $\mathfrak{S}_{\mathcal{A}} \models \exists^N x\, \psi(x)$ iff the cardinality of $\{n \in \mathsf{tem}(\mathcal{A}) \mid \mathfrak{S}_{\mathcal{A}} \models \psi(n)\}$ is divisible by $N$ (note that $x \equiv 0 \,(\mathrm{mod}\ N)$ is definable as $\exists^N y\, (y < x)$);

**FO(RPR):** FO($<$) with relational primitive recursion [23].

As well-known, FO($<, \equiv$) is strictly more expressive than FO($<$) and strictly less expressive than FO($<$, MOD), which is illustrated by the examples below.

▶ **Example 4.** ($i$) An FO($<$)-rewriting of $\boldsymbol{q}_1(x)$ is

$$\boldsymbol{Q}_1(x) = D(x) \wedge [C(x) \vee \exists y\, (A(y) \wedge \forall z\, ((x < z \leq y) \to B(z)))],$$

$\exists x\, \boldsymbol{Q}_1(x)$ is an FO($<$)-rewriting of $\boldsymbol{q}_1$.

($ii$) An FO($<, \equiv$)-rewriting of $\boldsymbol{q}_2(x)$ is

$$\begin{aligned}\boldsymbol{Q}_2(x) = \ & C(x) \vee \exists x, y\, [(A(x) \wedge A(y) \wedge \mathsf{odd}(x,y)) \vee \\ & (B(x) \wedge B(y) \wedge \mathsf{odd}(x,y)) \vee (A(x) \wedge B(y) \wedge \neg\mathsf{odd}(x,y))],\end{aligned}$$

where $\mathsf{odd}(x,y) = \big(x \equiv 0 \,(\mathrm{mod}\ 2) \leftrightarrow y \not\equiv 0 \,(\mathrm{mod}\ 2)\big)$ implies that $|x - y|$ is odd, and an FO($<, \equiv$)-rewriting of $\boldsymbol{q}_2$ is $\exists x\, \boldsymbol{Q}_2(x)$. Recall that odd is not expressible in FO($<$) [39].

($iii$) The OMQ $\boldsymbol{q}_3$ is not rewritable to an FO-formula with any numeric predicates as PARITY is not in $\mathrm{AC}^0$ [29]; the following sentence is an FO($<$, MOD)-rewriting of $\boldsymbol{q}_3$:

$$\begin{aligned}\boldsymbol{Q}_3 = \exists x, y\, \big[& E(x) \wedge (y \leq x) \wedge \forall z\, \big((y < z \leq x) \to A_0(z) \vee A_1(z)\big) \wedge \\ & \big((B_0(y) \wedge \exists^2 z\, ((y < z \leq x) \wedge A_1(z))) \vee (B_1(y) \wedge \neg\exists^2 z\, ((y < z \leq x) \wedge A_1(z)))\big)\big].\end{aligned}$$

($iv$) An FO($<$)-rewriting of $\boldsymbol{q}_4(x)$ is $B(x) \vee A(x-1)$; an FO($<$)-rewriting of $\boldsymbol{q}_4$ is $\boldsymbol{Q}_4 = \exists x\, (A(x) \vee B(x))$.

($v$) The same $\boldsymbol{Q}_4$ is an FO($<$)-rewriting of $\boldsymbol{q}_5$, and $B(x)$ is an FO($<$)-rewriting of $\boldsymbol{q}_5(x)$.

It has been shown in [7] that all (Boolean and specific) $LTL$ OMQs are $\mathsf{FO}(\mathsf{RPR})$-rewritable and that specific OMPQs can be classified syntactically by their rewritability type as shown in Table 2. This means, for example, that all $LTL_{core}^{\Box\bigcirc}$ OMPQs are $\mathsf{FO}(<,\equiv)$-rewritable, with some of them being not $\mathsf{FO}(<)$-rewritable. It is to be noted that $\mathsf{FO}(<,\mathsf{MOD})$-rewritable OMQs such as $\boldsymbol{q}_3$ in Example 2 are not captured by these syntactic classes.

| | OMAQs | | OMPQs | |
|---|---|---|---|---|
| $\boldsymbol{c}$ | $LTL_{\boldsymbol{c}}^{\Box}$ | $LTL_{\boldsymbol{c}}^{\bigcirc}$ and $LTL_{\boldsymbol{c}}^{\Box\bigcirc}$ | $LTL_{\boldsymbol{c}}^{\Box}$ | $LTL_{\boldsymbol{c}}^{\bigcirc}$ and $LTL_{\boldsymbol{c}}^{\Box\bigcirc}$ |
| *bool* | | $\mathsf{FO}(\mathsf{RPR})$ | $\mathsf{FO}(\mathsf{RPR})$ | |
| *krom* | $\mathsf{FO}(<)$ | $\mathsf{FO}(<,\equiv)$ | | $\mathsf{FO}(\mathsf{RPR})$ |
| *horn* | | $\mathsf{FO}(\mathsf{RPR})$ | $\mathsf{FO}(<)$ | |
| *core* | | $\mathsf{FO}(<,\equiv)$ | | $\mathsf{FO}(<,\equiv)$ |

**Table 2** Rewritability of specific $LTL$ OMQs.

In this paper, our aim is to understand how (complex it is) to decide the optimal type of FO-rewritability for a given $LTL$ OMQ $\boldsymbol{q}$ over $\Xi$-ABoxes. We begin by observing an intimate connection between $\mathcal{L}$-rewritability of OMQs and $\mathcal{L}$-definability of certain regular languages.

A language $\boldsymbol{L}$ over an alphabet $\Sigma$ is *$\mathcal{L}$-definable* if there is an $\mathcal{L}$-sentence $\varphi$ in the signature $\Sigma$, whose symbols are treated as unary predicates, such that, for any $w \in \Sigma^*$, we have $w = a_0 \dots a_n \in \boldsymbol{L}$ iff $\mathfrak{S}_w \models \varphi$, where $\mathfrak{S}_w$ is a structure with domain $\{0,\dots,n\}$, in which $\mathfrak{S}_w \models a(i)$ iff $a = a_i$, for $i \leq n$.

For any OMQ $\boldsymbol{q}$ and $\Xi \subseteq \mathsf{sig}(\boldsymbol{q})$, we regard $\Sigma_\Xi = 2^\Xi$ as an *alphabet*. Any $\Xi$-ABox $\mathcal{A}$ can be given as a $\Sigma_\Xi$-word $w_\mathcal{A} = a_0 \dots a_n$ with $a_i = \{A \mid A(i) \in \mathcal{A}\}$. Conversely, any $\Sigma_\Xi$-word $w = a_0 \dots a_n$ gives the ABox $\mathcal{A}_w$ with $\mathsf{tem}(\mathcal{A}_w) = [0,n]$ and $A(i) \in \mathcal{A}_w$ iff $A \in a_i$. The word $\emptyset$ corresponds to $\mathcal{A}_\emptyset = \emptyset$ with $\mathsf{tem}(\mathcal{A}_\emptyset) = [0,0]$.

The *language* $\boldsymbol{L}_\Xi(\boldsymbol{q})$, for a Boolean $\boldsymbol{q}$, is defined to be the set of $\Sigma_\Xi$-words $w_\mathcal{A}$ such that the certain answer to $\boldsymbol{q}$ over $\mathcal{A}$ is $\mathsf{yes}$. For a specific $\boldsymbol{q}(x)$, we take $\Gamma_\Xi = \Sigma_\Xi \cup \Sigma'_\Xi$ with a disjoint copy $\Sigma'_\Xi$ of $\Sigma_\Xi$ and represent a pair $(\mathcal{A}, i)$ with a $\Xi$-ABox $\mathcal{A}$ and $i \in \mathsf{tem}(\mathcal{A})$ as a $\Gamma_\Xi$-word $w_{\mathcal{A},i} = a_0 \dots a'_i \dots a_n$, where $a'_i = \{A \mid A(i) \in \mathcal{A}\} \in \Sigma'_\Xi$ and $a_j = \{A \mid A(j) \in \mathcal{A}\} \in \Sigma_\Xi$, for $j \neq i$. The *language* $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$ is the set of $\Gamma_\Xi$-words $w_{\mathcal{A},i}$ such that $i$ is a certain answer to $\boldsymbol{q}(x)$ over $\mathcal{A}$. The following result is proved in a way similar to [58, Theorem 2.1].

▶ **Proposition 5.** *Both $\boldsymbol{L}_\Xi(\boldsymbol{q})$ and $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$ are regular languages.*

**Proof.** Let $\mathsf{sub}_{\boldsymbol{q}}$ (or $\mathsf{sub}_\mathcal{O}$) be the set of temporal concepts in $\boldsymbol{q}$ (respectively, $\mathcal{O}$) and their negations. A *type for* $\boldsymbol{q}$ (respectively, $\mathcal{O}$) is any maximal subset $\tau \subseteq \mathsf{sub}_{\boldsymbol{q}}$ (respectively, $\tau \subseteq \mathsf{sub}_\mathcal{O}$) consistent with $\mathcal{O}$. Let $\boldsymbol{T}$ be the set of all types for $\boldsymbol{q}$. Define an NFA $\mathfrak{A}$ over $\Sigma_\Xi$ whose language $\boldsymbol{L}(\mathfrak{A})$ is $\Sigma_\Xi^* \setminus \boldsymbol{L}_\Xi(\boldsymbol{q})$. Its states are $Q_{\neg\varkappa} = \{\tau \in \boldsymbol{T} \mid \neg\varkappa \in \tau\}$. The transition relation $\to_a$, for $a \in \Sigma_\Xi$, is defined by taking $\tau_1 \to_a \tau_2$ if the following conditions hold:
(a) $a \subseteq \tau_2$,
(b) $\bigcirc_{\scriptscriptstyle F} C \in \tau_1$ iff $C \in \tau_2$,
(c) $\Box_{\scriptscriptstyle F} C \in \tau_1$ iff $C \in \tau_2$ and $\Box_{\scriptscriptstyle F} C \in \tau_2$,
(d) $\Diamond_{\scriptscriptstyle F} C \in \tau_1$ iff $C \in \tau_2$ or $\Diamond_{\scriptscriptstyle F} C \in \tau_2$,
and symmetrically for the corresponding past-time operators. The initial (accepting) states are those $\tau \in Q_{\neg\varkappa}$, for which $\tau \cup \{\Box_{\scriptscriptstyle P} \neg\varkappa\}$ (respectively, $\tau \cup \{\Box_{\scriptscriptstyle F} \neg\varkappa\}$) is consistent with $\mathcal{O}$. Then $w \in \boldsymbol{L}(\mathfrak{A})$ iff $(\mathcal{O}, \mathcal{A}_w) \not\models \exists x\, \varkappa(x)$, for any $w \in \Sigma_\Xi^*$. Indeed, if $w \in \boldsymbol{L}(\mathfrak{A})$, we take an

accepting run $\tau_0, \dots, \tau_n$ of $\mathfrak{A}$ on $w$, a model $\mathcal{I}^-$ of $\mathcal{O}$ with $\mathcal{I}^-, k \models \tau_0 \cup \{\Box_P \neg \varkappa\}$, a model $\mathcal{I}^+$ of $\mathcal{O}$ with $\mathcal{I}^+, l \models \tau_n \cup \{\Box_F \neg \varkappa\}$, for some $k, l \in \mathbb{Z}$, and construct a new interpretation $\mathcal{I}$ that has the types $\tau_0, \dots, \tau_n$ in the interval $[0, n]$, before (after) which it has the same types as in $\mathcal{I}^-$ in $(-\infty, k)$ (respectively, $\mathcal{I}^+$ on $(l, \infty)$). One can readily check that $\mathcal{I}$ is a model of $\mathcal{O}$ and $\mathcal{A}_w$ such that $\varkappa^{\mathcal{I}} = \emptyset$, and so $(\mathcal{O}, \mathcal{A}_w) \not\models \exists x\, \varkappa(x)$. The opposite direction is obvious.

To show that $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$ is regular, we observe first that the language $\boldsymbol{L}$ over $\Gamma_\Xi$ comprising words of the form $w_{\mathcal{A},i}$, for all non-empty $\Xi$-Aboxes $\mathcal{A}$ and $i \in \mathsf{tem}(\mathcal{A})$, is regular. Thus, it suffices to define an NFA $\mathfrak{A}$ over $\Gamma_\Xi$ such that $\boldsymbol{L}_\Xi(\boldsymbol{q}(x)) = \boldsymbol{L} \setminus \boldsymbol{L}(\mathfrak{A})$. The set of states in $\mathfrak{A}$ is $\boldsymbol{T} \cup \boldsymbol{T}'$ with a disjoint copy $\boldsymbol{T}'$ of $\boldsymbol{T}$. The set of initial states is $\boldsymbol{T}$ and the set of accepting states is $\boldsymbol{T}'$. The transition relation $\to_a$, for $a \in \Sigma_\Xi$, is defined by taking $\tau_1 \to_a \tau_2$ if either $\tau_1, \tau_2 \in \boldsymbol{T}$ or $\tau_1, \tau_2 \in \boldsymbol{T}'$ and conditions (a)–(d) are satisfied; for $a' \in \Sigma'_\Xi$, we set $\tau_1 \to_{a'} \tau_2$ if $\tau_1 \in \boldsymbol{T}$, $\tau_2 \in \boldsymbol{T}'$, $\neg\varkappa \in \tau_2$, $a' \subseteq \tau_2$, and (b)–(d) hold. It is easy to see that, for any $\Xi$-ABox $\mathcal{A}$ and $i \in \mathsf{tem}(\mathcal{A})$, there exists a model $\mathcal{I}$ of $\mathcal{O}$ and $\mathcal{A}$ with $i \notin \varkappa^{\mathcal{I}}$ iff $w_{\mathcal{A},i} \in \boldsymbol{L}(\mathfrak{A})$. ❑

Note that the number of states in the NFAs in the proof above is $2^{O(|\boldsymbol{q}|)}$ and that they can be constructed in exponential time in the size $|\boldsymbol{q}|$ of $\boldsymbol{q}$ as *LTL*-satisfiability is in PSPACE.

In Section 5, we show that, in fact, the type of $\mathcal{L}$-rewritability of $\boldsymbol{q}$ coincides with the type of $\mathcal{L}$-definability of the regular languages $\boldsymbol{L}_\Xi(\boldsymbol{q})$ and $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$. But before that, we revisit the well-known problem of deciding $\mathcal{L}$-definability of regular languages.

## 3 Preliminaries: Monoids, Groups and Automata

In this section, we first briefly remind the reader of the basic algebraic and automata-theoretic notions required in the remainder of the paper, and then prove the criteria of $\mathcal{L}$-definability of regular languages we need to obtain our complexity results.

### 3.1 Semigroups, monoids, groups

A *semigroup* is a structure $\mathfrak{S} = (S, \cdot)$ where $\cdot$ is an associative binary operation. Given $s, s' \in S$ and $n > 0$, we write $s^n$ for $s \cdot \ \dots\ \cdot s$ $n$-times, and often write $ss'$ for $s \cdot s'$. An element $s$ in a semigroup $\mathfrak{S}$ is called *idempotent* if $s^2 = s$. An element $e$ in a semigroup $\mathfrak{S}$ is called an *identity element* if $e \cdot x = x \cdot e = x$ for every $x \in S$. (It is easy to see that such an $e$, if exists, must be unique.) The identity element is clearly idempotent. A *monoid* is a semigroup that has an identity element. (We don't put it to the signature.) For any element $s$ in a monoid, we let $s^0 = e$. A monoid $\mathfrak{S} = (S, \cdot)$ is called a *group* if for every $x \in S$ there is some $x^- \in S$ such that $x \cdot x^- = x^- \cdot x = e$ for the identity element $e$ of $\mathfrak{S}$. Then $x^-$ is called the *inverse of* $x$. (It is easy to see that in a group every element has a unique inverse.) A group is called *trivial* if it has only one element, and *nontrivial* otherwise.

Given two groups $\mathfrak{G} = (G, \cdot)$ and $\mathfrak{G}' = (G', \cdot')$, a map $h \colon G \to G'$ is a *group homomorphism from* $\mathfrak{G}$ *to* $\mathfrak{G}'$ if for all $g_1, g_2 \in G$, $h(g_1 \cdot g_2) = h(g_1) \cdot' h(g_2)$. (It is easy to see that any group homomorphism maps the identity element of $\mathfrak{G}$ to the identity element of $\mathfrak{G}'$ and preserves all inverses as well. Also, the set $\{h(g) \mid g \in G\}$ is closed under $\cdot'$ and so it is a group, called the *image of* $\mathfrak{G}$ *under* $h$.) $\mathfrak{G}$ is a *subgroup of* $\mathfrak{G}'$ if $G \subseteq G'$ and the identity map $\mathsf{id}_G$ is a group homomorphism. Given $X \subseteq G$, the *subgroup of* $\mathfrak{G}$ *generated by* $X$ is the smallest subgroup of $\mathfrak{G}$ containing all elements from $X$. If $\mathfrak{G}$ is finite then every element of the subgroup generated by $X$ can be expressed as a combination (under $\cdot$) of elements of $X$.

Given a finite group $\mathfrak{G}$ with identity element $e$, the *order* $o_{\mathfrak{G}}(g)$ of an element $g$ in $\mathfrak{G}$ is the smallest positive number $n$ such that $g^n = e$. It is easy to see that $o_{\mathfrak{G}}(g)$ exists, and for any $k$, if $g^k = e$ then $o_{\mathfrak{G}}(g)$ divides $k$. Also, $o_{\mathfrak{G}}(g) = o_{\mathfrak{G}}(g^-)$ holds for every $g$. Also

$$\text{if } g \text{ is a nonidentity element in a group } \mathfrak{G}, \text{ then } g^k \neq g^{k+1} \text{ for any } k. \tag{6}$$

Given two semigroups $\mathfrak{S} = (S, \cdot)$, $\mathfrak{S}' = (S', \cdot')$, we say that $\mathfrak{S}'$ is a *subsemigroup* of $\mathfrak{S}$ if $S' \subseteq S$ and $\cdot'$ is the restriction of $\cdot$ to $S'$. Given a monoid $\boldsymbol{M} = (M, \cdot)$ and a set $S \subseteq M$, we say that $S$ *contains the group* $\mathfrak{G} = (G, \cdot')$, if $G \subseteq S$ and $\mathfrak{G}$ is a subsemigroup of $\boldsymbol{M}$. (We do **not** require that the identity element of $\boldsymbol{M}$ is in $\mathfrak{G}$, even if it is in $S$.) If $S = M$ then we also say that $\boldsymbol{M}$ *contains the group* $\mathfrak{G}$, or $\mathfrak{G}$ *is in* $\boldsymbol{M}$. We call a monoid $\boldsymbol{M}$ *aperiodic* if it does not contain any nontrivial groups.

Suppose $\mathfrak{S} = (S, \cdot)$ is a finite semigroup, and take any $s \in S$. Then, by the pigeonhole principle, there exist $i, j \geq 1$ such that $i + j \leq |S| + 1$ and $s^i = s^{i+j}$. Take the minimal such numbers, that is, let $i_s, j_s \geq 1$ be such that $i_s + j_s \leq |S| + 1$ and $s^{i_s} = s^{i_s+j_s}$ but $s^{i_s}, s^{i_s+1}, \dots, s^{i_s+j_s-1}$ are all different. Then clearly $\mathfrak{G}_s = (G_s, \cdot)$, where $G_s = \{s^{i_s}, s^{i_s+1}, \dots, s^{i_s+j_s-1}\}$, is a subsemigroup of $\mathfrak{S}$. It is easy to see that there is some $m \geq 1$ such that $i_s \leq m \cdot j_s < i_s + j_s \leq |S| + 1$, and so $s^{m \cdot j_s}$ is idempotent. Thus, for every element $s$ in a semigroup $\mathfrak{S}$, we have the following:

$$\text{there is } n \geq 1 \text{ such that } s^n \text{ is idempotent;} \tag{7}$$

$$\mathfrak{G}_s \text{ is a group in } \mathfrak{S} \text{ (isomorphic to the cyclic group } \mathbb{Z}_{j_s}); \tag{8}$$

$$\mathfrak{G}_s \text{ is nontrivial iff } s^n \neq s^{n+1} \text{ for any } n. \tag{9}$$

One can apply these to a particular setting. Let $\delta$ be a $Q \to Q$ function for some nonempty finite set $Q$. For any $p \in Q$, the subset $\{\delta^k(p) \mid k < \omega\}$ with the obvious multiplication is a finite semigroup, and so we have:

$$\text{For every } p \in Q \text{ there is } n_p \geq 1 \text{ such that } \delta^{n_p}\big(\delta^{n_p}(p)\big) = \delta^{n_p}(p). \tag{10}$$

$$\text{There exist } q \in Q \text{ and } n \geq 1 \text{ such that } q = \delta^n(q). \tag{11}$$

$$\text{For every } q \in Q, \text{ if } q = \delta^k(q) \text{ for some } k \geq 1, \text{ then there is } 1 \leq n \leq |Q| \text{ with } q = \delta^n(q). \tag{12}$$

We will also consider *solvable* groups and not solvable (*unsolvable*) groups, see [46] for a definition. We will only use the following facts about them:

- Any homomorphic image of a solvable group is solvable.
- The criterion of Kaplan and Levy [36] (generalising Thompson's [52, Cor.3]): A finite group $\mathfrak{G}$ is unsolvable iff it contains three elements $a, b, c$, such that $o_{\mathfrak{G}}(a) = 2$, $o_{\mathfrak{G}}(b)$ is an odd prime, $o_{\mathfrak{G}}(c) > 1$ and coprime to both 2 and $o_{\mathfrak{G}}(b)$, and $abc$ is the identity element of $\mathfrak{G}$.

A one-to-one and onto function on a finite set $S$ is called a *permutation on* $S$. The *order of a permutation* $\delta$ is its order in the group of all permutations on $S$ (whose operation is composition, and its identity element is the identity permutation $\mathsf{id}_S$). We will use the usual cycle notation for permutations.

Now suppose that $\mathfrak{G}$ is a monoid of $Q \to Q$ functions for some nonempty finite set $Q$. Let $S = \{q \in Q \mid e_{\mathfrak{G}}(q) = q\}$, where $e_{\mathfrak{G}}$ the identity element in $\mathfrak{G}$. For every function $\delta$ in $\mathfrak{G}$, let $\delta\restriction_S$ denote the restriction of $\delta$ to $S$. Then we have the following:

$$\mathfrak{G} \text{ is a group iff } \delta\restriction_S \text{ is a permutation on } S, \text{ for every } \delta \text{ in } \mathfrak{G}. \tag{13}$$

$$\text{If } \mathfrak{G} \text{ is a group and } \delta \text{ is a nonindentity element in it, then } \delta\restriction_S \neq \mathsf{id}_S, \text{ and the order of the permutation } \delta\restriction_S \text{ divides } o_{\mathfrak{G}}(\delta). \tag{14}$$

## 3.2 Automata: DFAs, NFAs, 2NFAs

A *two-way nondeterministic finite automaton* is a quintuple $\mathfrak{A} = (Q, \Sigma, \delta, Q_0, F)$ that consists of an alphabet $\Sigma$, a finite set of states $Q$ with a subset $Q_0 \neq \emptyset$ of initial states and a subset $F$ of accepting states, and a transition function $\delta\colon Q \times \Sigma \to 2^{Q\times\{-1,0,1\}}$ indicating the next state and whether the head should move left $(-1)$, right $(1)$, or stay put $(0)$. If $Q_0 = \{q_0\}$ and $|\delta(q, a)| = 1$, for all $q \in Q$ and $a \in \Sigma$, then $\mathfrak{A}$ is *deterministic*, in which case we write $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$. If $\delta(q, a) \subseteq Q \times \{1\}$, for all $q \in Q$ and $a \in \Sigma$, then $\mathfrak{A}$ is a *one-way* automaton, and we write $\delta\colon Q \times \Sigma \to 2^Q$. As usual, DFA and NFA refer to one-way deterministic and non-deterministic finite automata, respectively, while 2DFA and 2NFA to the corresponding two-way automata. Given a 2NFA $\mathfrak{A}$, we write $q \to_{a,d} q'$ if $(q', d) \in \delta(q, a)$; given an NFA $\mathfrak{A}$, we write $q \to_a q'$ if $q' \in \delta(q, a)$. A *run* of a 2NFA $\mathfrak{A}$ is a word in $(Q \times \mathbb{N})^*$. A run $(q_0, i_0), \dots, (q_m, i_m)$ is a *run of* $\mathfrak{A}$ *on a word* $w = a_0 \dots a_n \in \Sigma^*$ if $q_0 \in Q_0$, $i_0 = 0$ and there exist $d_0, \dots, d_{m-1} \in \{-1, 0, 1\}$ such that $q_j \to_{a_j,d_j} q_{j+1}$ and $i_{j+1} = i_j + d_j$ for all $j$, $0 \leq j < m$. The run is *accepting* if $q_m \in F$, $i_m = n + 1$. $\mathfrak{A}$ *accepts* $w \in \Sigma^*$ if there is an accepting run of $\mathfrak{A}$ on $w$; the language $\boldsymbol{L}(\mathfrak{A})$ of $\mathfrak{A}$ is the set of all words accepted by $\mathfrak{A}$.

Given an NFA $\mathfrak{A}$, states $q, q' \in Q$, and $w = a_0 \dots a_n \in \Sigma^*$, we write $q \to_w q'$ if either $w = \varepsilon$ and $q' = q$ or there is a run of $\mathfrak{A}$ on $w$ that starts with $(q_0, 0)$ and ends with $(q', n+1)$. We say that a state $q \in Q$ is *reachable* if $q' \to_w q$, for some $q' \in Q_0$ and $w \in \Sigma^*$.

Given a DFA $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$, for any word $w \in \Sigma^*$, we define a function $\delta_w\colon Q \to Q$ by taking $\delta_w(q) = q'$ iff $q \to_w q'$. We define an equivalence relation $\sim$ on the set $Q^r \subseteq Q$ of reachable states by taking $q \sim q'$ iff for every $w \in \Sigma^*$ we have $\delta_w(q) \in F$ iff $\delta_w(q') \in F$. We denote the $\sim$-class of $q$ by $q/_\sim$, and let $X/_\sim = \{q/_\sim \mid q \in X\}$ for any $X \subseteq Q^r$. Define $\tilde{\delta}_w\colon Q^r/_\sim \to Q^r/_\sim$ by taking $\tilde{\delta}_w(q/_\sim) = \delta_w(q)/_\sim$. Then $\big(Q^r/_\sim, \Sigma, \tilde{\delta}, q_0/_\sim, (F \cap Q^r)/_\sim\big)$ is the minimal DFA whose language coincides with the language of $\mathfrak{A}$. Given a regular language $\boldsymbol{L}$, we denote by $\mathfrak{A}_{\boldsymbol{L}}$ the minimal DFA whose language is $\boldsymbol{L}$.

The *transition monoid of* a DFA $\mathfrak{A}$ takes the form $M(\mathfrak{A}) = \big(\{\delta_w \mid w \in \Sigma^*\}, \cdot\big)$, where $\cdot$ is the composition $\circ$ of functions, that is, $\delta_v \cdot \delta_w = \delta_w \circ \delta_v = \delta_{vw}$, for any $v, w$. The *syntactic monoid* $M(\boldsymbol{L})$ *of* $\boldsymbol{L}$ is the transition monoid $M(\mathfrak{A}_{\boldsymbol{L}})$ of $\mathfrak{A}_{\boldsymbol{L}}$. The map $\eta_{\boldsymbol{L}}$ from $\Sigma^*$ to the domain of $M(\boldsymbol{L})$ defined by taking $\eta_{\boldsymbol{L}}(w) = \tilde{\delta}_w$ is called the *syntactic morphism of* $\boldsymbol{L}$. Given a set $W \subseteq \Sigma^*$, we set $\eta_{\boldsymbol{L}}(W) = \{\eta_{\boldsymbol{L}}(w) \mid w \in W\}$. We call $\eta_{\boldsymbol{L}}$ *quasi-aperiodic* if $\eta_L(\Sigma^t)$ is aperiodic for every $t < \omega$.

A language $\boldsymbol{L}$ over $\Sigma$ is $\mathcal{L}$*-definable* if there is an $\mathcal{L}$-sentence $\varphi$ in the signature $\Sigma$, whose symbols are treated as unary predicates, such that, for any $w \in \Sigma^*$, we have $w = a_0 \dots a_n \in \boldsymbol{L}$ iff $\mathfrak{S}_w \models \varphi$, where $\mathfrak{S}_w$ is an FO-structure with domain $\{0, \dots, n\}$ ordered by $<$, in which $\mathfrak{S}_w \models a(i)$ iff $a = a_i$, for $0 \leq i \leq n$.

Table 3 summarises the known results that connect definability of a regular language $\boldsymbol{L}$ with properties of the syntactic monoid $M(\boldsymbol{L})$ and syntactic morphism $\eta_{\boldsymbol{L}}$ (see [10] for details) and with its circuit complexity under a reasonable binary encoding of $\boldsymbol{L}$'s alphabet (see, e.g., [14, Lemma 2.1]) and the assumption that $\mathrm{ACC}^0 \neq \mathrm{NC}^1$. We also remind the reader that a regular language is $\mathsf{FO}(<)$-definable iff it is star-free (see [51] and references therein) and that $\mathrm{AC}^0 \subsetneqq \mathrm{ACC}^0 \subseteq \mathrm{NC}^1$ (see, e.g., [34, 51]).

From now on, we assume that $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<, \equiv), \mathsf{FO}(<, \mathsf{MOD})\}$.

We conclude the preliminaries by proving algebraic criteria of $\mathcal{L}$-definability of regular languages that are used in what follows.

| definability of $\boldsymbol{L}$ | algebraic characterisation of $\boldsymbol{L}$ | circuit complexity |
|---|---|---|
| $\mathsf{FO}(<)$ | $M(\boldsymbol{L})$ is aperiodic | in $\mathrm{AC}^0$ |
| $\mathsf{FO}(<,\equiv)$ | $\eta_{\boldsymbol{L}}$ is quasi-aperiodic | |
| $\mathsf{FO}(<,\mathsf{MOD})$ | all groups in $M(\boldsymbol{L})$ are solvable | in $\mathrm{ACC}^0$ |
| $\mathsf{FO}(\mathsf{RPR})$ | arbitrary $M(\boldsymbol{L})$ | in $\mathrm{NC}^1$ |
| not in $\mathsf{FO}(<,\mathsf{MOD})$ | $M(\boldsymbol{L})$ contains an unsolvable group | $\mathrm{NC}^1$-hard |

**Table 3** Definability, algebraic characterisations, and circuit complexity of regular languages.

## 3.3 Criteria of $\mathcal{L}$-definability

Our aim now is to prove Theorem 6 below. Note that the equivalence $(i)$, which follows from [47], was used to show that deciding $\mathsf{FO}(<)$-definability is in PSPACE [49]. Criteria $(ii)$ and $(iii)$ appear to be new.

▶ **Theorem 6.** *For any DFA $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$, the following criteria hold:*

$(i)$ [47, 49] *$\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<)$-definable iff $\mathfrak{A}$ contains a nontrivial cycle, that is, there exist a word $u \in \Sigma^*$, a state $q \in Q^r$, and a number $k \leq |Q|$ such that $q \not\sim \delta_u(q)$ and $q = \delta_{u^k}(q)$;*

$(ii)$ *$\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<,\equiv)$-definable iff there exist words $u, v \in \Sigma^*$, a state $q \in Q^r$, and a number $k \leq |Q|$ such that $q \not\sim \delta_u(q)$, $q = \delta_{u^k}(q)$, $|v| = |u|$, and $\delta_{u^i}(q) = \delta_{u^i v}(q)$, for every $i < k$;*

$(iii)$ *$\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<,\mathsf{MOD})$-definable iff there exist words $u, v \in \Sigma^*$, a state $q \in Q^r$ and numbers $k, l \leq |Q|$ such that $k$ is an odd prime, $l > 1$ and coprime to both $2$ and $k$, $q \not\sim \delta_u(q)$, $q \not\sim \delta_v(q)$, $q \not\sim \delta_{uv}(q)$, and $\delta_x(q) \sim \delta_{xu^2}(q) \sim \delta_{xv^k}(q) \sim \delta_{x(uv)^l}(q)$, for all $x \in \{u, v\}^*$.*

**Proof.** Throughout, we consider the minimal DFA $\mathfrak{A}_{\boldsymbol{L}(\mathfrak{A})}$, with transition function $\tilde{\delta}$.

$(i)(\Rightarrow)$: Suppose that $\mathfrak{G}$ is a nontrivial group in $M(\mathfrak{A}_{\boldsymbol{L}(\mathfrak{A})})$. Let $u \in \Sigma^*$ be such that $\tilde{\delta}_u$ is a nonidentity element in $\mathfrak{G}$. We claim that there is $p \in Q^r$ such that $\tilde{\delta}_{u^n}(p/_\sim) \neq \tilde{\delta}_{u^{n+1}}(p/_\sim)$ for any $n > 0$. Indeed, otherwise for every $p \in Q^r$ there is $n_p > 0$ with $\tilde{\delta}_{u^{n_p}}(p/_\sim) = \tilde{\delta}_{u^{n_p+1}}(p/_\sim)$. Let $n = \max\{n_p \mid p \in Q^r\}$. Then $\tilde{\delta}_{u^n} = \tilde{\delta}_{u^{n+1}}$, contradicting (6).

