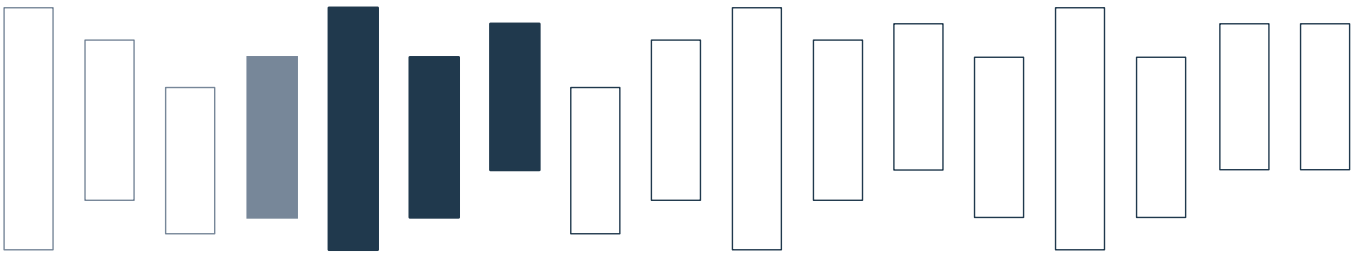
Technical Report AC-TR-15-008

April 2015

First-Order Queries on Finite Abelian Groups

Simone Bova and Barnaby Martin



This is the authors' copy of a paper that appears in the Proceedings of CSL, 2015.

www.ac.tuwien.ac.at/tr

First-Order Queries on Finite Abelian Groups*

Simone Bova[1] and Barnaby Martin[2]

1 **Vienna University of Technology**
Vienna, Austria
simone.bova@tuwien.ac.at
2 **Middlesex University**
London, United Kingdom
b.martin@mdx.ac.uk

Abstract

We study the computational problem of checking whether a logical sentence is true in a finite abelian group. We prove that model checking first-order sentences on finite abelian groups is fixed-parameter tractable, when parameterized by the size of the sentence. We also prove that model checking monadic second-order sentences on finite abelian groups finitely presented by integer matrices is not fixed-parameter tractable (under standard assumptions in parameterized complexity).

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems – Computations on Discrete Structures, F.4.1 Mathematical Logic – Model theory, G.2.1 Combinatorics – Combinatorial Algorithms.

Keywords and phrases Finite Abelian Groups, First-Order Logic, Monadic Second-Order Logic

Digital Object Identifier 10.4230/LIPIcs.CSL.2015.41

1 Introduction

The *model checking* problem for first-order logic is the problem of deciding whether a given first-order sentence is true in a given finite structure; it encompasses a wide range of fundamental combinatorial problems. The problem is trivially decidable in $O(n^k)$ time, where n is the size of the structure and k is the size of the sentence, but it is not polynomial-time decidable or even *fixed-parameter tractable* when parameterized by k (under complexity assumptions in classical and parameterized complexity, respectively).

Restrictions of the model checking problem to fixed classes of structures have been intensively investigated from the perspective of parameterized algorithms and complexity. Starting from seminal work by Courcelle [4] and Seese [18], structural properties of *graphs* sufficient for fixed-parameter tractability of model checking have been identified, culminating in the recent result by Grohe, Kreutzer, and Siebertz that model checking first-order logic on classes of *nowhere dense* graphs is fixed-parameter tractable [10]. On graph classes closed under subgraphs the result is known to be tight; at the same time, there are classes of *somewhere dense* graphs (not closed under subgraphs) with fixed parameter tractable first-order (and even monadic second-order) logic model checking; the prominent examples are graph classes of bounded clique-width solved by Courcelle, Makowsky, and Rotics [5].

In contrast to its mature understanding on graphs, the model checking problem has been very little investigated on classes of structures characterized by mathematical properties,

* The first author was supported by FWF grant P26200. The second author was supported by EPSRC grant EP/L005654/1.

24th EACSL Annual Conference on Computer Science Logic (CSL 2015).
Editor: Stephan Kreutzer; pp. 41–59

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

such as ordered structures or algebraic structures [9]. Recent work has redressed the former, ordered case [1, 2, 8], but hitherto little has been done in the latter, algebraic case.

Finite *groups* are fundamental in mathematics and computer science, and are perhaps the most prominent candidate to propose an investigation in this domain. Computational problems on finite groups are important and challenging. The notorious group isomorphism problem has long been known to be solvable in quasipolynomial time; it remains a huge open problem whether this can be improved to polynomial [16].

In contrast to general finite groups, the nice structure of finite *abelian* groups makes their associated problems simpler, both technically and computationally; isomorphism queries can be answered in linear time [13]. Yet, abelian groups remain a very important subclass; in finite model theory they appear in the literature on constraint satisfaction problems since the seminal work of Feder and Vardi [6].

Contribution. In this paper, we study the problem of model checking first-order logic on finite abelian groups. Our first contribution is a positive answer to a question posed by Grohe [9, Problem 8.2].

▶ **Result 1.** *Model checking first-order sentences on finite abelian groups, parameterized by the size of the sentence, is fixed-parameter tractable in linear time with a nonelementary parameter dependence.*

The proof is based on a revisiting of Baur-Monk's theorem on quantifier elimination in modules [11, Theorem A.1.1], which provides fresh insight into this classical result, both on important computational aspects of the class of sentences where quantifiers are eliminated, and on the mechanics the elimination procedure itself.

The theorem provides an effective procedure for reducing a first-order sentence ψ to a boolean combination of *invariant sentences* that is equivalent to ψ on abelian groups; formally, invariant sentences are first-order sentences, in the prefix class Σ_2, of the form

$$\exists x_1 \dots \exists x_k \Big(\bigwedge_{1 \le i \le k} \phi_1(x_i) \wedge \bigwedge_{1 \le i < j \le k} \neg \phi_2(x_i - x_j) \Big)$$

where $\phi_1 = \exists y_1 \dots \exists y_l \bigwedge_i \alpha_i$ and $\phi_2 = \exists z_1 \dots \exists z_m \bigwedge_j \beta_j$ are primitive positive formulas in one free variable (and k, l, and m grow with ψ).

It is unclear whether invariant sentences can be model checked in polynomial time on finite abelian groups;[1] if true, this would immediately imply a fixed-parameter tractable algorithm for model checking first-order logic on finite abelian groups. However, invariant sentences express bounds on the index of primitively positively definable subgroups (of an abelian group) into each other; for instance, the example above states that the index of the subgroup defined by $\phi_1 \wedge \phi_2$ in the subgroup defined by ϕ_1 is at least k. Therefore, if the underlying (abelian) group is finite, by Lagrange's theorem checking an invariant sentence reduces to computing the ratio between the orders of ϕ_1 and $\phi_1 \wedge \phi_2$, which is in turn the problem of counting the number of elements satisfying a primitive positive formula in one free variable in a finite abelian group.

The latter is feasible in polynomial time, and indeed in two ways: either by reducing to a linear number of calls to the algorithm by Bulatov and Dalmau for constraint satisfaction

[1] Owing to Szmielew [20], we can even assume $l = m = 1$. In this case, we can eliminate y_1 and z_1 by instantiating on all elements in the target structure, and then reduce to (a disjunction of) existentially closed conjunctions of equalities and inequalities. But this syntactic form is readily verified to be computationally hard in general.

problems on Maltsev constraints [3]; or, more directly, by reducing to a quadratic check of a formula in 2-variable logic (that is, built using only two variable symbols) using algebraic techniques.

We conclude the commentary of our first result remarking that the actual implementation of the elimination procedure, described in Section 4, is technically nontrivial, and is explicit enough to enlighten an upper bound (albeit a nonelementary one) on its complexity (which remains fairly hidden in the rather concise presentations of Baur-Monk elimination available in the literature).

In a quest to provide a measure of tightness for our first result, we investigated the problem of model checking monadic second-order logic on finite abelian groups, yet another question posed by Grohe [9]. Unfortunately we cannot answer this question, but at least we can prove the following.

▶ **Result 2.** *Model checking monadic second-order sentences on* succinctly presented *finite abelian groups, parameterized by the size of the sentence, is not fixed-parameter tractable (unless* $\mathrm{W}[1] \subseteq \mathrm{FPT}$*).*

In this setup, the group is not given as usual by its multiplication table (whose size is quadratic in the order of the group), but instead it is given by what we call a *succinct presentation*. This is a finite presentation in the usual sense [17], encoded by an integer matrix whose entries are encoded in binary as it is customary, for instance, in computational group theory and computer algebra systems. Roughly, a finite presentation is a formula of size $O(\log n)$ capable of representing a group of size n; in succinct presentations, such a representation power is already attained by formulas of size $O(\log \log n)$.

It is clear that checking formulas on structures represented succinctly is, in principle, harder. Indeed, we establish our second result by giving a fixed-parameter tractable reduction from the clique problem (parameterized by the size of the clique) to the problem of model checking monadic second-order sentences on succinctly presented finite abelian groups (parameterized by the size of the sentence).

The idea of the reduction is as follows. By the fundamental theorem [17], every finite abelian group admits a canonical decomposition as a direct sum of prime power order cyclic groups. Now, each vertex of the given graph is associated to a prime number and each edge to a positive integer; and the finite abelian group derived from the graph has a direct summand for each edge leaving each vertex (hence the direct summands are twice as much as the edges), whose prime power order is equal to the prime associated to the vertex raised to the positive integer associated to the edge (this group has a succinct presentation of linear size).

Then the key technical observation is that, despite monadic second-order logic cannot express that two sets have the same size, it can indeed express that two subsets of two cyclic subgroups of a group have the same size. Building on this, we can express by monadic second-order formulas that two direct summands of the group have the same base, or the same exponent, and therefore easily reduce a clique query on the given graph to an equivalent monadic second-order query (only depending on the size of the clique) on the derived group.

Organization. The paper is organized as follows. In Section 2, we prepare terminology and notation. In Section 3, we establish the crucial lemmas in preparation of the result on first-order logic, presented in Section 4. In Section 5, we present the result on monadic second-order logic.

2 Preliminaries

We recall some basic terminology and notation on logic, groups, and complexity, and refer the reader to any standard textbook for further details [17, 7].

For $n \geq 1$ integer, we let $[n]$ denote $\{1, \ldots, n\}$.

Logical Formulas. Throughout the paper, we work on the vocabulary $\gamma = \{+, -, 0\}$, where $+$ is a binary operation symbol, $-$ is a unary operation symbol, and 0 is a constant symbol. An *atom* has the form $t = s$ where t and s are terms built using the operation symbols in γ. We freely use the shortcut nx for the term $x + \cdots + x$ if $n > 0$, or the term $-(x + \cdots + x)$ if $n < 0$, where x occurs n times; we also write $x - y$ instead of $x + (-y)$. A *literal* is an atom or a negated atom. For every set $\{x_1, \ldots, x_l\}$ of variables, we let $\mathcal{FO}(x_1, \ldots, x_l)$ denote the class of all first-order formulas (with equality) built over γ and having free variables among $x_1, \ldots, x_l$. We let $\mathcal{FO}$ denote the class of all first-order sentences (with equality) built over γ. A first-order formula in $\mathcal{FO}(x_1, \ldots, x_l)$ is *primitive positive* if it is built from atoms using conjunction ($\wedge$) and existential quantification ($\exists$). We let $\mathcal{PP}(x_1, \ldots, x_l)$ denote the class of all primitive positive formulas in $\mathcal{FO}(x_1, \ldots, x_l)$, and $\mathcal{PP}$ denote the class of all primitive positive sentences in $\mathcal{FO}$. Similarly, for every set $\{X_1, \ldots, X_m\}$ of set variables and every set $\{x_1, \ldots, x_l\}$ of individual variables, we let $\mathcal{MSO}(X_1, \ldots, X_m, x_1, \ldots, x_l)$ denote the class of all monadic second-order formulas (with equality) built over γ and having free variables among $X_1, \ldots, X_m, x_1, \ldots, x_l$. We let $\mathcal{MSO}$ denote the class of all monadic second-order sentences (with equality) built over γ. We freely use standard shortcuts, for instance $X \subseteq Y$ instead of $(\forall x)(x \in X \to x \in Y)$, et cetera, and occasionally write

$$\begin{pmatrix} \phi_1 \\ \vdots \\ \phi_n \end{pmatrix}$$

instead of $(\phi_1 \wedge \cdots \wedge \phi_n)$.

If $\mathbb{A}$ is a structure and $\psi(X_1, \ldots, X_m, x_1, \ldots, x_l)$ is a formula, both on the same vocabulary, and f is an assignment of $X_1, \ldots, X_m$ in $\mathcal{P}(A)$ and $x_1, \ldots, x_l$ in A, we write $\mathbb{A}, f \models \psi$ if ψ is true in $\mathbb{A}$ under the assignment f. We also liberally write $\mathbb{A} \models \psi(A_1, \ldots, A_m, a_1, \ldots, a_l)$ to indicate that ψ is true in $\mathbb{A}$ under the assignment sending X_i to $A_i \in \mathcal{P}(A)$ and x_i to $a_i \in A$. Moreover, we write $\psi(X_1, \ldots, X_m, x_1, \ldots, x_l)^{\mathbb{A}}$, or $\psi^{\mathbb{A}}$ in short, to denote the set of all tuples $((A_1, \ldots, A_m), (a_1, \ldots, a_l))$ in $\mathcal{P}(A)^m \times A^l$ such that $\mathbb{A} \models \psi(A_1, \ldots, A_m, a_1, \ldots, a_l)$.

Group Theory. We view a group as a structure $\mathbb{G} = (G, +^{\mathbb{G}}, -^{\mathbb{G}}, 0^{\mathbb{G}})$ on vocabulary γ where $+^{\mathbb{G}}$ is an operation satisfying the group axioms, $0^{\mathbb{G}}$ denotes its identity element, and $-^{\mathbb{G}}g$ denotes the inverse element of $g \in G$. The group is *finite* if its order, $|G|$, is finite.

Let $\mathbb{G}$ be a group. A nonempty subset $S \subseteq G$ is (the universe of) a *subgroup* $\mathbb{S}$ of $\mathbb{G}$ if $0^{\mathbb{G}} \in S$, $-^{\mathbb{G}}s \in S$ for all $s \in S$, and $s +^{\mathbb{G}} s' \in S$ for all $s, s' \in S$. It is known that $S \subseteq G$ is a subgroup of $\mathbb{G}$ if and only if S is nonempty and $s -^{\mathbb{G}} s' \in S$ for all $s, s' \in S$; in the finite, $S \subseteq G$ is a subgroup of $\mathbb{G}$ if and only if S is nonempty and $s +^{\mathbb{G}} s' \in S$ for all $s, s' \in S$.

Let $\mathbb{G}$ be a group, let $\mathbb{S}$ be a subgroup of $\mathbb{G}$, and let $g \in G$. The *(right) coset* of $\mathbb{S}$ in $\mathbb{G}$ with respect to g, denoted by $S + g$, is the set $\{s +^{\mathbb{G}} g \colon s \in S\}$. It is known that the cosets of $\mathbb{S}$ in $\mathbb{G}$ are either identical or disjoint, and all have the same size (equal to the order of S, as S is itself a coset). Hence, the set of all cosets of $\mathbb{S}$ in $\mathbb{G}$ forms a partition of G. Consider the case where G is finite. Then, by Lagrange's theorem, the order of S divides the order of G, and $|G|/|S|$ is the number of cosets of $\mathbb{S}$ partitioning $\mathbb{G}$, known as the *index* of $\mathbb{S}$ in $\mathbb{G}$.

A group $\mathbb{G}$ is *abelian* if the operation $+^{\mathbb{G}}$ is commutative. We let $\mathcal{AG}_{\text{fin}}$ denote the class of finite abelian groups. Let $\mathbb{Z}(p,e)$ denote the cyclic group of order p^e (or equivalently the additive group modulo p^e, that is $\{0,1,\ldots,p^e-1\}$ equipped with addition modulo p^e), where p is a prime number and e a positive integer. Every finite abelian group is isomorphic to a direct sum of prime power order cyclic groups, called *primary decomposition*,

$$\mathbb{Z}(p_1,e_{1,1}) \oplus \cdots \oplus \mathbb{Z}(p_1,e_{1,n_1}) \oplus \cdots \oplus \mathbb{Z}(p_m,e_{m,1}) \oplus \cdots \oplus \mathbb{Z}(p_m,e_{m,n_m}),$$

where the p_i are prime numbers and the exponents $e_{i,j}$ are positive integers uniquely determined by the isomorphism type of the group.

A *succinct presentation* of an abelian group is a finite presentation of an abelian group encoded by an integer matrix, whose entries are encoded in binary, as customary in computational group theory. The abelian group finitely presented by the $m \times n$ integer matrix $A \in \mathbb{Z}^{m\times n}$ is the abelian group generated by the n generators $x_1,\ldots,x_n$, subject to the m relations $a_{i,1}x_1 + \ldots + a_{i,n}x_n = 0$ for $i \in [m]$. Intuitively, a binary (instead of a unary) encoding for the integer entries of the matrix corresponds to encode a term ax in size logarithmic (instead of linear) in the absolute value of a, which motivates our terminology. We let $\mathcal{AG}_{\text{spfin}}$ denote the class of all succinctly presented finite abelian groups.

Model Checking. We study the parameterized complexity of the following two computational problems. First, the problem of model checking first-order logic on finite abelian groups, in symbols $\text{MC}(\mathcal{AG}_{\text{fin}},\mathcal{FO})$, that is the problem of deciding, given $\mathbb{A} \in \mathcal{AG}_{\text{fin}}$ and $\psi \in \mathcal{FO}$, whether $\mathbb{A} \models \psi$. Second, the problem of model checking monadic second-order logic on succinctly presented finite abelian groups, in symbols $\text{MC}(\mathcal{AG}_{\text{spfin}},\mathcal{MSO})$, that is the problem of deciding, given a succinct presentation $A \in \mathbb{Z}^{m\times n}$ of a finite abelian group $\mathbb{A}$ and a sentence $\psi \in \mathcal{MSO}$, whether $\mathbb{A} \models \psi$. We regard both problems as parameterized problems, where instance $(\mathbb{A},\psi)$ is parameterized by the size of ψ.

3 Basic Facts

In this section we collect some crucial facts about the combinatorics of cosets in finite groups and about primitive positive logic over abelian groups.

We start mining, from the proof of Baur-Monk quantifier elimination theorem [11, Theorem A.1.1], a nice combinatorial property of cosets in finite groups. Roughly, in a finite group, the size of a union of cosets equals the size of the corresponding union of subgroups, hence computing the size of a union of cosets reduces to an elementary counting problem on the corresponding subgroups.

▶ **Lemma 1.** *Let $\mathbb{A}$ be a finite group. Let $\mathbb{G}$ and $\mathbb{H}_i$ $(i \in I)$ be subgroups of $\mathbb{A}$. Let C be a coset of $\mathbb{G}$ in $\mathbb{A}$ and let D_i be a coset of $\mathbb{H}_i$ in $\mathbb{A}$ $(i \in I)$. Then $C \subseteq \bigcup_{i\in I} D_i$ if and only if*

$$0 = \sum_J (-1)^{|J|} \frac{|G \cap \bigcap_{i\in J} H_i|}{|G \cap \bigcap_{i\in I} H_i|}$$

where J ranges over all subsets of I such that $C \cap \bigcap_{i\in J} D_i \neq \emptyset$.

Proof. Let $\mathbb{N}$ denote the subgroup of $\mathbb{A}$ with universe $N = G \cap \bigcap_{i\in I} H_i$. Let $C/N = \{N + c\colon c \in C\}$. In words, C/N is the set of (right) cosets of $\mathbb{N}$ in $\mathbb{A}$ with respect to elements in $C \subseteq A$. Similarly, let $D_i/N = \{N + d\colon d \in D_i\}$, $i \in I$.