By (10), there is $m \geq 1$ with $\tilde{\delta}_{u^{2m}}(p/_\sim) = \tilde{\delta}_{u^m}(p/_\sim)$. Let $s/_\sim = \tilde{\delta}_{u^m}(p/_\sim)$. Then $s/_\sim = \tilde{\delta}_{u^m}(s/_\sim)$, and so the restriction of $\delta_{u^m}$ to the subset $s/_\sim$ of $Q^r$ is an $s/_\sim \to s/_\sim$ function. By (11), there exist $q \in s/_\sim$ and $n \geq 1$ such that $(\delta_{u^m})^n(q) = q$. Thus, $\delta_{u^{mn}}(q) = q$, and so by (12), there is $k \leq |Q|$ with $\delta_{u^k}(q) = q$. As $s/_\sim \neq \tilde{\delta}_u(s/_\sim)$, we also have $q \not\sim \delta_u(q)$, as required.

$(i)(\Leftarrow)$: Suppose the condition holds for $\mathfrak{A}$. Then there exists $u \in \Sigma^*$, $q \in Q^r/_\sim$, and $k < \omega$ are such that $q \neq \tilde{\delta}_u(q)$ and $q = \tilde{\delta}_{u^k}(q)$. Then $\tilde{\delta}_{u^n} \neq \tilde{\delta}_{u^{n+1}}$ for any $n > 0$. Indeed, otherwise we have some $n > 0$ with $\tilde{\delta}_{u^n}(q) = \tilde{\delta}_{u^{n+1}}(q)$. Let $i, j$ be such that $n = i \cdot k + j$ and $j < k$. Then

$$q = \tilde{\delta}_{u^k}(q) = \tilde{\delta}_{u^{(i+1)k}}(q) = \tilde{\delta}_{u^n u^{k-j}}(q) = \tilde{\delta}_{u^{n+1} u^{k-j}}(q) = \tilde{\delta}_{u^{(i+1)k} u}(q) = \tilde{\delta}_u(q).$$

So by (8) and (9), the group $\mathfrak{G}_{\tilde{\delta}_u}$ is a nontrivial group in $M(\boldsymbol{L})$.

$(ii)(\Rightarrow)$: Suppose that $\mathfrak{G}$ is a nontrivial group in $\eta_{\boldsymbol{L}}(\Sigma^t)$ for some $t < \omega$. Let $u \in \Sigma^t$ be such that $\tilde{\delta}_u$ is a nonidentity element in $\mathfrak{G}$. As is shown in the proof of the $\Rightarrow$ direction of $(i)$, there exist $s \in Q^r$ and $m \geq 1$ such that $s/_\sim \neq \tilde{\delta}_u(s/_\sim)$ and $s/_\sim = \tilde{\delta}_{u^m}(s/_\sim)$. Now let $v \in \Sigma^t$ be such that $\tilde{\delta}_v$ is the identity element in $\mathfrak{G}$, and consider $\delta_v$. By (7), there is $\ell \geq 1$ such that $\delta_{v^\ell}$ is idempotent. Then $\delta_{v^{2\ell-1}v^{2\ell}} = \delta_{v^{2\ell-1}}$. Thus, if we let $\bar{u} = uv^{2\ell-1}$ and

$\bar{v} = v^{2\ell}$, then $|\bar{u}| = |\bar{v}|$ and $\delta_{\bar{u}^i} = \delta_{\bar{u}^i\bar{v}}$ for any $i < \omega$. Also, $\tilde{\delta}_{u^i} = \tilde{\delta}_{\bar{u}^i}$ for every $i \geq 1$, and so the restriction of $\delta_{\bar{u}^m}$ to $s/_\sim$ is an $s/_\sim \to s/_\sim$ function. By (11), there exist $q \in s/_\sim$ and $n \geq 1$ such that $(\delta_{\bar{u}^m})^n(q) = q$. Thus, $\delta_{\bar{u}^{mn}}(q) = q$, and so by (12), there is some $k \leq |Q|$ with $\delta_{\bar{u}^k}(q) = q$. As $s/_\sim \neq \tilde{\delta}_u(s/_\sim) = \tilde{\delta}_{\bar{u}}(s/_\sim)$, we also have $q \not\sim \delta_{\bar{u}}(q)$, as required.

*(ii)*($\Leftarrow$): Suppose the condition holds for $\mathfrak{A}$. Then there exist $u, v \in \Sigma^*$, $q \in Q^r/_\sim$, and $k < \omega$ are such that $q \neq \tilde{\delta}_u(q)$, $q = \tilde{\delta}_{u^k}(q)$, $|v| = |u|$, and $\tilde{\delta}_{u^i}(q) = \tilde{\delta}_{u^iv}(q)$, for every $i < k$. As $M(\mathfrak{A}_{\boldsymbol{L}(\mathfrak{A})})$ is finite, it has finitely many subsets. So there exists $i, j \geq 1$ such that $\eta_{\boldsymbol{L}}(\Sigma^{i|u|}) = \eta_{\boldsymbol{L}}(\Sigma^{(i+j)|u|})$. Let $z$ be a multiple of $j$ with $i \leq z < i + j$. Then $\eta_{\boldsymbol{L}}(\Sigma^{z|u|}) = \eta_{\boldsymbol{L}}(\Sigma^{(z|u|)^2})$, and so $\eta_{\boldsymbol{L}}(\Sigma^{z|u|})$ is closed under the composition of functions (that is, the semigroup operation of $M(\mathfrak{A}_{\boldsymbol{L}(\mathfrak{A})})$). Let $w = uv^{z-1}$ and consider the group $\mathfrak{G}_{\tilde{\delta}_w}$ (defined above (7)–(9)). Then $G_{\tilde{\delta}_w} \subseteq \eta_{\boldsymbol{L}}(\Sigma^{z|u|})$. We claim that $\mathfrak{G}_{\tilde{\delta}_w}$ is nontrivial. Indeed, on the one hand, $\tilde{\delta}_w(q) = \tilde{\delta}_{uv^{z-1}}(q) = \tilde{\delta}_u(q) \neq q$. On the other hand, $\tilde{\delta}_{w^k}(q) = \tilde{\delta}_{u^k}(q) = q$. As is shown in the proof of the $\Leftarrow$ direction of *(i)*, $\mathfrak{G}_{\tilde{\delta}_w}$ is nontrivial.

*(iii)*($\Rightarrow$): Suppose $\mathfrak{G}$ is an unsolvable group in $M(\mathfrak{A}_{\boldsymbol{L}(\mathfrak{A})})$. By the Kaplan–Levy criterion, $\mathfrak{G}$ contains three functions $a, b, c$, such that $o_{\mathfrak{G}}(a) = 2$, $o_{\mathfrak{G}}(b)$ is an odd prime, $o_{\mathfrak{G}}(c) > 1$ and coprime to both 2 and $o_{\mathfrak{G}}(b)$, and $c \circ b \circ a = e_{\mathfrak{G}}$ for the identity element $e_{\mathfrak{G}}$ of $\mathfrak{G}$. Let $u, v \in \Sigma^*$ be such that $a = \tilde{\delta}_u$, $b = \tilde{\delta}_v$ and $c = (\tilde{\delta}_{uv})^-$, and let $k = o_{\mathfrak{G}}(\tilde{\delta}_v)$ and $r = o_{\mathfrak{G}}(c) = o_{\mathfrak{G}}(\tilde{\delta}_{uv})$. Then $r > 1$ and coprime to both 2 and $k$. Let $S = \{p \in Q^r/_\sim \mid e_{\mathfrak{G}}(p) = p\}$. As $\tilde{\delta}_x$ is $\mathfrak{G}$ for every $x \in \{u, v\}^*$, we have $e_{\mathfrak{G}} \circ \tilde{\delta}_x = \tilde{\delta}_x$. Thus,

$$\tilde{\delta}_{xu^2}(q) = \tilde{\delta}_{u^2}\big(\tilde{\delta}_x(q)\big) = e_{\mathfrak{G}}\big(\tilde{\delta}_x(q)\big) = (e_{\mathfrak{G}} \circ \tilde{\delta}_x)(q) = \tilde{\delta}_x(q), \quad \text{and}$$
$$\tilde{\delta}_{xv^k}(q) = \tilde{\delta}_{v^k}\big(\tilde{\delta}_x(q)\big) = e_{\mathfrak{G}}\big(\tilde{\delta}_x(q)\big) = (e_{\mathfrak{G}} \circ \tilde{\delta}_x)(q) = \tilde{\delta}_x(q), \quad \text{for every } q \in S.$$

Then by (13),, each of $\tilde{\delta}_u\!\restriction_S$, $\tilde{\delta}_v\!\restriction_S$ and $\tilde{\delta}_{uv}\!\restriction_S$ is a permutation on $S$. By (14), the order of $\tilde{\delta}_u\!\restriction_S$ is 2, the order of $\tilde{\delta}_v\!\restriction_S$ is $k$, and the order $l$ of $\tilde{\delta}_{uv}\!\restriction_S$ is a $> 1$ divisor of $r$, and so it is coprime to both 2 and $k$. Also, we have $k, l \leq |S| \leq |Q|$. Further, for every $x$, if $q$ is in $S$ then $\tilde{\delta}_x(q) \in S$ as well. So we have

$$\tilde{\delta}_{x(uv)^l}(q) = \tilde{\delta}_{(uv)^l}\big(\tilde{\delta}_x(q)\big) = (\tilde{\delta}_{uv}\!\restriction_S)^l\big(\tilde{\delta}_x(q)\big) = \mathsf{id}_S\big(\tilde{\delta}_x(q)\big) = \tilde{\delta}_x(q), \quad \text{for every } q \in S.$$

It remains to show that there is some $q \in S$ such that $q \neq \tilde{\delta}_u(q)$, $q \neq \tilde{\delta}_u(q)$, and $q \neq \tilde{\delta}_{uv}(q)$. We will use that the length of any cycle in a permutation divides the order of the permutation. First, we show there is $q \in S$ with $q \neq \tilde{\delta}_u(q)$ and $q \neq \tilde{\delta}_u(q)$. Indeed, as $\tilde{\delta}_u\!\restriction_S \neq \mathsf{id}_S$, there is $q \in S$ such that $\tilde{\delta}_u(q) = q' \neq q$. As the order of $\tilde{\delta}_u\!\restriction_S$ is 2, $\tilde{\delta}_u(q') = q$. If both $\tilde{\delta}_v(q) = q$ and $\tilde{\delta}_v(q') = q'$ were the case, then $\tilde{\delta}_{uv}(q) = q'$ and $\tilde{\delta}_{uv}(q') = q$ would hold, and so $(qq')$ would be a cycle in $\tilde{\delta}_{uv}\!\restriction_S$, contradicting that $l$ is coprime to 2. So take some $q \in S$ such that $\tilde{\delta}_u(q) = q' \neq q$ and $\tilde{\delta}_v(q) \neq q$. If $\tilde{\delta}_v(q') \neq q$ then $\tilde{\delta}_{uv}(q) \neq q$, and so $q$ is a good choice. So suppose that $\tilde{\delta}_v(q') = q$, and let $q'' = \tilde{\delta}_v(q)$. Then $q'' \neq q'$, as $k$ is odd. Thus, $\tilde{\delta}_{uv}(q') \neq q'$, and so $q'$ is a good choice.

*(iii)*($\Leftarrow$): Suppose $u, v \in \Sigma^*$, $q \in Q^r$, and $k, l < \omega$ are satisfying the conditions. For every $x \in \{u, v\}^*$, we define an equivalence relation $\approx_x$ on $Q^r/_\sim$ by taking $p \approx_x p'$ iff $\tilde{\delta}_x(p) = \tilde{\delta}_x(p')$. Then we clearly have that $\approx_x \subseteq \approx_{xy}$, for all $x, y \in \{u, v\}^*$. As $Q$ is finite, there is $z \in \{u, v\}^*$ such that $\approx_z = \approx_{zy}$ for all $y \in \{u, v\}^*$. Take such a $z$. By (7), $\tilde{\delta}_z^n$ is idempotent for some $n \geq 1$. We let $w = z^n$. Then $\tilde{\delta}_w$ is idempotent and we also have that

$$\approx_w = \approx_{wy} \quad \text{for all } y \in \{u, v\}^*. \tag{15}$$

Now let $G_{\{u,v\}} = \{\tilde{\delta}_{wxw} \mid x \in \{u, v\}^*\}$. Then $G_{\{u,v\}}$ is closed under composition. Let $\mathfrak{G}_{\{u,v\}}$ be the subsemigroup of $M(\mathfrak{A}_{\boldsymbol{L}(\mathfrak{A})})$ with universe $G_{\{u,v\}}$. Then $\tilde{\delta}_w = \tilde{\delta}_{w\varepsilon w}$ is an identity element in $\mathfrak{G}_{\{u,v\}}$. Let $S = \{p \in Q^r/_\sim \mid \tilde{\delta}_w(p) = p\}$. We show that

$$\text{for every } \tilde{\delta} \text{ in } \mathfrak{G}_{\{u,v\}},\ \tilde{\delta}\restriction_S \text{ is a permutation on } S, \tag{16}$$

and so $\mathfrak{G}_{\{u,v\}}$ is a group by (13). Indeed, take some $x \in \{u,v\}^*$. As $\tilde{\delta}_w\big(\tilde{\delta}_{wxw}(p)\big) = \tilde{\delta}_{wxww}(p) = \tilde{\delta}_{wxw}(p)$ for any $p \in Q^r/_\sim$, $\tilde{\delta}_{wxw}\restriction_S$ is an $S \to S$ function. Also, if $p, p' \in S$ and $\tilde{\delta}_{wxw}(p) = \tilde{\delta}_{wxw}(p')$ then $p \approx_{wxw} p'$. Thus, by (15), $p \approx_w p'$, that is, $p = \tilde{\delta}_w(p) = \tilde{\delta}_w(p') = p'$, proving (16).

We show that the group $\mathfrak{G}_{\{u,v\}}$ is unsolvable by finding an unsolvable homomorphic image of it. To this end, let $R = \big\{p \in Q^r/_\sim \mid p = \tilde{\delta}_x(q) \text{ for some } x \in \{u,v\}^*\big\}$. We claim that for every $\tilde{\delta}$ in $\mathfrak{G}_{\{u,v\}}$, $\tilde{\delta}\restriction_R$ is a permutation on $R$, and so the function $h$ mapping every $\tilde{\delta}$ to $\tilde{\delta}\restriction_R$ is a group homomorphism from $\mathfrak{G}_{\{u,v\}}$ to the group of all permutations on $R$. Indeed, by (16), it is enough to show that $R \subseteq S$. To this end, we let $\overline{w} = \overline{z}_m \dots \overline{z}_1$, where $w = z_1 \dots z_m$ for some $z_i \in \{u,v\}$, $\overline{u} = u$ and $\overline{v} = v^{k-1}$. By using that $\tilde{\delta}_x(q) = \tilde{\delta}_{x(u)^2}(q) = \tilde{\delta}_{x(v)^k}(q)$ for all $x \in \{u,v\}^*$, we obtain that

$$\tilde{\delta}_{yw\overline{w}}(q) = \tilde{\delta}_{\overline{z}_{m-1}\dots\overline{z}_1}\big(\tilde{\delta}_{yz_1\dots z_m\overline{z}_m}(q)\big) = \tilde{\delta}_{\overline{z}_{m-1}\dots\overline{z}_1}\big(\tilde{\delta}_{yz_1\dots z_{m-1}}(q)\big) = \dots$$
$$\dots = \tilde{\delta}_{\overline{z}_1}\big(\tilde{\delta}_{yz_1}(q)\big) = \tilde{\delta}_{xz_1\overline{z}_1}(q) = \tilde{\delta}_y(q), \quad \text{for all } y \in \{u,v\}^*. \tag{17}$$

Now suppose that $p \in R$, that is, $p = \tilde{\delta}_x(q)$ for some $x \in \{u,v\}^*$. Then, by (17),

$$\tilde{\delta}_w(p) = \tilde{\delta}_w\big(\tilde{\delta}_x(q)\big) = \tilde{\delta}_{xw}(q) = \tilde{\delta}_{xww\overline{w}}(q) = \tilde{\delta}_{xw\overline{w}}(q) = \tilde{\delta}_x(q) = p,$$

and so $p \in S$, as required.

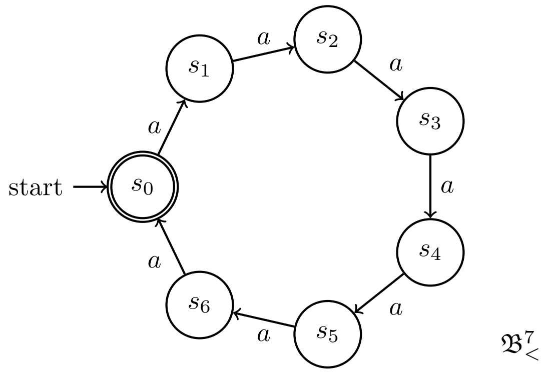
Now let $\mathfrak{G}$ be the image of $\mathfrak{G}_{\{u,v\}}$ under $h$. We prove that $\mathfrak{G}$ is unsolvable by finding three elements $a, b, c$ in it such that $o_{\mathfrak{G}}(a) = 2$, $o_{\mathfrak{G}}(b) = k$, $o_{\mathfrak{G}}(c)$ is coprime to both 2 and $o_{\mathfrak{G}}(b)$, and $c \circ b \circ a = \mathsf{id}_R$ (the identity element of $\mathfrak{G}$). So let $a = h(\tilde{\delta}_{wuw})$, $b = h(\tilde{\delta}_{wvw})$, and $c = h(\tilde{\delta}_{wuvw})^-$. Observe that for every $x \in \{u,v\}^*$, $h(\tilde{\delta}_{wxw}) = \tilde{\delta}_x\restriction_R$, and so $c \circ b \circ a = \mathsf{id}_R$. Also, for any $\tilde{\delta}_x(q) \in R$, $a^2\big(\tilde{\delta}_x(q)\big) = (\tilde{\delta}_u\restriction_R)^2\big(\tilde{\delta}_x(q)\big) = \tilde{\delta}_{xu^2}(q) = \tilde{\delta}_x(q)$ by our assumption, and so $a^2 = \mathsf{id}_R$. On the other hand, $q \in R$ as $\tilde{\delta}_\varepsilon(q) = q$, and $\mathsf{id}_R(q) = q \neq \tilde{\delta}_u(q)$ by our assumption, so $a \neq \mathsf{id}_R$. As $o_{\mathfrak{G}}(a)$ divides 2, $o_{\mathfrak{G}}(a) = 2$ follows. Similarly, we can show that $o_{\mathfrak{G}}(b) = k$ (using that $\tilde{\delta}_{xv^k}(q) = \tilde{\delta}_x(q)$ for every $x \in \{u,v\}^*$, and $u \neq \tilde{\delta}_v(q)$). Finally (using that $\tilde{\delta}_{x(uv)^l}(q) = \tilde{\delta}_x(q)$ for every $x \in \{u,v\}^*$, and $u \neq \tilde{\delta}_{uv}(q)$), we obtain that $h(\tilde{\delta}_{wuvw})^l = \mathsf{id}_R$ and $h(\tilde{\delta}_{wuvw}) \neq \mathsf{id}_R$. Therefore, it follows that $o_{\mathfrak{G}}(c) = o_{\mathfrak{G}}\big(h(\tilde{\delta}_{wuvw})^-\big) = o_{\mathfrak{G}}\big(h(\tilde{\delta}_{wuvw})\big) > 1$ and divides $l$, and so coprime to both 2 and $k$, as required. ❏

## 4 Deciding FO-definability of regular languages

We now settle the complexity of deciding $\mathcal{L}$-definability of the language of a given (minimal) DFA or 2NFA, for each $\mathcal{L}$ in question. Deciding $\mathsf{FO}(<)$-definability for the languages of DFAs and NFAs is known to be PSpace-complete [14, 21, 49]. For other FO-languages $\mathcal{L}$, the problem has been recorded as decidable in [10] but the exact complexity seems to remain open. We start with the lower bound.

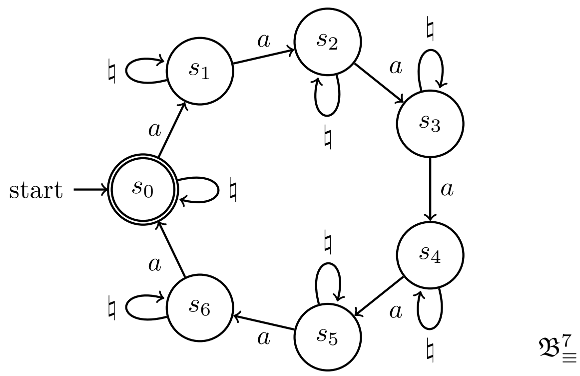
### 4.1 PSpace-hardness

We require three families of DFAs $\mathfrak{B}^p_<$, $\mathfrak{B}^p_\equiv$ and $\mathfrak{B}^p_{\mathsf{MOD}}$, where $p > 5$ is a prime number with $p \not\equiv \pm 1 \pmod{10}$. The first one, shown below for $p = 7$,



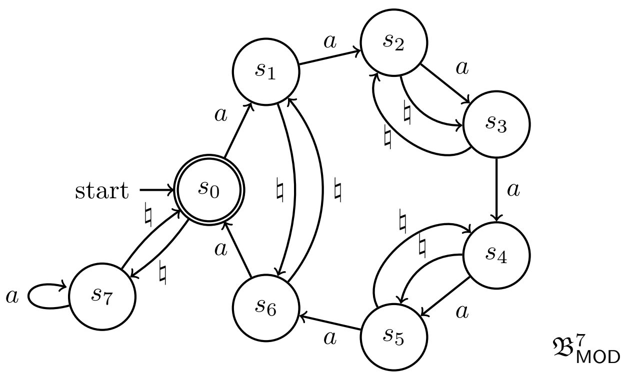

is defined in general as $\mathfrak{B}^p_< = \big(\{s_i \mid i < p\}, \{a\}, \delta^{\mathfrak{B}^p_<}, s_0, \{s_0\}\big)$, where $\delta_a^{\mathfrak{B}^p_<}(s_i) = s_j$ whenever $i, j < p$ and $j \equiv i + 1 \pmod p$. It is straightforward to check that the language $\boldsymbol{L}(\mathfrak{B}^p_<)$ consists of all words of the form $(a^p)^*$, $\mathfrak{B}^p_<$ is the minimal DFA for this language, and the syntactic monoid $M(\mathfrak{B}^p_<)$ is the cyclic group of order $p$ (generated by the permutation $\delta_a^{\mathfrak{B}^p_<}$).

The second family of DFAs, shown below for $p = 7$,




is defined in general as $\mathfrak{B}^p_\equiv = \big(\{s_i \mid i < p\}, \{a, \natural\}, \delta^{\mathfrak{B}^p_\equiv}, s_0, \{s_0\}\big)$, where $\delta_\natural^{\mathfrak{B}^p_\equiv}(s_i) = s_i$ and $\delta_a^{\mathfrak{B}^p_\equiv}(s_i) = s_j$ whenever $i, j < p$ and $j \equiv i + 1 \pmod p$. It is straightforward to check that the language $\boldsymbol{L}(\mathfrak{B}^p_\equiv)$ consists of all words of $a$'s and $\natural$'s whose number of $a$'s is divisible by $p$, $\mathfrak{B}^p_\equiv$ is the minimal DFA for this language, and the syntactic monoid $M(\mathfrak{B}^p_\equiv)$ is also the cyclic group of order $p$ (generated by the permutation $\delta_a^{\mathfrak{B}^p_\equiv}$).

Finally, the DFAs in the third family, depicted below for $p = 7$,




is defined in general as $\mathfrak{B}^p_{\mathsf{MOD}} = \big(\{s_i \mid i \le p\}, \{a, \natural\}, \delta^{\mathfrak{B}^p_{\mathsf{MOD}}}, s_0, \{s_0\}\big)$, where

– $\delta_a^{\mathfrak{B}^p_{\mathsf{MOD}}}(s_p) = s_p$, and $\delta_a^{\mathfrak{B}^p_{\mathsf{MOD}}}(s_i) = s_j$ whenever $i, j < p$ and $j \equiv i + 1 \pmod p$;
– $\delta_\natural^{\mathfrak{B}^p_{\mathsf{MOD}}}(s_0) = s_p$, $\delta_\natural^{\mathfrak{B}^p_{\mathsf{MOD}}}(s_p) = s_0$, and $\delta_\natural^{\mathfrak{B}^p_{\mathsf{MOD}}}(s_i) = s_j$ whenever $1 \leq i, j < p$ and $i \cdot j \equiv p - 1 \pmod p$, that is, $j = -1/i$ in the finite field $\mathbb{F}_p$.

It is straightforward to check that $\mathfrak{B}^p_{\mathsf{MOD}}$ is the minimal DFA for its language, and the syntactic monoid $M(\mathfrak{B}^p_{\mathsf{MOD}})$ is the permutation group generated by the permutations $\delta_a^{\mathfrak{B}^p_{\mathsf{MOD}}}$ and $\delta_\natural^{\mathfrak{B}^p_{\mathsf{MOD}}}$.

▶ **Lemma 7.** *For any prime $p > 5$ with $p \not\equiv \pm 1 \pmod{10}$, the group $M(\mathfrak{B}^p_{\mathsf{MOD}})$ is unsolvable, but all of its proper subgroups are solvable.*

**Proof.** It is straightforward to check that the order of the permutation $\delta_\natural^{\mathfrak{B}^p_{\mathsf{MOD}}}$ is 2, the order of $\delta_a^{\mathfrak{B}^p_{\mathsf{MOD}}}$ is $p$, while the order of the inverse of $\delta_{\natural a}^{\mathfrak{B}^p_{\mathsf{MOD}}}$ is the same as the order of $\delta_{\natural a}^{\mathfrak{B}^p_{\mathsf{MOD}}}$, which is 3. So $M(\mathfrak{B}^p_{\mathsf{MOD}})$ is unsolvable, for any prime $p$, by the Kaplan–Levy criterion. In order to show that all proper subgroups of $M(\mathfrak{B}^p_{\mathsf{MOD}})$ are solvable, we show that $M(\mathfrak{B}^p_{\mathsf{MOD}})$ is a subgroup of the *projective special linear group* $\mathrm{PSL}_2(p)$. If $p$ is a prime with $p > 5$ and $p \not\equiv \pm 1 \pmod{10}$, then all proper subgroups of $\mathrm{PSL}_2(p)$ are solvable; see, e.g., [37, Theorem 2.1]. (So $M(\mathfrak{B}^p_{\mathsf{MOD}})$ is in fact isomorphic to the unsolvable group $\mathrm{PSL}_2(p)$.)

Consider the set $P = \{0, 1, \dots, p-1, \infty\}$ of all points of the projective line over the field $\mathbb{F}_p$. By identifying $s_i$ with $i$ for $i < p$, and $s_p$ with $\infty$, we may regard the elements of $M(\mathfrak{B}^p_{\mathsf{MOD}})$ as $P \to P$ functions. The group $\mathrm{PSL}_2(p)$ consists of all $P \to P$ functions of the form

$$i \mapsto \frac{w \cdot i + x}{y \cdot i + z}, \quad \text{where } w \cdot z - x \cdot y = 1, \text{ with the field arithmetic of } \mathbb{F}_p \text{ being extended}$$
$$\text{by, for any } i \in P,\ i + \infty = \infty,\ 0 \cdot \infty = 1 \text{ and } i \cdot \infty = \infty \text{ for } i \neq 0.$$

Then it is easy to check that the two generators of $M(\mathfrak{B}^p_{\mathsf{MOD}})$ are in $\mathrm{PSL}_2(p)$: take $w = 1$, $x = 1$, $y = 0$, $z = 1$ for $\delta_a^{\mathfrak{B}^p_{\mathsf{MOD}}}$, and $w = 0$, $x = 1$, $y = p - 1$, $z = 0$ for $\delta_\natural^{\mathfrak{B}^p_{\mathsf{MOD}}}$. ❑

We are now in a position to establish the PSPACE-lower bound:

▶ **Theorem 8.** *For $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<, \equiv), \mathsf{FO}(<, \mathsf{MOD})\}$, deciding $\mathcal{L}$-definability of the language $\boldsymbol{L}(\mathfrak{A})$ of a given minimal DFA $\mathfrak{A}$ is PSPACE-hard.*

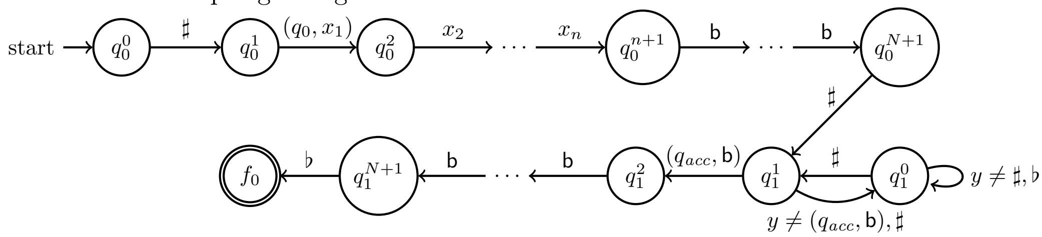
**Proof.** That deciding $\mathsf{FO}(<)$-definability of $\boldsymbol{L}(\mathfrak{A})$ is PSPACE-hard was established by Cho and Huynh [21]. We modify and generalise their construction to cover $\mathsf{FO}(<, \equiv)$- and $\mathsf{FO}(<, \mathsf{MOD})$-definability, too.

Suppose $\boldsymbol{M}$ is a deterministic Turing machine that decides a language using at most $N = P_{\boldsymbol{M}}(n)$ tape cells on any input of size $n$, for some polynomial $P_{\boldsymbol{M}}$. Given such an $\boldsymbol{M}$ and some input $\boldsymbol{x}$, our aim is to define three minimal DFAs whose languages are, respectively, $\mathsf{FO}(<)$-, $\mathsf{FO}(<, \equiv)$-, and $\mathsf{FO}(<, \mathsf{MOD})$-definable iff $\boldsymbol{M}$ rejects $\boldsymbol{x}$, and whose sizes are polynomial in $N$ and the size $|\boldsymbol{M}|$ of $\boldsymbol{M}$.

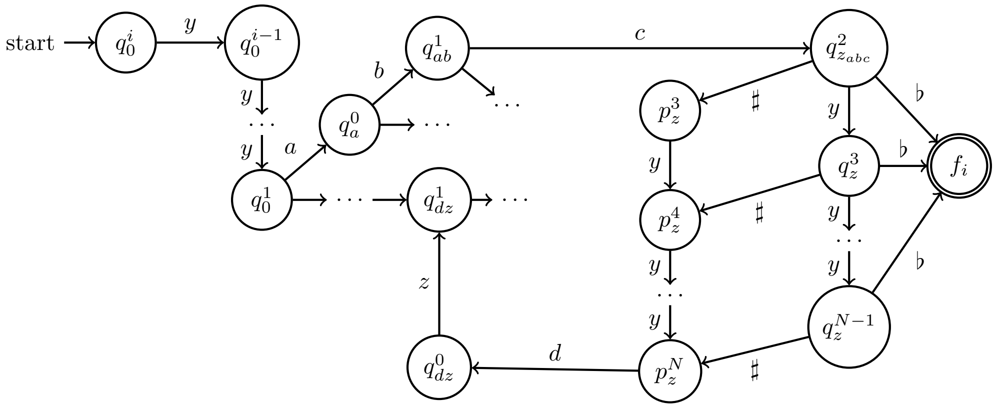
To this end, suppose that $\boldsymbol{M}$ is of the form $\boldsymbol{M} = (Q, \Gamma, \gamma, \mathsf{b}, q_0, q_{acc})$ with a set $Q$ of states, tape alphabet $\Gamma$ with $\mathsf{b}$ for blank, transition function $\gamma$, initial state $q_0$ and accepting state $q_{acc}$. Without loss of generality we assume that $\boldsymbol{M}$ erases the tape before accepting and has its head at the left-most cell in an accepting configuration, and if $\boldsymbol{M}$ does not accept the input, it runs forever. Given an input word $\boldsymbol{x} = x_1 \dots x_n$ over $\Gamma$, we represent configurations $\mathfrak{c}$ of the computation of $\boldsymbol{M}$ on $\boldsymbol{x}$ by the $N$-long word written on the tape (with sufficiently many blanks at the end) in which the symbol $y$ in the active cell is replaced by the pair $(q, y)$ for the current state $q$. The accepting computation of $\boldsymbol{M}$ on $\boldsymbol{x}$ is encoded by a word $\sharp\,\mathfrak{c}_1\,\sharp\,\mathfrak{c}_2\,\sharp\,\dots\,\sharp\,\mathfrak{c}_{k-1}\,\sharp\,\mathfrak{c}_k\flat$ over the alphabet $\Sigma = \Gamma \cup (Q \times \Gamma) \cup \{\sharp, \flat\}$, with $\mathfrak{c}_1, \mathfrak{c}_2, \dots, \mathfrak{c}_k$ being the subsequent configurations. In particular, $\mathfrak{c}_1$ is the initial configuration on $\boldsymbol{x}$ (so it is of the form $(q_0, x_1)x_2 \dots x_n\mathsf{b} \dots \mathsf{b}$), and $\mathfrak{c}_k$ is the accepting configuration (so it is of the form $(q_{acc}, \mathsf{b})\mathsf{b} \dots \mathsf{b}$). As usual for this representation of computations, we may regard $\gamma$ as a partial function from $\big(\Gamma \cup (Q \times \Gamma)\big)^3$ to $\Gamma \cup (Q \times \Gamma)$.

Let $p_{\boldsymbol{M},\boldsymbol{x}} = p$ be the first prime such that $p \geq N + 2$ and $p \not\equiv \pm 1 \pmod{10}$. By [13, Corollary 1.6], $p$ is polynomial in $N$. Our first aim is to construct a $p + 1$-long sequence $\mathfrak{A}_i$ of pairwise disjoint minimal DFAs over the alphabet $\Sigma$. Each $\mathfrak{A}_i$ has size polynomial in $N$ and $|\boldsymbol{M}|$, and it checks certain properties of an accepting computation on $\boldsymbol{x}$ such that $\boldsymbol{M}$ accepts $\boldsymbol{x}$ iff the intersection of the $\boldsymbol{L}(\mathfrak{A}_i)$ is not empty and consists of the single word encoding the accepting computation on $\boldsymbol{x}$.

The DFA $\mathfrak{A}_0$ checks that an input word starts with the initial configuration on $\boldsymbol{x}$ and ends with the accepting configuration:




When $1 \leq i \leq N$, the DFA $\mathfrak{A}_i$ checks, for all $j$, whether $\gamma(\sigma^j_{i-1}, \sigma^j_i, \sigma^j_{i+1}) = \sigma^{j+1}_i$, where $\sigma^k_l$ denotes the $l$th symbol of $\mathfrak{c}^k$.

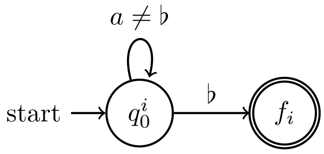



Formally $\delta_i$ consists of the following transitions for $a, b, c \in \Sigma' \setminus \{\flat\}$ and $b, c \neq \sharp$:

$$(q^j_0, b, q^{j-1}_0),\ (q^1_0, a, q^0_a),\ (q^0_a, b, q^1_{ab}),\ (q^1_{ab}, c, q^2_{z_{abc}}),\ (q^1_{ab}, \sharp, p^2_{z_{ab\sharp}}),$$
$$(q^j_a, b, q^{j+1}_a), \quad \text{for } a \neq \sharp \text{ and } 1 < j < N - 1,$$
$$(q^j_a, \sharp, p^{j+1}_a), \quad \text{for } a \neq \sharp \text{ and } 1 < j < N - 1,$$
$$(p^j_a, b, p^{j+1}_a), \quad \text{for } a \neq \sharp, \text{ and } 1 < j < N - 1,$$
$$(p^N_a, b, q^0_{ba}),\ (q^N_a, \sharp, q^0_{\sharp a}),\ (q^0_{ab}, b, q^1_{ab}),$$
$$(q^j_a, \flat, f_i), \quad 1 \leq j \leq N,$$
$$(q^1_{ab}, \flat, f_i).$$

Here, $z_{abc} = \gamma(a, b, c)$ for $a, b, c \in \Gamma \cup (Q \times \Gamma)$.

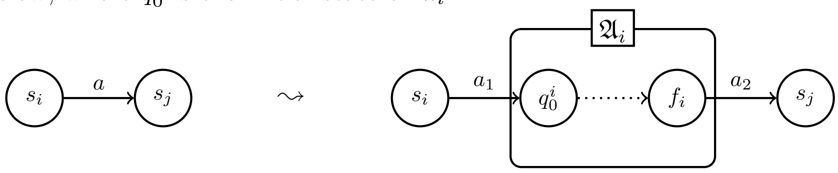
Finally, if $N + 1 \leq i \leq p$ then $\mathfrak{A}_i$ accepts all words with a single occurrence of $\flat$, which is the input's last character:




Note that $\mathfrak{A}_{p-1} = \mathfrak{A}_p$ as $p \geq N+2$. It is not hard to check that each of the $\mathfrak{A}_i$ is a minimal DFA that does not contain nontrivial cycles and the following holds:

▶ **Lemma 9.** *$\boldsymbol{M}$ accepts $\boldsymbol{x}$ iff $\bigcap_{i=0}^{p} \boldsymbol{L}(\mathfrak{A}_i) \neq \emptyset$, in which case this language consists of a single word that encodes the accepting computation of $\boldsymbol{M}$ on $\boldsymbol{x}$.*

Now take some fresh symbols $a_1, a_2$. We define three automata $\mathfrak{A}_<$, $\mathfrak{A}_\equiv$, $\mathfrak{A}_{\mathsf{MOD}}$ over the same tape alphabet $\Sigma_+ = \Sigma \cup \{a_1, a_2, \natural\}$ by taking, respectively, $\mathfrak{B}^p_<$, $\mathfrak{B}^p_\equiv$, $\mathfrak{B}^p_{\mathsf{MOD}}$ and replacing each transition $s_i \to_a s_j$ in them by a fresh copy of $\mathfrak{A}_i$, for $i \leq p$, as shown in the picture below, where $q_0^i$ is the initial state of $\mathfrak{A}_i$.




We make each of $\mathfrak{A}_<$, $\mathfrak{A}_\equiv$, $\mathfrak{A}_{\mathsf{MOD}}$ deterministic by adding a trash state *tr* looping on itself with every $y \in \Sigma_+$, and then adding the missing transitions leading to *tr*. It follows from the construction that $\mathfrak{A}_<$, $\mathfrak{A}_\equiv$, and $\mathfrak{A}_{\mathsf{MOD}}$ are minimal DFAs, and they are of size polynomial in $N$ and $|\boldsymbol{M}|$.

▶ **Lemma 10.** *(i) $\boldsymbol{L}(\mathfrak{A}_<)$ is $\mathsf{FO}(<)$-definable iff $\bigcap_{i=0}^{p} \boldsymbol{L}(\mathfrak{A}_i) = \emptyset$.*
*(ii) $\boldsymbol{L}(\mathfrak{A}_\equiv)$ is $\mathsf{FO}(<,\equiv)$-definable iff $\bigcap_{i=0}^{p} \boldsymbol{L}(\mathfrak{A}_i) = \emptyset$.*
*(iii) $\boldsymbol{L}(\mathfrak{A}_{\mathsf{MOD}})$ is $\mathsf{FO}(<,\mathsf{MOD})$-definable iff $\bigcap_{i=0}^{p} \boldsymbol{L}(\mathfrak{A}_i) = \emptyset$.*

**Proof.** In both directions we use that each of the DFAs $\mathfrak{A}_<, \mathfrak{A}_\equiv, \mathfrak{A}_{\mathsf{MOD}}$ is minimal, and so we can replace $\sim$ by $=$ in the conditions of Theorem 6. For the ($\Rightarrow$) directions, given some $w \in \bigcap_{i=0}^{p} \boldsymbol{L}(\mathfrak{A}_i)$, in each case we show how to satisfy the corresponding condition of Theorem 6:

*(i)*: Take $u = a_1 w a_2$, $q = s_0$, and $k = p$.
*(ii)*: Take $u = a_1 w a_2$, $v = \natural^{|u|}$, $q = s_0$, and $k = p$.
*(iii)*: Take $u = \natural$, $v = a_1 w a_2$, $q = s_0$, $k = p$ and $l = 3$.

For the ($\Leftarrow$) directions, in each case we show that the corresponding condition of Theorem 6 implies that $\bigcap_{i=0}^{p} \boldsymbol{L}(\mathfrak{A}_i)$ is not empty. To this end, we define a $\Sigma_+^* \to \{a, \natural\}^*$ homomorphism by taking $h(\natural) = \natural$, $h(a_1) = a$, and $h(b) = \varepsilon$ for all other $b \in \Sigma_+$.

*(i)* and *(ii)*: Let $\circ \in \{<, \equiv\}$ and suppose $q$ is a state in $\mathfrak{A}^p_\circ$ and $u' \in \Sigma_+^*$ such that $q \neq \delta_{u'}^{\mathfrak{A}^p_\circ}(q)$ and $q = \delta_{(u')^k}^{\mathfrak{A}^p_\circ}(q)$ for some $k$. Let $S = \{s_0, s_1, \dots, s_{p-1}\}$. We claim that there exist $s \in S$ and $u \in \Sigma_+^*$ such that

$$s \neq \delta_u^{\mathfrak{A}^p_\circ}(s), \tag{18}$$

$$\delta_x^{\mathfrak{A}^p_\circ}(s) \in S, \quad \text{for every } x \in \{u\}^*. \tag{19}$$

Indeed, observe that none of the states along the cyclic $q \to_{(u')^k} q$ path $\Pi$ in $\mathfrak{A}^p_\circ$ is *tr*. So there is some state along $\Pi$ that is in $S$, as otherwise one of the $\mathfrak{A}_i$ would contain a nontrivial cycle. Therefore, $u'$ must be of the form $w \natural^n a_1 w'$ for some $w \in \Sigma^*$, $n < \omega$ and $w' \in \Sigma_+^*$. It is easy to see that $s = \delta_{(u')^{k-1} w}^{\mathfrak{A}^p_\circ}(q)$ and $u = \natural^n a_1 w' w$ is as required in (18) and (19).

As $M(\mathfrak{B}^p_\circ)$ is a finite group, the set $\{\delta^{\mathfrak{B}^p_\circ}_{h(x)} \mid x \in \{u\}^*\}$ forms a subgroup $\mathfrak{G}$ in it (the subgroup generated by $\delta^{\mathfrak{B}^p_\circ}_{h(u)}$). We show that $\mathfrak{G}$ is nontrivial by finding a nontrivial homomorphic image of it. To this end, (19) implies that, for every $x \in \{u\}^*$, the restriction $\delta^{\mathfrak{A}^p_\circ}_x\restriction_{S'}$ of $\delta^{\mathfrak{A}^p_\circ}_x$ to the set $S' = \{\delta^{\mathfrak{A}^p_\circ}_y(s) \mid y \in \{u\}^*\}$ is an $S' \to S'$ function and $\delta^{\mathfrak{A}^p_\circ}_x\restriction_{S'} = \delta^{\mathfrak{B}^p_\circ}_{h(x)}\restriction_{S'}$. As $M(\mathfrak{B}^p_\circ)$ is a group of permutations on a set containing $S'$, $\delta^{\mathfrak{B}^p_\circ}_{h(x)}\restriction_{S'}$ is a permutation of $S'$, for every $x \in \{u\}^*$. Thus, $\{\delta^{\mathfrak{B}^p_\circ}_{h(x)}\restriction_{S'} \mid x \in \{u\}^*\}$ is a homomorphic image of $\mathfrak{G}$ that is nontrivial by (18).

Finally, as $\mathfrak{G}$ is a nontrivial subgroup of the cyclic group $M(\mathfrak{B}^p_\circ)$ of order $p$ and $p$ is a prime, it follows that $\mathfrak{G} = M(\mathfrak{B}^p_\circ)$. Therefore, there is $x \in \{u\}^*$ with $\delta^{\mathfrak{B}^p_\circ}_{h(x)} = \delta^{\mathfrak{B}^p_\circ}_a$ (a permutation containing the $p$-cycle $(s_0 s_1 \dots s_{p-1})$ 'around' all elements of $S$), and so $S' = S$ and $x = \natural^n a_1 w a_2 w'$ for some $n < \omega$, $w \in \Sigma^*$, and $w' \in \Sigma^*_+$. As $n = 0$ when $\circ =<$ and $\delta^{\mathfrak{A}^p_=}_{\natural^n}(s)$ for every $s \in S$, $S' = S$ implies that $w \in \bigcap_{i=0}^{p-1} \boldsymbol{L}(\mathfrak{A}_i) = \bigcap_{i=0}^{p} \boldsymbol{L}(\mathfrak{A}_i)$.

$(iii)$: Suppose $q$ is a state in $\mathfrak{A}^p_{\mathsf{MOD}}$ and $u', v' \in \Sigma^*_+$ such that $q \neq \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{u'}(q)$, $q \neq \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{v'}(q)$, $q \neq \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{u'v'}(q)$, and $\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_x(q) = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{x(u')^2}(q) = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{x(v')^k}(q) = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{x(u'v')^l}(q)$ for some odd prime $k$ and number $l$ that is coprime to both 2 and $k$. Let $S = \{s_0, s_1, \dots, s_p\}$. We claim that there exist $s \in S$ and $u, v \in \Sigma^*_+$ such that

$$s \neq \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_u(s),\ s \neq \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_v(s),\ s \neq \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{uv}(s), \tag{20}$$

$$\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_x(s) \in S, \quad \text{for every } x \in \{u, v\}^*, \tag{21}$$

$$\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_x(s) = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{xu^2}(s) = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{xv^k}(s) = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{x(uv)^l}(s), \quad \text{for every } x \in \{u, v\}^*. \tag{22}$$

Indeed, by an argument similar to the one in the proof of $(i)$ and $(ii)$ above, we must have $u' = w_u \natural^n a_1 w'_u$ and $v' = w_v \natural^m a_1 w'_v$ for some $w_u, w_v \in \Sigma^*$, $n, m < \omega$ and $w'_u, w'_v \in \Sigma^*_+$. For every $x \in \{u, v\}^*$, as both $\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{xw_u}(q)$ and $\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{xw_v}(q)$ are in $S$, they must be the same state. Using this it is not hard to see that $s = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_{u'w_u}(q)$, $u = \natural^n a_1 w'_u w_u$ and $v = \natural^m a_1 w'_v w_v$ are as required in (20)–(22).

As $M(\mathfrak{B}^p_{\mathsf{MOD}})$ is a finite group, the set $\{\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(x)} \mid x \in \{u, v\}^*\}$ forms a subgroup $\mathfrak{G}$ in it (the subgroup generated by $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(u)}$ and $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(v)}$). We show that $\mathfrak{G}$ is unsolvable by finding an unsolvable homomorphic image of it. To this end, we let $S' = \{\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_y(s) \mid y \in \{u, v\}^*\}$. Then (21) implies that $S' \subseteq S$ and

$$\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(x)}(s') = \delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_x(s') \in S', \quad \text{for all } s' \in S \text{ and } x \in \{u, v\}^*, \tag{23}$$

and so the restriction $\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_x\restriction_{S'}$ of $\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_x$ to $S'$ is an $S' \to S'$ function and $\delta^{\mathfrak{A}^p_{\mathsf{MOD}}}_x\restriction_{S'} = \delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(x)}\restriction_{S'}$. As $M(\mathfrak{B}^p_{\mathsf{MOD}})$ is a group of permutations on a set containing $S'$, $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(x)}\restriction_{S'}$ is a permutation of $S'$, for every $x \in \{u, v\}^*$. Thus, $\{\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(x)}\restriction_{S'} \mid x \in \{u, v\}^*\}$ is a homomorphic image of $\mathfrak{G}$ that is unsolvable by the Kaplan–Levy criterion: By (20), (22), and 2 and $k$ being primes, the order of the permutation $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(u)}\restriction_{S'}$ is 2, the order of $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(v)}\restriction_{S'}$ is $k$, and the order of $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(uv)}\restriction_{S'}$ (which is the same as the order of its inverse) is a $> 1$ divisor of $l$, and so coprime to both 2 and $k$.

As $\mathfrak{G}$ is an unsolvable subgroup of $M(\mathfrak{B}^p_{\mathsf{MOD}})$, it follows from Lemma 7 that $\mathfrak{G} = M(\mathfrak{B}^p_{\mathsf{MOD}})$, and so $\{u, v\}^* \not\subseteq \natural^*$. We claim that $S' = S$ also follows. Indeed, let $x \in \{u, v\}^*$ be such that $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(x)} = \delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_a$. As $|S'| \geq 2$ by (20), $s \in \{s_0, \dots, s_{p-1}\}$ must hold, and so

$\{s_0, \dots, s_{p-1}\} \subseteq S'$ follows by (23). As there is $y \in \{u, v\}^*$ with $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{h(y)} = \delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_{\natural}$, $s_p \in S'$ also follows by (23).

Finally, as $\{u, v\}^* \not\subseteq \natural^*$, there is $x \in \{u, v\}^*$ of the form $\natural^n a_1 w a_2 w'$ for some $n < \omega$, $w \in \Sigma$ and $w' \in \Sigma^*_+$. As $S' = S$, $\delta^{\mathfrak{B}^p_{\mathsf{MOD}}}_x(s_i) \in S$ for every $i \leq p$, and so $w \in \bigcap_{i=0}^p \boldsymbol{L}(\mathfrak{A}_i)$. ❑

As $\mathfrak{A}_<$, $\mathfrak{A}_\equiv$, and $\mathfrak{A}_{\mathsf{MOD}}$ are all of size polynomial in $N$ and $|\boldsymbol{M}|$, Theorem 8 clearly follows from Lemmas 9 and 10. ❑

## 4.2 Deciding $\mathcal{L}$-definability of 2NFAs in PSpace

In this section, we give a PSPACE-algorithm deciding whether the language of any given 2NFA is $\mathcal{L}$-definable, for $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<, \equiv), \mathsf{FO}(<, \mathsf{MOD})\}$, which matches the lower bound established in the previous section.

Let $\mathfrak{A} = (Q, \Sigma, \delta, Q_0, F)$ be a 2NFA. Following [20], for any $w \in \Sigma^+$, we introduce four binary relations $\mathsf{b}_{lr}(w)$, $\mathsf{b}_{rl}(w)$, $\mathsf{b}_{rr}(w)$, and $\mathsf{b}_{ll}(w)$ on $Q$ describing the *left-to-right*, *right-to-left*, *right-to-right*, and *left-to-left behaviour of* $\mathfrak{A}$ *on* $w$. Namely,

- $(q, q') \in \mathsf{b}_{lr}(w)$ if there is a run of $\mathfrak{A}$ on $w$ from $(q, 0)$ to $(q', |w|)$;
- $(q, q') \in \mathsf{b}_{rr}(w)$ if there is a run of $\mathfrak{A}$ on $w$ from $(q, |w| - 1)$ to $(q', |w|)$;
- $(q, q') \in \mathsf{b}_{rl}(w)$ if, for some $a \in \Sigma$, there is a run on $aw$ from $(q, |aw| - 1)$ to $(q', 0)$ such that no $(q'', 0)$ occurs in it before $(q', 0)$;
- $(q, q') \in \mathsf{b}_{ll}(w)$ if, for some $a \in \Sigma$, there is a run on $aw$ from $(q, 1)$ to $(q', 0)$ such that no $(q'', 0)$ occurs in it before $(q', 0)$.

For $w = \varepsilon$ (the empty word), we define the $\mathsf{b}_{ij}(w)$ as the identity relation on $Q$.

Let $\mathsf{b} = (\mathsf{b}_{lr}, \mathsf{b}_{rl}, \mathsf{b}_{rr}, \mathsf{b}_{ll})$, where the $\mathsf{b}_{ij}$ are the behaviours of $\mathfrak{A}$ on some $w \in \Sigma^*$, in which case we can also write $\mathsf{b}(w)$, and let $\mathsf{b}' = \mathsf{b}(w')$, for some $w' \in \Sigma^*$. We define the composition $\mathsf{b} \cdot \mathsf{b}' = \mathsf{b}''$ with components $\mathsf{b}''_{ij}$ as follows. Let $X$ be the transitive closure of $\mathsf{b}'_{ll} \circ \mathsf{b}_{rr}$, and let $Y$ be the transitive closure of $\mathsf{b}_{rr} \circ \mathsf{b}'_{ll}$. Then, we set:

$$\mathsf{b}''_{lr} = \mathsf{b}_{lr} \circ \mathsf{b}'_{lr} \cup \mathsf{b}_{lr} \circ X \circ \mathsf{b}'_{lr},$$
$$\mathsf{b}''_{rl} = \mathsf{b}'_{rl} \circ \mathsf{b}_{rl} \cup \mathsf{b}'_{rl} \circ Y \circ \mathsf{b}_{rl},$$
$$\mathsf{b}''_{rr} = \mathsf{b}'_{rr} \cup \mathsf{b}'_{rl} \circ Y \circ \mathsf{b}_{rr} \circ \mathsf{b}'_{lr},$$
$$\mathsf{b}''_{ll} = \mathsf{b}_{ll} \cup \mathsf{b}_{lr} \circ X \circ \mathsf{b}'_{ll} \circ \mathsf{b}_{rl}.$$

One can readily check that $\mathsf{b}'' = \mathsf{b}(ww')$.

We define the DFA $\mathfrak{A}' = (Q', \Sigma, \delta', q'_0, F')$ by taking

$$Q' = \big\{(B_{lr}, B_{rr}) \mid B_{lr} \subseteq Q_0 \times Q,\ B_{rr} \subseteq Q \times Q\big\},$$
$$q'_0 = \big(\{(q, q) \mid q \in Q_0\}, \emptyset\big),$$
$$F' = \big\{(B_{lr}, B_{rr}) \mid (q_0, q) \in B_{lr},\ \text{for some } q_0 \in Q_0 \text{ and } q \in F\big\},$$
$$\text{for any } a \in \Sigma,\ \delta'_a\big((B_{lr}, B_{rr})\big) = (B'_{lr}, B'_{rr}),\ \text{where } B'_{lr} = B_{lr} \circ X(a) \circ \mathsf{b}_{lr}(a),$$
$$B'_{rr} = B_{rr} \cup \mathsf{b}_{rl}(a) \circ Y(a) \circ \mathsf{b}_{lr}(a),\ \text{and } X(a) \text{ and } Y(a) \text{ are the}$$
$$\text{reflexive transitive closures of, respectively, } \mathsf{b}_{ll}(a) \circ B_{rr} \text{ and } B_{rr} \circ \mathsf{b}_{ll}(a).$$

It is not hard to see that

$$\text{for any } w \in \Sigma^*,\ \delta'_w\big((B_{lr}, B_{rr})\big) = (B'_{lr}, B'_{rr}) \text{ iff } B'_{lr} = B_{lr} \circ X(w) \circ \mathsf{b}_{lr}(w),$$
$$B'_{rr} = B_{rr} \cup \mathsf{b}_{rl}(w) \circ Y(w) \circ \mathsf{b}_{lr}(w),\ \text{where } X(w) \text{ and } Y(w) \text{ are the}$$
$$\text{reflexive transitive closures of, respectively, } \mathsf{b}_{ll}(w) \circ B_{rr} \text{ and } B_{rr} \circ \mathsf{b}_{ll}(w). \quad (24)$$

Also, it can be shown in a way similar to [48, 56] that

$$\boldsymbol{L}(\mathfrak{A}) = \boldsymbol{L}(\mathfrak{A}'). \tag{25}$$

▶ **Theorem 11.** *For $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<,\equiv), \mathsf{FO}(<,\mathsf{MOD})\}$, deciding $\mathcal{L}$-definability of the language $\boldsymbol{L}(\mathfrak{A})$ of any given 2NFA $\mathfrak{A}$ can be done in PSPACE.*

**Proof.** Let $\mathfrak{A}'$ be the DFA defined above for the given 2NFA $\mathfrak{A}$. First, we consider $\mathsf{FO}(<)$-definability. By Theorem 6 (*i*) and (25), $\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<)$-definable iff there exist a word $u \in \Sigma^*$, a reachable state $q \in Q'$, and a number $k \leq |Q'|$ such that $q \not\sim \delta'_u(q)$ and $q = \delta'_{u^k}(q)$. We guess the required $k$ in binary, $q$, and some quadruple of binary relations $\mathsf{b}(u)$ on $Q$. Clearly, they all can be stored in polynomial space in the size of $\mathfrak{A}$. To check that our guesses are correct, we first check that the quadruple $\mathsf{b}(u)$ indeed corresponds to some $u \in \Sigma^*$. This is done by guessing a sequence $\mathsf{b}_0, \dots, \mathsf{b}_n$ of pairwise distinct quadruples of binary relations on $Q$ such that $\mathsf{b}_0 = \mathsf{b}(u_0)$ and $\mathsf{b}_{i+1} = \mathsf{b}_i \cdot \mathsf{b}(u_{i+1})$, for some characters $u_0, \dots, u_n \in \Sigma$. (Any sequence with a subsequence starting after $\mathsf{b}_i$ and ending with $\mathsf{b}_{i+m}$, for some $i$ and $m$ such that $\mathsf{b}_i = \mathsf{b}_{i+m}$, is equivalent, in the context of this proof, to the sequence with such a subsequence removed.) Therefore, we can assume that $n \leq 2^{O(|Q|)}$, and so $n$ can be guessed in binary and stored in PSPACE. So, the stage of our algorithm that checks that $\mathsf{b}(u)$ corresponds to some $u \in \Sigma^*$ makes $n$ iterations and continues to the next stage if $\mathsf{b}_n = \mathsf{b}(u)$ or terminates with an answer $\mathsf{no}$ otherwise. Now, using $\mathsf{b}(u)$, we are able to compute $\mathsf{b}(u^k)$ by means of a sequence $\mathsf{b}_0, \dots, \mathsf{b}_k$, where $\mathsf{b}_0 = \mathsf{b}(u)$ and $\mathsf{b}_{i+1} = \mathsf{b}_i \cdot \mathsf{b}(u)$. With $\mathsf{b}(u)$ ($\mathsf{b}(u^k)$), we are able to compute $\delta'_u(q)$ (respectively, $\delta'_{u^k}(q)$) in PSPACE using (24). If $\delta'_{u^k}(q) \neq q$, the algorithm terminates with an answer $\mathsf{no}$. Otherwise, in the final stage of the algorithm, we check that $\delta'_u(q) \not\sim q$. This is done by guessing $v \in \Sigma^*$, such that $\delta'_v(q) = q_1$, $\delta'_v\big(\delta'_u(q)\big) = q_2$, and $q_1 \in F'$ iff $q_1 \notin F'$. We guess such a $v$ (if exists) in the form of $\mathsf{b}(v)$ using an algorithm analogous to that for guessing $u$ above.

We next consider $\mathsf{FO}(<,\equiv)$-definability. By Theorem 6 (*ii*) and (25), $\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<,\equiv)$-definable iff there there exist words $u, v \in \Sigma^*$, a reachable state $q \in Q'$, and a number $k \leq |Q'|$ such that $q \not\sim \delta'_u(q)$, $q = \delta'_{u^k}(q)$, $|v| = |u|$, and $\delta'_{u^i}(q) = \delta'_{u^i v}(q)$, for every $i < k$. We outline how to modify the algorithm for $\mathsf{FO}(<)$-definability above to check $\mathsf{FO}(<,\equiv)$-definability. First, we need to guess and check $v$ in the form of $\mathsf{b}(v)$ in parallel with guessing and checking $u$ in the form of $\mathsf{b}(u)$, making sure that $|v| = |u|$. For that, we guess a sequence of pairwise distinct pairs $(\mathsf{b}_0, \mathsf{b}'_0), \dots, (\mathsf{b}_n, \mathsf{b}'_n)$ such that the $\mathsf{b}_i$ are as above, $\mathsf{b}'_0 = \mathsf{b}(v_0)$ and $\mathsf{b}'_{i+1} = \mathsf{b}'_i \cdot \mathsf{b}(v_{i+1})$, for some $v_0, \dots, v_n \in \Sigma$. (Any such sequence of pairs with a subsequence starting after $(\mathsf{b}_i, \mathsf{b}'_i)$ and ending with $(\mathsf{b}_{i+m}, \mathsf{b}'_{i+m})$, for some $i$ and $m$ such that $(\mathsf{b}_i, \mathsf{b}'_i) = (\mathsf{b}_{i+m}, \mathsf{b}'_{i+m})$, is equivalent to the sequence with that subsequence removed.) So $n \leq 2^{O(|Q|)}$. For each $i < k$, we can then compute $\delta'_{u^i}(q)$ and $\delta'_{u^i v}(q)$, using (24), and check whether whether they are equal.

Finally, we consider the case of $\mathsf{FO}(<,\mathsf{MOD})$-definability. By Theorem 6 (*iii*) and (25), $\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<,\mathsf{MOD})$-definable iff there exist words $u, v \in \Sigma^*$, a reachable state $q \in Q'$ and numbers $k, l \leq |Q'|$ such that $k$ is an odd prime, $l > 1$ and coprime to both 2 and $k$, $q \not\sim \delta'_u(q)$, $q \not\sim \delta'_v(q)$, $q \not\sim \delta'_{uv}(q)$, and $\delta'_x(q) \sim \delta'_{xu^2}(q) \sim \delta'_{xv^k}(q) \sim \delta'_{x(uv)^l}(q)$, for all $x \in \{u, v\}^*$. We start by guessing $u, v \in \Sigma^*$ in the form of, respectively, $\mathsf{b}(u)$ and $\mathsf{b}(u)$. Also, we guess $k$ and $l$ in binary and check that $k$ is an odd prime and $l$ is coprime to both 2 and $k$. By (24), $\delta'_x$ is determined by $\mathsf{b}(x)$, for every $x \in \{u, v\}^*$. Thus, we can proceed as follows to verify that $u$, $v$, $k$ and $l$ are as required. We perform the following steps, for *each* quadruple $\mathsf{b}$ of binary relations on $Q$. First, we check whether $\mathsf{b} = \mathsf{b}(x)$, for some $x \in \{u, v\}^*$ (we discuss the algorithm for this in the next paragraph). If this is not the case, we construct the *next* quadruple $\mathsf{b}'$ and process it as this $\mathsf{b}$. If it is the case, we compute all the states $\delta'_x(q)$,

$\delta'_{xu^2}(q)$, $\delta'_{xv^k}(q)$, $\delta'_{x(uv)^l}(q)$, $\delta'_u(q)$, $\delta'_v(q)$, $\delta'_{uv}(q)$, and check their required (non)equivalences w.r.t. $\sim$, using the same method as for checking $\delta'_u(q) \not\sim q$ above. If they do not hold as required, our algorithm terminates with an answer no. Otherwise, we construct the *next* quadruple $\mathsf{b}'$ and process it as this $\mathsf{b}$. When all possible quadruples $\mathsf{b}$ of binary relations of $Q$ have been processed, the algorithm terminates with an answer yes.

Thus, it remains to explain how to check that a given quadruple $\mathsf{b}$ is equal to $\mathsf{b}(x)$, for some $x \in \{u, v\}^*$. We simply guess a sequence $\mathsf{b}_0, \dots, \mathsf{b}_n$ of quadruples of binary relations on $Q$ such that $\mathsf{b}_0 = \mathsf{b}(w_0)$, $\mathsf{b}_n = \mathsf{b}$ and $\mathsf{b}_{i+1} = \mathsf{b}_i \cdot \mathsf{b}(w_{i+1})$, where $w_i \in \{u, v\}$. It follows from the argument above that it is enough to consider $n \leq 2^{O(|Q|)}$. ❏

## 5 Deciding FO-rewritability of *LTL* OMQs

In this section, using results and constructions from the previous one, we establish the complexity of recognising the type of FO-rewritability of any given *LTL* OMQ $\boldsymbol{q}$. The following proposition formalises the connection between $\mathcal{L}$-rewritability of $\boldsymbol{q}$ and $\mathcal{L}$-definability of the corresponding regular languages $\boldsymbol{L}_\Xi(\boldsymbol{q})$ and $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$.

▶ **Proposition 12.** *Let* $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<,\equiv), \mathsf{FO}(<,\mathsf{MOD})\}$ *and* $\Xi \subseteq \mathsf{sig}(\boldsymbol{q})$.

*(i) A Boolean LTL OMQ* $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ *is* $\mathcal{L}$*-rewritable over* $\Xi$*-ABoxes iff the language* $\boldsymbol{L}_\Xi(\boldsymbol{q})$ *is* $\mathcal{L}$*-definable.*

*(ii) A specific LTL OMQ* $\boldsymbol{q}(x) = (\mathcal{O}, \varkappa(x))$ *is* $\mathcal{L}$*-rewritable over* $\Xi$*-ABoxes iff the language* $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$ *is* $\mathcal{L}$*-definable.*

**Proof.** $(i)$ For every $A \in \Xi$, let $\chi_A(y) = \bigvee_{A \in a \in \Sigma_\Xi} a(y)$, where $a(y)$ is a unary predicate associated with $a \in \Sigma_\Xi$. Conversely, for every $a \in \Sigma_\Xi$, let $\chi_a(y) = \bigwedge_{A \in a} A(y) \wedge \bigwedge_{A \notin a} \neg A(y)$. For any $\Xi$-ABox $\mathcal{A} \in \Sigma_\Xi^*$ and any $n \in \mathsf{tem}(\mathcal{A})$, we have $\mathfrak{S}_\mathcal{A} \models A(n)$ iff $\mathfrak{S}_{w_\mathcal{A}} \models \chi_A(n)$, and $\mathfrak{S}_{w_\mathcal{A}} \models a(n)$ iff $\mathfrak{S}_\mathcal{A} \models \chi_a(n)$. Thus, we obtain an $\mathcal{L}$-sentence defining $\boldsymbol{L}_\Xi(\boldsymbol{q})$ by taking an $\mathcal{L}$-rewriting of $\boldsymbol{q}$ and replacing all atoms $A(y)$ in it with $\chi_A(y)$. Conversely, we obtain an $\mathcal{L}$-rewriting of $\boldsymbol{q}$ by taking an $\mathcal{L}$-sentence defining $\boldsymbol{L}_\Xi(\boldsymbol{q})$ and replacing all $a(y)$ in it with $\chi_a(y)$.