Since C is a coset of $\mathbb{G}$ in $\mathbb{A}$, and $\mathbb{N}$ is a subgroup of $\mathbb{G}$, C is a (disjoint) union of cosets of $\mathbb{N}$ in $\mathbb{A}$. Similarly, D_i is a (disjoint) union of cosets of $\mathbb{N}$ in $\mathbb{A}$ ($i \in I$), and hence $\bigcup_{i\in I} D_i$ is a (disjoint) union of cosets of $\mathbb{N}$ in $\mathbb{A}$. Therefore, $C \subseteq \bigcup_{i\in I} D_i$ if and only if

$$C/N \subseteq \bigcup_{i\in I} D_i/N.$$

Since $\mathbb{A}$ is finite, C/N and D_i/N for all $i \in I$ are finite. By elementary combinatorics, if $B, B_1, \ldots, B_n$ are finite sets, then $B \subseteq \bigcup_{i\in[n]} B_i$ if and only if $0 = \sum_{I\subseteq[n]} (-1)^{|I|} |B \cap \bigcap_{i\in I} B_i|$ [12, Proposition 3.2]. Hence, $C/N \subseteq \bigcup_{i\in I} D_i/N$ if and only if

$$0 = \sum_{J\subseteq I} (-1)^{|J|} |C/N \cap \bigcap_{i\in J} D_i/N|.$$

Moreover, $C/N \cap \bigcap_{i\in J} D_i/N = (C \cap \bigcap_{i\in J} D_i)/N$ for all $J \subseteq I$, hence we reduce to

$$0 = \sum_{J\subseteq I} (-1)^{|J|} |(C \cap \bigcap_{i\in J} D_i)/N|.$$

If $C \cap \bigcap_{i\in J} D_i = \emptyset$ for some $J \subseteq I$, then the corresponding term does not contribute to the sum. Otherwise, $|(C \cap \bigcap_{i\in J} D_i)/N| = |(G \cap \bigcap_{i\in J} H_i)/N|$, and by Lagrange's theorem $|(G \cap \bigcap_{i\in J} H_i)/N| = |G \cap \bigcap_{i\in J} H_i|/|N|$, thus reducing to

$$0 = \sum_{J} (-1)^{|J|} \frac{|G \cap \bigcap_{i\in J} H_i|}{|G \cap \bigcap_{i\in I} H_i|}$$

where J ranges over all subsets of I such that $C \cap \bigcap_{i\in J} D_i \neq \emptyset$. ◀

We now make a few observations about primitive positive logic on abelian groups, starting from the folklore fact that, on abelian groups, primitive positive formulas in one free variable (respectively, with parameters) define subgroups (respectively, cosets).

▶ **Proposition 1.** *Let $\mathbb{A}$ be an abelian group.*

- *Let $\pi \in \mathcal{PP}(x)$. Then $\pi^{\mathbb{A}}$ is a subgroup of $\mathbb{A}$.*
- *Let $\pi \in \mathcal{PP}(x_1, \ldots, x_l)$ and $f\colon \{x_1, \ldots, x_{l-1}\} \to A$. If $\pi^{\mathbb{A},f} \neq \emptyset$, then $\pi^{\mathbb{A},f}$ is a coset in $\mathbb{A}$ of the subgroup $\pi(0, \ldots, 0, x_l)^{\mathbb{A}}$ of $\mathbb{A}$.*

We conclude the section describing an algorithm that, given a primitive positive formula in one free variable, returns a primitive positive formula, equivalent on abelian groups, written using only two distinct variable symbols. The algorithm is based on the computation of the Smith normal form of an integer matrix [15]; this algebraic technique is known to improve the syntactic form of primitive positive formulas [11, Lemma A.2.1], but its link with 2-variable logic is firstly and fruitfully observed here.

▶ **Proposition 2.** *There exists a single exponential time algorithm that, given a formula $\pi \in \mathcal{PP}(x)$, returns a formula $\rho \in \mathcal{PP}(x)$ of the form*

$$\rho = \bigwedge_i \exists y(c_i x = d_i y), \tag{1}$$

$c_i, d_i \in \mathbb{Z}$, such that ρ is equivalent to π on abelian groups.

Proof. Note that $\pi \in \mathcal{PP}(x)$ is equivalent on abelian groups to

$$\exists z_1 \dots \exists z_m \bigwedge_{i\in[n]} (r_i x = \sum_{j\in[m]} s_{ij} z_j)$$

where $r_i, s_{ij} \in \mathbb{Z}$, which can be displayed in matrix notation as

$$\exists z_1 \dots \exists z_m \left(R \begin{pmatrix} x \end{pmatrix} = S \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \right) \tag{2}$$

where $R \in \mathbb{Z}^{n\times 1}$ and $S \in \mathbb{Z}^{n\times m}$. By Smith's theorem, there exist invertible (square) matrices X and Y of orders m and n respectively such that XSY is diagonal. Therefore, upon replacing R by $XR = C$, S by $XSY = D$, and $(z_1, \dots, z_m)^T$ by $Y^{-1}(z_1, \dots, z_m)^T$, we have that (2) is equivalent on abelian groups to

$$\exists w_1 \dots \exists w_m \left(C \begin{pmatrix} x \end{pmatrix} = D \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \right) \tag{3}$$

where $C \in \mathbb{Z}^{n\times 1}$ and $D \in \mathbb{Z}^{n\times m}$ is diagonal.

Putting (3) back in formula notation and proceeding by logical principles, we have the following chain of equivalences on abelian groups, leading to the desired form:

$$\begin{aligned}
\pi &\equiv \exists z_1 \dots \exists z_m \bigwedge_{i\in[n]} (r_i x = \sum_{j\in[m]} s_{ij} z_j) \\
&\equiv \exists w_1 \dots \exists w_m \bigwedge_{i\in[n]} (c_i x = \sum_{j\in[m]} d_{ij} w_j) \\
&\equiv \exists w_1 \dots \exists w_m \bigwedge_{i\in[n]} (c_i x = d_{ii} w_i) \\
&\equiv \bigwedge_{i\in[n]} \exists w_i (c_i x = d_{ii} w_i) \\
&\equiv \bigwedge_{i\in[n]} \exists y (c_i x = d_{ii} y).
\end{aligned}$$

We conclude showing that ρ is computable in time single exponential in the size of π. There is an algorithm that computes D, X, and Y in time polynomial in m, n, and $s_* = \max_{i\in[m],j\in[n]} |s_{ij}|$; the integer entries in D and X have (absolute) value bounded above singly exponentially in $\max\{m, n\}$ and $\log s_*$ [19, Proposition 7.20 and Proposition 8.10]. [2] Since m, n, and s_*, as well as the entries in R, are bounded above by the size of π, it follows that the integers entries in C and D are bounded singly exponentially by the size of π. Hence ρ has size single exponential in the size of π, and is computable in time single exponential in the size of π. ◀

The nice algorithmic consequence of Proposition 2 is that we reduce the problem of computing $|\pi^{\mathbb{A}}|$, where π is a primitive positive formula on one free variable, to the problem

[2] The model of computation is an arithmetic RAM, but the algorithm translates into a polynomial-time algorithm on a standard RAM.

of computing $|\rho^{\mathbb{A}}|$, where ρ is a formula using only two variables. The latter merely requires quadratic work in the size of the structure: namely, there exists an algorithm that, given a finite abelian group $\mathbb{A}$ and a primitive positive formula $\rho(x)$ as in (1), computes the size of $\rho^{\mathbb{A}} = \{a \in A \colon \mathbb{A} \models \rho(a)\}$ in $O(k|A|^2)$ time, where k is the size of ρ.

Alternatively, for primitive positive formulas $\pi(x)$ on one free variable, it is possible to show that the problem of determining $|\pi^{\mathbb{A}}|$ is solvable in time polynomial in the size of $\mathbb{A}$ and π by calling $|A|$ times the algorithm by Bulatov and Dalmau for constraint satisfaction problems on Maltsev constraints [3]. We prefer the elementary approach of Proposition 2, as for our algorithmic result the exponential increase in size of ρ with respect to π is negligible.

4 First-Order Queries

In this section, we prove that model checking first-order logic on finite abelian groups is fixed-parameter tractable. Let $\mathbb{A}$ be a finite abelian group and let ψ be a first-order sentence in prenex form,

$$\psi = Q_1 x_1 \dots Q_m x_m \phi \tag{4}$$

where the Q_i are quantifiers, $\exists$ or $\forall$, and ϕ is a boolean combination of atoms.

We describe the algorithm referring to the pseudocode below (where $\triangleright$ denotes a comment, and $\leftarrows$ denotes an assignment). The subprocedure $\mathrm{FO}^2(\cdot)$ on Lines 7 and 10 is the algorithm in the statement of Proposition 2. The input is a pair $(\mathbb{A}, \psi)$, where $\mathbb{A}$ is a finite abelian group and ψ is a first-order sentence specified as above (Line 1).

The algorithm loops on $l = m, \dots, 1$ and constructs a first-order sentence

$$\psi_{l-1} = Q_1 x_1 \dots R_{l-1} x_{l-1} \phi_{l-1},$$

where $R_{l-1} \in \{\exists, \neg\exists\}$, such that $\mathbb{A} \models \psi_{l-1}$ if and only if $\mathbb{A} \models \psi$, and ϕ_{l-1} is a boolean combination of primitive positive formulas with free variables among $x_1, \dots, x_{l-1}$ (Lines 2-23). Intuitively, the algorithm computes ψ_{l-1} from ψ_l by "eliminating" the quantifier on variable x_l (Lines 6-14).

It follows that ψ_0 is a boolean combination, denote it by $\mathrm{bool}(\mu_1, \dots, \mu_L)$, of primitive positive sentences $\mu_1, \dots, \mu_L$ (Line 24). Moreover, $\mathbb{A} \models \psi$ if and only if $\mathbb{A} \models \psi_0$. Since each primitive positive sentence is true in $\mathbb{A}$, it holds that $\mathbb{A} \models \psi_0$ if and only if $\mathbb{A} \models \mathrm{bool}(\top, \dots, \top)$, which is easily checked (Lines 25-26).