$(ii)$ $(\Rightarrow)$ Let $\varphi(x)$ be an $\mathcal{L}$-rewriting of $\boldsymbol{q}(x)$ and let $\varphi'(x)$ be the result of replacing atoms $A(y)$ in $\varphi(x)$ with $\chi'_A(y) = \bigvee_{A \in a \in \Gamma_\Xi} a(y)$. Given an ABox $\mathcal{A}$ and $i \in \mathsf{tem}(\mathcal{A})$, we have $\mathfrak{S}_\mathcal{A} \models \varphi(i)$ iff $\mathfrak{S}_{w_{\mathcal{A},i}} \models \varphi'(i)$. A word $w = a_0 \dots a_n \in \Gamma_\Xi^*$ is in $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$ iff (a) there is $i$ such that $a_i \in \Sigma'_\Xi$, (b) $a_j \in \Sigma_\Xi$ for all $j \neq i$, and (c) $\mathfrak{S}_w \models \varphi'(i)$. Therefore, for the sentence

$$\varphi'' = \exists x \Big(\varphi'(x) \wedge \forall y \, \big[\big((y = x) \to \bigvee_{a' \in \Sigma'_\Xi} a'(y)\big) \wedge \big((y \neq x) \to \bigvee_{a \in \Sigma_\Xi} a(y)\big)\big]\Big)$$

and a word $w \in \Gamma_\Xi^*$, we have $\mathfrak{S}_w \models \varphi''$ iff $w = w_{\mathcal{A},i}$ for some $\mathcal{A}$ and $i$ such that $\mathfrak{S}_\mathcal{A} \models \varphi(i)$. It follows that $\varphi''$ defines $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$.

$(\Leftarrow)$ Suppose $\psi$ is an $\mathcal{L}$-sentence defining $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$. Let $\psi'(x)$ be the result of replacing atoms $a(y)$ in $\varphi$, for $a \in \Sigma_\Xi$, with $a(y) \wedge (x \neq y)$ and atoms $a'(y)$, for $a' \in \Sigma'_\Xi$, with $a(y) \wedge (x = y)$. For $w = a_0 \dots a_n \in \Sigma_\Xi^*$, we have $\mathfrak{S}_w \models \psi'(i)$ iff $\mathfrak{S}_{w_i} \models \psi$, where $w_i$ is $w$ with $a_i$ replaced by $a'_i$. Let $\psi''(x)$ be the result of replacing $a(y)$ in $\psi'(x)$ with $\chi_a(y)$. Then, for any ABox $\mathcal{A}$ and $i \in \mathsf{tem}(\mathcal{A})$, we have $\mathfrak{S}_\mathcal{A} \models \psi''(i)$ iff $\mathfrak{S}_{w_\mathcal{A}} \models \psi'(i)$ iff $\mathfrak{S}_{w_{\mathcal{A},i}} \models \psi$, and so $\psi''(x)$ is a rewriting of $\boldsymbol{q}$. ❏

In view of Proposition 12, we can reformulate the evaluation problem for $\boldsymbol{q}$ and $\boldsymbol{q}(x)$ over $\Xi$-ABoxes as the *word problem* for the languages $\boldsymbol{L}_\Xi(\boldsymbol{q})$ and $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$, both of which are regular by Proposition 5. Furthermore, to make circuit complexity applicable to our languages, we can assume that the alphabets $\Sigma_\Xi$ and $\Gamma_\Xi$ of $\boldsymbol{L}_\Xi(\boldsymbol{q})$ and $\boldsymbol{L}_\Xi(\boldsymbol{q}(x))$ are encoded in binary in a way preserving the properties of languages from Table 3. For example, one can take an encoding similar to that in [14, Lemma 2.1]. Then Table 3 yields the following correspondences between the data complexity of answering and FO-rewritability of Boolean and specific *LTL* OMQs $\boldsymbol{q}$:

- $\boldsymbol{q}$ is $\mathsf{FO}(<, \equiv)$-rewritable iff it can be answered in $\mathrm{AC}^0$;
- $\boldsymbol{q}$ is $\mathsf{FO}(<, \mathsf{MOD})$-rewritable iff it can be answered in $\mathrm{ACC}^0$;
- $\boldsymbol{q}$ is not $\mathsf{FO}(<, \mathsf{MOD})$-rewritable iff answering $\boldsymbol{q}$ in $\mathrm{NC}^1$-complete (unless $\mathrm{ACC}^0 = \mathrm{NC}^1$);
- $\boldsymbol{q}$ is $\mathsf{FO}(<, \mathsf{RPR})$-rewritable iff it can be answered in $\mathrm{NC}^1$.

As a consequence of Theorem 11, which is applied to the exponential-size NFAs constructed in the proof of Proposition 5, we immediately obtain the following upper bound:

▶ **Theorem 13.** *Deciding $\mathcal{L}$-rewritability of both Boolean and specific LTL OMQs over $\Xi$-ABoxes can be done in* ExpSpace.

Before establishing a matching lower bound, we prove two technical results, which allow us to reduce, in certain cases, $\mathcal{L}$-rewritability of specific OMQs to $\mathcal{L}$-rewritability of Boolean OMQs. Call two OMQs $\Xi$-*equivalent* (or simply *equivalent*) if they have the same certain answers over every $\Xi$-ABox (respectively, over every ABox). Our first useful observation allows one to remove axioms with $\bot$ from $LTL^{\Box\bigcirc}_{bool}$ ontologies:

▶ **Lemma 14.** *Let $\mathcal{O}$ be an $LTL^{\Box\bigcirc}_{bool}$ ontology, let $\mathcal{O}'$ result from $\mathcal{O}$ by removing every axiom of the form $C_1 \land \dots \land C_k \to \bot$, and let $\mathcal{O}''$ result from $\mathcal{O}$ by replacing every axiom of the form $C_1 \land \dots \land C_k \to \bot$ with $C_1 \land \dots \land C_k \to A'$, $A' \to \bigcirc_F A'$, $A' \to \bigcirc_P A'$, $A' \to A$, for a fresh atom $A'$. Let $\Xi$ be a signature that does not contain the newly introduced atoms $A'$.*

*(i) Every Boolean OMAQ $\boldsymbol{q} = (\mathcal{O}, A)$ is $\Xi$-equivalent to the OMAQ $\boldsymbol{q}' = (\mathcal{O}'', A)$. Every specific OMAQ $\boldsymbol{q}(x) = (\mathcal{O}, A(x))$ is $\Xi$-equivalent to the OMAQ $\boldsymbol{q}'(x) = (\mathcal{O}'', A(x))$.*

*(ii) Every Boolean OMPQ $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ is equivalent to the OMPQ $\boldsymbol{q}'' = (\mathcal{O}', \varkappa')$, where*

$$\varkappa' = \varkappa \lor \bigvee_{C_1 \land \dots \land C_k \to \bot \in \mathcal{O}} \Diamond_F \Diamond_P (C_1 \land \dots \land C_k)$$

*Every specific OMPQ $\boldsymbol{q}(x) = (\mathcal{O}, \varkappa(x))$ is equivalent to the OMPQ $\boldsymbol{q}''(x) = (\mathcal{O}', \varkappa'(x))$.*

**Proof.** We only show the first claim in $(i)$; other claims are similar and left to the reader. Let $\mathcal{A}$ be any $\Xi$-ABox. Suppose the certain answer to $\boldsymbol{q}'$ over $\mathcal{A}$ is no. This means that there is a model $\mathcal{I}$ of $\mathcal{O}''$ and $\mathcal{A}$ such that $\mathcal{I}, n \not\models A$ for all $n \in \mathbb{Z}$. Then $\mathcal{I}$ is also a model of $\mathcal{O}$ and $\mathcal{A}$. Indeed, if $\mathcal{I}, n \models C_1 \land \dots \land C_k$ for some $n \in \mathbb{Z}$, then $\mathcal{I}, n \models A'$, and so $\mathcal{I}, n \models A$, which is a contradiction. It follows that the answer to $\boldsymbol{q}$ over $\mathcal{A}$ is no. Conversely, suppose the answer to $\boldsymbol{q}$ over $\mathcal{A}$ is no. Let $\mathcal{I}$ be a model of $\mathcal{O}$ and $\mathcal{A}$ such that $\mathcal{I}, n \not\models A$ for all $n \in \mathbb{Z}$. Extend $\mathcal{I}$ to the fresh atoms $A'$ by setting $\mathcal{I}, n \not\models A'$. Then $\mathcal{I}$ is a model of $\mathcal{O}''$ and $\mathcal{A}$, as required. ❑

The next statement, which will be used in Theorems 16, 20, 27, and 29, shows that deciding $\mathcal{L}$-rewritability of specific $LTL^{\bigcirc}_{horn}$-OMAQs $\boldsymbol{q}(x)$ is polynomially reducible to deciding $\mathcal{L}$-rewritability of Boolean $LTL^{\bigcirc}_{horn}$-OMAQs $\boldsymbol{q}$:

▶ **Proposition 15.** *Let $\mathcal{O}$ be an $LTL^{\Box\bigcirc}_{horn}$-ontology without occurrences of $\bot$, $A$ an atom, $\varkappa$ a positive LTL formula, and $\Xi$ a signature. Let $X, X'$ be fresh atomic concepts and $\Xi_X = \Xi \cup \{X\}$. Then the following hold:*

*(i) The specific OMAQ $\boldsymbol{q}(x) = (\mathcal{O}, A(x))$ is $\mathcal{L}$-rewritable over $\Xi$-ABoxes iff the Boolean OMAQ $\boldsymbol{q}' = (\mathcal{O} \cup \{A \land X \to X'\}, X')$ is $\mathcal{L}$-rewritable over $\Xi_X$-ABoxes.*

(*ii*) *The specific OMPQ* $\boldsymbol{q}_\varkappa(x) = (\mathcal{O}, \varkappa(x))$ *is* $\mathcal{L}$*-rewritable over* $\Xi$*-ABoxes iff the Boolean OMPQ* $\boldsymbol{q}_X = (\mathcal{O}, X \land \varkappa)$ *is* $\mathcal{L}$*-rewritable over* $\Xi_X$*-ABoxes.*

**Proof.** We only show $(i)$ as the proof of $(ii)$ is analogous. Recall from [7] that, since $\mathcal{O}$ is a Horn ontology, for any ABox $\mathcal{A}$ consistent with $\mathcal{O}$, there is a *canonical model* $\mathcal{C}_{\mathcal{O},\mathcal{A}}$ of $\mathcal{O}$ and $\mathcal{A}$ such that for any OMPQ $\varkappa$,

$$(\mathcal{O}, \mathcal{A}) \models \exists x \varkappa(x) \text{ iff } \mathcal{C}_{\mathcal{O},\mathcal{A}} \models \varkappa(k) \text{ for some } k \in \mathbb{Z}$$
$$\mathcal{C}_{\mathcal{O},\mathcal{A}} \models \varkappa(k) \text{ iff } (\mathcal{O}, \mathcal{A}) \models \varkappa(k) \text{ for all } k \in \mathbb{Z}. \tag{26}$$

($\Rightarrow$) We show that if $\boldsymbol{Q}(x)$ is an $\mathcal{L}$-rewriting of $\boldsymbol{q}(x)$ over $\Xi$-ABoxes, then $\exists x\,(\boldsymbol{Q}(x) \land X(x))$ is an $\mathcal{L}$-rewriting of $\boldsymbol{q}_X$ over $\Xi_X$-ABoxes, that is, $\mathfrak{S}_\mathcal{A} \models \exists x\,(\boldsymbol{Q}(x) \land X(x))$ iff the answer to $\boldsymbol{q}_X$ over $\mathcal{A}$ is yes, for every $\Xi_X$-ABox $\mathcal{A}$. ($\Rightarrow$) Suppose $\mathfrak{S}_\mathcal{A} \models \exists x\,(\boldsymbol{Q}(x) \land X(x))$. As $X$ does not occur in $\mathcal{O}$, we then have $\mathfrak{S}_\mathcal{A} \models \boldsymbol{Q}(n)$ and $\mathfrak{S}_\mathcal{A} \models X(n)$, for some $n \in \mathsf{tem}(\mathcal{A})$. Since $\boldsymbol{Q}(x)$ is a rewriting of $\boldsymbol{q}(x)$, it follows that $n$ is a certain answer to $\boldsymbol{q}(x)$ over $\mathcal{A}$, and so $\mathcal{I}, n \models \varkappa$ for every model $\mathcal{I}$ of $(\mathcal{O}, \mathcal{A})$. Since $\mathcal{I}, n \models X$, for every such model $\mathcal{I}$, it follows that $\mathcal{I}, n \models X \land \varkappa$ for every model $\mathcal{I}$ of $(\mathcal{O}, \mathcal{A})$, as required. ($\Leftarrow$) Suppose the answer to $\boldsymbol{q}_X$ over $\mathcal{A}$ is yes. As $\boldsymbol{q}_X$ is Horn, it follows that $\mathcal{I}, n \models X \land \varkappa$ for the *canonical model* $\mathcal{I}$ of $(\mathcal{O}, \mathcal{A})$. Since $X$ does not occur in $\mathcal{O}$, there exists $n$ in $\mathsf{tem}(\mathcal{A})$ such that $\mathfrak{S}_\mathcal{A} \models X(n)$ and $\mathcal{I}, n \models \varkappa$. Thus, $n$ is a certain answer to $\boldsymbol{q}(x)$ over $\mathcal{A}$, and so $\mathfrak{S}_\mathcal{A} \models \exists x\,(\boldsymbol{Q}(x) \land X(x))$.

($\Leftarrow$) Suppose $\boldsymbol{Q}$ is an $\mathcal{L}$-rewriting of $\boldsymbol{q}_X$ over $\Xi_X$-ABoxes. Fix a variable $x$ that does not occur in $\boldsymbol{Q}$ and let $\boldsymbol{Q}^-$ be the result of replacing every occurrence of $X(y)$ in $\boldsymbol{Q}$ with $(x = y)$. We show that $\boldsymbol{Q}^-$ is an $\mathcal{L}$-rewriting of $\boldsymbol{q}(x)$ over $\Xi$-ABoxes. Given a $\Xi$-ABox $\mathcal{A}$, construct the $\Xi_X$-ABox $\mathcal{A}_X^k = \mathcal{A} \cup \{X(k)\}$, for any $k \in \mathsf{tem}(\mathcal{A})$. Note that $\mathfrak{S}_\mathcal{A} \models \boldsymbol{Q}^-(k)$ iff $\mathfrak{S}_{\mathcal{A}_X^k} \models \boldsymbol{Q}$, for every $k \in \mathsf{tem}(\mathcal{A})$. Indeed, $\mathfrak{S}_{\mathcal{A}_X^k} \models X(y) \leftrightarrow (k = y)$, and so $\mathfrak{S}_{\mathcal{A}_X^k} \models \boldsymbol{Q} \leftrightarrow \boldsymbol{Q}^-(k)$. It remains to recall that $X$ does not occur in $\boldsymbol{Q}^-$, from which $\mathfrak{S}_{\mathcal{A}_X^k} \models \boldsymbol{Q}^-(k)$ iff $\mathfrak{S}_\mathcal{A} \models \boldsymbol{Q}^-(k)$. Now, suppose $k$ is a certain answer to $\boldsymbol{q}(x)$ over $\mathcal{A}$. Then the certain answer to $\boldsymbol{q}_X$ over $\mathcal{A}_X^k$ is yes, and so $\mathfrak{S}_{\mathcal{A}_X^k} \models \boldsymbol{Q}$, which implies $\mathfrak{S}_\mathcal{A} \models \boldsymbol{Q}^-(k)$. Conversely, if $k$ is not a certain answer to $\boldsymbol{q}$ over $\mathcal{A}$, then the answer to $\boldsymbol{q}_X$ over $\mathcal{A}_X^k$ is no. We then have $\mathfrak{S}_{\mathcal{A}_X^k} \not\models \boldsymbol{Q}$, and so $\mathfrak{S}_\mathcal{A} \not\models \boldsymbol{Q}^-(k)$. ❑

In the remainder of this section, we establish a matching ExpSpace lower bound, which holds already for $LTL_{horn}^{\bigcirc}$ OMAQs and $LTL_{krom}^{\bigcirc}$ OMPEQs.

A *counter* is a set $\mathbb{A} = \{A_j^i \mid i = 0, 1,\ j = 1, \dots, k\}$ of atomic concepts that will be used to store values between 0 and $2^k - 1$, which can be different at different time points. The counter $\mathbb{A}$ is *well-defined* at a time point $n \in \mathbb{Z}$ in an interpretation $\mathcal{I}$ if $\mathcal{I}, n \models A_j^0 \land A_j^1 \to \bot$ and $\mathcal{I}, n \models A_j^0 \lor A_j^1$, for any $j = 1, \dots, k$. In this case, the *value of* $\mathbb{A}$ at $n$ in $\mathcal{I}$ is given by the unique binary number $b_k \dots b_1$ for which $\mathcal{I}, n \models A_1^{b_1} \land \dots \land A_k^{b_k}$. We require the following formulas, for $c = b_k \dots b_1$:

- $[\mathbb{A} = c] = A_1^{b_1} \land \dots \land A_k^{b_k}$ with $\mathcal{I}, n \models [\mathbb{A} = c]$ iff the value of $\mathbb{A}$ is $c$ (provided that $\mathbb{A}$ is well-defined);
- $[\mathbb{A} < c] = \bigvee_{\substack{k \ge i \ge 1\\ b_i = 1}} \big(A_i^0 \land \bigwedge_{j=i+1}^k A_j^{b_j}\big)$ with $\mathcal{I}, n \models [\mathbb{A} < c]$ iff the value of $\mathbb{A}$ is smaller than $c$ (provided that $\mathbb{A}$ is well-defined);
- $[\mathbb{A} > c] = \bigvee_{\substack{k \ge i \ge 1\\ b_i = 0}} \big(A_i^1 \land \bigwedge_{j=i+1}^k A_j^{b_j}\big)$ with $\mathcal{I}, n \models [\mathbb{A} > c]$ iff the value of $\mathbb{A}$ is greater than $c$ (provided that $\mathbb{A}$ is well-defined).

We regard the set $(\bigcirc_{\!F} \mathbb{A}) = \{\bigcirc_{\!F} A_j^i \mid i = 0, 1,\ j = 1, \dots, k\}$ as another counter that stores at $n$ in $\mathcal{I}$ the value stored by $\mathbb{A}$ at $n + 1$ in $\mathcal{I}$. This allows us to use formulas such as $[\mathbb{A} > c_1] \to [(\bigcirc_{\!F} \mathbb{A}) = c_2]$, which says that if the value of $\mathbb{A}$ at $n$ in $\mathcal{I}$ is greater than $c_1$, then the value of $\mathbb{A}$ at $n + 1$ in $\mathcal{I}$ is $c_2$.

Given two counters $\mathbb{A}$ and $\mathbb{B}$, we set

$$[\mathbb{A} = \mathbb{B}] = \bigwedge_{j=1}^{k} \big((B_j^0 \to A_j^0) \land (B_j^1 \to A_j^1)\big),$$

$$[\mathbb{A} = \mathbb{B}+1] = \bigwedge_{i=1}^{k} \big((B_i^0 \land B_{i-1}^1 \land \dots \land B_1^1 \to A_i^1 \land A_{i-1}^0 \land \dots \land A_1^0) \land \bigwedge_{j<i}((B_i^0 \land B_j^0 \to A_i^0) \land (B_i^1 \land B_j^0 \to A_i^1))\big).$$

We have $\mathcal{I}, n \models [\mathbb{A} = \mathbb{B}]$ iff the values of $\mathbb{A}$ and $\mathbb{B}$ at $n$ in $\mathcal{I}$ coincide, and $\mathcal{I}, n \models [\mathbb{A} = \mathbb{B}+1]$ iff the value of $\mathbb{A}$ at $n$ is equal to the value of $\mathbb{B}$ at $n$ plus one. In a similar way, we define the formula $[\mathbb{A} = \mathbb{B} - 1]$.

▶ **Theorem 16.** *For any $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<,\equiv), \mathsf{FO}(<,\mathsf{MOD})\}$, deciding $\mathcal{L}$-rewritability of $LTL_{horn}^{\bigcirc}$ Boolean or specific OMAQs over $\Xi$-ABoxes is* ExpSpace*-hard.*

**Proof.** Consider a deterministic Turing machine $\boldsymbol{M}$ with exponential space bound, which behaves as described in the proof of Theorem 8. Given an input word $\boldsymbol{x} = x_1 \dots x_n$, let $N$ be the space needed for the computation of $\boldsymbol{M}$ on $\boldsymbol{x}$, and let $N'$ be the first prime exceeding $N+1$ and such that $N' \not\equiv \pm 1 \mod 10$. Our aim is to construct $LTL_{horn}^{\bigcirc}$ ontologies $\mathcal{O}_<$, $\mathcal{O}_\equiv$ and $\mathcal{O}_{\mathsf{MOD}}$ of polynomial size that simulate the exponential-size, $O(N')$, DFAs $\mathfrak{A}_<$, $\mathfrak{A}_\equiv$ and $\mathfrak{A}_{\mathsf{MOD}}$ from the proof Theorem 8, whose languages are $\mathcal{L}$-definable (for the corresponding $\mathcal{L}$) iff $\boldsymbol{M}$ rejects $\boldsymbol{x}$.

First we define $\mathcal{O}_<$. Let $k = \lceil \log_2 N' \rceil + 1$.

The ontology $\mathcal{O}_<$ uses the following atomic concepts: the symbols in $\Sigma''$ from the proof of Theorem 8, $S$, $Q_0$, $Q_1$, $Q_a$, $Q_{ab}$, $P_a$ for $a, b \in \Sigma'$, $F$, $X$, $Y$, and $F_{end}$; we also use counters $\mathbb{A}$ and $\mathbb{L}$ with atomic concepts $A_j^i$ and $L_j^i$, for $i = 0, 1$, $j = 1, \dots, k$. Set $\Xi = \Sigma'' \cup \{X, Y\}$, where $\Sigma''$ is defined in the proof of Theorem 8.

In the DFA $\mathfrak{A}_i$, we represent

- each state $q_y^j$ of $\mathfrak{A}_i$ as $[\mathbb{A} = i] \land Q_y \land [\mathbb{L} = j]$;
- each state $p_a^j$ of $\mathfrak{A}_i$ as $[\mathbb{A} = i] \land P_a \land [\mathbb{L} = j]$;
- $f_i$ as $[\mathbb{A} = i] \land F$;
- $s_i$ as $[\mathbb{A} = i] \land S$.

To make the ontology $\mathcal{O}_<$ simulate the automaton $\mathfrak{A}_<$ (see Lemma 17) we require the following axioms (which are equivalent to polynomially-many $LTL_{horn}^{\bigcirc}$ axioms):

- $a \land b \to \bot$, for $a, b \in \Xi$; $(\star_1)$
- $X \to [(\bigcirc_F \mathbb{A}) = 0] \land \bigcirc_F S$ to simulate the initial state of $\mathfrak{A}_<$; $(\star_2)$
- $[\mathbb{A} = 0] \land S \land Y \to F_{end}$ to simulate the accepting state of $\mathfrak{A}_<$; $(\star_3)$
- the axioms

$$[\mathbb{A} = 0] \land S \land a_1 \to [(\bigcirc_F \mathbb{A}) = 0] \land \bigcirc_F Q_0 \land [(\bigcirc_F \mathbb{L}) = \mathbb{A}],$$
$$[\mathbb{A} < N' - 1] \land F \land a_2 \to [(\bigcirc_F \mathbb{A}) = \mathbb{A} + 1] \land \bigcirc_F S,$$
$$[\mathbb{A} = N' - 1] \land F \land a_2 \to [(\bigcirc_F \mathbb{A}) = 0] \land \bigcirc_F S;$$

  describing the behaviour of $\mathfrak{A}_<$ in states $s_i$ and $f_i$;

- the axioms

$$[\mathbb{A}=0]\land Q_0\land[\mathbb{L}=0]\land\sharp\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_0\land[(\bigcirc_F\mathbb{L})=1],$$
$$[\mathbb{A}=0]\land Q_0\land[\mathbb{L}=1]\land(q_1,x_1)\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_0\land[(\bigcirc_F\mathbb{L})=2],$$
$$\dots$$
$$[\mathbb{A}=0]\land Q_0\land[\mathbb{L}=n]\land x_n\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_0\land[(\bigcirc_F\mathbb{L})=n+1],$$
$$[\mathbb{A}=0]\land Q_0\land[\mathbb{L}>n]\land[\mathbb{L}< N+1]\land\mathsf{b}\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_0\land[(\bigcirc_F\mathbb{L})=\mathbb{L}+1],$$
$$[\mathbb{A}=0]\land Q_0\land[\mathbb{L}=N+1]\land\sharp\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_1\land[(\bigcirc_F\mathbb{L})=1],$$
$$[\mathbb{A}=0]\land Q_1\land[\mathbb{L}=1]\land a\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_1\land[(\bigcirc_F\mathbb{L})=0],\quad\text{for } a\neq(q_{acc},\mathsf{b}),\sharp,$$
$$[\mathbb{A}=0]\land Q_1\land[\mathbb{L}=0]\land a\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_1\land[(\bigcirc_F\mathbb{L})=0],\quad\text{for } a\neq\sharp,$$
$$[\mathbb{A}=0]\land Q_1\land[\mathbb{L}=0]\land\sharp\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_1\land[(\bigcirc_F\mathbb{L})=1],$$
$$[\mathbb{A}=0]\land Q_1\land[\mathbb{L}=1]\land(q_{acc},\mathsf{b})\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_1\land[(\bigcirc_F\mathbb{L})=2],$$
$$[\mathbb{A}=0]\land Q_1\land[\mathbb{L}>1]\land[\mathbb{L}<N+1]\land\mathsf{b}\to[(\bigcirc_F\mathbb{A})=0]\land\bigcirc_F Q_1\land[(\bigcirc_F\mathbb{L})=\mathbb{L}+1)],$$
$$[\mathbb{A}=0]\land Q_1\land[\mathbb{L}=N+1]\land\flat\to[\mathbb{A}=0]\land\bigcirc_F F$$

describing the transitions of $\mathfrak{A}_0$;

- the axioms for $a,b,c\in\Sigma'\setminus\{\flat\}$, $b,c\neq\sharp$

$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_0\land[\mathbb{L}>1]\land a\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_0\land[(\bigcirc_F\mathbb{L})=\mathbb{L}-1],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_0\land[\mathbb{L}=1]\land a\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_a\land\bigcirc_F[\mathbb{L}=0],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_a\land[\mathbb{L}=0]\land b\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_{ab}\land\bigcirc_F[\mathbb{L}=1],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_{ab}\land[\mathbb{L}=1]\land c\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_{z_{abc}}\land\bigcirc_F[\mathbb{L}=2],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_{ab}\land[\mathbb{L}=1]\land\sharp\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F P_{z_{ab\sharp}}\land\bigcirc_F[\mathbb{L}=2],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_a\land[\mathbb{L}>1]\land[\mathbb{L}<N]\land b\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_a\land[(\bigcirc_F\mathbb{L})=\mathbb{L}+1],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_a\land[\mathbb{L}>1]\land[\mathbb{L}<N]\land\sharp\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F P_a\land[(\bigcirc_F\mathbb{L})=\mathbb{L}+1]$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land P_a\land[\mathbb{L}>1]\land[\mathbb{L}<N]\land b\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F P_a\land[(\bigcirc_F\mathbb{L})=\mathbb{L}+1],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land P_a\land[\mathbb{L}=N]\land b\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_{ba}\land[(\bigcirc_F\mathbb{L})=0],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_a\land[\mathbb{L}=N]\land\sharp\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_{\sharp a}\land[(\bigcirc_F\mathbb{L})=0],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_{ab}\land[\mathbb{L}=0]\land b\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_{ab}\land[(\bigcirc_F\mathbb{L})=1],$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_b\land[\mathbb{L}<N+1]\land\flat\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F F,$$
$$[\mathbb{A}>0]\land[\mathbb{A}<N+1]\land Q_{bc}\land[\mathbb{L}=1]\land\flat\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F F,$$

simulating the transitions of $\mathfrak{A}_i$, for $0<i\le N+1$;

- the axioms

$$[\mathbb{A}>N+1]\land[\mathbb{A}<N'+1]\land Q_0\land a\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F Q_0\land[(\bigcirc_F\mathbb{L})=\mathbb{L}],\ \text{for } a\neq\flat,$$
$$[\mathbb{A}>N+1]\land[\mathbb{A}<N'+1]\land Q_0\land\flat\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F F$$

simulating the transitions of $\mathfrak{A}_i$, for $N+1\le i\le N'$.

Next, we define the ontology $\mathcal{O}_\equiv$ by adding to $\mathcal{O}_<$ the axiom

$$[\mathbb{A}<N'+1]\land S\land\natural\to[(\bigcirc_F\mathbb{A})=\mathbb{A}]\land\bigcirc_F S$$

simulating the $\natural$-transitions in $\mathfrak{A}_\equiv$. We also we extend $\Xi$ with the atomic concept $\natural$.

To define the ontology $\mathcal{O}_{\mathsf{MOD}}$ more work is needed. First, we extend $\mathcal{O}_<$ with

- the following axioms regarding $\mathfrak{A}_{N'}$:

$$[\mathbb{A} = N'] \wedge S \wedge a_1 \to [(\bigcirc_F \mathbb{A}) = N'] \wedge \bigcirc_F Q_0,$$
$$[\mathbb{A} = N'] \wedge F \wedge a_2 \to [(\bigcirc_F \mathbb{A}) = N'] \wedge \bigcirc_F S,$$

- the following axioms handling $\natural$:

$$[\mathbb{A} = 0] \wedge S \wedge \natural \to [(\bigcirc_F \mathbb{A}) = N'] \wedge \bigcirc_F S,$$
$$[\mathbb{A} = N'] \wedge S \wedge \natural \to [(\bigcirc_F \mathbb{A}) = 0] \wedge S,$$
$$[\mathbb{A} > 0] \wedge [\mathbb{A} < N'] \wedge S \wedge \natural \to [(\bigcirc_F \mathbb{A}) = \mathbb{J}] \wedge \bigcirc_F S.$$

Here, $\mathbb{J}$ is a new counter that stores the value $j = -1/i$ in the field $\mathbb{F}_{N'}$, which is required to make sure that, for $i \neq 0, N'$, we have

$$\mathcal{O}_{\mathsf{MOD}} \models [\mathbb{A} = i] \wedge S \wedge \natural \to [(\bigcirc_F \mathbb{A}) = j] \wedge \bigcirc_F S.$$

We achieve this as follows. We compute the number $r$ such that $ir = 1 \mod N'$ using the following modified version of Penk's algorithm; see, e.g., [38, Exercise 4.5.2.39]. The algorithm starts with $u = N'$, $v = i$, $r = 0$, $s = 1$. In the course of the algorithm, $u$ and $v$ decrease, with the following conditions being met: $\mathrm{GCD}(u, v) = 1$, $u = ri \mod N'$, and $v = si \mod N'$. The algorithm repeats the following steps until $v = 0$:

- if $v$ is even, replace it with $v/2$, and replace $s$ with either $s/2$ or $(s + N')/2$, whichever is a whole number;
- if $u$ is even, replace it with $u/2$, and replace $r$ with either $r/2$ or $(r + N')/2$, whichever is a whole number;
- if $u, v$ are odd and $u > v$, replace $u$ with $(u - v)/2$ and $r$ with either $(r - s)/2$ or $(r - s + N')/2$, whichever is a whole number;
- if $u, v$ are odd and $v \geq u$, replace $v$ with $(v - u)/2$ and $s$ with either $(s - r)/2$ or $(s - r + N')/2$, whichever is a whole number.

The binary length of the larger of $u$ and $v$ is reduced by at least one bit, guaranteeing that the procedure terminates in at most $2k$ iterations while maintaining the conditions. At termination, $v = 0$ as otherwise a reduction is still possible. If $u = 1$, we get $1 = ri \bmod N'$ and $r = 1/i$ in the field $\mathbb{F}_{N'}$, so we can set $j = N' - r$.

For two counters $\mathbb{X}$ and $\mathbb{Y}$, set

$$[\mathbb{X} = \mathbb{Y}/2] \;=\; X_k^0 \wedge \bigwedge_{l=2}^{k} \big((Y_l^0 \to X_{l-1}^0) \wedge (Y_l^1 \to X_{l-1}^1)\big)\,.$$

We have $\mathcal{I}, n \models [\mathbb{X} = \mathbb{Y}/2]$ iff the values $x$ of $\mathbb{X}$ and $y$ of $\mathbb{Y}$ at $n$ in $\mathcal{I}$ satisfy $x = \lfloor y/2 \rfloor$. We define three new counters $\mathbb{C}^{=}_{\mathbb{XY}}$, $\mathbb{C}^{-}_{\mathbb{XY}}$, and $\mathbb{C}^{+}_{\mathbb{XY}}$, which come with the following axioms, for all $\iota_1, \iota_2, \iota_3 \in \{0, 1\}$, that should be added to the ontology:

$$X_i^{\iota_1} \wedge Y_i^{\iota_2} \to (C^{=}_{\mathbb{XY}})_i^{(\iota_1 + \iota_2 + 1) \bmod 2}, \qquad \text{for all } i \in [1, k],$$
$$X_1^{\iota_1} \wedge Y_1^{\iota_2} \to (C^{+}_{\mathbb{XY}})_1^{0},$$
$$X_{i-1}^{\iota_1} \wedge Y_{i-1}^{\iota_2} \wedge (C^{+}_{\mathbb{XY}})_{i-1}^{\iota_3} \to (C^{+}_{\mathbb{XY}})_i^{(\iota_1\iota_2 + \iota_1\iota_3 + \iota_2\iota_3) \bmod 2}, \qquad \text{for all } i \in [2, k],$$
$$X_1^{\iota_1} \wedge Y_1^{\iota_2} \to (C^{-}_{\mathbb{XY}})_1^{0},$$
$$X_{i-1}^{\iota_1} \wedge Y_{i-1}^{\iota_2} \wedge (C^{-}_{\mathbb{XY}})_{i-1}^{\iota_3} \to (C^{-}_{\mathbb{XY}})_i^{(\iota_1\iota_2 + \iota_1\iota_3 + \iota_2\iota_3 + \iota_2 + \iota_3) \bmod 2}, \qquad \text{for all } i \in [2, k].$$

Define the following formulas, where $\mathbb{W}$ is some counter:

$$[\mathbb{X} > \mathbb{Y}] = \bigvee_{i=1}^{k} \big(X_i^1 \wedge Y_i^0 \wedge \bigwedge_{j=i+1}^{k} (C^{=}_{\mathbb{X}\mathbb{Y}})^1_i\big),$$

$$[\mathbb{X} \geq \mathbb{Y}] = [\mathbb{X} > \mathbb{Y}] \vee \bigwedge_{i=1}^{k} (C^{=}_{\mathbb{X}\mathbb{Y}})^1_i,$$

$$[\mathbb{W} = \mathbb{X} + \mathbb{Y}] = \bigwedge_{i=1}^{k} \bigwedge_{\iota_{1,2,3} \in \{0,1\}} \big(X_i^{\iota_1} \wedge Y_i^{\iota_2} \wedge (C^{+}_{\mathbb{X}\mathbb{Y}})^{\iota_3}_i \to W_i^{\iota_1+\iota_2+\iota_3 \bmod 2}\big),$$

$$[\mathbb{W} = \mathbb{X} - \mathbb{Y}] = \bigwedge_{i=1}^{k} \bigwedge_{\iota_{1,2,3} \in \{0,1\}} \big(X_i^{\iota_1} \wedge Y_i^{\iota_2} \wedge (C^{-}_{\mathbb{X}\mathbb{Y}})^{\iota_3}_i \to W_i^{\iota_1+\iota_2+\iota_3 \bmod 2}\big).$$

We have $\mathcal{I}, n \models [\mathbb{X} > \mathbb{Y}]$, $\mathcal{I}, n \models [\mathbb{X} \geq \mathbb{Y}]$, $\mathcal{I}, n \models [\mathbb{W} = \mathbb{X} + \mathbb{Y}]$, or $\mathcal{I}, n \models [\mathbb{W} = \mathbb{X} - \mathbb{Y}]$ iff the values $x$ of $\mathbb{X}$, $y$ of $\mathbb{Y}$, and $w$ of $\mathbb{W}$ at $n$ in $\mathcal{I}$ satisfy, respectively, the following conditions: $x > y$, $x \geq y$, $(x + y < 2^k) \to (w = x + y)$, or $(x > y) \to (w = x - y)$.

In our ontology $\mathcal{O}_{\mathsf{MOD}}$, we use counters $\mathbb{U}_l$, $\mathbb{V}_l$, $\mathbb{R}_l$, $\mathbb{R}^+_l$, $\mathbb{R}^-_l$, $\mathbb{S}_l$, $\mathbb{S}^-_l$, $\mathbb{S}^+_l$, $\mathbb{D}_l$, $\mathbb{G}_l$, $\mathbb{H}_l$, for $l \in [0, \dots, 2k]$, along with some auxiliary counters $\mathbb{C}_{\mathbb{X}\mathbb{Y}}$. Intuitively, the counters with the index $l$ hold the values of the corresponding expressions after the $l$-th step of the algorithm according to the table below:

| | |
|---|---|
| $\mathbb{U}_l, \mathbb{V}_l, \mathbb{R}_l, \mathbb{S}_l$ | $u, v, r, s$ |
| $\mathbb{R}^+_l, \mathbb{S}^+_l$ | $r + N', s + N'$ |
| $\mathbb{R}^-_l, \mathbb{S}^-_l$ | $-r \bmod N', -s \bmod N'$ |
| $\mathbb{D}_l$ | $\lvert u - v \rvert$ |
| $\mathbb{G}_l$ | the even number from the pair $((r - s) \bmod N')$, $((r - s) \bmod N') + N'$ |
| $\mathbb{H}_l$ | the even number from the pair $((s - r) \bmod N')$, $((s - r) \bmod N') + N'$ |

We add the following axioms (simulating the algorithm above) to the ontology $\mathcal{O}_{\mathsf{MOD}}$ constructed so far:

$$[\mathbb{A} > 0] \wedge [\mathbb{A} < N'] \wedge S \wedge \natural \to [\mathbb{U}_0 = N'] \wedge [\mathbb{V}_0 = \mathbb{A}] \wedge [\mathbb{R}_0 = 0] \wedge [\mathbb{S}_0 = 1],$$

$$[\mathbb{U}_l > \mathbb{V}_l] \to [\mathbb{D}_l = \mathbb{U}_l - \mathbb{V}_l],$$

$$[\mathbb{V}_l \geq \mathbb{U}_l] \to [\mathbb{D}_l = \mathbb{V}_l - \mathbb{U}_l],$$

$$[\mathbb{R}^+_l = \mathbb{R}_l + \mathbb{U}_0] \wedge [\mathbb{R}^-_l = \mathbb{U}_0 - \mathbb{R}_l] \wedge [\mathbb{S}^+_l = \mathbb{S}_l + \mathbb{U}_0] \wedge [\mathbb{S}^-_l = \mathbb{U}_0 - \mathbb{S}_l],$$

$$[\mathbb{R}_l \geq \mathbb{S}_l] \wedge (((R_l)^0_1 \wedge (S_l)^0_1) \vee ((R_l)^1_1 \wedge (S_l)^1_1)) \to [\mathbb{G}_l = \mathbb{R}_l - \mathbb{S}_l] \wedge [\mathbb{H}_l = \mathbb{S}^+_l + \mathbb{R}^-_l],$$

$$[\mathbb{R}_l \geq \mathbb{S}_l] \wedge (((R_l)^1_1 \wedge (S_l)^0_1) \vee ((R_l)^0_1 \wedge (S_l)^1_1)) \to [\mathbb{G}_l = \mathbb{R}_l + \mathbb{S}^-_l] \wedge [\mathbb{H}_l = \mathbb{S}^+_l - \mathbb{R}_l],$$

$$[\mathbb{S}_l > \mathbb{R}_l] \wedge (((R_l)^0_1 \wedge (S_l)^0_1) \vee ((R_l)^1_1 \wedge (S_l)^1_1)) \to [\mathbb{G}_l = \mathbb{R}^+_l + \mathbb{S}^-_l] \wedge [\mathbb{H}_l = \mathbb{S}_l - \mathbb{R}_l],$$

$$[\mathbb{S}_l > \mathbb{R}_l] \wedge (((R_l)^1_1 \wedge (S_l)^0_1) \vee ((R_l)^0_1 \wedge (S_l)^1_1)) \to [\mathbb{G}_l = \mathbb{R}^+_l - \mathbb{S}_l] \wedge [\mathbb{H}_l = \mathbb{S}_l + \mathbb{R}^-_l],$$

$$[\mathbb{V}_l > 0] \wedge (V_l)^0_1 \wedge (S_l)^0_1 \to [\mathbb{U}_{l+1} = \mathbb{U}_l] \wedge [\mathbb{V}_{l+1} = \mathbb{V}_l/2] \wedge [\mathbb{R}_{l+1} = \mathbb{R}_l] \wedge [\mathbb{S}_{l+1} = \mathbb{S}_l/2],$$

$$[\mathbb{V}_l > 0] \wedge (V_l)^0_1 \wedge (S_l)^1_1 \to [\mathbb{U}_{l+1} = \mathbb{U}_l] \wedge [\mathbb{V}_{l+1} = \mathbb{V}_l/2] \wedge [\mathbb{R}_{l+1} = \mathbb{R}_l] \wedge [\mathbb{S}_{l+1} = \mathbb{S}^+_l/2],$$

$$(V_l)^1_1 \wedge (U_l)^0_1 \wedge (R_l)^0_1 \to [\mathbb{U}_{l+1} = \mathbb{U}_l/2] \wedge [\mathbb{V}_{l+1} = \mathbb{V}_l] \wedge [\mathbb{R}_{l+1} = \mathbb{R}_l/2] \wedge [\mathbb{S}_{l+1} = \mathbb{S}_l],$$

$$(V_l)^1_1 \wedge (U_l)^0_1 \wedge (R_l)^1_1 \to [\mathbb{U}_{l+1} = \mathbb{U}_l/2] \wedge [\mathbb{V}_{l+1} = \mathbb{V}_l] \wedge [\mathbb{R}_{l+1} = \mathbb{R}^+_l/2] \wedge [\mathbb{S}_{l+1} = \mathbb{S}_l],$$

$$(V_l)^1_1 \wedge (U_l)^1_1 \wedge [\mathbb{U}_l > \mathbb{V}_l] \to [\mathbb{U}_{l+1} = \mathbb{D}_l/2] \wedge [\mathbb{V}_{l+1} = \mathbb{V}_l] \wedge [\mathbb{R}_{l+1} = \mathbb{H}_l/2] \wedge [\mathbb{S}_{l+1} = \mathbb{S}_l],$$

$$(V_l)^1_1 \wedge (U_l)^1_1 \wedge [\mathbb{V}_l \geq \mathbb{U}_l] \to [\mathbb{U}_{l+1} = \mathbb{U}_l] \wedge [\mathbb{V}_{l+1} = \mathbb{D}_l/2] \wedge [\mathbb{R}_{l+1} = \mathbb{R}_l] \wedge [\mathbb{S}_{l+1} = \mathbb{G}_j/2],$$

$$[\mathbb{V}_l = 0] \to [\mathbb{J} = \mathbb{R}^-_l].$$

Here, as before, $\Xi = \Sigma'' \cup \{X, Y\}$. We call $\Psi$ a *state-formula* if it takes one of the following forms: $\big([\mathbb{A} = i] \wedge Q_y \wedge [\mathbb{L} = j]\big)$, $\big([\mathbb{A} = i] \wedge P_a \wedge [\mathbb{L} = j]\big)$, $\big([\mathbb{A} = i] \wedge S\big)$, or $\big([\mathbb{A} = i] \wedge F\big)$, in which case we refer to, respectively, $q_y^j$ of $\mathfrak{A}_i$, $p_a^j$ of $\mathfrak{A}_i$, $s_i$, or $f_i$ as *the state corresponding to* $\Psi$.

For $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<, \equiv), \mathsf{FO}(<, \mathsf{MOD})\}$, use $\mathfrak{A}_\mathcal{L}$ and $\mathcal{O}_\mathcal{L}$ to denote the corresponding automaton and ontology defined above.

▶ **Lemma 17.** *Let $\mathcal{A}$ be a $\Xi$-ABox and let $\Psi$ be a state-formula. Then the following hold:*
*$(i)$ $\mathcal{A}$ is inconsistent with $\mathcal{O}_\mathcal{L}$ iff there is $i$ such that $a(i), b(i) \in \mathcal{A}$ for different $a, b \in \Xi$.*
*$(ii)$ If $\mathcal{A}$ is consistent with $\mathcal{O}_\mathcal{L}$, then $\mathcal{O}_\mathcal{L}, \mathcal{A} \models \Psi(l)$ iff $\mathcal{A}$ contains a subset of the form*

$$\{X(l-m-1), a_1(l-m), a_2(l-m+1), a_3(l-m+2), \dots, a_m(l-1)\}, \tag{27}$$

*where $m \geq 0$, $a_h \in \Sigma''$ for all $h \in [1, m]$, and $\mathfrak{A}_\mathcal{L}$, having read the word $a_1 \dots a_m$, is in the state corresponding to $\Psi$.*

**Proof.** $(i)$ This is so because the only axiom that can lead to inconsistency is $(\star_1)$ and, for consistent $\mathcal{A}$ and $\mathcal{O}_\mathcal{L}$, $b \in \Xi$ and $n \in \mathbb{Z}$, we have $\mathcal{O}, \mathcal{A} \models b(n)$ iff $b(n) \in \mathcal{A}$.

$(ii)$ $(\Leftarrow)$ If there is such a subset of $\mathcal{A}$, then $\mathcal{O}_\mathcal{L}, \mathcal{A} \models \big([\mathbb{A} = 0] \wedge S\big)(l-m)$. One can check by induction on $j$ that if the automaton is in a state $q$ after reading $a_1 \dots a_{j-1}$, then $\mathcal{O}, \mathcal{A} \models \Psi'(l-m+j)$, where $q$ is the state corresponding to the state-formula $\Psi'$.

$(\Rightarrow)$ If $\mathcal{O}_\mathcal{L}, \mathcal{A} \models A_{j_1}^{\iota_1}(l)$, for some $A_{j_1}^{\iota_1} \in \mathbb{A}$, then $\mathcal{O}_\mathcal{L}, \mathcal{A} \models b(l-1)$, for some $b \in \Xi$. There are two possibilities: either $b = X$ or $b \in \Sigma''$ and there is $A_{j_2}^{\iota_2} \in \mathbb{A}$ such that $\mathcal{O}_\mathcal{L}, \mathcal{A} \models A_{j_2}^{\iota_2}(l-1)$. Therefore there is a unique subset of $\mathcal{A}$ of the form (27). By induction on $j \in [1, m+1]$ we can prove that there is a unique state-formula $\Psi_j$ such that $\mathcal{O}_\mathcal{L}, \mathcal{A} \models \Psi_j(l-m+j)$ and it corresponds to the state $\mathfrak{A}_\mathcal{L}$ is in after reading $a_1 \dots a_{j-1}$. ❑

▶ **Lemma 18.** *For $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<, \equiv), \mathsf{FO}(<, \mathsf{MOD})\}$, the $LTL_{horn}^{\bigcirc}$ OMAQ $(\mathcal{O}_\mathcal{L}, F_{end})$ is $\mathcal{L}$-rewritable over $\Xi$-ABoxes iff the language $\boldsymbol{L}(\mathfrak{A}_\mathcal{L})$ is $\mathcal{L}$-definable.*

**Proof.** $(\Rightarrow)$ For $w = a_1 \dots a_m \in \Sigma''$, let $\mathcal{A}_w = \{X(0), a_1(1), \dots, a_m(m), Y(m+1)\}$. By Lemma 17 and $(\star_2)$, we see that $w \in \boldsymbol{L}(\mathfrak{A}_\mathcal{L})$ iff the answer to $(\mathcal{O}_\mathcal{L}, F_{end})$ over $\mathcal{A}_w$ is $\mathsf{yes}$.

$(\Leftarrow)$ Suppose $\boldsymbol{L}(\mathfrak{A}_\mathcal{L})$ is $\mathcal{L}$-definable and $\mathcal{A}$ is a $\Xi$-ABox. If the certain answer to $(\mathcal{O}_\mathcal{L}, F_{end})$ is $\mathsf{yes}$, then either $\mathcal{A}$ is inconsistent with $\mathcal{O}_\mathcal{L}$, or $\mathcal{O}_\mathcal{L}, \mathcal{A} \models \big([\mathbb{A} = 0] \wedge S \wedge Y\big)(x)$ for some $x$. By Lemma 17 $(i)$, inconsistency is $\mathcal{L}$-definable. Suppose that $\mathcal{A}$ is consistent with $\mathcal{O}_\mathcal{L}$. If $\mathcal{O}_\mathcal{L}, \mathcal{A} \models \big([\mathbb{A} = 0] \wedge S \wedge Y\big)(x)$ then $\mathcal{A}$ contains a subset of the form

$$\{X(i-1), a_1(i), a_2(i+1), a_3(i+2), \dots, a_{k-i}(k-1), Y(k)\}$$

with $a_1 a_2 \dots a_{k-i} \in \boldsymbol{L}(\mathfrak{A}_\mathcal{L})$. As $\boldsymbol{L}(\mathfrak{A}_\mathcal{L})$ is definable by an $\mathcal{L}$-formula this condition is also $\mathcal{L}$-definable. ❑

Theorem 16 is a direct consequence of Lemma 18 and the properties of $\mathfrak{A}_\mathcal{L}$. ❑

▶ **Theorem 19.** *For any $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<, \equiv), \mathsf{FO}(<, \mathsf{MOD})\}$, deciding $\mathcal{L}$-rewritability of Boolean and specific $LTL_{krom}^{\bigcirc}$ OMPEQs over $\Xi$-ABoxes is ExpSpace-complete.*

**Proof.** The upper bound follows from Proposition 5 and Theorem 8. We show the matching lower bound by reduction of $LTL_{horn}^{\bigcirc}$ OMAQs to $LTL_{krom}^{\bigcirc}$ OMPEQs and using Theorem 16. Consider an $LTL_{horn}^{\bigcirc}$ OMAQ $\boldsymbol{q} = (\mathcal{O}, A)$. We can assume that all of the axioms in $\mathcal{O}$ take the form $\boldsymbol{C} \to \bot$ or $\boldsymbol{C} \to B$, for some $\boldsymbol{C} = C_1 \wedge \dots \wedge C_n$ and an atomic concept $B$. We construct an $LTL_{krom}^{\bigcirc}$ OMPQ $\boldsymbol{q}' = (\mathcal{O}', \varkappa)$ that is $\mathcal{L}$-rewritable over $\Xi$-ABoxes iff $\boldsymbol{q}$ is

$\mathcal{L}$-rewritable. Using the atomic concepts $\{B, \bar{B} \mid B \in \mathsf{sig}(\boldsymbol{q})\}$, we define $\mathcal{O}'$ to contain the axioms $B \wedge \bar{B} \to \bot$ and $\top \to B \vee \bar{B}$, for all $B \in \mathsf{sig}(\boldsymbol{q})$, and set

$$\varkappa \;=\; A \;\;\vee \bigvee_{\boldsymbol{C}\to\bot \text{ in } \mathcal{O}} \Diamond_F \Diamond_P \boldsymbol{C} \;\;\vee \bigvee_{\boldsymbol{C}\to B \text{ in } \mathcal{O}} \Diamond_F \Diamond_P (\boldsymbol{C} \wedge \bar{B}).$$

It is readily seen that, for any $\Xi$-ABox $\mathcal{A}$, the certain answer to $\boldsymbol{q}$ over $\mathcal{A}$ is yes iff the answer to $\boldsymbol{q}'$ over $\mathcal{A}$ is yes, and $k$ is a certain answer to $\boldsymbol{q}(x)$ over $\mathcal{A}$ iff it is also a certain answer to $\boldsymbol{q}'(x)$. It follows that $\boldsymbol{q}'$ is $\mathcal{L}$-rewritable over $\Xi$-ABoxes iff $\boldsymbol{q}$ is $\mathcal{L}$-rewritable. ❑

# 6 Deciding $\mathcal{L}$-rewritability of linear positive $LTL^{\bigcirc}_{horn}$ OMQs

As well known, deciding FO-rewritability of (classical) monadic datalog queries is 2ExpTime-complete [12, 24], which goes down to PSpace-complete for the important class of linear monadic queries [24, 54].

In this section, we focus on linear $LTL^{\bigcirc}_{horn}$ OMPQs. First, in Section 6.1, for any linear $LTL^{\bigcirc}_{horn}$ OMAQ $\boldsymbol{q}$, we construct in polynomial space a DFA $\mathfrak{A}'$ such that $\boldsymbol{q}$ is $\mathcal{L}$-rewritable iff $\boldsymbol{L}(\mathfrak{A}')$ is $\mathcal{L}$-definable, for any $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<,\equiv), \mathsf{FO}(<,\mathsf{MOD})\}$. So, by Theorem 11, deciding $\mathcal{L}$-rewritability of linear $LTL^{\bigcirc}_{horn}$ OMAQs $\boldsymbol{q}$ can be done in PSpace. An essential part of this proof is the construction of a (polynomial-size) 2NFA $\mathfrak{A}^{\Xi}_{\boldsymbol{q}}$ that recognises a certain encoding of the language of $\boldsymbol{q}$. Further in this section, we show that any DFA can be simulated by a linear $LTL^{\bigcirc}_{horn}$ OMAQ, which gives a PSpace lower bound for deciding $\mathcal{L}$-rewritability. In Section 6.2, we give semantic criteria of $\mathcal{L}$-rewritiability, for $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<,\equiv)\}$, of $LTL^{\bigcirc}_{horn}$ OMPQs and a PSpace algorithm for checking their $\mathcal{L}$-rewritability, which is based on the 2NFA $\mathfrak{A}^{\Xi}_{\boldsymbol{q}}$.

## 6.1 Linear OMAQs

▶ **Theorem 20.** *For any $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<,\equiv), \mathsf{FO}(<,\mathsf{MOD})\}$, deciding $\mathcal{L}$-rewritability of linear $LTL^{\bigcirc}_{horn}$ OMAQs over $\Xi$-ABoxes can be done in PSpace.*

**Proof.** By $(i)$ of Lemma 14 and Proposition 15, it suffices to prove this result for Boolean OMAQs in the given class without occurrences of $\bot$. Let $\boldsymbol{q} = (\mathcal{O}, B)$ be such an OMAQ and $\Xi$ a signature. By possibly adding new IDB predicates, we convert $\mathcal{O}$ to the form with axioms of two types:

$(\varrho_1)$ $C_1 \wedge \dots \wedge C_k \to A'$,

$(\varrho_2)$ $C_1 \wedge \dots \wedge C_k \wedge \bigcirc^i A \to A'$,

where $k \geq 0$, $C_1, \dots, C_k$ contain no IDB atomic concepts, $A, A' \in \mathit{idb}(\mathcal{O})$, $i \in \{-1, 0, 1\}$, and

$$\bigcirc^j A = \begin{cases} A, & \text{if } j = 0,\\ \underbrace{\bigcirc_P \dots \bigcirc_P}_{j} A, & \text{if } j < 0,\\ \underbrace{\bigcirc_F \dots \bigcirc_F}_{j} A, & \text{if } j > 0.\end{cases}$$

First, we define a quadruple $\mathfrak{A}^{\Xi}_{\mathcal{O}} = (2^{\Xi}, Q, \{q_0\}, \delta)$ (which is in essence a 2NFA without final states), where *the set of states* $Q = \bigcup_{\varrho \in \mathcal{O}} Q_\varrho \cup \{q_0, q_h\} \cup \{q_A \mid A \in \mathit{idb}(\mathcal{O})\}$, $Q_0 = \{q_0\}$, and the *transition function* $\delta = \bigcup_{\varrho \in \mathcal{O}} \delta_\varrho \cup \{q_0 \to_{a,1} q_0 \mid a \in 2^{\Xi}\}$, where $Q_\varrho$ and $\delta_\varrho$ are defined as follows. If $\varrho$ is of the form $(\varrho_1)$ and $C_i = \bigcirc^{j_i} A_i$, $1 \leq i \leq k$, then $Q_\varrho = \{q_\varrho\} \cup Q'_\varrho$ and $\delta_\varrho = \{q_0 \to_{a,0} q_\varrho \mid a \in 2^{\Xi}\} \cup \delta'_\varrho$, where $Q'_\varrho$ and $\delta'_\varrho$ are described below. If $j_1 < 0$ (the cases

$j_1 = 0$ and $j_1 > 0$ are analogous), then $\delta'_\varrho$ is such that $\mathfrak{A}^{\Xi}_{\mathcal{O}}$ makes $j_1 - 1$ steps to the left, by reading any symbols from $2^{\Xi}$. After that, if we read any symbol $a$ with $A_1 \notin a$, $\mathfrak{A}^{\Xi}_{\mathcal{O}}$ comes to a fixed 'dead-end' state $q_h$. Otherwise, it makes $j_1 - 1$ steps to the right (i.e., to where it was originally before executing any transitions for $i = 1$) and repeats the same process for $C_2 = \bigcirc^{j_2} A_2$, etc. After executing the transitions for $C_k = \bigcirc^{j_k} A_k$ and provided that $q_h$ was avoided, we come to the state $q_{A'}$. If $\varrho$ is of the form ($\varrho_2$), then $Q_\varrho$ is the same as above and $\delta_\varrho = \{q_A \to_{a,0} q_\varrho \mid a \in 2^{\Xi}\} \cup \delta'_\varrho$ is the same as above, finishing in either $q_h$ or $q_{A'}$.

By an *atomic type* $\upsilon_{\mathcal{O}}$ *for* $\mathcal{O}$, we mean a restriction of some type $\tau$ for $\mathcal{O}$ (see Proposition 5) to atomic concepts (or their negations). Given a model $\mathcal{I}$ of $\mathcal{O}$, we denote by $\upsilon_{\mathcal{I},\mathcal{O}}(n)$, for $n \in \mathbb{Z}$, the atomic type for $\mathcal{O}$ that holds in $\mathcal{I}$ at $n$. We omit $\mathcal{I}$ from $\upsilon_{\mathcal{I},\mathcal{O}}(n)$ when it is clear from the context. Recall that $\mathcal{C}_{\mathcal{O},\mathcal{A}}$ denotes the canonical model of $\mathcal{O}$ and $\mathcal{A}$, which exists because $\mathcal{O}$ is $\bot$-free. Let $N = M + 2M^2$, where $M$ is the number of occurrences of $\bigcirc_F$ and $\bigcirc_P$ in $\mathcal{O}$.

▶ **Lemma 21.** *Let $\mathcal{A}$ be any ABox of the form $\emptyset^N \mathcal{B} \emptyset^N$ and $\mathcal{O}$ a linear LTL$^{\bigcirc}_{horn}$ ontology. Then we have: $A \in \upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{A}}}(\ell)$ iff there exists a run $(q_0, 0), \dots, (q, \ell), (q_A, i)$ of $\mathfrak{A}^{\Xi}_{\mathcal{O}}$ on $\mathcal{A}$, for all $N \le \ell < |\mathcal{A}| - N$.*

**Proof.** We call a sequence $\mathfrak{D}$ of the form

$$(C^0_1 \wedge \dots \wedge C^0_{k_0} \to A_1, n_1), (C^1_1 \wedge \dots \wedge C^1_{k_1} \wedge \bigcirc^{i_1} A_1 \to A_2, n_2), \dots,$$
$$(C^m_1 \wedge \dots \wedge C^m_{k_m} \wedge \bigcirc^{i_m} A_m \to A, n_{m+1}) \quad (28)$$

a *derivation* of $A$ from $\mathcal{O}$ and $\mathcal{A}$ if the axioms are from $\mathcal{O}$ and the numbers $n_1, \dots, n_m, n_{m+1}$ are such that $n_{j+1} = n_j + i_j$ and $\mathcal{A} \models C^j_1 \wedge \dots \wedge C^j_{k_j}(n_{j+1})$. We say that such a derivation ends at $n$ if $n_{m+1} = n$. It is straightforward to verify that $A \in \upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{A}}}(\ell)$ iff there is a derivation of $A$ at $\ell$, for any $\ell \in \mathbb{Z}$.

Let $\mathcal{A}$ be of the form $\emptyset^N \mathcal{B} \emptyset^N$. Our next aim is to prove that ($a$) for any $N \le \ell < |\mathcal{A}| - N$, if is a derivation of $A$ at $\ell$, then there is a derivation (28) of $A$ at $\ell$ such that $0 \le n_j < |\mathcal{A}|$, for all numbers $n_j$ in this derivation.

▶ **Proposition 22.** *Let $\mathfrak{D}_1$, $\mathfrak{D}_2$, $\mathfrak{D}_3$ be derivations from $\mathcal{O}$ and $\mathcal{A}$ of the form:*

$$\mathfrak{D}_1 = \dots, (C_1 \wedge \dots \wedge C_k \wedge \bigcirc^i A \to A_0, n_0),$$
$$\mathfrak{D}_2 = (\bigcirc^{i_0} A_0 \to A_1, n_1), \dots, (\bigcirc^{i_{m-1}} A_{m-1} \to A_m, n_m),$$
$$\mathfrak{D}_3 = (C'_1 \wedge \dots \wedge C'_{k'} \wedge \bigcirc^i A_m \to A_{m+1}, n_{m+1}), \dots$$

*If $\mathfrak{D}_1\mathfrak{D}_2\mathfrak{D}_3$ is a derivation of $A$ at $\ell$, then there exists a derivation $\mathfrak{D}_1\mathfrak{D}'_2\mathfrak{D}_3$ of $A$ at $\ell$ from $\mathcal{O}$ and $\mathcal{A}$ such that* $\min\{n_0, n_{m+1}\} - 2M^2 \le n_j \le \max\{n_0, n_{m+1}\} + 2M^2$ *for all numbers $n_j$ in $\mathfrak{D}'_2$.*

**Proof.** Suppose $n_{m+1} > n_0$ (the opposite case is analogous). Let $j$ be the earliest number in $\mathfrak{D}_2$ such that

- either $n_j = n_{m+1}$ and $n_{j+k} = n_{m+1}$ for some $k \ge 0$,
- or $n_j = n_0$ and $n_{j+k} = n_0$ for some $k \ge 0$.

If such $j$ does not exist, then clearly, (d) holds with $\mathfrak{D}'_2 = \mathfrak{D}_2$ and we are done. Suppose the former case holds for the earliest $j$. Let $\mathfrak{D}_2 = \mathfrak{D}_4\mathfrak{D}_5\mathfrak{D}_6$, where $\mathfrak{D}_5$ is the subsequence of $\mathfrak{D}_2$ between $j$ (not inclusive) and $j + k$. In $\mathfrak{D}_5$, consider any quadruple $((A_{j'}, n_{j'}), (A_{j''}, n_{j''}), (A_{k''}, n_{k''}), (A_{k'}, n_{k'}))$ such that $j' \le j'' \le k'' \le k'$, $n_{j'} = n_{k'}$,

$n_{j''} = n_{k''}$, $A_{j'} = A_{j''}$ and $A_{k'} = A_{k''}$. Clearly, $\mathfrak{D}_1(\mathfrak{D}_4\mathfrak{D}'_5\mathfrak{D}_6)\mathfrak{D}_3$ is also a derivation $L$ at $\ell$ from $\mathcal{O}$ and $\mathcal{A}$, where

$$\begin{aligned}\mathfrak{D}'_5 =& (\bigcirc^{i_j} A_j \to A_{j+1}, n_{j+1}), \dots, (\bigcirc^{i_{j'}-1} A_{j'-1} \to A_{j'}, n_{j'}), (\bigcirc^{i_{j''}} A_{j''} \to A_{j''+1}, n_{j''+1} - d), \dots\\ & (\bigcirc^{i_{k''}-1} A_{k''-1} \to A_{k''}, n_{k''} - d), (\bigcirc^{i_{k'}} A_{k'} \to A_{k'+1}, n_{k'+1}), \dots,\\ & (\bigcirc^{i_{j+k-1}} A_{j+k-1} \to A_{j+k}, n_{j+k}),\end{aligned}$$

and $d = n_{j''} - n_{j'}$. After recursively applying to $\mathfrak{D}_5$ the transformation above for each quadruple $((A_{j'}, n_{j'}), (A_{j''}, n_{j''}), (A_{k''}, n_{k''}), (A_{k'}, n_{k'}))$ as above, we obtain $\mathfrak{D}'_5$. It is easy to check that there exist no $n_1 \neq n_2$ and atoms $A, B$ such that both $(\bigcirc^{i_1} A_1 \to A, n_1), \dots, (\bigcirc^{i_2} A_2 \to B, n_1)$ and $(\bigcirc^{i_3} A_3 \to A, n_2), \dots, (\bigcirc^{i_4} A_4 \to B, n_2)$ are in $\mathfrak{D}'_5$. Therefore, $|n_{j'} - n_{m+1}| \leq 2M^2$ for all numbers $n_{j'}$ in $\mathfrak{D}'_5$. If the latter case holds for the earliest $j$, analogously, we can transform the subsequence $\mathfrak{D}_5$ of $\mathfrak{D}_2$ between $j$ (not inclusive) and $j+k$ into the subsequence $\mathfrak{D}'_5$ with all numbers $|n_{j'} - n_0| \leq 2M^2$. Then, we find $j$ in $\mathfrak{D}_6$ satisfying one of the two cases above and transform $\mathfrak{D}_6$ analogously. We proceed until there are no more $j$ satisfying either of the two cases and the result $\mathfrak{D}'_2$ of transformation is, clearly, as required by the proposition. ❑

Now, to show $(a)$, consider a derivation $\mathfrak{D}$ of $A$ at $\ell$, for $N \leq \ell < |\mathcal{A}| - N$ with the numbers $n_j$. Take the first $n_j$, such that $n_j \geq |\mathcal{B}| + M$ or $n_j < 2M^2$. Suppose the former was the case. Since $\mathcal{A}_i = \emptyset$ for $|\emptyset^N \mathcal{B}| \leq i < |\mathcal{A}|$, it follows that there exists $n_{j'}$, for $j' < j$, such that $2M^2 \leq n_{j'} < |\mathcal{B}| + M$ and a (sub)sequence $(\bigcirc^{i_{j'}} A_{j'} \to A_{j'+1}, n_{j'+1}), \dots, (\bigcirc^{i_{j-1}} A_{j-1} \to A_j, n_j)$ is in $\mathfrak{D}$. We expand this subsequence by taking all $(\bigcirc^{i_j} A_j \to A_{j+1}, n_j), \dots, (\bigcirc^{i_{j''}-1} A_{j''-1} \to A_{j''}, n_{j''})$, such that $j''$ is the first after $j$ such that $n_{j''} = n_{j'}$. Let $\mathfrak{D} = \mathfrak{D}_1\mathfrak{D}_2\mathfrak{D}_3$, where $\mathfrak{D}_2$ is the expanded sequence above. By applying Proposition 22, we obtain a derivation $\mathfrak{D}_1\mathfrak{D}'_2\mathfrak{D}_3$ of $A$ at $\ell$, where all numbers $n_j$ in $\mathfrak{D}_1\mathfrak{D}'_2$ are $2M^2 \leq n_j \leq n_{j'} + 2M^2 < |\mathcal{A}|$. If the latter above was the case, i.e., $n_j < 2M^2$, we analogously obtain a derivation of $A$ at $\ell$, where all numbers $n_j$ in $\mathfrak{D}_1\mathfrak{D}'_2$ are $0 \leq n_{j'} - 2M^2 \leq n_j < |\mathcal{B}| + M$. By continuing to apply Proposition 22 to $D_3$ a required number of times, we obtain the derivation of $A$ at $\ell$ with all the numbers as required in $(a)$.