ModelCheck$(\mathbb{A}, \psi)$
1 $\quad \triangleright\ \psi$ as in (4)
2 $\quad$ **if** $Q_m = \exists$ **then** $\psi_m \leftarrows Q_1 x_1 \dots Q_{m-1} x_{m-1} \exists x_m \mathrm{dnf}(\phi)$
3 $\quad$ **else** $\psi_m \leftarrows Q_1 x_1 \dots Q_{m-1} x_{m-1} \neg\exists x_m \mathrm{dnf}(\neg\phi)$
4 $\quad$ **for** $l = m, \dots, 1$
5 $\quad\quad \triangleright\ \psi_l = Q_1 x_1 \dots Q_{l-1} x_{l-1} R_l x_l \bigvee_{i \in I} (\pi_i \wedge \bigwedge_{j \in J_i} \neg\pi_{ij})$ where $\pi_i, \pi_{ij} \in \mathcal{PP}(x_1, \dots, x_l)$
6 $\quad\quad$ **forall** $i \in I, M \subseteq J_i, X \subseteq \mathcal{P}(M)$
7 $\quad\quad\quad \sigma_{i,M} \leftarrows \mathrm{FO}^2((\pi_i \wedge \bigwedge_{j \in M} \pi_{ij})(0, \dots, 0, x_l))$
8 $\quad\quad\quad C_{i,M} \leftarrows |\sigma_{i,M}^{\mathbb{A}}|$
9 $\quad\quad\quad$ **forall** $Y \in X$
10 $\quad\quad\quad\quad \rho_{i,Y} \leftarrows \mathrm{FO}^2((\pi_i \wedge \bigwedge_{j \in Y} \pi_{ij})(0, \dots, 0, x_l))$
11 $\quad\quad\quad\quad C_{i,Y} \leftarrows |\rho_{i,Y}^{\mathbb{A}}|$
12 $\quad\quad\quad$ **if** $0 = \sum_{Y \in X} (-1)^{|Y|} (C_{i,Y} / C_{i,M})$ **then** $\theta_{i,M,X} \leftarrows \top$ **else** $\theta_{i,M,X} \leftarrows \bot$

13 $\quad \theta_{i,M} \leftrightharpoons \bigwedge_{X \subseteq \mathcal{P}(M)} \left(\begin{pmatrix} \bigwedge_{Y \in X} \exists x_l(\pi_i \wedge \bigwedge_{j \in Y} \pi_{ij}) \\ \bigwedge_{Y \in \mathcal{P}(M) \setminus X} \neg\exists x_l(\pi_i \wedge \bigwedge_{j \in Y} \pi_{ij}) \end{pmatrix} \to \theta_{i,M,X} \right)$

14 $\quad \phi_{l-1} \leftrightharpoons \neg \bigwedge_{i \in I} \bigwedge_{M \subseteq J_i} \left(\begin{pmatrix} \exists x_l \pi_i \\ \bigwedge_{j \in M} \exists x_l \pi_{ij} \\ \bigwedge_{j \in J_i \setminus M} \neg\exists x_l \pi_{ij} \end{pmatrix} \to \theta_{i,M} \right)$

15 $\quad \triangleright$ ϕ_{l-1} boolean combination of primitive positive formulas with $x_1, \ldots, x_{l-1}$ free

16 $\quad$ **case** $Q_{l-1} = \exists, R_l = \exists$ **:**

17 $\quad\quad \psi_{l-1} \leftrightharpoons Q_1 x_1 \ldots \exists x_{l-1} \mathrm{dnf}(\phi_{l-1})$

18 $\quad$ **case** $Q_{l-1} = \exists, R_l = \neg\exists$ **:**

19 $\quad\quad \psi_{l-1} \leftrightharpoons Q_1 x_1 \ldots \exists x_{l-1} \mathrm{dnf}(\neg\phi_{l-1})$

20 $\quad$ **case** $Q_{l-1} = \forall, R_l = \exists$ **:**

21 $\quad\quad \psi_{l-1} \leftrightharpoons Q_1 x_1 \ldots \neg\exists x_{l-1} \mathrm{dnf}(\neg\phi_{l-1})$

22 $\quad$ **case** $Q_{l-1} = \forall, R_l = \neg\exists$ **:**

23 $\quad\quad \psi_{l-1} \leftrightharpoons Q_1 x_1 \ldots \neg\exists x_{l-1} \mathrm{dnf}(\phi_{l-1})$

24 $\triangleright$ $\psi_0 = \mathrm{bool}(\mu_1, \ldots, \mu_L)$ boolean combination of primitive positive sentences

25 **if** $\mathbb{A} \models \mathrm{bool}(\top, \ldots, \top)$ **then accept**

26 **reject**

We now prove that the algorithm is correct.

▶ **Lemma 2.** *Let $\mathbb{A}$ be a finite abelian group and ψ be a first-order sentence specified as in* (4)*. Then $\mathbb{A} \models \psi$ if and only if* ModelCheck$(\mathbb{A}, \psi)$ *accepts.*

Proof. Let $\psi = Q_1 x_1 \ldots Q_m x_m \phi$, where ϕ is a boolean combination of atoms. For $l \in \{0, 1, \ldots, m\}$, let

$$\psi_l = Q_1 x_1 \ldots Q_{l-1} x_{l-1} R_l x_l \phi'_l,$$

be the formula computed by ModelCheck$(\mathbb{A}, \psi)$ either on Line 2 or 3 ($l = m$), or on Line 16, 18, 20, or 22 ($l < m$). Here, $R_l \in \{\exists, \neg\exists\}$.

By induction on $l = m, \ldots, 0$, we prove that:

(I1) $\mathbb{A} \models \psi$ if and only if $\mathbb{A} \models \psi_l$;

(I2) $\phi'_l = \bigvee_{i \in I}(\pi_i \wedge \bigwedge_{j \in J_i} \neg\pi_{ij})$, where the π_i and π_{ij} are primitive positive formulas on free variables $x_1, \ldots, x_l$.

It follows that $\mathbb{A} \models \psi$ if and only if $\mathbb{A} \models \psi_0$. Since ψ_0 is a boolean combination of primitive positive sentences, each true in $\mathbb{A}$, the correctness of the algorithm follows (Lines 24-26). We now give the inductive argument.

Base Case ($l = m$). Invariants (I1) and (I2) clearly hold if ψ_m is set as in Line 2 or 3. The operator $\mathrm{dnf}(\cdot)$, given a boolean combination of atoms, returns a logically equivalent disjunctive normal form.

Inductive Step ($l - 1$, $l \leq m$). By (I1) and (I2), we have inductively

$$\psi_l = Q_1 x_1 \ldots Q_{l-1} x_{l-1} R_l x_l \phi'_l,$$

$R_l \in \{\exists, \neg\exists\}$, such that $\mathbb{A} \models \psi$ if and only if $\mathbb{A} \models \psi_l$. Intuitively, the algorithm constructs the first-order sentence

$$\psi_{l-1} = Q_1 x_1 \ldots R_{l-1} x_{l-1} \phi'_{l-1},$$

satisfying invariants (I1) and (I2), by "eliminating" the quantifier on variable x_l in ψ_l, as follows.

Consider the case where $Q_{l-1} = R_l = \exists$ (Line 16), so that

$$\begin{aligned}\psi_l &= Q_1 x_1 \ldots \exists x_{l-1} \exists x_l \phi'_l \\ &= Q_1 x_1 \ldots \exists x_{l-1} \exists x_l \bigvee_{i \in I} (\pi_i \wedge \bigwedge_{j \in J_i} \neg \pi_{ij})\end{aligned}$$

where the π_i and π_{ij} are formulas in $\mathcal{PP}(x_1, \ldots, x_l)$ by the induction hypothesis on ψ_l. The remaining cases (Line 18, Line 20, and Line 22) reduce to this case by handling negations as described in the pseudocode (Lines 18-23).

For readability, we first introduce the following notation. The operator $\mathcal{P}(\cdot)$, given a finite set, returns its powerset. For $i \in I$, $M \subseteq J_i$, and $X \subseteq \mathcal{P}(M)$ let:

$$\alpha_{i,M} = \exists x_l \pi_i \wedge \bigwedge_{j \in M} \exists x_l \pi_{ij} \wedge \bigwedge_{j \in J_i \setminus M} \neg \exists x_l \pi_{ij} \tag{5}$$

$$\beta_{i,M,X} = \bigwedge_{Y \in X} \exists x_l (\pi_i \wedge \bigwedge_{j \in Y} \pi_{ij}) \wedge \bigwedge_{Y \in \mathcal{P}(M) \setminus X} \neg \exists x_l (\pi_i \wedge \bigwedge_{j \in Y} \pi_{ij}) \tag{6}$$

We now claim that,

$$\exists x_l \phi'_l = \exists x_l \bigvee_{i \in I} (\pi_i \wedge \bigwedge_{j \in J_i} \neg \pi_{ij}) \tag{7}$$

$$\equiv \bigvee_{i \in I} \exists x_l (\pi_i \wedge \bigwedge_{j \in J_i} \neg \pi_{ij}) \tag{8}$$

$$\equiv \neg \bigwedge_{i \in I} \forall x_l (\pi_i \to \bigvee_{j \in J_i} \pi_{ij}) \tag{9}$$

$$\equiv \neg \bigwedge_{i \in I} \bigwedge_{M \subseteq J_i} (\alpha_{i,M} \to \forall x_l (\pi_i \to \bigvee_{j \in M} \pi_{ij})) \tag{10}$$

$$\equiv_{\mathbb{A}} \neg \bigwedge_{i \in I} \bigwedge_{M \subseteq J_i} (\alpha_{i,M} \to \bigwedge_{X \subseteq \mathcal{P}(M)} (\beta_{i,M,X} \to \theta_{i,M,X})) \tag{11}$$

$$= \phi_{l-1} \tag{12}$$

where $\theta_{i,M,X} \in \{\bot, \top\}$.

Before proving the claim, note that ϕ_{l-1} in (12) is the formula on Line 14. By the above chain of equivalences, ϕ_{l-1} is equivalent in $\mathbb{A}$ to $\exists x_l \phi'_l$. Therefore, the formula ψ_{l-1} defined on Line 17 is equivalent to ψ on $\mathbb{A}$. Hence ψ_{l-1} satisfies invariant (I1). Moreover, since $\theta_{i,M,X}$ is either $\bot$ or $\top$ (Line 12), by inspection of Lines 13 and 14 (or (5) and (6), where we observe that the variable x_l is existentially quantified in each π_i and π_{ij}), ϕ_{l-1} is a boolean combination of formulas in $\mathcal{PP}(x_1, \ldots, x_{l-1})$. Therefore, the formula ψ_{l-1} defined on Line 17 by taking the disjunctive normal form of ϕ_{l-1} also satisfies invariant (I2), as desired.