Now the proof of Lemma 21 is complete. Indeed, there is an immediate correspondence between runs of $\mathfrak{A}^{\Xi}_{\mathcal{O}}$ on $\mathcal{A}$ and derivations of $L$ by $\mathcal{O}$ and $\mathcal{A}$ whose all numbers $n_j$ are such that $0 \leq n_j < |\mathcal{A}|$. ❑

We now return to the proof of Theorem 20. Define a 2NFA $\mathfrak{A}^{\Xi}_{\boldsymbol{q}} = (2^{\Xi}, Q', Q_0, \delta', F)$, where $Q' = Q \cup \{q_1\}$, $F = \{q_1\}$, and $\delta' = \delta \cup \{q_B \to_{a,0} q_1, q_1 \to_{a,1} q_1 \mid a \in 2^{\Xi}\}$. Using Lemma 21, we obtain:

$$\boldsymbol{L}_{\Xi}(\boldsymbol{q}) = \{\boldsymbol{a} \in \Sigma^*_{\Xi} \mid \emptyset^N \boldsymbol{a} \emptyset^N \in \boldsymbol{L}(\mathfrak{A}^{\Xi}_{\boldsymbol{q}})\}. \tag{29}$$

However, we need an automaton $\mathfrak{A}'$, which can be constructed in polynomial space, such that $\boldsymbol{L}_{\Xi}(\boldsymbol{q}) = \boldsymbol{L}(\mathfrak{A}')$ and $\mathcal{L}$-definability of $\mathfrak{A}'$ can be decided in PSpace. Consider the DFA $\mathfrak{A}'$ from Section 4.2 that recognises the language of a 2NFA $\mathfrak{A}$. We construct $\mathfrak{A}'$ from $\mathfrak{A}^{\Xi}_{\boldsymbol{q}}$ as in that section except the definition of $q'_0$ and $F'$, which is now as follows: $q'_0 = (\{(q_0, q) \in \mathsf{b}_{rr}(\emptyset^N)\}, \mathsf{b}_{rr}(\emptyset^N))$ and $F' = \{(B_{lr}, B_{rr}) \mid (q_0, q_1) \in B_{lr} \circ X\}$, where $X$ is the reflexive and transitive closure of $\mathsf{b}_{ll}(\emptyset^N) \circ B_{rr}$. Using (29), it is easily shown that $\boldsymbol{L}_{\Xi}(\boldsymbol{q}) = \boldsymbol{L}(\mathfrak{A}')$ and $\mathfrak{A}'$ is clearly constructible from $\boldsymbol{q}$ in PSpace. That $\mathcal{L}$-definability of $\mathfrak{A}'$ is decidable in PSpace, follows from the proof of Theorem 11. ❑

▶ **Theorem 23.** *For any $\mathcal{L} \in \{\mathsf{FO}(<), \mathsf{FO}(<, \equiv), \mathsf{FO}(<, \mathsf{MOD})\}$, deciding $\mathcal{L}$-rewritability of linear $LTL^{\bigcirc}_{horn}$ OMAQs over $\Xi$-ABoxes is PSpace-complete.*

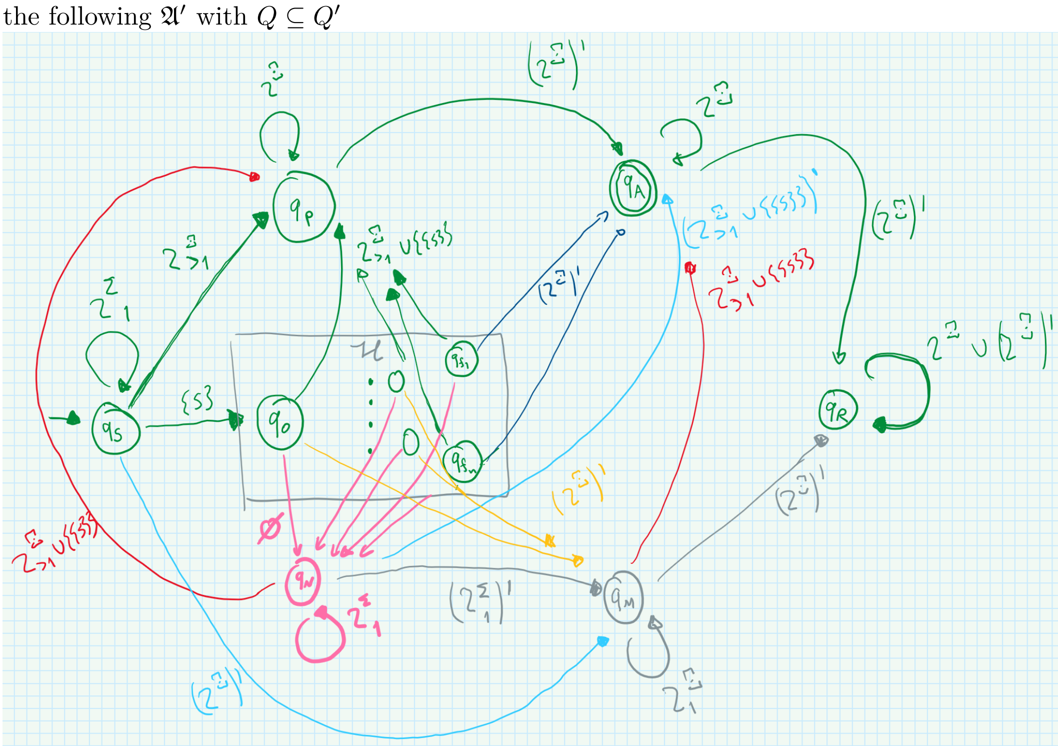
**Proof.** By Proposition 15 $(i)$, it is sufficient to show the lower bound result for *specific* linear $\mathit{LTL}^{\bigcirc}_{\mathit{horn}}$ OMAQs $\boldsymbol{q}(x) = (\mathcal{O}, A(x))$. We provide a reduction from the problem of deciding $\mathcal{L}$ rewritability of a DFA $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$. We set $\Xi = \Sigma \cup \{s\}$, for a fresh symbol $s$, and construct $\mathcal{O}$ with $\mathit{idb}(\mathcal{O}) \subseteq \{\bar{q} \mid q \in Q\} \cup \{A, X\}$ such that

$$\boldsymbol{L}(\mathfrak{A}) \text{ is } \mathcal{L}\text{-definable iff } \boldsymbol{L}_{\Xi}(\boldsymbol{q}(x)) \text{ is } \mathcal{L}\text{-definable.} \tag{30}$$

Let $\mathcal{O}$ contain the axioms $s \to \bigcirc_{\!F}\bar{q}_0$, $\bar{q} \to A$, for all $q \in F$, $\bar{q} \wedge a \to \bigcirc_{\!F}\bar{r}$, for all $q \to_a r$ in $\delta$, $a \wedge b \to \bot$ for all distinct $a, b \in \Xi$, and $s \to \bigcirc_{\!F}X$, $X \to \bigcirc_{\!F}X$, $X \wedge s \to \bot$. Let $2^{\Xi}_1$ be the set of all $B \in 2^{\Xi}$ with $|B| \leq 1$, i..e., $2^{\Xi}_1 = \{\emptyset\} \cup \bigcup_{a \in \Xi}\{\{a\}\}$, and let $2^{\Xi}_{>1}$ be $2^{\Xi} \setminus 2^{\Xi}_1$. We analogously define $2^{\Sigma}_1$ and $2^{\Sigma}_{>1}$. To prove (30), observe that (recall that the alphabet of $\boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$ is $2^{\Xi} \cup (2^{\Xi})'$):

$$\begin{aligned}\boldsymbol{L}_{\Xi}(\boldsymbol{q}(x)) = &\{U\{s\}\{u_0\}\dots\{u_n\}B'V \mid U, V \in (2^{\Sigma}_1)^*,\ u_0 \dots u_n \in \boldsymbol{L}(\mathfrak{A}), B' \in (2^{\Sigma}_1)'\} \cup \\ &\{UB'V \mid |U_i| > 1 \text{ for some } i, |B| > 1, \text{ or } |V_i| > 1 \text{ for some } i\} \cup \\ &\{UB'V \mid s \text{ occurs at distinct positions of } UB'V\}\end{aligned}$$

We now construct a DFA $\mathfrak{A}'$ such that $\boldsymbol{L}(\mathfrak{A}') = \boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$. It is straightforward to verify that the following $\mathfrak{A}'$ with $Q \subseteq Q'$



can be taken. (The grey box shows $Q$, where the state $q_0$ is initial in $\mathfrak{A}$ and the states $q_{f_1}, \dots,$ $q_{f_n}$ are final. The transitions between $q, q' \in Q$ in $\mathfrak{A}'$ are defined by taking $q \to_{\{a\}} q'$ iff $q \to_a q'$ in $\mathfrak{A}$, while all the other transitions in $\mathfrak{A}'$ are shown in the picture. As usual, when an arrow is marked by a sets of symbols from $2^{\Xi} \cup (2^{\Xi})'$, the corresponding transition holds for each symbol in the set.) We also observe that:

$$q \sim q' \text{ in } \mathfrak{A} \text{ iff } q \sim q' \text{ in } \mathfrak{A}', \text{ for all } q, q' \in Q, \tag{31}$$

$$q \not\sim q' \text{ in } \mathfrak{A}', \text{ for all } q \in Q,\ q' \in Q' \setminus Q. \tag{32}$$

We now show that $\boldsymbol{L}(\mathfrak{A})$ is $\mathcal{L}$-definable iff $\boldsymbol{L}(\mathfrak{A}')$ is $\mathcal{L}$-definable. We prove the direction $(\Rightarrow)$, while the opposite direction is easier and left to the reader. Let first $\mathcal{L} = \mathsf{FO}(<)$ and suppose $\boldsymbol{L}(\mathfrak{A}')$ is not $\mathsf{FO}(<)$ definable. By Theorem 6 $(i)$, there exists a reachable state $q$ in $\mathfrak{A}'$, a word $U \in (2^\Xi \cup (2^\Xi)')^*$ and $k$, satisfying the corresponding conditions. By the structure of $\mathfrak{A}'$, it is clear that the state $q$ is in $Q$ and $U = \{u_0\} \dots \{u_n\}$, for some $u \in \Sigma^*$, and $\delta'_{U^i}(q) \in Q$, for all $i \le k$. Therefore, we have $q$ in $\mathfrak{A}$ such that $\delta_{u^k}(q) = q$. By (31), it also follows that $q \not\sim \delta_u(q)$ in $\mathfrak{A}$, and so $\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<)$-definable. The proof for $\mathcal{L} = \mathsf{FO}(<, \equiv)$ is analogous and left to the reader. Let now $\mathsf{FO}(<, \mathsf{MOD})$ and suppose $\boldsymbol{L}(\mathfrak{A}')$ is not $\mathsf{FO}(<, \mathsf{MOD})$ definable. By Theorem 6 $(iii)$, there exists a reachable state $q$ in $\mathfrak{A}'$ and $U, V \in (2^\Xi \cup (2^\Xi)')^*$ such that the corresponding conditions are satisfied. Consider the sequence of states $q, \delta'_U(q), \delta'_{U^2}(q), \dots$ and observe $\delta'_{U^i}(q) \sim \delta'_{U^{i+2}}(q)$ and $\delta'_{U^i}(q) \not\sim \delta'_{U^{i+1}}(q)$ (in $\mathfrak{A}'$), for all $i \ge 0$. By the structure of $\mathfrak{A}'$ and (32), it follows that all $\delta'_{U^i}(q)$ are in $Q$ and $U = \{u_0\} \dots \{u_n\}$, for some $u \in \Sigma^*$. Also, because $q \sim \delta'_{V^k}(q) \sim \delta'_{(UV)^l}(q)$ and (32), it follows that $\delta'_V(q), \delta'_{(UV)}(q) \in Q$ and $V = \{v_0\} \dots \{v_m\}$, for some $v \in \Sigma^*$. Finally, using (31) and an observation that $\delta'_X(q) = \delta_x(q)$, for all words $X = \{x_0\} \dots \{x_n\}$ and $x \in \Sigma^*$, we conclude that $\mathfrak{A}$ satisfies condition $(iii)$ of Theorem 6, and so $\boldsymbol{L}(\mathfrak{A})$ is not $\mathsf{FO}(<, \mathsf{MOD})$-definable. ❑

## 6.2 Linear OMPQs

By Lemma 14 and Proposition 15, it suffices to prove this result for Boolean OMPQs in the given class without occurrences of $\bot$. Let $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ be a such an OMPQ. We start with the criterion and algorithm for $\mathsf{FO}(<)$-definability, and address $\mathsf{FO}(<, \equiv)$-definability after. The set of all types for $\boldsymbol{q}$ is denoted by $\boldsymbol{T_q}$. Given a model $\mathcal{I}$ of $\mathcal{O}$, we denote by $\tau_{\mathcal{I}}(n)$, for $n \in \mathbb{Z}$, the type for $\boldsymbol{q}$ that holds in $\mathcal{I}$ at $n$. In the rest of this section, we assume and $\varkappa$ of the form $\Diamond_P \Diamond_F \varkappa'$. This is w.l.o.g. by (26).

▶ **Lemma 24.** *Let $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ be an OMPQ with an $LTL^{\Box\bigcirc}_{horn}$-ontology $\mathcal{O}$. Then $\boldsymbol{q}$ is not $\mathsf{FO}(<)$-rewritable over $\Xi$-Aboxes iff there exist such ABoxes $\mathcal{A}$, $\mathcal{B}$, $\mathcal{D}$ and $k \ge 2$ such that the following conditions hold:*
*$(i)$ $\neg\varkappa \in \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}| - 1)$ and $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}\mathcal{B}^k| - 1)$;*
*$(ii)$ $\varkappa \in \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}| - 1)$ and $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}^{k+1}| - 1)$.*

**Proof.** Consider the DFA $\mathfrak{A}$ over the alphabet $2^\Xi$ with the set of states $Q = 2^{\boldsymbol{T_q}}$, where $q_{-1} = \boldsymbol{T_q}$ is the initial state and the set of final states is $F = \{q \mid \varkappa \in \tau, \text{ for all } \tau \in q\}$. We expand the relation $\to_a$ defined on $\boldsymbol{T_q}$ in Proposition 5 to $Q$ by setting $\delta(q, a) = \{\tau \mid \tau' \to_a \tau \text{ for some } \tau \in q\}$. Clearly, $\mathfrak{A}$ is deterministic. In fact, $\mathfrak{A}$ is a determinasation of the NFA used in Proposition 5 with some simplifications. We write $q \Rightarrow_{\mathcal{A}} q'$ to say that, having started in state $q$ and having read an ABox $\mathcal{A}$, the DFA $\mathfrak{A}$ is in state $q'$. We observe the following important property of $\mathfrak{A}$. Let $q_{-1} \Rightarrow_{\mathcal{A}_0} q_0 \dots q_{n-1} \Rightarrow_{\mathcal{A}_n} q_n$ be a run of $\mathfrak{A}$ on $\mathcal{A} = \mathcal{A}_0 \dots \mathcal{A}_n$, and let $\bar{q}_i = \{\tau \in q_i \mid \tau \to_{\mathcal{A}_{i+1} \dots \mathcal{A}_n} \tau', \text{ for some } \tau' \in q_n\}$. Then

$$\tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}}}(i) = \bigcap \bar{q}_i, \text{ for } -1 \le i \le n. \tag{33}$$

Similarly to the proof of Proposition 5, one can check that $\boldsymbol{L}_\Xi(\boldsymbol{q}) = \boldsymbol{L}(\mathfrak{A})$.

($\Rightarrow$) Suppose $\boldsymbol{q}$ is not $\mathsf{FO}(<)$-rewritable. By Lemma 6 $(i)$, it follows that there exist ABoxes $\mathcal{A}$, $\mathcal{B}$, $\mathcal{D}$ and $k \ge 2$ such that $q_{-1} \Rightarrow_{\mathcal{A}} q_0$, $q_0 \Rightarrow_{\mathcal{B}} q_1$, $q_0 \Rightarrow^k_{\mathcal{B}} q_0$ and $q_0 \Rightarrow_{\mathcal{D}} q'_0$, $q_1 \Rightarrow_{\mathcal{D}} q'_1$, for some $q'_0, q'_1 \in Q$ such that $q'_0 \notin F$ and $q'_1 \in F$. Since $q'_0 \notin F$, by (33), we have $\neg\varkappa \in \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}\mathcal{B}^k| - 1)$ as required in $(i)$. To show $(ii)$, we observe that $q'_1 \in F$ by (33) implies $\varkappa \in \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}^{k+1}| - 1)$, as required.

($\Leftarrow$) Assuming $(i)$ and $(ii)$, let $q_0, q_1, q_2$ be states in $\mathfrak{A}$ with $q_{-1} \Rightarrow_{\mathcal{A}} q_0 \Rightarrow_{\mathcal{B}} q_1 \Rightarrow_{\mathcal{B}^{k-1}} q_2 \Rightarrow_{\mathcal{B}} q'_2$. Let $q_3, q'_3$ be such that $q_2 \Rightarrow_{\mathcal{D}} q_3$ and $q'_2 \Rightarrow_{\mathcal{D}} q'_3$. It follows by (33) that $q_3 \notin F$ and

$q_3' \in F$. Observe that, if we had $q_0 = q_2$, we could conclude that $\boldsymbol{q}$ is not $\mathsf{FO}(<)$-rewritable, as the conditions of aperiodicity for $\mathfrak{A}$ (see the proof of ($\Rightarrow$)) would be satisfied. Since we are not guaranteed that, we use the following property of the canonical models that follow from ($i$) and ($ii$): (a) $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{AB}^k|-1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{kj}\mathcal{D}}}(|\mathcal{AB}^{kj}|-1)$, for any $j \geq 1$; (b) $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}}(|\mathcal{AB}^{k+1}|-1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{kj+1}\mathcal{D}}}(|\mathcal{AB}^{kj+1}|-1)$, for any $j \geq 1$. Take some $i, j \geq 1$ that satisfy $q_0 \Rightarrow_{\mathcal{AB}^{ki}} q_4 \Rightarrow_{\mathcal{B}} q_4' \Rightarrow_{\mathcal{B}^{kj}} q_4 \Rightarrow_{\mathcal{B}} q_4'$, for some $q_4, q_4'$. By ($i$), ($ii$), (a) and (b), we have that $q_5 \notin F$ and $q_5' \in F$ for such $q_5$ and $q_5'$ that $q_4 \Rightarrow_{\mathcal{D}} q_5$ and $q_4' \Rightarrow_{\mathcal{D}} q_5'$. Therefore, $\boldsymbol{q}$ is not $\mathsf{FO}(<)$-rewritable, as the conditions of aperiodicity for $\mathfrak{A}$ are satisfied (as in the ($\Rightarrow$)-proof with $\mathcal{A}$, $\mathcal{B}$, $\mathcal{D}$ and $k$ being, respectively, $\mathcal{AB}^{ki}$, $\mathcal{B}$, $\mathcal{D}$ and $kj$). ❏

► **Corollary 25.** *Let $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ be an OMPQ with an $LTL^{\Box\bigcirc}_{horn}$-ontology $\mathcal{O}$. If there exist $\Xi$-ABoxes $\mathcal{A}, \mathcal{B}, \mathcal{D}$ and $k \geq 2$ satisfying conditions ($i$) and ($ii$) above, then there exist $\mathcal{A}, \mathcal{B}, \mathcal{D}$ and $k$ with $|\mathcal{A}|, |\mathcal{D}|, k \leq 2^{O(|\boldsymbol{q}|)}$ satisfying these conditions.*

**Proof.** First, we show that there is $\mathcal{A}$ with $|\mathcal{A}| \leq 2|\boldsymbol{T_q}|^2$. Indeed, consider the sequence

$$(\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}(0), \mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}(0)), \dots, (\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}(|\mathcal{A}|-2), \mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}(|\mathcal{A}|-2)).$$

Suppose, the $i$-th member of this sequence is equal to its $j$-th member, for $i < j$, and denote $\mathcal{A}^{<i}\mathcal{A}^{\geq j}$ by $\mathcal{A}'$. We clearly have $\mathcal{C}_{\mathcal{O},\mathcal{A}'\mathcal{B}^k\mathcal{D}}(|\mathcal{A}'|-1) = \mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}(|\mathcal{A}|-1)$ and $\mathcal{C}_{\mathcal{O},\mathcal{A}'\mathcal{B}^{k+1}\mathcal{D}}(|\mathcal{A}'\mathcal{B}|-1) = \mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}(|\mathcal{AB}|-1)$, and conditions ($i$) and ($ii$) are satisfied with $\mathcal{A}'$ in place of $\mathcal{A}$. The rest of the argument is straightforward. Similarly it is shown that there exists $\mathcal{D}$ with $|\mathcal{D}| \leq 2|\boldsymbol{T_q}|^2$. To show that there exists $k \leq 2|\boldsymbol{T_q}|^2$, we consider the sequence

$$(\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}(|\mathcal{AB}|-1), \mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}(|\mathcal{AB}^2|-1)), \dots,$$
$$(\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}(|\mathcal{AB}^{k-1}|-1), \mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}(|\mathcal{AB}^k|-1)).$$

Clearly, if the $i$-th member of this sequence is equal to its $j$-th member, for $i < j$, then conditions ($i$) and ($ii$) are satisfied with $k - (j - i)$ in place of $k$. ❏

Observe that we do not claim that there exists $\mathcal{B}$ with $|\mathcal{B}| \leq 2^{O(|\boldsymbol{q}|)}$ However, this is the case for *linear $LTL^{\Box\bigcirc}_{horn}$*-ontologies, as follows from the proof of Theorem 27.

Let $\mathcal{O}$ be in normal form, as in the proof of Theorem 20. Consider the 2NFA $\mathfrak{A}^{\Xi}_{\mathcal{O}}$ from that proof. Throughout this section, $\mathsf{b}_\bullet$, for $\bullet \in \{lr, rr, rl, ll\}$, and $\mathsf{b}$ are defined with respect to $\mathfrak{A}^{\Xi}_{\mathcal{O}}$. It will be convenient to define each $\mathsf{b}_\bullet(w)$ as an identity relation on $Q$, for the empty string $w$, and $\mathsf{b}(w)$ is defined accordingly.

► **Lemma 26.** *Let $\mathcal{A}$ be an ABox of the form $\emptyset^N\mathcal{B}\emptyset^N$ and $\mathcal{O}$ a linear $LTL^{\Box\bigcirc}_{horn}$-ontology. Let $X(\ell)$ be the reflexive and transitive closure of $\mathsf{b}_{ll}(\mathcal{A}^{>\ell}) \circ \mathsf{b}_{rr}(\mathcal{A}^{\leq\ell})$. Then $\upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{A}}}(\ell) = \{A \mid (q_0, A) \in \mathsf{b}_{lr}(\mathcal{A}^{\leq\ell}) \circ X(\ell)\}$, for any $N \leq k < |\mathcal{A}| - N$.*

**Proof.** Easily follows from Lemma 21. Observe that there exists a run $(q_0, 0), \dots, (q, \ell), (q_L, i)$ of $\mathfrak{A}^{\Xi}_{\mathcal{O}}$ on $\mathcal{A}$ iff $(q_0, q_L) \in \mathsf{b}_{lr}(\mathcal{A}^{\leq\ell}) \circ X(\ell)$, for all $\ell < |\mathcal{A}|$. ❏

► **Theorem 27.** *Deciding $\mathsf{FO}(<)$-rewritability of OMPQs $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ with a linear $LTL^{\bigcirc}_{horn}$-ontology $\mathcal{O}$ over $\Xi$-ABoxes can be done in PSPACE.*

**Proof.** By Theorem 24 and Corollary 25, we need to check the existence of $\mathcal{A}, \mathcal{B}, \mathcal{D}$, $k \geq 2$, such that $|\mathcal{A}|, |\mathcal{D}|, k \leq 2^{O(|\boldsymbol{q}|)}$ and conditions ($i$) and ($ii$) hold. Without loss of generality, we assume that $\mathcal{A}$ has a prefix $\emptyset^N$ and $\mathcal{D}$ has a suffix $\emptyset^N$.

We start by guessing numbers $N_{\mathcal{A}} = |\mathcal{A}|$, $N_{\mathcal{D}} = |\mathcal{D}|$ and $k$. We guess two types $\tau_0$ and $\tau_1$ that represent, respectively, $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(N)$ and $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{A}|-1)$, and three types $\tau_0'$, $\tau_0''$, $\tau_1'$

that represent, respectively, $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}}(N)$, $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}}(|\mathcal{A}|-1)$, and $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}}(|\mathcal{AB}|-1)$. Next, we compute $\mathsf{b}(\emptyset^N)$ and guess $\mathsf{b}(\mathcal{A})$, $\mathsf{b}(\mathcal{B})$, $\mathsf{b}(\mathcal{D})$. Note that, given $\mathsf{b}(\mathcal{B})$, we are able to compute $\mathsf{b}(\mathcal{X})$ for each $\mathcal{X} \in \{\mathcal{B}^i \mid 1 \leq i \leq k+1\}$. Now, we guess $\mathcal{A}$—symbol by symbol—by means of a sequence of pairs

$$(\mathsf{b}(\mathcal{A}^{\leq 0}), \mathsf{b}(\mathcal{A}^{>0})), \dots, (\mathsf{b}(\mathcal{A}^{\leq N_{\mathcal{A}}-1}), \mathsf{b}(\mathcal{A}^{>N_{\mathcal{A}}-1}))$$

such that $\mathsf{b}(\mathcal{A}^{\leq i})\cdot\mathsf{b}(\mathcal{A}^{>i}) = \mathsf{b}(\mathcal{A})$, for all $i$, and there are $a_i \in 2^{\Xi}$ with $\mathsf{b}(\mathcal{A}^{\leq i+1}) = \mathsf{b}(\mathcal{A}^{\leq i})\cdot\mathsf{b}(a_i)$ and $\mathsf{b}(\mathcal{A}^{>i}) = \mathsf{b}(a_i)\cdot\mathsf{b}(\mathcal{A}^{>i+1})$. Moreover, we require that $a_i = \emptyset$ for $i < N$. Observe that the pairs of the sequence with $i \geq N$ together with $\mathsf{b}(\mathcal{B})$ and $\mathsf{b}(\mathcal{D})$, by Lemma 26, give us $\upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i)$. When we compute $\upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(N)$, we check whether it is subsumed by $\tau_0$ (if not, the algorithm terminates with an answer $\mathsf{no}$). Furthermore, we need to check the following condition:

$$\varkappa' \in \tau_{\mathcal{C}_{\mathcal{O},\{A(0) \mid A \in \tau_0\}}}(0) \quad \text{implies} \quad \varkappa' \in \tau_0,$$

for each $\varkappa'$ of the form $\Box_P \varkappa''$, $\Diamond_P \varkappa''$ from $\mathsf{sub}_{\boldsymbol{q}}$ (if not, the algorithm terminates and returns $\mathsf{no}$). We have now checked that the type $\tau_0$ is potentially guessed correctly (subject to further checks). We can apply the same method to check that $\tau_0'$ is potentially guessed correctly. For the remaining $N < i < N_{\mathcal{A}}$, since $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i)$ is determined by $\upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i)$ and $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i-1)$, we are able to compute $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{A}|-1)$ or obtain a conflict, e.g., $\Box_F A \in \tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i-1)$ and $\neg A \in \upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i)$. In the latter case, the algorithm terminates answering $\mathsf{no}$. In the former case, we check if $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{A}|-1)$ is equal to $\tau_1$, in which case $\tau_1$ is guessed correctly, and if not, the algorithm terminates answering $\mathsf{no}$. Analogously it is checked if $\tau_0''$ is guessed correctly using $\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}$.

Now, we show how to check that all the types $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i)$, for $|\mathcal{A}| \leq i < |\mathcal{AB}^k|$, are correct, that $\tau_1'$ is guessed correctly, and that all the types $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}}(i)$, for $|\mathcal{AB}| \leq i < |\mathcal{AB}^{k+1}|$ are correct. We only demonstrate the algorithm for $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i)$. Observe that $\varkappa' \in \tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(i)$ iff $\varkappa' \in \tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(j)$ iff $\varkappa' \in \tau_1$, for each $\varkappa'$ of the form $\Box\varkappa''$, $\Diamond\varkappa''$ from $\mathsf{sub}_{\boldsymbol{q}}$ and all $|\mathcal{A}|-1 \leq i,j < |\mathcal{AB}^k|$. To do the required check, we need to guess a sequence of pairs

$$(\mathsf{b}(\mathcal{B}^{\leq 0}), \mathsf{b}(\mathcal{B}^{>0})), \dots, (\mathsf{b}(\mathcal{B}^{\leq |\mathcal{B}|-1}), \mathsf{b}(\mathcal{B}^{>|\mathcal{B}|-1})) \tag{34}$$

such that $\mathsf{b}(\mathcal{B}^{\leq i}) \cdot \mathsf{b}(\mathcal{B}^{>i}) = \mathsf{b}(\mathcal{B})$, for all $i$, and there are $a \in 2^{\Xi}$ with $\mathsf{b}(\mathcal{B}^{\leq i+1}) = \mathsf{b}(\mathcal{B}^{\leq i}) \cdot \mathsf{b}(a)$ and $\mathsf{b}(\mathcal{B}^{>i}) = \mathsf{b}(a) \cdot \mathsf{b}(\mathcal{B}^{>i+1})$. While we do not have any bound on $|\mathcal{B}|$ yet (unlike on $|\mathcal{A}|$, $|\mathcal{D}|$ and $k$), we can easily observe that any sequence (34) with repeating members at positions $0 \leq i' < i'' \leq |\mathcal{B}|-1$ is equivalent for the purposes of this proof to the sequence with all the members $i', \dots, i''-1$ removed. Since there are $\leq 2^{O(|\boldsymbol{q}|)}$ distinct pairs as above, it follows that $|\mathcal{B}| \leq 2^{O(|\boldsymbol{q}|)}$, if $\mathcal{B}$ required by Lemma 24 exists. By Lemma 26, using an element $i$ of this sequence, we are able to compute $\upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{AB}^j|+i)$, for all $0 \leq j < k$. We only need to check that such an atomic type is not in conflict with the modal formulas in $\tau_1$, e.g., $\Box_P A \in \tau_1$ and $\neg A \in \upsilon_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{AB}^j|+i)$. If a conflict is detected for some $i$ and $j$, the algorithm terminates answering $\mathsf{no}$. Here, we also verify that $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{AB}^k|-1) = \tau_1$ (respectively, if $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}}(|\mathcal{AB}^{k+1}|-1) = \tau_1'$). Finally, we need to check that all the types $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^k\mathcal{D}}}(|\mathcal{AB}^k|+i)$ (respectively, in $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{AB}^{k+1}\mathcal{D}}}(|\mathcal{AB}^{k+1}|+i)$), are correct, for $0 \leq i < N_{\mathcal{D}}-N$. The details are left to the reader.

❑

We now turn to $\mathsf{FO}(<,\equiv)$-definability.