We now prove the claim. The equivalences (8)-(9) hold by logical principles, and the equivalence (10) is readily verified. It remains to show that (11) holds, which is the crucial step of the construction. Here, the notation $\equiv_{\mathbb{A}}$ means that this equivalence is relative to the structure $\mathbb{A}$ (as opposed to the previous equivalences, that are logical equivalences holding for all structures).

By inspection of (11), it is sufficient to show that for all $i \in I$, $M \subseteq J_i$, and all

$f\colon \{x_1,\dots,x_{l-1}\} \to A$ such that $\mathbb{A}, f \models \alpha_{i,M}$, the following are equivalent:

$$\mathbb{A}, f \models \forall x_l(\pi_i \to \bigvee_{j\in M} \pi_{ij}) \tag{13}$$

$$\mathbb{A}, f \models \bigwedge_{X\subseteq\mathcal{P}(M)} (\beta_{i,M,X} \to \theta_{i,M,X}) \tag{14}$$

First we show that (13) is equivalent to a certain combinatorial statement involving cosets of primitive positive definable subgroups of $\mathbb{A}$, next we show that (14) is equivalent to $\mathbb{A}, f \models \theta_{i,M,X_*}$ for a suitably chosen $X \in \mathcal{P}(\mathcal{P}(M))$, and we conclude showing the equivalence of the combinatorial statement and $\mathbb{A}, f \models \theta_{i,M,X_*}$.

First, since $\mathbb{A}, f \models \alpha_{i,M}$, we have that $\pi_i^{\mathbb{A},f}$ and $\pi_{ij}^{\mathbb{A},f}$ are nonempty for all $j \in M$. Hence, by Proposition 1, $\pi_i^{\mathbb{A},f}$ is a coset in $\mathbb{A}$ of the subgroup $\pi_i(0,\dots,0,x_l)^{\mathbb{A}}$, and $\pi_{ij}^{\mathbb{A},f}$ is a coset in $\mathbb{A}$ of the subgroup $\pi_{ij}(0,\dots,0,x_l)^{\mathbb{A}}$ for all $j \in M$. We therefore have that (13) is equivalent to

$$\pi_i^{\mathbb{A},f} \subseteq \bigcup_{j\in M} \pi_{ij}^{\mathbb{A},f} \tag{15}$$

where $\pi_i^{\mathbb{A},f}$ and $\pi_{ij}^{\mathbb{A},f}$ are the described cosets in $\mathbb{A}$.

Next, observe that there exists exactly one $X \subseteq \mathcal{P}(M)$ such that $\mathbb{A}, f \models \beta_{i,M,X}$. Indeed, note that $\mathcal{P}(M)$ is partially ordered by the inclusion relation. Then the unique choice of X in $\mathcal{P}(\mathcal{P}(M))$ is determined as follows: X contains exactly those $Y \in \mathcal{P}(M)$ contained in some $Y' \in \mathcal{P}(M)$ that is maximal with the property that $\mathbb{A}, f \models \exists x_l(\pi_i \wedge \bigwedge_{j\in Y'} \pi_{ij})$. Let X_* denote this unique choice of X in $\mathcal{P}(\mathcal{P}(M))$. It follows that (14) is equivalent to

$$\mathbb{A}, f \models \theta_{i,M,X_*} \tag{16}$$

We are now in a position to conclude the argument. By Lemma 1, it holds that (15) is equivalent to

$$0 = \sum_{Y} (-1)^{|Y|} \frac{|(\pi_i \wedge \bigwedge_{j\in Y} \pi_{ij})(0,\dots,0,x_l)^{\mathbb{A}}|}{|(\pi_i \wedge \bigwedge_{j\in M} \pi_{ij})(0,\dots,0,x_l)^{\mathbb{A}}|} \tag{17}$$

where Y ranges on all subsets of M such that $(\pi_i \wedge \bigwedge_{j\in Y} \pi_{ij})^{\mathbb{A},f} \neq \emptyset$. Since $\mathbb{A}, f \models \beta_{i,M,X_*}$, it holds that X_* is exactly the set of all subsets Y of M such that $(\pi_i \wedge \bigwedge_{j\in Y} \pi_{ij})^{\mathbb{A},f} \neq \emptyset$. Hence (17) is equivalent to

$$0 = \sum_{Y\in X_*} (-1)^{|Y|} \frac{|(\pi_i \wedge \bigwedge_{j\in Y} \pi_{ij})(0,\dots,0,x_l)^{\mathbb{A}}|}{|(\pi_i \wedge \bigwedge_{j\in M} \pi_{ij})(0,\dots,0,x_l)^{\mathbb{A}}|} \tag{18}$$

By Proposition 2, the subprocedure $\mathrm{FO}^2(\cdot)$, given a primitive positive formula in one free variable, returns a primitive positive formula written using only 2 distinct variable symbols that is equivalent on abelian groups. Then, $\sigma_{i,M}$ on Line 7 is equivalent in $\mathbb{A}$ to $(\pi_i \wedge \bigwedge_{j\in M} \pi_{ij})(0,\dots,0,x_l)$, and $\rho_{i,Y}$ on Line 10 is equivalent in $\mathbb{A}$ to $(\pi_i \wedge \bigwedge_{j\in Y} \pi_{ij})(0,\dots,0,x_l)$. It follows that, on Line 8 and 11, we have that $C_{i,M} = |(\pi_i \wedge \bigwedge_{j\in M} \pi_{ij})(0,\dots,0,x_l)^{\mathbb{A}}|$ and $C_{i,Y} = |(\pi_i \wedge \bigwedge_{j\in Y} \pi_{ij})(0,\dots,0,x_l)^{\mathbb{A}}|$. Hence (18) is equivalent to

$$0 = \sum_{Y\in X_*} (-1)^{|Y|} (C_{i,Y}/C_{i,M}) \tag{19}$$

which happens exactly when θ_{i,M,X_*} is settled to $\top$ on Line 12, which is in turn equivalent to (16).

Summarizing, (13) is equivalent to (14), which settles (11), and hence the claim. The proof is complete. ◀

We analyze the runtime of the algorithm. We let $\exp_b^{i+1}(\cdot) = \exp_b(\exp_b^i(\cdot)) = b^{\exp_b^i(\cdot)}$.

▶ **Lemma 3.** *Let $\mathbb{A}$ be a finite abelian group and ψ be a first-order sentence in prenex form with m quantifiers. Then* ModelCheck$(\mathbb{A}, \psi)$ *runs in $\exp_2^{m+2}(O(k)) \cdot |A|^2$ time, where k is the size of ψ.*

Proof. For $l = m, \dots, 1$, let

$$\psi_l = Q_1 x_1 \dots Q_{l-1} x_{l-1} R_l x_l \bigvee_{i \in I_l} (\pi_i \wedge \bigwedge_{j \in J_{l,i}} \neg \pi_{ij}) \tag{20}$$

where $R_l \in \{\exists, \neg\exists\}$, and π_i and π_{ij} are in $\mathcal{PP}(x_1, \dots, x_l)$ for all $i \in I_l$ and $j \in J_{l,i}$. Note that ψ_l is the formula created on Lines 2-3 and Lines 16-23. For $l = m, \dots, 1$, we define a set $E_l \subseteq \mathcal{PP}(x_1, \dots, x_l)$ as follows

$$E_l = \{\pi_i, \pi_{ij} \colon i \in I_l, j \in J_{l,i}\},$$

and we let S_l be the size of the largest formula in E_l. We now prove by induction on $l = m, \dots, 1$ that

$$|E_l| \leq \exp_2^{m-l}(k) \tag{21}$$

$$S_l \leq k \prod_{j=l+1}^{m} |E_j| \tag{22}$$

where as usual the empty product equals 1 and $\exp_2^0(k) = k$.

The size of E_m is bounded above by the number of atoms in the sentence ψ given in input and the size of a formula in E_m is bounded above by the size of ψ, hence

$$|E_m| \leq k$$
$$S_m \leq k$$

For $l \leq m$, let $|E_l| = \exp_2^{m-l}(k)$ and $S_l = k \prod_{j=l+1}^{m} |E_j|$. Suffices to show that the following inequalities hold:

$$|E_{l-1}| \leq 2^{|E_l|}$$
$$S_{l-1} \leq |E_l| S_l$$

Indeed, the formula ϕ_{l-1} obtained in Line 15 (used to build ψ_{l-1} on Lines 16-23) is a boolean combination of the primitive positive formulas with free variables among $x_1, \dots, x_{l-1}$ created on Line 13 and 14 by existentially quantify the variable x_l in conjunctions of the form

$$\pi_i \wedge \bigwedge_{j \in S} \pi_{ij}$$

where S is a subset (of the index set) of E_l. Thus there are at most $2^{|E_l|}$ formulas in E_{l-1}. Moreover, by the same token, the size of a formula in E_{l-1} is bounded above by the number of formulas in E_l times the size S_l of the largest such formula.

We now analyze the runtime of the algorithm. Lines 2-3 are feasible in time single exponential in the size of the input sentence ψ. We claim that, for $l = m, \dots, 1$, the time spent on the corresponding iteration of the loop on Lines 4-23 is in $O(\exp_2^{m-l+3}(k) \cdot |A|^2)$. It follows that the whole loop on Lines 4-23 is feasible in $\exp_2^{m+2}(O(k)) \cdot |A|^2$ time (note that $m \leq k$), and the statement is settled.

We conclude the proof showing that, for $l = m, \dots, 1$, the corresponding iteration of the loop on Lines 4-23 is feasible in $O(\exp_2^{m-l+3}(k) \cdot |A|^2)$ time.