▶ **Lemma 28.** *Let $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ be an OMPQ with an $LTL_{horn}^{\Box\bigcirc}$-ontology $\mathcal{O}$. Then $\boldsymbol{q}$ is not $\mathsf{FO}(<,\equiv)$-rewritable over $\Xi$-Aboxes iff there exist such ABoxes $\mathcal{A}, \mathcal{B}, \mathcal{D}$ and $k \geq 2$, such*

*that (i) and (ii) from Lemma 24 hold and there exist ABoxes* $\mathcal{W}$, $\mathcal{U}$, *such that* $\mathcal{B} = \mathcal{U}\mathcal{W}$, $|\mathcal{W}| = |\mathcal{U}|$,
*(iii)* $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}\mathcal{B}^i| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}\mathcal{B}^i\mathcal{U}| - 1)$, *for all* $i < k$, *and*
*(iv)* $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}^i| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}^i\mathcal{U}| - 1)$, *for all* $i$, $1 \leq i \leq k$.

**Proof.** ($\Rightarrow$) Suppose $\boldsymbol{q}$ is not $\mathsf{FO}(<,\equiv)$-rewritable. By Theorem 6 $(ii)$, there exist the ABoxes $\mathcal{A}, \mathcal{W}, \mathcal{U}, \mathcal{D}$ with $|\mathcal{W}| = |\mathcal{U}|$ and $k \geq 2$ such that

$$q_{-1} \Rightarrow_{\mathcal{A}} q_0 \Rightarrow_{\mathcal{U}} q_0 \Rightarrow_{\mathcal{W}} q_1 \Rightarrow_{\mathcal{U}} q_1 \Rightarrow_{\mathcal{W}} \cdots \Rightarrow_{\mathcal{W}} q_{k-1} \Rightarrow_{\mathcal{U}} q_{k-1} \Rightarrow_{\mathcal{W}} q_0,$$

$q_0 \Rightarrow_{\mathcal{D}} r_0$, $q_1 \Rightarrow_{\mathcal{D}} r_1$ for some $r_0, r_1 \in Q$ such that $r_0 \notin F$. That $(i)$ and $(ii)$ are satisfied for $\mathcal{B} = \mathcal{U}\mathcal{W}$ is shown as in the proof of Lemma 24. Then $(iii)$ and $(iv)$ easily follow from (33).

($\Leftarrow$) Suppose $(i)$–$(iv)$ hold and $\mathcal{E}(i_0, \dots, i_j) = \mathcal{U}^{i_0}\mathcal{W} \dots \mathcal{U}^{i_j}\mathcal{W}$. Let $\mathcal{F}_{j'}(i_0, \dots, i_j)$ be the prefix of $\mathcal{E}(i_0, \dots, i_j)$ of the form $\mathcal{U}^{i_0}\mathcal{W} \dots \mathcal{U}^{i_{j'-1}}\mathcal{W}\mathcal{U}^{i_{j'}}$, for $j' \leq j$. By the properties of the canonical models, we then obtain the following, for $0 \leq n \leq m$ and $0 \leq \ell < k$:

**(a)** $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{E}(i_0,\dots,i_{km+k-1})\mathcal{D}}}(|\mathcal{A}\mathcal{F}_{kn+\ell}(i_0, \dots, i_{km+k-1})| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^k\mathcal{D}}}(|\mathcal{A}\mathcal{B}^\ell| - 1)$, for all $n, \ell \geq 0$;

**(b)** $\tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{E}(i_0,\dots,i_{km+k-1},i_0)\mathcal{D}}}(|\mathcal{A}\mathcal{F}_{kn+\ell+1}(i_0, \dots, i_{km+k-1}, i_0)| - 1) = \tau_{\mathcal{C}_{\mathcal{O},\mathcal{A}\mathcal{B}^{k+1}\mathcal{D}}}(|\mathcal{A}\mathcal{B}^{\ell+1}| - 1)$.

Take the DFA $\mathfrak{A}$ from the proof of Lemma 24, assume without loss of generality that $|Q| \geq 3$, and, for $m \geq 0$, consider the sequence

$$q_{-1} \Rightarrow_{\mathcal{A}\mathcal{U}^{|Q|!-1}} q_0 \Rightarrow_{\mathcal{U}^{|Q|!}} q_0' \Rightarrow_{\mathcal{W}} q_0'' \Rightarrow_{\mathcal{U}^{|Q|!-1}} q_1 \Rightarrow_{\mathcal{U}^{|Q|!}} q_1' \Rightarrow_{\mathcal{W}} q_1'' \Rightarrow_{\mathcal{U}^{|Q|!-1}} \dots$$
$$q_{km+k-1} \Rightarrow_{\mathcal{U}^{|Q|!}} q_{km+k-1}' \Rightarrow_{\mathcal{W}} q_{km+k}.$$

Clearly, $q_i = q_i'$ for $0 \leq i < km + k$. By taking an appropriate $m$, as in the proof of Lemma 24, we can find $i$ and $j$, such that

$$q_{-1} \Rightarrow_{\mathcal{A}\mathcal{U}^{|Q|!-1}(\mathcal{W}\mathcal{U}^{|Q|!-1})^{ik}} r_0 \Rightarrow_{\mathcal{W}\mathcal{U}^{|Q|!-1}} r_1 \Rightarrow_{\mathcal{W}\mathcal{U}^{|Q|!-1}} \cdots \Rightarrow_{\mathcal{W}\mathcal{U}^{|Q|!-1}} r_{jk+k-1} \Rightarrow_{\mathcal{W}\mathcal{U}^{|Q|!-1}} r_0$$

and $r_\ell \Rightarrow_{\mathcal{U}^{|Q|!}} r_\ell$, for $0 \leq \ell < jk + k$. It can be readily shown using $(a)$ and $(b)$ that $q_0' \notin F$ and $q_1' \in F$ for such $q_0'$ and $q_1'$ that $r_0 \Rightarrow_{\mathcal{D}} q_0'$ and $r_1 \Rightarrow_{\mathcal{D}} q_1'$. Now, we have found a set of states in $\mathfrak{A}$ that satisfies the condition of Theorem 6 $(ii)$ with $w = \mathcal{W}\mathcal{U}^{|Q|!-1}$ and $u = \mathcal{U}^{|Q|!}$. Therefore, $\boldsymbol{q}$ is not $\mathsf{FO}(<,\equiv)$-rewritable. ❑

▶ **Theorem 29.** *Deciding* $\mathsf{FO}(<,\equiv)$*-rewritability of OMPQs* $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ *with a linear* $LTL^{\bigcirc}_{horn}$*-ontology* $\mathcal{O}$ *over* $\Xi$*-ABoxes can be done in* PSPACE.

**Proof.** The proof relies on Theorem 6 $(ii)$. Clearly, Corollary 25 holds providing the bound of $2^{O(|\boldsymbol{q}|)}$ on $|\mathcal{A}|$, $|\mathcal{D}|$ and $k$. The same bound on $|\mathcal{W}|$, $|\mathcal{U}|$ and $|\mathcal{B}|$ follows from the same argument as in the proof of Theorem 27 and a straightforward modification of that proof gives a PSPACE algorithm we are after. ❑

The criterion of Theorem 6 $(iii)$ is harder to transform to a PSPACE-checkable condition on canonical models and ABoxes, and the complexity of deciding $\mathsf{FO}(<,\mathsf{MOD})$-rewritability of linear OMPQs remains open at the moment.

## 7 $\mathsf{FO}(<)$-rewritability of $LTL^{\bigcirc}_{krom}$ OMAQs and $LTL^{\bigcirc}_{core}$ OMPQs

Our next aim is to look for non-trivial classes of OMQs deciding FO-rewritability of which could be 'easier' than PSPACE. Syntactically, the simplest type of axioms (5) are binary clauses: $C_1 \to C_2$ and $C_1 \land C_2 \to \bot$, known as *core* axioms, which together with $C_1 \lor C_2$ form the class Krom. In the atemporal case, the W3C standard language *OWL 2 QL*, designed specifically for ontology-based data access, allows core clauses only and uniformly guarantees FO-rewritability [3, 19].

As we saw in the proof of Theorem 19, OMPEQs with disjunctive axioms can simulate $LTL^{\bigcirc}_{horn}$ OMAQs, and so are too complex for the purposes of this section. On the other hand, $LTL^{\bigcirc}_{krom}$ OMAQs and $LTL^{\bigcirc}_{core}$ OMPQs are all $\mathsf{FO}(<, \equiv)$-rewritable [7]. Below, we focus on deciding $\mathsf{FO}(<)$-rewritability of OMQs in these classes.

▶ **Theorem 30.** *Deciding $\mathsf{FO}(<)$-rewritability of Boolean and specific $LTL^{\bigcirc}_{krom}$ OMAQs over $\Xi$-ABoxes is* coNP*-complete.*

**Proof.** Suppose $\boldsymbol{q} = (\mathcal{O}, A)$ is an $LTL^{\bigcirc}_{krom}$ OMAQ and $\mathcal{O}$ is consistent. Using the form of Krom axioms, one can show [7] that, for any ABox $\mathcal{A}$ and $l \in \mathbb{Z}$, we have $(\mathcal{O}, \mathcal{A}) \models A(l)$ iff one of the following holds: $(i)$ there are $k \le l$ and $B(k) \in \mathcal{A}$ such that $\mathcal{O} \models B \to \bigcirc_{F}^{l-k} A$; $(ii)$ there are $k > l$ and $B(k) \in \mathcal{A}$ such that $\mathcal{O} \models B \to \bigcirc_{P}^{k-l} A$; $(iii)$ $\mathcal{O}$ and $\mathcal{A}$ are inconsistent, i.e., there exist $k_1 \le k_2$, $B(k_1) \in \mathcal{A}$ and $C(k_2) \in \mathcal{A}$ such that $\mathcal{O} \models B \to \bigcirc_{F}^{k_2-k_1} \neg C$.

Let $lit(\boldsymbol{q}) = \{C, \neg C \mid C \in \mathsf{sig}(q)\}$. For any $L_1, L_2 \in lit(\boldsymbol{q})$, we can construct a unary NFA $\mathfrak{A}_{L_1L_2}$ of size $O(|\boldsymbol{q}|)$ that accepts $\boldsymbol{L}_{L_1L_2} = \{a^n \mid \mathcal{O} \models L_1 \to \bigcirc_{F}^{n} L_2,\ n \ge 0\}$. The set of its states is $lit(\boldsymbol{q})$, $L_1$ is the initial state, the set of accepting states is $\{L_2\}$, and the transitions are the following:

- $L \to_a L'$ if $\mathcal{O} \models L \to \bigcirc_F L'$;
- $L \to_\varepsilon L'$ if $\mathcal{O} \models L \to L'$.

Let $\Xi^{\exists}_{A} = \{B \in \Xi \mid \mathcal{O}, \{B(0)\} \models \exists x A(x)\}$ and $\Xi^{\forall}_{A} = \{B \in \Xi \mid \mathcal{O}, \{B(0)\} \models \forall x A(x)\}$.

▶ **Lemma 31.** $(i)$ *The language $\boldsymbol{L}_{\Xi}(\boldsymbol{q})$ is $\mathsf{FO}(<)$-definable iff, for all $B, C \in \Xi \setminus \Xi^{\exists}_{A}$, the language $\boldsymbol{L}_{B\neg C}$ is $\mathsf{FO}(<)$-definable.*

$(ii)$ *The language $\boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$ is $\mathsf{FO}(<)$-definable iff the following holds:*
- *for all $B \in \Xi$, the languages $\boldsymbol{L}_{BA}$ and $\boldsymbol{L}_{\neg A\neg B}$ are $\mathsf{FO}(<)$-definable;*
- *for all $B, C \in \Xi \setminus \Xi^{\forall}_{A}$ such that one of the $\boldsymbol{L}_{BA}$ and $\boldsymbol{L}_{\neg A\neg C}$ is finite, the language $\boldsymbol{L}_{B\neg C}$ is $\mathsf{FO}(<)$-definable.*

**Proof.** $(i)$ $(\Rightarrow)$ If $\boldsymbol{L}_{\Xi}(\boldsymbol{q})$ is $\mathsf{FO}(<)$-definable then so is $\boldsymbol{L}_{\Xi}(\boldsymbol{q}) \cap \boldsymbol{L}(\{B\}\emptyset^*\{C\})$, for any $B, C$. For $B, C \notin \Xi^{\exists}_{A}$, we have $\{B\}\emptyset^n\{C\} \in \boldsymbol{L}_{\Xi}(\boldsymbol{q})$ iff $\mathcal{O} \models B \to \bigcirc_{F}^{n+1} \neg C$.

$(\Leftarrow)$ For a $\Xi$-ABox $\mathcal{A}$, we have $w_{\mathcal{A}} \in \boldsymbol{L}_{\Xi}(\boldsymbol{q})$ iff either there is $B(k) \in \mathcal{A}$, for some $B \in \Xi^{\exists}_{A}$, or there are $B, C \in \Xi \setminus \Xi^{\exists}_{A}$ and $k \le l$ such that $B(k), C(l) \in \mathcal{A}$ and $\mathcal{O} \models B \to \bigcirc_{F}^{k-l} \neg C$. By assumption, all of these conditions are $\mathsf{FO}(<)$-definable.

$(ii)$ $(\Rightarrow)$ If $\boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$ is $\mathsf{FO}(<)$-definable, then so is $\boldsymbol{L}_{\Xi}(\boldsymbol{q}(x)) \cap \boldsymbol{L}(\{B\}\emptyset^*\emptyset')$ (see the definition of $\boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$ in Section 2) and $\boldsymbol{L}_{\Xi}(\boldsymbol{q}(x)) \cap \boldsymbol{L}(\emptyset'\emptyset^*\{B\})$, for any $B \in \Xi$. We have $\{B\}\emptyset^n\emptyset' \in \boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$ iff $\mathcal{O} \models B \to \bigcirc_{F}^{n+1} A$ and $\emptyset'\emptyset^*\{B\} \in \boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$ iff $\mathcal{O} \models B \to \bigcirc_{P}^{n+1} A$. Suppose $B, C \in \Xi \setminus \Xi^{\forall}_{A}$ and $\boldsymbol{L}_{BA}$ is finite. There is $l \in \mathbb{Z}$ such that $\mathcal{O}, \{C(0)\} \not\models A(l)$ and there is $k$ such that $k > n$ for all $a^n \in \boldsymbol{L}_{BA}$. For $m > k + |l|$, we have $\mathcal{O}, \{B(0), C(m)\} \models A(m + l)$ iff $\mathcal{O} \models B \to \bigcirc_{F}^{m} \neg C$. The case when $\boldsymbol{L}_{\neg A\neg C}$ is finite is similar.

$(\Leftarrow)$ One can prove by induction on the construction of star-free expressions that every star-free language over a unary alphabet is either finite or cofinite. Since, for all $B \in \Xi$, the languages $\boldsymbol{L}_{BA}$ and $\boldsymbol{L}_{\neg A\neg B}$ are $\mathsf{FO}(<)$-definable, they all are star-free. Therefore, there is $n \in \mathbb{N}$ such that, for any $B$ and $n_1, n_2 > n$, we have $a^{n_1} \in \boldsymbol{L}_{BA}$ iff $a^{n_2} \in \boldsymbol{L}_{BA}$ and similarly for $\boldsymbol{L}_{\neg A\neg B}$.

For a $\Xi$-ABox $\mathcal{A}$ and $k \in \mathbb{Z}$, we have $w_{\mathcal{A},k} \in \boldsymbol{L}_{\Xi}(\boldsymbol{q}(x))$ iff either there is $B(l) \in \mathcal{A}$ with $l \le k$ and $\mathcal{O} \models B \to \bigcirc_{F}^{l-k} A$, or there is $B(l) \in \mathcal{A}$ with $l > k$ and $\mathcal{O} \models B \to \bigcirc_{P}^{k-l} A$, or there are $B(k), C(l) \in \mathcal{A}$ such that $k - l < 2n$ and $\mathcal{O} \models B \to \bigcirc_{F}^{k-l} \neg C$, or there are $B(k), C(l) \in \mathcal{A}$

such that $k - l \geq 2n$, $\boldsymbol{L}_{BA}$ and $\boldsymbol{L}_{\neg A \neg C}$ are infinite, or $B(k), C(l) \in \mathcal{A}$ such that $k - l \geq 2n$, one of $\boldsymbol{L}_{BA}$ and $\boldsymbol{L}_{\neg A \neg C}$ is finite and $\mathcal{O} \models B \to \bigcirc_F^{k-l} \neg C$. All of these conditions are $\mathsf{FO}(<)$ definable. (In the fourth case, since $\boldsymbol{L}_{BA}$ is $\mathsf{FO}(<)$-definable and infinite, $\mathcal{O} \models B \to \bigcirc_F^n \Box_F A$ and, similarly, $\mathcal{O} \models C \to \bigcirc_P^n \Box_P A$; therefore, $\mathcal{O}, \{B(k), C(l)\} \models \forall x A(x)$ and we do not need to check for inconsistency.) ❑

Thus, to check $\mathsf{FO}(<)$-rewritability of $\boldsymbol{q}$ and $\boldsymbol{q}(x)$, it suffices to check $\mathsf{FO}(<)$-definability, emptiness and finiteness of the languages of the form $\boldsymbol{L}_{L_1 L_2}$. Emptiness and finiteness can be checked in NL. Using [50, Theorem 6.1], one can show that deciding $\mathsf{FO}(<)$-definability of the language of a unary NFA is coNP-complete, which gives the required upper bound for deciding $\mathsf{FO}(<)$-rewritability of both Boolean and specific $LTL_{krom}^{\bigcirc}$ OMAQs.

To show the matching lower bound, for any unary NFA $\mathfrak{A} = (Q, \{a\}, \delta, q_0, F)$ without $\varepsilon$-transitions, we define an $LTL_{core}^{\bigcirc}$ ontology $\mathcal{O}_{\mathfrak{A}}$ with the axioms $X \to \bigcirc_F q_0$, $q \wedge Y \to \bot$, for every $q \in F$, and $q \to \bigcirc_F p$, for every transition $q \to_a p$. The OMAQs $\boldsymbol{q} = (\mathcal{O}_{\mathfrak{A}}, A)$ for $A \notin Q \cup \{X, Y\}$ and $\boldsymbol{q}(x) = (\mathcal{O}_{\mathfrak{A}}, A(x))$ are $\mathsf{FO}(<)$-rewritable over $\{X, Y\}$-ABoxes iff $\boldsymbol{L}(\mathfrak{A})$ is star-free because $\mathcal{O}, \mathcal{A} \models A(l)$, for an $\{X, Y\}$-ABox $\mathcal{A}$, iff $\mathcal{A}$ is inconsistent with $\mathcal{O}_{\mathfrak{A}}$. An $\{X, Y\}$-ABox $\mathcal{A}$ is inconsistent iff there are $X(i), Y(j) \in \mathcal{A}$ with $a^{j-i-1} \in \boldsymbol{L}(\mathfrak{A})$. ❑

Our next result deals with a weaker (Horn∩Krom) ontology language but more expressive queries.

▶ **Theorem 32.** *Deciding* $\mathsf{FO}(<)$*-rewritability of Boolean and specific* $LTL_{core}^{\bigcirc}$ *OMPEQs over* $\Xi$*-ABoxes is* $\Pi_2^p$ *-complete.*

**Proof.** By Proposition 15 (*ii*) and Lemma 14, it is enough to consider Boolean $LTL_{core}^{\bigcirc}$ OMPEQs $\boldsymbol{q} = (\mathcal{O}, \boldsymbol{q})$ with $\bot$-free $\mathcal{O}$. We further assume, without loss of generality, that all of the axioms have the following forms: $A \to B$, $A \to \bigcirc_F B$, or $A \to \bigcirc_P B$, for atomic $A$ and $B$.

▶ **Lemma 33.** *For* $v \in \Sigma_\Xi^*$*, deciding whether* $v \in \boldsymbol{L}_\Xi(\boldsymbol{q})$ *can be done in NP.*

**Proof.** We prove that, given an ABox $\mathcal{A}$ and $j \in \mathbb{Z}$, checking $\mathcal{O}, \mathcal{A} \models \varkappa(j)$ is in NP.

The proof is by induction on the construction of $\varkappa$. If $\varkappa$ is atomic and $\mathcal{O}, \mathcal{A} \models \varkappa(j)$ then there is $B(i) \in \mathcal{A}$ such that $\mathcal{O} \models B \to \bigcirc_F^{j-i} A$ or $\mathcal{O} \models B \to \bigcirc_P^{i-j} A$, which can be checked in polynomial time. The cases $\varkappa = \varkappa_1 \wedge \varkappa_2$ and $\varkappa = \varkappa_1 \vee \varkappa_2$ are obvious.

Let $\varkappa = \Diamond_F \varkappa_1$. If $\mathcal{O}, \mathcal{A} \models \varkappa(j)$, then $\mathcal{O}, \mathcal{A} \models \varkappa_1(i)$ for some $i > j$. By the structure of the canonical models [7], the required $i$ can be found in the interval $j < i < |j| + \max \mathcal{A} + 2^{O(|\mathcal{O}|)}$. So it is of polynomial length and we can non-deterministically guess it along with the necessary certificate proving that $\mathcal{O}, \mathcal{A} \models \varkappa_1(i)$, which exists by IH. The case of $\varkappa = \Diamond_P \varkappa_1$ is symmetric.

It remains to recall from [7] that the certain answer to $\boldsymbol{q}$ over $\mathcal{A}$ is $\mathsf{yes}$ iff there exists $j \in [-O(2^{\mathcal{O}}), \max \mathcal{A} + O(2^{\mathcal{O}})]$ such that $\mathcal{O}, \mathcal{A} \models \varkappa(j)$. ❑

Using criteria (*i*)–(*iii*) from the proof of Theorem 30, the assumption above, and the structure of $\varkappa$, we obtain that $\mathcal{O}, \mathcal{A} \models \exists \varkappa(x)$ iff $\mathcal{O}, \mathcal{A}' \models \exists \varkappa(x)$, for some $\mathcal{A}' \subseteq \mathcal{A}$ with $|\mathcal{A}'| \leq |\varkappa|$. We reformulate this observation in slightly different terms. Let $\mathcal{B}$ be the set of words $w = w_1 \dots w_k \in \Sigma_\Xi^*$ such that, for every $i$, we have $|w_i| \geq 1$ and $|w_1| + \dots + |w_k| \leq |\varkappa|$. With every such $w$ we associate the language $L_w = \boldsymbol{L}(\emptyset^* w_1 \emptyset^* \dots \emptyset^* w_k \emptyset^*) \cap \boldsymbol{L}_\Xi(\boldsymbol{q})$. For $\Sigma_{\boldsymbol{q}}^*$-words $v$ and $v'$, we write $v' \leq v$ if they are of the same length and $v_i' \subseteq v_i$, for all $i$.

▶ **Lemma 34.** *For every* $v \in \Sigma_\Xi^*$*, we have* $v \in \boldsymbol{L}_\Xi(\boldsymbol{q})$ *iff there is* $v' \leq v$ *such that* $v' \in \boldsymbol{L}_w$ *for some* $w \in \mathcal{B}$.

We also require the following:

▶ **Lemma 35.** *A regular language*

$$\boldsymbol{L} \subseteq \boldsymbol{L}(a^*b_1a^*b_2a^* \dots a^*b_ka^*)$$

*with $a \notin \{b_1, \dots, b_k\}$ is star-free iff $\boldsymbol{L}$ can be defined by a regular expression of the form*

$$\alpha = \bigcup_{i=1}^{n} \alpha_{i,0}b_1\alpha_{i,1} \dots \alpha_{i,k-1}b_k\alpha_{i,k}$$

*for some $n \in \mathbb{N}$, where each $\alpha_{i,j}$ is either $a^{l_{ij}}$ or $a^{l_{ij}}a^*$, for some $l_{ij} \in \mathbb{N}$.*

**Proof.** ($\Leftarrow$) All individual members of the union are concatenations of star-free languages. Therefore, $\boldsymbol{L}$ is star-free because star-free languages are closed under concatenation and union.

($\Rightarrow$) The proof is by induction on $k$. For $k = 0$, $\boldsymbol{L} \subseteq \boldsymbol{L}(a^*)$ is either finite or cofinite. If it is finite, then $\boldsymbol{L} = \bigcup_{j=1}^{m} a^{i_j}$; otherwise, $\boldsymbol{L} = \bigcup_{j=1}^{m} a^{i_j} \cup \{a^n \mid n > i_m\}$, and so $\boldsymbol{L} = \boldsymbol{L}(a^{i_m}a^* \cup \bigcup_{j=1}^{m-1} a^{i_j})$.

Let $k > 0$. Let $\mathfrak{A} = (Q, \Sigma, \delta, q_0, F)$ be a minimal DFA accepting $\boldsymbol{L}$. Let $B = \{q \in Q \mid \exists i\, \delta_{a^i}(q_0) = q\}$ and let $B' = \{q \in B \mid \delta(q, b_1) \text{ is defined}\}$. For a non-trash $p \in B'$, let $\boldsymbol{L}_p$ be the language accepted by the automaton $(B, \{a\}, \delta|_B, I, \{p_B\})$ and let $\boldsymbol{L}'_p$ be the language accepted by the automaton $(Q/B, \Sigma, \delta|_{Q/B}, \delta(p, b_1), F)$. Clearly, $\boldsymbol{L}'_p \subseteq \boldsymbol{L}(a^*b_2a^*b_3a^* \dots a^*b_ka^*)$ and both $L_p$ and $L'_p$ are star-free. Therefore, by IH, there are a regular expression $\bigcup_{i=1}^{n_p} \alpha_{i,0}^p$ defining $\boldsymbol{L}_p$ and a regular expression $\bigcup_{j=1}^{n'_p} \alpha_{j,1}^p b_2 \alpha_{j,2}^p \dots \alpha_{j,k-1}^p b_k \alpha_{j,k}^p$ defining $\boldsymbol{L}'_p$. Since $\boldsymbol{L} = \bigcup_{p \in B}(\boldsymbol{L}_p \cdot \{b_1\} \cdot \boldsymbol{L}'_p)$, the language $\boldsymbol{L}$ is defined by

$$\bigcup_{p \in B} \bigcup_{i=1}^{n_p} \bigcup_{j=1}^{n'_p} \alpha_{i,0}^p b_1 \alpha_{j,1}^p b_2 \alpha_{j,2}^p \dots \alpha_{j,k-1}^p b_k \alpha_{j,k}^p.$$

This completes the proof of the lemma. ❑

▶ **Lemma 36.** *The language $\boldsymbol{L}_\Xi(\boldsymbol{q})$ is star-free iff $\boldsymbol{L}_w$ is star-free, for every $w \in \mathcal{B}$.*

**Proof.** ($\Rightarrow$) If $\boldsymbol{L}_\Xi(\boldsymbol{q})$ is star-free, then so is $\boldsymbol{L}_w$ because $\boldsymbol{L}(\emptyset^*w_1\emptyset^* \dots \emptyset^*w_k\emptyset^*)$ is star-free and star-free languages are closed under intersection.

($\Leftarrow$) Suppose the language $\boldsymbol{L}_w$ is star-free. By Lemma 35, $\boldsymbol{L}_w$ is defined by the expression $\alpha_w = \bigcup_{i=1}^{n_w} \alpha_{i,0}w_1\alpha_{i,1} \dots \alpha_{i,k-1}w_k\alpha_{i,k}$ for some $n_w \in \mathbb{N}$, where each $\alpha_{i,j}$ is either $\emptyset^l$ or $\emptyset^l\emptyset^*$. Let $\alpha'_{i,j} = \sigma^l$ or $\sigma^l\varnothing^c$ (we use $\emptyset$ to denote the letter of $\Sigma_\Xi$ and $\varnothing$ to denote the empty language), respectively, where $\sigma = \bigcup_{a \in \Sigma_{\boldsymbol{q}}} a$. Let

$$\alpha'_w = \bigcup_{j=1}^{n_w} \Big(\alpha'_{j,0}(\bigcup_{w_1 \subseteq a} a)\alpha'_{j,1} \dots \alpha'_{j,k-1}(\bigcup_{w_k \subseteq a} a)\alpha'_{j,k}\Big).$$

We see that $\alpha'_w$ is star-free and $\boldsymbol{L}(\alpha'_w) = \{v \in \Sigma^*_{\boldsymbol{q}} \mid \exists v' \in \boldsymbol{L}_w\; v' < v\}$. It follows that $\boldsymbol{L}(\bigcup_{w \in \mathcal{B}} \alpha'_w) = \boldsymbol{L}_\Xi(\boldsymbol{q})$ and $\boldsymbol{L}_\Xi(\boldsymbol{q})$ is star-free. ❑

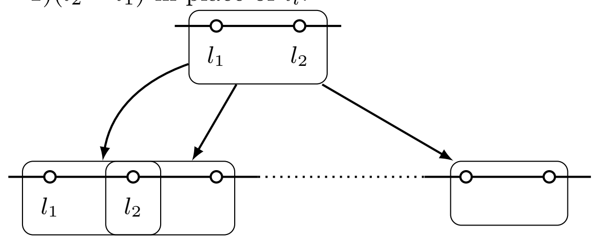
For $w = w_1 \dots w_k \in \mathcal{B}$ and $I = (i_0, \dots, i_k)$, let $v_{w,I} = \emptyset^{i_0}w_1\emptyset^{i_1} \dots w_k\emptyset^{i_k}$. For $c \in \mathbb{N}$, let $I_{\leq c}$ be $I$ with all $i_j > c$ replaced with $c$.

▶ **Lemma 37.** *$\boldsymbol{L}_w$ is star-free iff $v_{w,I} \in \boldsymbol{L}_\Xi(\boldsymbol{q})$ just in case $v_{w,I_{\leq c}} \in \boldsymbol{L}_\Xi(\boldsymbol{q})$, for all $I$, where $c = 2^{|\mathsf{sig}(\boldsymbol{q})|+|\varkappa|} + 1$.*

**Proof.** ($\Leftarrow$) For $w = w_1 \dots w_k$, let $\mathcal{I}_w = \{I = (i_0, \dots, i_k) \mid \max i_l \leq c, v_{w,I} \in \boldsymbol{L}(\boldsymbol{q})\}$. It is a finite set. For each $I \in \mathcal{I}_w$, let $\alpha_I = \alpha_{I,0} b_1 \alpha_{I,1} \dots b_k \alpha_{I,k}$ where $\alpha_{I,j} = \emptyset^j$ if $j < c$ and $\emptyset^c \emptyset^*$ if $j = c$. We see that $\boldsymbol{L}_w$ is defined by $\bigcup_{I \in \mathcal{I}_w} \alpha_I$, and so it is star-free.

($\Rightarrow$) Consider $\alpha_w$ from Lemma 36. Each $\alpha_{i,j}$ is either $\emptyset^l$ or $\emptyset^l \emptyset^*$. Choose $l_{max}$ to be bigger than all of the $l$. We see that $v_{w,I} \in \boldsymbol{L}_\Xi(\boldsymbol{q})$ iff $v_{w,I_{\leq l_{max}}} \in \boldsymbol{L}_\Xi(\boldsymbol{q})$.

Consider ABox $\mathcal{A}$ corresponding to $v_{w,I_{\leq c}}$ and choose $l$ such that $i_l = c$. There are two places in the part of the canonical model corresponding to $i_l$ where exactly the same atomic concepts and subformulas of $\varkappa$ are true. Let them be $l_1$ and $l_2$. If we 'repeat' the $[l_1+1, l_2]$ part $m$ times, we obtain exactly the canonical model for the ABox corresponding to $v_{w,I'}$ where $I'$ has $c + (m-1)(l_2 - l_1)$ in place of $i_l$.




We can choose $m$ so that $c + (m-1)(l_2 - l_1) > l_{max}$. We can do the same for all $i_j = c$ in $I_{<c}$ and all $i_j \geq c$ in $I$. So the words $v_{w,I_{<l_{max}}}$, $v_{w,I_{<c}}$ and $v_{w,I}$ are in or out of $\boldsymbol{L}_w$ simultaneously. ❑

We are now in a position to show that deciding $\mathsf{FO}(<)$-rewritability of $\boldsymbol{q}$ can be done in $\Pi_2^p$. Indeed, $\boldsymbol{q}$ is not $\mathsf{FO}(<)$-rewritable iff we can guess $w \in \mathcal{B}$ and $I$ such that $\max(I) < 2c$ and only one of $v_I$ and $v_{I<c}$ belongs to $\boldsymbol{L}_\Xi(\boldsymbol{q})$. By Lemma 33, we can check membership in $\boldsymbol{L}_\Xi(\boldsymbol{q})$ using an NP-oracle, so the problem is in $\text{coNP}^{\text{NP}} = \Pi_2^p$.

We show the matching lower bound by reduction of $\forall\exists$3CNF. Suppose we are given a QBF $\forall X \exists Y \varphi$ with a 3CNF $\varphi$, $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_m\}$. We construct an $LTL_{core}^{\bigcirc}$ OMPEQ $\boldsymbol{q}_\varphi = (\mathcal{O}_\varphi, \varkappa_\varphi)$ such that $\boldsymbol{q}_\varphi$ is $\mathsf{FO}(<)$-rewritable iff $\forall X \exists Y \varphi(X, Y)$ is true.