In view of (20), first note the following ($i \in I_l$ and $j \in J_{l,i}$):

- $|I_l| \leq 3^{|E_l|}$, as there are at most 3^c clauses on c distinct variables (here, a formula in E_l plays the role of a variable);
- $|J_{i,l}| \leq |E_l|$, because $J_{i,l}$ is a subset (of indices) of formulas in E_l.

As $M \subseteq J_{i,l}$ and $X \in \mathcal{P}(\mathcal{P}(M))$, it follows that the loop on Line 6 is executed at most

$$|I_l| \cdot |\mathcal{P}(J_{i,l})| \cdot |\mathcal{P}(\mathcal{P}(J_{i,l}))| \leq 3^{|E_l|} \cdot 2^{|E_l|} \cdot 2^{2^{|E_l|}}$$

times which is in $O(\exp_2^{m-l+2}(k))$.

For each iteration, $\mathrm{FO}^2(\cdot)$ in Line 7 requires time single exponential in the size of the formula given in input, by Proposition 2; the latter is a formula in E_l, hence its size is bounded above by $|E_{l+1}|S_{l+1}$, in turn in $O(|E_{l+1}|)$ by (22). Hence $\sigma_{i,M}$ has size single exponential in $|E_{l+1}|$. Then Line 8 requires time single exponential in $|E_{l+1}|$ and quadratic in $|A|$, by the remark following Proposition 2.

Line 9 iterates at most $2^{|E_l|}$ times as $|X| \leq |\mathcal{P}(M)| \leq |\mathcal{P}(J_{i,l})| \leq 2^{|E_l|}$. Each iteration requires as above time single exponential in $|E_{l+1}|$ and quadratic in $|A|$ on Lines 10 and 11, again by Proposition 2 and the surrounding discussion.

Line 12 sums at most $|X| \leq 2^{|E_l|}$ integer numbers not larger than $|A|$.

The formula $\theta_{i,M}$ on Line 13 has size at most $|\mathcal{P}(\mathcal{P}(M))| \leq 2^{2^{|E_l|}}$ (the size of the index set of the outermost conjunction), times $|\mathcal{P}(M)| \leq 2^{|E_l|}$ (the number of formulas on the left of the implication symbol in each conjunct), times S_l (the size of the largest such formula, as they belong in E_l). Thus $\theta_{i,M}$ has size at most $2^{2^{|E_l|}}2^{|E_l|}S_l$, which is in $O(\exp_2^{m-l+2}(k))$, and is computable in the same time.

The formula ϕ_{l-1} on Line 14 has size at most $|I_l| \leq 3^{|E_l|}$ times $|\mathcal{P}(M)| \leq 2^{|E_l|}$ (the sizes of the index sets of the two outermost conjunctions) times an upper bound on the size of the conjuncts. Each conjunct has one part on the left and one part on the right of the implication symbol. The part on the right is $\theta_{i,M}$ of size $O(\exp_2^{m-l+2}(k))$ by the above argument. The part on the left has size at most $|M| \leq |E_l|$ (the number of formulas on the left of the implication symbol in each conjunct) times S_l (the size of the largest such formula, as they belong in E_l). Hence, each conjunct has size at most $O(\exp_2^{m-l+2}(k))$. Therefore, ϕ_{l-1} has size at most $O(\exp_2^{m-l+2}(k))$, and is computable in the same time.

Line 17 (or 19, or 21, or 23) are feasible in time single exponential in the size of the formula ϕ_{l-1} on Line 14, hence in $O(\exp_2^{m-l+3}(k))$ time. Summarizing, iteration l is feasible in $O(\exp_2^{m-l+3}(k) \cdot |A|^2)$ time. ◀

As the encoding of $\mathbb{A}$ has size quadratic in $|A|$, we conclude the following.

▶ **Theorem 4.** $\mathrm{MC}(\mathcal{AG}_{\mathrm{fin}}, \mathcal{FO})$ *is fixed-parameter tractable in linear time (with a nonelementary parameter dependence).*

5 Monadic Second-Order Queries

In this section, we prove that model checking monadic second-order logic is not fixed-parameter tractable on *succinctly presented* finite abelian groups.

We proceed in two steps. First, we define a family of monadic second order formulas. Next, we use these formulas to define a suitable reduction.

In the scope of this section,

$$\mathbb{A} = \mathbb{Z}(p_1, e_{1,1}) \oplus \cdots \oplus \mathbb{Z}(p_1, e_{1,d_1}) \oplus \cdots \oplus \mathbb{Z}(p_n, e_{n,1}) \oplus \ldots \oplus \mathbb{Z}(p_n, e_{n,d_n}) \tag{23}$$

is a finite abelian group, presented by its primary decomposition, where the p_i are pairwise distinct prime numbers, and the $e_{i,j}$ are positive integers.

We now introduce a family of monadic second order formulas, and describe their meaning in $\mathbb{A}$. First, we identify subgroups of $\mathbb{A}$ as follows. Let:

$$\mathrm{Sg}(X) \leftrightharpoons 0 \in X \wedge (\forall x, y \in X)(x + y \in X)$$
$$\mathrm{Sg}(X, Y) \leftrightharpoons X \subseteq Y \wedge \mathrm{Sg}(X) \wedge \mathrm{Sg}(Y)$$

The following is readily verified.

- $\mathbb{A} \models \mathrm{Sg}(S)$ if and only if $S \subseteq A$ is (the universe of) a subgroup of $\mathbb{A}$ (a nonempty subset of a *finite* group is a subgroup if and only if it is closed under the group operation).
- $\mathbb{A} \models \mathrm{Sg}(R, S)$ if and only if $R \subseteq S \subseteq A$ and R and S are (universes of) subgroups $\mathbb{R}$ and $\mathbb{S}$ of $\mathbb{A}$. It follows that $\mathbb{R}$ is a subgroup of $\mathbb{S}$.

We now identify cyclic groups and their generators in $\mathbb{A}$ as follows. Let:

$$\mathrm{Cycl}(X, x) \leftrightharpoons (\forall Y \subseteq X)((0 \in Y \wedge (\forall y \in Y)(y + x \in Y)) \to Y = X)$$
$$\mathrm{Cycl}(X) \leftrightharpoons (\exists x \in X)(\forall Y)((x \in Y \wedge \mathrm{Sg}(Y, X)) \to Y = X)$$

▶ **Claim 1.** *Let $\mathbb{S}$ be a subgroup of $\mathbb{A}$ with universe $S \subseteq A$ and let $g \in S$. Then $\mathbb{A} \models \mathrm{Cycl}(S, g)$ if and only if $\mathbb{S}$ is cyclic generated by g.*

If $\mathbb{S}$ is a subgroup of $\mathbb{A}$ with universe $S \subseteq A$, it follows that $\mathbb{A} \models \mathrm{Cycl}(S)$ if and only if $\mathbb{S}$ is cyclic. Among cyclic subgroups of $\mathbb{A}$, we identify prime power order cyclic subgroups of $\mathbb{A}$ as follows. Let:

$$\mathrm{PrPow}(X) \leftrightharpoons (\forall Y, Z)((\mathrm{Sg}(Y, X) \wedge \mathrm{Sg}(Z, X)) \to (Y \subseteq Z \vee Z \subseteq Y))$$

▶ **Claim 2.** *Let $\mathbb{S}$ be a nontrivial* cyclic *subgroup of $\mathbb{A}$ with universe $S \subseteq A$. Then $\mathbb{A} \models \mathrm{PrPow}(S)$ if and only if $|S| = p^e$ for some prime number p and some positive integer e.*

Call the prime power order cyclic subgroups of $\mathbb{A}$ that do not have proper prime power order cyclic supergroups in $\mathbb{A}$ *prime terms* of $\mathbb{A}$. Let $\mathrm{PrPowCyclSg}(X) \leftrightharpoons \mathrm{Sg}(S) \wedge \mathrm{Cycl}(S) \wedge \mathrm{PrPow}(S)$ and

$$\mathrm{PrTerm}(X) \leftrightharpoons \left(\begin{array}{c} \mathrm{PrPowCyclSg}(X) \\ (\forall Y) \left(\left(\begin{array}{c} \mathrm{PrPowCyclSg}(Y) \\ X \subseteq Y \end{array} \right) \to Y = X \right) \end{array} \right)$$

By the above, it follows immediately that the $\mathbb{A} \models \mathrm{PrPowCyclSg}(S)$ if and only if $\mathbb{S}$ is a prime power order cyclic subgroup of $\mathbb{A}$, where $S \subseteq A$. Moreover, for $S \subseteq A$, it holds that $\mathbb{A} \models \mathrm{PrTerm}(S)$ if and only if $\mathbb{S}$ is a prime term of $\mathbb{A}$.

We now make a key observation. Despite monadic second order logic cannot express that two sets have the same size [14, along the lines of Proposition 7.12], indeed it can express that two subsets of two cyclic subgroups of a group have the same size. The details follow. Let:

$$\mathrm{Eq}(X, Y) = (\exists Z) \left(\begin{array}{l} (\forall x \in X)(\exists! y \in Y)(x + y \in Z) \\ (\forall y \in Y)(\exists! x \in X)(x + y \in Z) \end{array} \right)$$

▶ **Claim 3.** *Let $C \subseteq A$ and $D \subseteq A$ be subsets (of universes) of prime terms of $\mathbb{A}$.*[3] *Then $\mathbb{A} \models \mathrm{Eq}(C, D)$ if and only if $|C| = |D|$.*

Proof. For the sake of notation, let $\mathbb{T}_1$ and $\mathbb{T}_2$ be respectively the first and second term in the primary decomposition of $\mathbb{A}$, of order l and m respectively (where l and m are prime powers), and let C and D be subsets of the prime terms of $\mathbb{A}$ isomorphic to $\mathbb{T}_1$ and $\mathbb{T}_2$, respectively. Then $C = \{(c_1, 0, \dots, 0), \dots, (c_{l'}, 0, \dots, 0)\}$ and $D = \{(0, d_1, \dots, 0), \dots, (0, d_{m'}, \dots, 0)\}$, where $\{c_1, \dots, c_{l'}\} \subseteq \{0, 1, \dots, l-1\}$ and $\{d_1, \dots, d_{m'}\} \subseteq \{0, 1, \dots, m-1\}$.

Assume $|C| = |D|$, and let $b' \colon C \to D$ be a bijection. Clearly, b' is completely characterized by a bijection $b \colon \{c_1, \dots, c_{l'}\} \to \{d_1, \dots, d_{m'}\}$; in particular, $l' = m'$. Let $f(Z) = \{(c_1, b(c_1), \dots, 0), \dots, (c_{l'}, b(c_{l'}), \dots, 0)\}$. We show that

$$\mathbb{A}, f \models (\forall x \in C)(\exists! y \in D)(x + y \in Z) \wedge (\forall y \in D)(\exists! x \in C)(x + y \in Z).$$

Let $(c, 0, \dots, 0) \in C$. Then, there exists exactly one $d \in D$ such that $c +^{\mathbb{A}} d \in f(Z)$, namely $d = (c, b(c), \dots, 0)$. Similarly, let $(0, d, \dots, 0) \in D$. Then, there exists exactly one $c \in C$ such that $c +^{\mathbb{A}} d \in f(Z)$, namely $c = (b^{-1}(d), d, \dots, 0)$.

Conversely, let $B \subseteq A$ be such that

$$\mathbb{A} \models (\forall x \in C)(\exists! y \in D)(x + y \in B) \wedge (\forall y \in D)(\exists! x \in C)(x + y \in B). \tag{24}$$

Then for all $c \in C$, there exists exactly one $d \in D$, such that $c +^{\mathbb{A}} d \in B$. Let $b \colon C \to D$ be the function defined by the above condition, that is $b(c) = d$ if and only if $c +^{\mathbb{A}} d \in B$. We show that b is a bijection.

For injectivity, let $c, c' \in C$ be such that $b(c) = b(c') = d \in D$. Then, $c +^{\mathbb{A}} d \in B$ and $c' +^{\mathbb{A}} d \in B$. By (24), there exists exactly one $c'' \in C$ such that $c'' +^{\mathbb{A}} d \in B$. Hence $c = c'$.

For surjectivity, let $d \in D$. By (24), there exists $c \in C$ such that $c +^{\mathbb{A}} d \in B$. Let $b(c) = d'$. Then, by definition of b, it holds that $c +^{\mathbb{A}} d' \in B$. Hence, $c +^{\mathbb{A}} d \in B$ and $c +^{\mathbb{A}} d' \in B$. By (24), there exists exactly one $d'' \in D$ such that $c +^{\mathbb{A}} d'' \in B$. Hence $d = d'$. Then $b(c) = d$, and b is surjective. ◀

Let $\mathbb{C}$ and $\mathbb{D}$ be prime terms of $\mathbb{A}$. We conclude defining formulas that establish whether the prime power order of $\mathbb{C}$ and $\mathbb{D}$ have the same base or the same exponent. First, we deal with the base:

$$\mathrm{Base}(X, Y) \leftrightharpoons \begin{pmatrix} \mathrm{Sg}(Y, X) \wedge Y \neq \{0\} \\ (\forall Z)((\mathrm{Sg}(Z, Y) \wedge Z \neq \{0\}) \to Z = Y) \end{pmatrix}$$

$$\mathrm{EqBase}(X, Y) \leftrightharpoons (\exists X', Y') \begin{pmatrix} \mathrm{Base}(X, X') \\ \mathrm{Base}(Y, Y') \\ \mathrm{Eq}(X', Y') \end{pmatrix}$$

▶ **Claim 4.** *Let $C \subseteq A$ be the universe of a prime term $\mathbb{C}$ of $\mathbb{A}$, say isomorphic to $\mathbb{Z}(p, e)$, and let $B \subseteq C$. Then $\mathbb{A} \models \mathrm{Base}(C, B)$ if and only if B is (the universe of) the subgroup of $\mathbb{C}$ is isomorphic of $\mathbb{Z}(p)$.*

Claim 3 and Claim 4 imply the following.

▶ **Claim 5.** *Let $C, D \subseteq A$ such that $\mathbb{C}$ and $\mathbb{D}$ are distinct prime terms of $\mathbb{A}$, say isomorphic to $\mathbb{Z}(p, e)$ and $\mathbb{Z}(q, d)$ respectively. Then $\mathbb{A} \models \mathrm{EqBase}(C, D)$ if and only if $p = q$.*

[3] Along similar lines, the statement can be proved more generally for cyclic subgroups of $\mathbb{A}$ whose intersection is trivial (contains only the identity).

Finally, we deal with exponents:

$$\mathrm{Exp}(X,Y) \leftrightharpoons (\forall Z)(\mathrm{Sg}(Z,X) \to (\exists! y \in Y)\mathrm{Cycl}(Z,y))$$
$$\mathrm{EqExp}(X,Y) \leftrightharpoons (\exists X', Y') \begin{pmatrix} \mathrm{Exp}(X,X') \\ \mathrm{Exp}(Y,Y') \\ \mathrm{Eq}(X',Y') \end{pmatrix}$$

Recall that every subgroup of a cyclic subgroup is cyclic. The following is clear.

▶ **Claim 6.** *Let $C \subseteq A$ such that $\mathbb{C}$ is a prime term of $\mathbb{A}$, say isomorphic to $\mathbb{Z}(p,e)$, and let $E \subseteq C$. Then $\mathbb{A} \models \mathrm{Exp}(C,E)$ if and only if E contains exactly one generator for each (necessarily, cyclic) subgroup of $\mathbb{C}$.*

Claim 3 and Claim 6 imply the following.

▶ **Claim 7.** *Let $C, D \subseteq A$ such that $\mathbb{C}$ and $\mathbb{D}$ are distinct prime terms of $\mathbb{A}$, say isomorphic to $\mathbb{Z}(p,e)$ and $\mathbb{Z}(q,d)$ respectively. Then $\mathbb{A} \models \mathrm{EqExp}(C,D)$ if and only if $e = d$.*

We now describe the reduction. A graph $\mathbf{G} = (G, E^{\mathbf{G}})$ is a relational structure on a binary relation symbol E, where $E^{\mathbf{G}} \subseteq G^2$ is symmetric and irreflexive; we liberally view $E^{\mathbf{G}}$ as a subset of 2-element subsets of G. The clique problem, Clique, is to decide, given a graph $\mathbf{G}$ and an integer $k \geq 0$, whether $\mathbf{G}$ contains a clique on k vertices. We regard Clique as a parameterized problem, where instance $(\mathbf{G}, k)$ is parameterized by k.

We give a fixed-parameter tractable reduction from Clique to $\mathrm{MC}(\mathcal{AG}_{\mathrm{spfin}}, \mathcal{MSO})$. Let $(\mathbf{G}, k)$ be an instance of Clique. Let $\mathbf{G} = (G, E^{\mathbf{G}})$, where $G = \{v_1, \ldots, v_n\}$ and $E^{\mathbf{G}} = \{e_1, \ldots, e_m\}$. For $v_i \in G$, let $\{f_{i,1}, \ldots, f_{i,d_i}\} = \{e \in E^{\mathbf{G}} \colon v_i \in e\}$, and let $\mathrm{degree}(v_i) = d_i$ denote the degree of v_i in $\mathbf{G}$. For each $i \in [n]$ and $j \in [d_i]$, let $m(i,j) \in [m]$ be such that $f_{i,j} = e_{m(i,j)}$.

We construct an instance (A, ϕ) of $\mathrm{MC}(\mathcal{AG}_{\mathrm{spfin}}, \mathcal{MSO})$, as follows. The succinct presentation A is a (square) diagonal integer matrix of order $\sum_{i \in [n]} d_i$ defined as follows. Let $p_1, \ldots, p_n$ be the first n prime numbers.

$$A = \mathrm{diag}(p_1^{m(1,1)}, \ldots, p_1^{m(1,d_1)}, p_2^{m(2,1)}, \ldots, p_2^{m(2,d_2)}, \ldots, p_n^{m(n,1)}, \ldots, p_n^{m(n,d_n)})$$

It is readily verified that the abelian group presented by A is (finite and) isomorphic to

$$\mathbb{A} = \mathbb{Z}(p_1, m(1,1)) \oplus \cdots \oplus \mathbb{Z}(p_1, m(1,d_1)) \oplus \cdots \oplus \mathbb{Z}(p_n, m(n,1)) \oplus \ldots \oplus \mathbb{Z}(p_n, m(n,d_n))$$

▶ **Example 5.** Let $\mathbf{G} = (G, E^{\mathbf{G}})$ where $G = \{v_1, v_2, v_3, v_4\}$, $E^{\mathbf{G}} = \{e_1, e_2, e_3, e_4, e_5\}$, $e_1 = \{v_1, v_2\}$, $e_2 = \{v_1, v_4\}$, $e_3 = \{v_2, v_3\}$, $e_4 = \{v_2, v_4\}$, and $e_5 = \{v_3, v_4\}$. Then $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, and $p_4 = 7$. Let $m(1,1) = 1$, $m(1,2) = 2$, $m(2,1) = 1$, $m(2,2) = 3$, $m(2,3) = 4$, $m(3,1) = 3$, $m(3,2) = 5$, $m(4,1) = 2$, $m(4,2) = 4$, and $m(4,3) = 5$. We therefore have that the succinct presentation presents the finite abelian group

$$\mathbb{Z}(2,1) \oplus \mathbb{Z}(2,2) \oplus \mathbb{Z}(3,1) \oplus \mathbb{Z}(3,3) \oplus \mathbb{Z}(3,4) \oplus \mathbb{Z}(5,3) \oplus \mathbb{Z}(5,5) \oplus \mathbb{Z}(7,2) \oplus \mathbb{Z}(7,4) \oplus \mathbb{Z}(7,5)$$

We now define a monadic second order formula ϕ, as follows. Let $K = [k] \times [k-1]$. Let $c_k \colon K \to K$ be the permutation of K uniquely determined by the following conditions:

- $c_k(i,j) = (i',j')$ if and only if $c_k(i',j') = (i,j)$, that is, c_k decomposes into $k(k-1)/2$ disjoint cycles of length 2;
- $c_k(i,j) = (j+1,i)$ for all $(i,j) \in K$ such that $1 \leq i \leq j < k$.