We use atomic concepts $A_i^j$, for $1 \leq i \leq m$, $0 \leq j \leq p_m - 1$, where $p_i$ is the $i$-th prime number, $z^0$ and $z^1$, for $z \in X \cup Y$, $A$ and $B$. The ontology $\mathcal{O}_\varphi$ comprises the axioms

$$A \to A_i^0, \quad A_i^j \to \bigcirc_F A_i^{(j+1) \bmod p_i}, \quad A_i^0 \to y_i^0, \quad A_i^1 \to y_i^1,$$
$$x_i^0 \to \bigcirc_F x_i^0, \quad x_i^1 \to \bigcirc_F x_i^1, \quad B \to \bigcirc_F \bigcirc_F B.$$

The size of the ontology $|\mathcal{O}_\varphi|$ is polynomial of $|X| + |Y|$ because $p_m = O(m \log m)$. Let $\varphi'$ result from $\varphi$ by replacing all $x_i$ with $x_i^1$, all $\neg x_i$ with $x_i^0$, and similarly for the $y_j$. We set

$$\varkappa_\varphi = A \wedge \bigwedge_{i=0}^{n} (x_i^0 \vee x_i^1) \wedge (B \vee \Diamond_F \varphi').$$

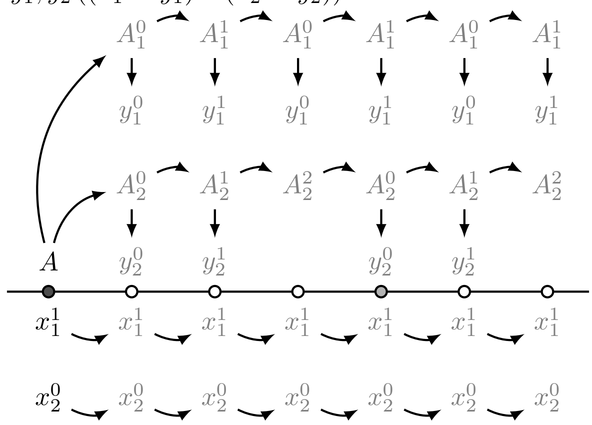
We now show that $\boldsymbol{q}_\varphi$ is as required. Suppose $\forall X \exists Y \varphi(X, Y)$ is true. Consider an ABox $\mathcal{A}$ with the answer yes. There is $t \in \mathbb{Z}$ such that $\mathcal{O}_\varphi, \mathcal{A} \models \varkappa_\varphi(t)$. We know that then $A(t) \in \mathcal{A}$, and $\mathcal{O}_\varphi, \mathcal{A} \models \bigwedge_{i=0}^n (x_i^0 \vee x_i^1)$. This means that, for every $i$, there is $x_i^0(s)$ or $x_i^1(s)$ in $\mathcal{A}$, for some $s \leq t$. There is an assignment for $as_1 \in 2^X$ such that $\mathcal{O}_\varphi, \mathcal{A} \models x_i^{as_1(i)}(s)$ for all $s > t$. For this assignment, there exists a corresponding assignment of $as_2 \in 2^Y$. There is a number $r$ such that $r \bmod p_i = as_2(i)$ for all $i \leq m$. Therefore $\mathcal{O}_\varphi, \mathcal{A} \models y_i^{as_2(i)}$, $\mathcal{O}_\varphi, \mathcal{A} \models \varphi'(t+r)$, and so $\mathcal{O}_\varphi, \mathcal{A} \models \Diamond_F \varphi'(j)$. Thus, the sentence

$$\exists t \left( A(t) \wedge \bigwedge_{i=0}^{n} \exists s \left( (s \leqslant t) \wedge (x_i^0(s) \vee x_i^1(s)) \right) \right)$$

is an FO(<)-rewriting of $\boldsymbol{q}_\varphi$.

If $\forall X \exists Y \varphi(X, Y)$ is false, then there is an assignment $as \in 2^X$ to the variables in $X$ such that $\varphi$ is false for any assignments to $Y$. Let $X_{as} = \{A\} \cup \bigcup_{i=1}^n \{x_i^{as(x_i)}\}$. Consider $\mathcal{A} = \{B(0)\} \cup \bigcup_{x \in X_{as}} x(l)$ for some $l > 0$. If the certain answer to $\boldsymbol{q}_\varphi$ over $\mathcal{A}$ is yes, then $\mathcal{O}_\varphi, \mathcal{A} \models \varkappa_\varphi(l)$. Therefore $\mathcal{O}_\varphi, \mathcal{A} \models B(l)$ since $\mathcal{O}_\varphi, \mathcal{A} \not\models \Diamond_F \varphi'(l)$. This means that, for $w = \{B\}X_{as}$, the language $\boldsymbol{L}_w$ is $\boldsymbol{L}(\emptyset^*\{B\}(\emptyset\emptyset)^* X_{as} \emptyset^*)$ and not star-free, and therefore $\boldsymbol{q}_\varphi$ is not FO(<)-rewritable by Lemma 36.

This picture illustrates the intended models of $\mathcal{O}_\varphi$ and $\mathcal{A} = \{A(0), x_1^1(0), x_2^0(0)\}$ for the formula $\varphi = \forall x_1, x_2 \exists y_1, y_2\, ((x_1 = y_1) \wedge (x_2 = y_2))$:




This completes the proof of Theorem 32. ❑

If we slightly increase the expressive power of $LTL_{core}^{\bigcirc}$ OMPEQs $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ by allowing $\Box$-operators in $\varkappa$, the problem of deciding FO(<)-rewritability becomes more complex:

▶ **Theorem 38.** *Deciding* FO(<)*-rewritability of Boolean and specific* $LTL_{core}^{\bigcirc}$ *OMPQs is* *PSPACE-complete*

**Proof.** By Proposition 15 and Lemma 14, it suffices to prove this theorem for Boolean $LTL_{core}^{\bigcirc}$ OMPQs. The upper bound follows from Theorem 27 as core OMQs are linear Horn OMQs.

To prove the matching lower bound, we reduce the PSPACE-complete DFA intersection problem (see, e.g., [14,21]) to OMQ rewritability. Let $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ with $\mathfrak{A}_i = (Q_i, \Sigma, \delta_i, q_0^i, F_i)$ be a sequence of DFAs that do not accept the empty word, have a common input alphabet, and disjoint sets of states.

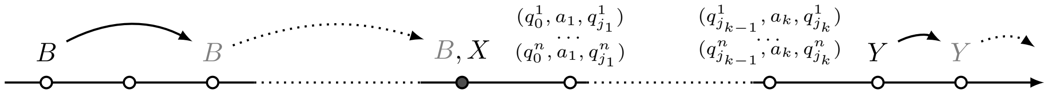
Let $Q_i = \{q_1^i, \dots, q_{j_i}^i\}$. Consider the following ontology $\mathcal{O}$ with atomic concepts $\{X, Y, B\} \cup \bigcup_{i \in [1,n]} \delta_i$:

$$(q_k^i, a, q_l^i) \wedge (q_m^i, b, q_n^i) \to \bot, \quad \text{if } k \neq m \text{ or } l \neq n,$$
$$(q_k^i, a, q_l^i) \wedge \bigcirc_F (q_m^i, b, q_n^i) \to \bot, \quad \text{if } l \neq m,$$
$$(q_k^i, a, q_l^i) \wedge (q_m^j, b, q_n^j) \to \bot, \quad \text{if } a \neq b,$$
$$X \wedge \bigcirc_F (q_k^i, a, q_l^i) \to \bot, \quad \text{for } k \neq 0,$$
$$(q_k^i, a, q_l^i) \wedge \bigcirc_F Y \to \bot, \quad \text{for } q_l^i \notin F_i,$$
$$X \wedge \bigcirc_F Y \to \bot,$$
$$Y \to \bigcirc_F Y,$$
$$B \to \bigcirc_F \bigcirc_F B.$$

Set

$$\varkappa \;=\; C \wedge X \wedge \Box_F \Big( ( \bigwedge_{i \in [1,n]} \bigvee_{(r,a,s) \in \delta_i} (r,a,s)) \vee Y \Big).$$

We claim that the OMQ $\boldsymbol{q} = (\mathcal{O}, \varkappa)$ is $\mathsf{FO}(<)$-rewritable over $\Xi$-ABoxes, for $\Xi = \mathsf{sig}(\boldsymbol{q})$, iff $\bigcap_{i\in[1,n]} L(\mathfrak{A}_i) = \emptyset$. The picture below illustrates the structure of the intended models:




($\Leftarrow$) If $\bigcap_{i\in[1,n]} L(\mathfrak{A}_i) = \emptyset$, then, for any ABox $\mathcal{A}$, we have $\mathcal{O}, \mathcal{A} \models \varkappa(k)$ iff the ABox $\mathcal{A}$ is inconsistent with $\mathcal{O}$. It follows that the disjunction $\mathcal{Q}$ of the following sentences (describing different cases of how $\mathcal{A}$ can be inconsistent with $\mathcal{O}$)

$$\bigvee_{i} \bigvee_{k \neq m, l \neq n} \exists s ((q_k^i, a, q_l^i)(s) \wedge (q_m^i, b, q_n^i)(s))$$

$$\bigvee_{i} \bigvee_{l \neq m} \exists s ((q_k^i, a, q_l^i)(s) \wedge (q_m^i, a, q_n^i)(s+1))$$

$$\bigvee_{i,j} \bigvee_{a \neq b} \exists s ((q_k^i, a, q_l^i)(s) \wedge (q_m^j, b, q_n^j)(s))$$

$$\bigvee_{i} \bigvee_{k > 0} \exists s \left( X(s) \wedge (q_k^i, a, q_l^i)(s+1) \right)$$

$$\bigvee_{A \in \{X\} \cup \{(r,a,s) \mid s \notin \bigcup_i F_i\}} \exists s, s' \left( (s \leq s' + 1) \wedge A(s') \wedge Y(s) \right)$$

is an $\mathsf{FO}(<)$-rewriting of $\boldsymbol{q}$.

($\Rightarrow$) Let $w = w_1 \dots w_k \in \bigcap_{i\in[1,n]} L(\mathfrak{A}_i)$. For $i \in [1,n]$ and $j \in [0,k]$, there exists $q_j^i \in Q_i$ such that $(q_{j-1}^i, w_j, q_j^i) \in \delta_i$. Let $w_{\mathcal{A}} = \{B\}$, $w_{\mathcal{B}} = \emptyset$ and $w_{\mathcal{C}}$ be the word corresponding to the ABox $\mathcal{C} = \{X(0)\} \cup \Big( \bigcup_{i\in[1,n]} \bigcup_{j\in[1,k]} \{(q_{j-1}^i, w_j, q_j^i)(j)\} \Big) \cup \{Y(k+1)\}$. We see that a word of the form $w_{\mathcal{A}} w_{\mathcal{B}}^n w_{\mathcal{C}}$ is in $L_\Xi(\boldsymbol{q})$ iff $n$ is odd. Therefore, $\boldsymbol{L}_\Xi(\boldsymbol{q})$ is not star-free, and $\boldsymbol{q}$ is not $\mathsf{FO}(<)$-rewritable. ❑

The reason causing the complexity gap between Theorems 32 and 38 can be explained by the rising combined complexity of answering $\mathit{LTL}_{core}^{\bigcirc}$ OMPQs, established by the following theorem, which should be compared with Lemma 33:

▶ **Theorem 39.** *Given an $\mathit{LTL}_{core}^{\bigcirc}$ OMPQ $\boldsymbol{q}(x) = (\mathcal{O}, \varkappa(x))$ and $x \in \mathbb{Z}$, checking whether $\mathcal{O}, \mathcal{A} \models \varkappa(x)$ is $P^{NP}[O(\log n)]$-complete.*

**Proof.** As we saw above, checking whether $\mathcal{O}, \mathcal{A} \models A(x)$, for atomic $A$, is in P. Therefore, for $\varphi$ without temporal operators, but possibly with atoms of the form $(x \geq k)$, for some $k \in \mathbb{Z}$, checking whether $\mathcal{O}, \mathcal{A} \models \varphi(x)$ is also in P. We call such formulas *simple*. For any simple $\varphi$ and any $\circ \in \{\Box_P, \Box_P, \Diamond_F, \Diamond_P\}$, the set of $x$ such that $\mathcal{O}, \mathcal{A} \models \circ\varphi(x)$ is either empty, the whole line, or a half-line. Therefore, in the canonical model of $\mathcal{O}$ and $\mathcal{A}$, either $\circ\varphi(x)$ is equivalent to $\top$, $\bot$, or there is $t \in [\min \mathcal{A} - c, \max \mathcal{A} + c]$, for some $c = 2^{O(|\mathcal{O}|)}$, such that $\circ\varphi(x)$ is equivalent to $x < t$ for $\circ = \Box_P, \Diamond_F$ or $t < x$ for $\circ = \Box_F, \Diamond_P$. We can find the precise equivalent (in the canonical model) atomic formula in NP. So we can find the equivalent formulas for the subformulas of $\varkappa$ of the form $\circ\varphi$, replace them with these atomic formulas, and repeat until we arrive to a single simple formula that can be evaluated in P at the given point. Therefore the combined complexity of $LTL^{\bigcirc}_{core}$ OMPQs belongs to the class TREES(NP), which is equivalent to $\mathrm{P}^{\mathrm{NP}}[O(\log n)]$ (see [32] for details).

To prove the matching lower bound, consider the $\mathrm{P}^{\mathrm{NP}}[O(\log n)]$-complete problem of checking validity in Carnap's modal logic. Carnap's modal logic is a nonstandard modal logic that differs substantially from the better-known Lewis' systems. In Carnap's modal logic, a subformula $\Diamond\psi$ of a formula $\varphi$ evaluates to true if $\psi$ is a consistent formula, and a subformula $\Box\psi$ evaluates to true iff $\psi$ is valid. Each modal subformula of $\varphi$ is evaluated independently of its context in $\varphi$.

The sentences true in Carnap's modal logic are precisely those sentences that are true in the fully connected Kripke structure, where each world corresponds to a finite set of propositional atoms made true, and each such set corresponds to precisely one world (see [32]).

Let $\mathsf{var}$ be a finite set of propositional variables. Let $S_{\mathsf{var}}$ be the fully connected Kripke structure, where each world corresponds to a finite set of propositional atoms from $\mathsf{var}$ made true, and each such set corresponds to precisely one world.

Let $p_i$ be the $i$-th prime number and let $P_n = \prod_{i=1}^{n} p_i$.

We construct an $LTL^{\bigcirc}_{core}$ ontology $\mathcal{O}_{\mathsf{var}}$ in the following way. The set of atomic concepts in it is

$$\{A^i_j \mid 1 \leqslant i \leqslant n, 0 \leqslant j \leqslant p_n - 1\} \cup \{X_i, \overline{X}_i \mid X_i \in \mathsf{var}\} \cup \{A, B\}.$$

The axioms of $\mathcal{O}_{\mathsf{var}}$ are

$$A \to A^i_0, \quad \text{for } 1 \leqslant i \leqslant n,$$
$$A^i_j \to \bigcirc_F A^i_{(j+1) \bmod p_i},$$
$$A^i_0 \to \overline{X}_i,$$
$$A^i_1 \to X_i,$$
$$A^i_j \to B, \quad \text{for } 1 \leqslant j \leqslant p_n - 2.$$

One can see that $|\mathcal{O}_{\mathsf{var}}|$ is polynomial in $|\mathsf{var}|$.

Let $\varphi(x_1, \dots, x_n)$ be a formula built from $x_i, 0, 1, \vee, \wedge, \neg, \Box, \Diamond$ in negation normal form with all the variables from $\mathsf{var}$. Define $\varkappa_\varphi$ inductively as follows:

$$\varkappa_{x_i} = X_i,$$
$$\varkappa_{\neg x_i} = \overline{X}_i,$$
$$\varkappa_{\varphi\vee\psi} = \varphi \vee \psi$$
$$\varkappa_{\varphi\wedge\psi} = \varphi \wedge \psi$$
$$\varkappa_{\Box\varphi} = \Box_F(B \vee \varphi)$$
$$\varkappa_{\Diamond\varphi} = \Diamond_F(\varphi).$$

Consider $\mathcal{A} = \{A(0)\}$. For any world $w \in S_{\mathsf{var}}$, there exists exactly one $n_w < P_n$ such that $n_w = 0 \bmod p_i$ iff $x_i \notin w$ and $n_w = 1 \bmod p_i$ iff $x_i \in w$. We see that, for any Boolean formula $\psi$, we have $\mathcal{O}_{\mathsf{var}}, \mathcal{A} \models \psi(n_w)$ iff $\psi$ is true in $w$. Then, for any $k > 0$, we have $\mathcal{O}_{\mathsf{var}}, \mathcal{A} \models \psi(k)$ iff $\mathcal{O}_{\mathsf{var}}, \mathcal{A} \models \psi(k + P_n)$ and if $\psi$ is a tautology then $\mathcal{O}_{\mathsf{var}}, \mathcal{A} \models \psi \vee B(k)$. By induction on the construction of $\varphi$ one can show that $\mathcal{O}_{\mathsf{var}}, \mathcal{A} \models \Box_F \varkappa_\varphi(0)$ iff $\varphi$ is valid in Carnap's logic. ❑

## 8 Conclusions

Motivated by ontology-based access to temporal data—a paradigm relying on FO-rewritability of ontology-mediated queries—we considered the problem of determining the optimal rewritability type and data complexity of answering any given $LTL$ OMQ. We showed that this problem is closely related to deciding $\mathsf{FO}(<)$-, $\mathsf{FO}(<, \equiv)$- and $\mathsf{FO}(<, \mathsf{MOD})$-definability of regular languages given by DFAs, NFAs and 2NFAs of different size. Various characterisations of $\mathsf{FO}(<)$-definability of the languages of DFAs/NFAs, deciding which is PSpace-complete, have long become classical results in automata theory. Here, we extended some of them to $\mathsf{FO}(<, \equiv)$, $\mathsf{FO}(<, \mathsf{MOD})$ and 2NFAs, establishing the same PSpace complexity bound. Based on these results, we showed how the clausal form of ontology axioms in OMQs, the temporal operators involved and the type of queries are reflected in the structure of automata accepting the OMQs' yes-data instances and the complexity of deciding their FO-definability.

Interesting open problems include understanding the impact of the $\Box$-operators in linear and core ontologies on the complexity of deciding FO-rewritability, extending our analysis to *MTL*-ontologies where OMQs are not necessarily $\mathsf{FO}(\mathsf{RPR})$-rewritable, and so are outside of $\mathrm{NC}^1$, and to 2D combinations of *LTL* with description logics, in particular *DL-Lite*.

## References


1. S. Abiteboul, R. Hull, and V. Vianu. *Foundations of Databases*. Addison-Wesley, 1995.
2. Foto N. Afrati and Christos H. Papadimitriou. The parallel complexity of simple logic programs. *J. ACM*, 40(4):891–916, 1993. doi:10.1145/153724.153752.
3. A. Artale, D. Calvanese, R. Kontchakov, and M. Zakharyaschev. The DL-Lite family and relations. *J. Artif. Intell. Res. (JAIR)*, 36:1–69, 2009.
4. A. Artale, R. Kontchakov, V. Ryzhikov, and M. Zakharyaschev. The complexity of clausal fragments of LTL. In *Proc. of the 19th Int. Conf. on Logic for Programming, Artificial Intelligence and Reasoning, LPAR 2013*, volume 8312 of *Lecture Notes in Computer Science*, pages 35–52. Springer, 2013.
5. Alessandro Artale, Roman Kontchakov, Alisa Kovtunova, Vladislav Ryzhikov, Frank Wolter, and Michael Zakharyaschev. First-order rewritability of temporal ontology-mediated queries. In *Proc. of the 24th Int. Joint Conference on Artificial Intelligence, IJCAI'15*, pages 2706–2712, 2015.
6. Alessandro Artale, Roman Kontchakov, Alisa Kovtunova, Vladislav Ryzhikov, Frank Wolter, and Michael Zakharyaschev. Ontology-mediated query answering over temporal data: A survey (invited talk). In Sven Schewe, Thomas Schneider, and Jef Wijsen, editors, *24th International Symposium on Temporal Representation and Reasoning, TIME 2017, October 16-18, 2017, Mons, Belgium*, volume 90 of *LIPIcs*, pages 1:1–1:37. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.TIME.2017.1.
7. Alessandro Artale, Roman Kontchakov, Alisa Kovtunova, Vladislav Ryzhikov, Frank Wolter, and Michael Zakharyaschev. First-order rewritability of ontology-mediated queries in linear temporal logic. *CoRR*, abs/2004.07221, 2020. URL: https://arxiv.org/abs/2004.07221, arXiv:2004.07221.
8. J.-F. Baget, M. Leclère, M.-L. Mugnier, and E. Salvat. On rules with existential variables: Walking the decidability line. *Artificial Intelligence*, 175(9–10):1620–1654, 2011.
9. David A. Mix Barrington. Bounded-width polynomial-size branching programs recognize exactly those languages in $\mathrm{NC}^1$. *J. Comput. Syst. Sci.*, 38(1):150–164, 1989.
10. David A. Mix Barrington, Kevin J. Compton, Howard Straubing, and Denis Thérien. Regular languages in $\mathrm{NC}^1$. *J. Comput. Syst. Sci.*, 44(3):478–499, 1992. doi:10.1016/0022-0000(92)90014-A.
11. David A. Mix Barrington and Denis Thérien. Finite monoids and the fine structure of $\mathrm{NC}^1$. *J. ACM*, 35(4):941–952, 1988.
12. Michael Benedikt, Balder ten Cate, Thomas Colcombet, and Michael Vanden Boom. The complexity of boundedness for guarded logics. In *30th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2015, Kyoto, Japan, July 6-10, 2015*, pages 293–304. IEEE Computer Society, 2015. doi:10.1109/LICS.2015.36.

13 M. Bennett, G. Martin, K. O'Bryant, and A. Rechnitzer. Explicit bounds for primes in arithmetic progressions. *Illinois Journal of Mathematics*, 62(1–4):427–532, 2018.

14 L. Bernátsky. Regular expression star-freeness is PSPACE-complete. *Acta Cybern.*, 13(1):1–21, 1997. URL: http://www.inf.u-szeged.hu/actacybernetica/edb/vol13n1/Bernatsky_1997_ActaCybernetica.xml.

15 M. Bienvenu, B. ten Cate, C. Lutz, and F. Wolter. Ontology-based data access: A study through disjunctive datalog, CSP, and MMSNP. *ACM Transactions on Database Systems*, 39(4):33:1–44, 2014.

16 Stefan Borgwardt, Walter Forkel, and Alisa Kovtunova. Finding new diamonds: Temporal minimal-world query answering over sparse aboxes. In Paul Fodor, Marco Montali, Diego Calvanese, and Dumitru Roman, editors, *Rules and Reasoning - Third International Joint Conference, RuleML+RR 2019, Bolzano, Italy, September 16-19, 2019, Proceedings*, volume 11784 of *Lecture Notes in Computer Science*, pages 3–18. Springer, 2019. doi:10.1007/978-3-030-31095-0_1.

17 Pierre Bourhis and Carsten Lutz. Containment in monadic disjunctive datalog, MMSNP, and expressive description logics. In Chitta Baral, James P. Delgrande, and Frank Wolter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016*, pages 207–216. AAAI Press, 2016. URL: http://www.aaai.org/ocs/index.php/KR/KR16/paper/view/12847.

18 A. Calì, G. Gottlob, and A. Pieris. Towards more expressive ontology languages: The query answering problem. *Artificial Intelligence*, 193:87–128, 2012.

19 D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. Autom. Reasoning*, 39(3):385–429, 2007.

20 Olivier Carton and Luc Dartois. Aperiodic Two-way Transducers and FO-Transductions. In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*, volume 41 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 160–174, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. URL: http://drops.dagstuhl.de/opus/volltexte/2015/5413, doi:10.4230/LIPIcs.CSL.2015.160.

21 Sang Cho and Dung T. Huynh. Finite-automaton aperiodicity is PSPACE-complete. *Theor. Comp. Sci.*, 88(1):99–116, 1991. URL: https://core.ac.uk/download/pdf/82662203.pdf.

22 C. Civili and R. Rosati. A broad class of first-order rewritable tuple-generating dependencies. In *Proc. of the 2nd Int. Datalog 2.0 Workshop*, volume 7494 of *Lecture Notes in Computer Science*, pages 68–80. Springer, 2012.

23 Kevin J. Compton and Claude Laflamme. An algebra and a logic for $NC^1$. *Inf. Comput.*, 87(1/2):240–262, 1990.

24 Stavros S. Cosmadakis, Haim Gaifman, Paris C. Kanellakis, and Moshe Y. Vardi. Decidable optimization problems for database logic programs (preliminary report). In *STOC*, pages 477–490, 1988.

25 Stéphane Demri, Valentin Goranko, and Martin Lange. *Temporal Logics in Computer Science*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2016.

26 Volker Diekert and Paul Gastin. First-order definable languages. In Jörg Flum, Erich Grädel, and Thomas Wilke, editors, *Logic and Automata: History and Perspectives [in Honor of Wolfgang Thomas]*, volume 2 of *Texts in Logic and Games*, pages 261–306. Amsterdam University Press, 2008.

27 Cristina Feier, Antti Kuusisto, and Carsten Lutz. Rewritability in monadic disjunctive datalog, MMSNP, and expressive description logics. *Logical Methods in Computer Science*, 15(2), 2019. doi:10.23638/LMCS-15(2:15)2019.

28 M. Fisher, C. Dixon, and M. Peim. Clausal temporal resolution. *ACM Trans. Comput. Logic*, 2(1):12–56, 2001.

**29** Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984.

**30** D. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Many-Dimensional Modal Logics: Theory and Applications*, volume 148 of *Studies in Logic*. Elsevier, 2003.

**31** Olga Gerasimova, Stanislav Kikot, Agi Kurucz, Vladimir V. Podolskii, and Michael Zakharyaschev. A data complexity and rewritability tetrachotomy of ontology-mediated queries with a covering axiom. In Diego Calvanese, Esra Erdem, and Michael Thielscher, editors, *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, pages 403–413, 2020. doi:10.24963/kr.2020/41.

**32** Georg Gottlob. NP trees and carnap's modal logic. *J. ACM*, 42(2):421–457, 1995. doi:10.1145/201019.201031.

**33** Gerd G. Hillebrand, Paris C. Kanellakis, Harry G. Mairson, and Moshe Y. Vardi. Undecidable boundedness problems for datalog programs. *J. Log. Program.*, 25(2):163–190, 1995. doi:10.1016/0743-1066(95)00051-K.

**34** Stasys Jukna. *Boolean Function Complexity - Advances and Frontiers*, volume 27 of *Algorithms and combinatorics*. Springer, 2012. doi:10.1007/978-3-642-24508-4.

**35** Hans W. Kamp. *Tense Logic and the Theory of Linear Order*. PhD thesis, Computer Science Department, University of California at Los Angeles, USA, 1968.

**36** Gil Kaplan and Dan Levy. Solvability of finite groups via conditions on products of 2-elements and odd p-elements. *Bulletin of the Australian Mathematical Society*, 82(2):265–273, 2010. doi:10.1017/S0004972710000079.

**37** Oliver H. King. The subgroup structure of finite classical groups in terms of geometric configurations. In Bridget S. Webb, editor, *Surveys in Combinatorics, 2005 [invited lectures from the Twentieth British Combinatorial Conference, Durham, UK, July 2005]*, volume 327 of *London Mathematical Society Lecture Note Series*, pages 29–56. Cambridge University Press, 2005.

**38** Donald Ervin Knuth. *The art of computer programming, Volume II: Seminumerical Algorithms, 3rd Edition*. Addison-Wesley, 1998. URL: https://www.worldcat.org/oclc/312898417.

**39** L. Libkin. *Elements Of Finite Model Theory*. Springer, 2004.

**40** C. Lutz, F. Wolter, and M. Zakharyaschev. Temporal description logics: A survey. In *Proc. of the 15th Int. Symposium on Temporal Representation and Reasoning (TIME 2008)*, pages 3–14, 2008.

**41** Carsten Lutz and Leif Sabellek. Ontology-mediated querying with the description logic EL: trichotomy and linear datalog rewritability. In *Proc. of the 26th Int. Joint Conf. on Artificial Intelligence (IJCAI 2017)*, pages 1181–1187, 2017.

**42** Jerzy Marcinkowski. DATALOG sirups uniform boundedness is undecidable. In *Proceedings, 11th Annual IEEE Symposium on Logic in Computer Science, New Brunswick, New Jersey, USA, July 27-30, 1996*, pages 13–24. IEEE Computer Society, 1996. doi:10.1109/LICS.1996.561299.

**43** Jerzy Marcinkowski. Achilles, turtle, and undecidable boundedness problems for small DATALOG programs. *SIAM J. Comput.*, 29(1):231–257, 1999. doi:10.1137/S0097539797322140.

**44** Antonella Poggi, Domenico Lembo, Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Riccardo Rosati. Linking data to ontologies. *J. on Data Semantics*, 10:133–173, 2008.

**45** Alexander Rabinovich. A proof of Kamp's theorem. *Logical Methods in Computer Science*, 10(1), 2014.

**46** Joseph Jonah Rotman. *An introduction to the theory of groups*. Springer-Verlag, New York; Berlin; Heidelberg [etc.], 1999. URL: http://www.worldcat.org/search?qt=worldcat_org_all&q=0387942858.

**47** Marcel Paul Schützenberger. On finite monoids having only trivial subgroups. *Inf. Control.*, 8(2):190–194, 1965. doi:10.1016/S0019-9958(65)90108-7.

48 J. C. Shepherdson. The reduction of two-way automata to one-way automata. *IBM Journal of Research and Development*, 3(2):198–200, 1959. doi:10.1147/rd.32.0198.

49 Jacques Stern. Complexity of some problems from the theory of automata. *Inf. Control.*, 66(3):163–176, 1985. doi:10.1016/S0019-9958(85)80058-9.

50 Larry J. Stockmeyer and Albert R. Meyer. Word problems requiring exponential time: Preliminary report. In Alfred V. Aho, Allan Borodin, Robert L. Constable, Robert W. Floyd, Michael A. Harrison, Richard M. Karp, and H. Raymond Strong, editors, *Proceedings of the 5th Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1973, Austin, Texas, USA*, pages 1–9. ACM, 1973. doi:10.1145/800125.804029.

51 Howard Straubing. *Finite Automata, Formal Logic, and Circuit Complexity.* Birkhauser Verlag, 1994.

52 John G. Thompson. Nonsolvable finite groups all of whose local subgroups are solvable. *Bull. Amer. Math. Soc.*, 74(3):383–437, 05 1968. URL: https://projecteuclid.org:443/euclid.bams/1183529612.

53 Jeffrey D. Ullman and Allen Van Gelder. Parallel complexity of logical query programs. *Algorithmica*, 3:5–42, 1988. doi:10.1007/BF01762108.

54 Ron van der Meyden. Predicate boundedness of linear monadic datalog is in PSPACE. *Int. J. Found. Comput. Sci.*, 11(4):591–612, 2000. doi:10.1142/S0129054100000351.

55 Moshe Y. Vardi. Decidability and undecidability results for boundedness of linear recursive queries. In Chris Edmondson-Yurkanan and Mihalis Yannakakis, editors, *Proceedings of the Seventh ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, March 21-23, 1988, Austin, Texas, USA*, pages 341–351. ACM, 1988. URL: http://doi.acm.org/10.1145/308386.308470, doi:10.1145/308386.308470.

56 Moshe Y. Vardi. A note on the reduction of two-way automata to one-way atuomata. *Inf. Process. Lett.*, 30(5):261–264, March 1989. doi:10.1016/0020-0190(89)90205-6.

57 Moshe Y. Vardi. Automata-theoretic techniques for temporal reasoning. In Patrick Blackburn, J. F. A. K. van Benthem, and Frank Wolter, editors, *Handbook of Modal Logic*, volume 3 of *Studies in logic and practical reasoning*, pages 971–989. North-Holland, 2007. doi:10.1016/s1570-2464(07)80020-6.

58 Moshe Y. Vardi and Pierre Wolper. An automata-theoretic approach to automatic program verification (preliminary report). In *Proc. of the Symposium on Logic in Computer Science (LICS'86)*, pages 332–344, 1986.

59 Przemyslaw Andrzej Walega, Bernardo Cuenca Grau, Mark Kaminski, and Egor V. Kostylev. Datalogmtl over the integer timeline. In Diego Calvanese, Esra Erdem, and Michael Thielscher, editors, *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, pages 768–777, 2020. doi:10.24963/kr.2020/79.

60 Przemyslaw Andrzej Walega, Bernardo Cuenca Grau, Mark Kaminski, and Egor V. Kostylev. Tractable fragments of datalog with metric temporal operators. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, pages 1919–1925. ijcai.org, 2020. doi:10.24963/ijcai.2020/266.

61 Guohui Xiao, Diego Calvanese, Roman Kontchakov, Domenico Lembo, Antonella Poggi, Riccardo Rosati, and Michael Zakharyaschev. Ontology-based data access: A survey. In Jérôme Lang, editor, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden.*, pages 5511–5519. ijcai.org, 2018. doi:10.24963/ijcai.2018/777.

62 Guohui Xiao, Linfang Ding, Benjamin Cogrel, and Diego Calvanese. Virtual knowledge graphs: An overview of systems and use cases. *Data Intell.*, 1(3):201–223, 2019. doi:10.1162/dint_a_00011.