Note that the number of pairs $(i,j) \in K$ such that $1 \leq i \leq j < k$ is equal to $\binom{k}{2}$, the number of edges in a clique on k vertices. The following example illustrates how c_k relates to a clique on k vertices.

▶ **Example 6.** We have $c_4(1,1) = (2,1)$, $c_4(1,2) = (3,1)$, $c_4(1,3) = (4,1)$, $c_4(2,1) = (1,1)$, $c_4(2,2) = (3,2)$, $c_4(2,3) = (4,2)$, $c_4(3,1) = (1,2)$, $c_4(3,2) = (2,2)$, $c_4(3,3) = (4,3)$. c_4 decomposes into 6 disjoint cycles,

$$c_4 = ((1,1)(2,1))((1,2)(3,1))((1,3)(4,1))((2,2)(3,2))((2,3)(4,2))((3,3)(4,3)),$$

and the edges of a 4-clique on vertices $\{1,2,3,4\}$ are obtained by projecting the pairs in each cycle onto their first coordinate: $\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}$.

We now define ϕ as follows:

$$\phi = (\underset{(i,j)\in K}{\exists} X_{i,j}) \left(\begin{array}{c} \bigwedge_{(i,j),(i',j')\in K,(i,j)\neq(i',j')} X_{i,j} \cap X_{i',j'} = \{0\} \\ \bigwedge_{(i,j)\in K} \mathrm{PrTerm}(X_{i,j}) \\ \bigwedge_{i\in[k]} \bigwedge_{j,j'\in[k-1]} \mathrm{EqBase}(X_{i,j}, X_{i,j'}) \\ \bigwedge_{(i,j)\in K} \mathrm{EqExp}(X_{i,j}, X_{c_k(i,j)}) \end{array} \right) \tag{25}$$

▶ **Claim 8.** $(\mathbf{G},k) \in$ Clique *if and only if* $(A,\phi) \in \mathrm{MC}(\mathcal{AG}_{\mathrm{spfin}}, \mathcal{MSO})$.

Proof. Recall that the number of pairs $(i,j) \in K = [k] \times [k-1]$ such that $1 \leq i \leq j < k$ is equal to $\binom{k}{2}$, the number of edges in a clique on k vertices.

($\Rightarrow$) Let the vertices $\{u_i : i \in [k]\} \subseteq G$ and the edges $\{f_{(i,j)} : (i,j) \in K\} \subseteq E^{\mathbf{G}}$ form a clique on k vertices in $\mathbf{G}$ such that the following holds:

- $f_{(i,j)} \cap f_{(i,j')} = u_i$,
- $f_{(i,j)} = \{u_i, c_k(i,j)_1\}$,

where $c_k(i,j)_1$ denotes the projection of $c_k(i,j)$ onto the first coordinate. For $i \in [k]$, let $n(i) \in [n]$ be such that $u_i = v_{n(i)}$, and for $(i,j) \in K$, let $m(i,j) \in [m]$ be such that $f_{(i,j)} = e_{m(i,j)}$.

For $(i,j) \in K$, let $C_{i,j}$ be (the universe of) the subgroup $\mathbb{C}_{i,j}$ of $\mathbb{A}$ satisfying the following:

- $\mathbb{C}_{i,j}$ is isomorphic to $\mathbb{Z}(p_{n(i)}, m(i,j))$;
- $\mathbb{C}_{i,j}$ has no prime power order cyclic proper supergroup in $\mathbb{A}$.

By construction such subgroup $\mathbb{C}_{i,j}$ of $\mathbb{A}$ exists and is unique, hence the above definition is sound. It is easy to verify that the family of $C_{i,j}$'s witnesses the truth of (25) in $\mathbb{A}$.

($\Leftarrow$) Let $C_{i,j} \subseteq A$ for $(i,j) \in K$ witnesses that ϕ holds in $\mathbb{A}$. Therefore, the $\mathbb{C}_{i,j}$ form a family of $\binom{k}{2}$ prime terms of $\mathbb{A}$ (by the first two lines in (25)). By the third line in (25), the $\mathbb{C}_{i,j}$'s partition into k blocks $V_1,\ldots,V_k$ such that the orders of groups in block V_l are all powers of the same prime $p_{i_l} \in \{p_1,\ldots,p_n\}$.

Let $v_{i_1},\ldots,v_{i_k} \subseteq G$ be the vertices of $\mathbf{G}$ corresponding to the primes $p_{i_1},\ldots,p_{i_k}$. We claim that there are $\binom{k}{2}$ edges between the vertices $v_{i_1},\ldots,v_{i_k}$; since $\mathbf{G}$ does not contain loops nor multiedges, it follows that the vertices $v_{i_1},\ldots,v_{i_k}$ form a clique of size k in $\mathbf{G}$.

We first observe that for all $l,l' \in [k]$, $l \neq l'$, there is an edge between v_{i_l} and $v_{i_{l'}}$. By the fourth line in (25) and the definition of c_k, we have that $V_1 \cup \cdots \cup V_k$ partitions into $\binom{k}{2}$ 2-element sets $\{\mathbb{C},\mathbb{C}'\}$ such that $\mathbb{C}$ in V_l and $\mathbb{C}'$ in $V_{l'}$ $(l,l' \in [k], l \neq l')$ such that the orders of $\mathbb{C}$ and $\mathbb{C}'$ have the same exponent. By construction, such exponent is the index of an edge between the vertices corresponding to the (prime) base of the orders of $\mathbb{C}$ and $\mathbb{C}'$. Since, by construction, the index of each edge is the exponent of exactly two prime terms of $\mathbb{A}$, distinct 2-element sets of the form above contribute distinct edges, thus contributing $\binom{k}{2}$ edges in total between vertices $v_{i_1},\ldots,v_{i_k}$. ◀

The construction of ϕ only depends on k. The complexity of constructing A is determined by:

- the time to generate the first $|G| = n$ prime numbers $p_1, \ldots, p_n$ which is roughly in $O(n^3)$ as the nth prime is bounded above by n^2 and the sieve of Eratosthenes finds all primes not larger than l in time $O(l \log \log l)$;
- the size of A, a square integer matrix of order at most $n(n-1)$ whose integer entries are bounded above by $p_n^m \leq n^{2n^2}$, thus at most n^4 entries each of size in $O(n^2 \log n)$.

Therefore (A, ϕ) is computable from $(\mathbf{G}, k)$ in time $f(k)\text{poly}(n)$ for some computable function f over the natural numbers. We thus conclude the following.

▶ **Theorem 7.** $\text{MC}(\mathcal{AG}_{\text{spfin}}, \mathcal{MSO})$ *is not fixed-parameter tractable (unless* $\text{W}[1] \subseteq \text{FPT}$*).*

6 Discussion

We proved that first-order logic is fixed-parameter tractable on finite abelian groups, and monadic second-order logic is W[1]-hard on succinctly presented finite abelian groups. What is the complexity of model checking monadic second-order logic on finitely presented abelian groups (without the succinctness condition)? On finite abelian groups?

Our work suggests some questions on general groups, reasonable in that they do not settle the isomorphism problem. For example, model checking the conjunctive positive fragment (first-order sentences on the group vocabulary built using $\forall$, $\exists$, $\wedge$, and $=$) on finite abelian groups is polynomial-time tractable; this fact can be derived from the literature or established directly by our elimination technique. How hard is model checking conjunctive positive queries on finite groups? Yet, the outstanding open question concerns the parameterized complexity of first-order (and monadic second-order) properties of finite groups.

Acknowledgments. The authors thank Carlo Toffalori for a clarification on Baur-Monk theorem.

References

1 S. Bova, R. Ganian, and S. Szeider. Model Checking Existential Logic on Partially Ordered Sets. In *CSL-LICS*, 2014.

2 S. Bova, R. Ganian, and S. Szeider. Quantified Conjunctive Queries on Partially Ordered Sets. In *IPEC*, 2014.

3 A. A. Bulatov and V. Dalmau. A Simple Algorithm for Mal'tsev Constraints. *SIAM J. Comput.*, 36(1):16–27, 2006.

4 B. Courcelle. The Monadic Second-Order Logic of Graphs. I. Recognizable Sets of Finite Graphs. *Inform. Comput.*, 85(1):12–75, 1990.

5 B. Courcelle, J. A. Makowsky, and U. Rotics. Linear Time Solvable Optimization Problems on Graphs of Bounded Clique-Width. *Theory Comput. Syst.*, 33(2):125–150, 2000.

6 T. Feder and M. Vardi. The Computational Structure of Monotone Monadic SNP and Constraint Satisfaction: A Study through Datalog and Group Theory. *SIAM J. Comput.*, 28:57–104, 1999.

7 J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer, 2010.

8 J. Gajarský, P. Hliněný, J. Obdržálek, and S. Ordyniak. Faster Existential FO Model Checking on Posets. In *ISAAC*, 2014.

9 M. Grohe. Logic, Graphs, and Algorithms. Technical Report in *ECCC*, TR07-091, 2007.

10 M. Grohe, S. Kreutzer, and S. Siebertz. Deciding First-Order Properties of Nowhere Dense Graphs. In *STOC*, 2014.

11 W. Hodges. *Model Theory*. Cambridge University Press, 1993.
12 S. Jukna. *Extremal Combinatorics*. Springer, 2001.
13 T. Kavitha. Linear Time Algorithms for Abelian Group Isomorphism and Related Problems. *J. Comput. Syst. Sci.*, 73:986–996, 2007.
14 L. Libkin. *Elements of Finite Model Theory*. Springer, 2010.
15 C.C. MacDuffee. *The Theory of Matrices*. Dover, 2004.
16 G. L. Miller. On the $n^{\log n}$ Isomorphism Technique: A Preliminary Report. In *STOC*, 1978.
17 J. Rotman. *An Introduction to the Theory of Groups*. Springer, 1999.
18 D. Seese. Linear Time Computable Problems and First-Order Descriptions. *Mathematical Structures in Computer Science*, 6(6):505–526, 1996.
19 A. Storjohann. Algorithms for Matrix Canonical Forms. Ph. D. Thesis, ETH Zürich, 2000.
20 W. Szmielew. Elementary Properties of Abelian Groups. *Fundamenta Math.*, 41:203–271, 1955.