Expressive power of linear algebra query languages

Expressive power of linear algebra query languages

Floris Geerts
University of Antwerp
floris.geerts@uantwerp.be

Thomas Muñoz
PUC Chile and IMFD Chile
tfmunoz@uc.cl

Cristian Riveros
PUC Chile and IMFD Chile
cristian.riveros@uc.cl

Domagoj Vrgoč
PUC Chile and IMFD Chile
dvrgoc@ing.puc.cl

ABSTRACT

Linear algebra algorithms often require some sort of iteration or recursion as is illustrated by standard algorithms for Gaussian elimination, matrix inversion, and transitive closure. A key characteristic shared by these algorithms is that they allow looping for a number of steps that is bounded by the matrix dimension. In this paper we extend the matrix query language MATLANG with this type of recursion, and show that this suffices to express classical linear algebra algorithms. We study the expressive power of this language and show that it naturally corresponds to arithmetic circuit families, which are often said to capture linear algebra. Furthermore, we analyze several sub-fragments of our language, and show that their expressive power is closely tied to logical formalisms on semiring-annotated relations.

1 INTRODUCTION

Linear algebra-based algorithms have become a key component in data analytic workflows. As such, there is a growing interest in the database community to integrate linear algebra functionalities into relational database management systems [5, 23, 25–27]. In particular, from a query language perspective, several proposals have recently been put forward to unify relational algebra and linear algebra. Two notable examples of this are: LARA [22], a minimalistic language in which a number of atomic operations on associative tables are proposed, and MATLANG, a query language for matrices [7].

Both LARA and MATLANG have been studied by the database theory community, showing interesting connections to relational algebra and logic. For example, fragments of LARA are known to capture first-order logic with aggregation [4], and MATLANG has been recently shown to be equivalent to a restricted version of the (positive) relational algebra on K-relations, RA^+_K [8], where K denotes a semiring. On the other hand, some standard constructions in linear algebra are out of reach for these languages. For instance, it was shown that under standard complexity-theoretic assumptions, LARA can not compute the inverse of a matrix or its determinant [4], and operations such as the transitive closure of a matrix are known to be inexpressible in MATLANG [7]. Given that these are fundamental constructs in linear algebra, one might wonder how to extend LARA or MATLANG in order to allow expressing such properties.

One approach would be to add these constructions explicitly to the language. Indeed, this was done for MATLANG in [7], and LARA in [4]. In these works, the authors have extended the core language with the trace, the inverse, the determinant, or the eigenvectors operators and study the expressive power of the result. However, one can argue that there is nothing special in these operators, apart they have been used historically in linear algebra textbooks and they extend the expressibility of the core language. The question here is whether these new operators form a sound and natural choice to extend the core language, or are they just some particular queries that we would like to support.

In this paper we take a more principled approach by studying what are the atomic operations needed to define standard linear algebra algorithms. Inspecting any linear algebra textbook, one sees that most linear algebra procedures heavily rely on the use of for-loops in which iterations happen over the dimensions of the matrices involved. To illustrate this, let us consider the example of computing the transitive closure of a graph. This can be done using a modification of the Floyd-Warshall algorithm [10], which takes as its input an $n \times n$ adjacency matrix A representing our graph, and operates according to the following pseudo-code:

```
for k = 1..n do
    for i = 1..n do
        for j = 1..n do
            A[i, j] := A[i, j] + A[i, k] · A[k, j]
```

After executing the algorithm, all of the non zero entries signify an edge in the (irreflexive) transitive closure graph.

By examining standard linear algebra algorithms such as Gaussian elimination, LU-decomposition, computing the inverse of a matrix, or its determinant, we can readily see that this pattern continues. Namely, we observe that there are two main components to such algorithms: (i) the ability to iterate up to the matrix dimension; and (ii) the ability to access a particular position in our matrix. In order to allow this behavior in a query language, we propose to extend MATLANG with limited recursion in the form of for-loops, resulting in the language for-MATLANG. To simulate the two components of standard linear algebra algorithms in a natural way, we simulate a loop of the form `for` $i = 1..n$ `do` by leveraging canonical vectors. In other words, we use the canonical vectors $b_1 = (1, 0, \ldots)$, $b_2 = (0, 1, \ldots), \ldots$, to access specific rows and columns, and iterate over these vectors. In this way, we obtain a language able to compute important linear algebra operators such as LU-decomposition, determinant, matrix inverse, among other things.

Of course, a natural question to ask now is whether this really results in a language suitable for linear algebra? We argue that the correct way to approach this question is to compare our language to arithmetic circuits, which have been shown to capture the vast majority of existing matrix algorithms, from basic ones such as computing the determinant and the inverse, to complex procedures such as discrete Fourier transformation, and Strassen's algorithm

(see [1, 30] for an overview of the area), and can therefore be considered to effectively capture linear algebra. In the main technical result of this paper, we show that for-MATLANG indeed computes the same class of functions over matrices as the ones computed by arithmetic circuit families of bounded degree. As a consequence, for-MATLANG inherits all expressiveness properties of circuits, and thus can simulate any linear algebra algorithm definable by circuits.

Having established that for-MATLANG indeed provides a good basis for a linear algebra language, we move to a more fine-grained analysis of the expressiveness of its different fragments. For this, we aim to provide a connection with logical formalisms, similarly as was done by linking LARA and MATLANG to first-order logic with aggregates [4, 7]. As we show, capturing different logics correspond to restricting how matrix variables are updated in each iteration of the for-loops allowed in for-MATLANG. For instance, if we only allow to add some temporary result to a variable in each iteration (instead of rewriting it completely like in any programming language), we obtain a language, called sum-MATLANG, which is equivalent to RA_K^+, directly extending an analogous result shown for MATLANG, mentioned earlier [8]. We then study updating matrix variables based on another standard linear algebra operator, the Hadamard product, resulting in a fragment called FO-MATLANG, which we show to be equivalent to weighted logics [13]. Finally, in prod-MATLANG we update the variables based on the standard matrix product, and link this fragment to the ones discussed previously.

Contribution and outline.

- After we recall MATLANG in Section 2, we show in Section 3 how for-loops can be added to MATLANG in a natural way. We also observe that for-MATLANG strictly extends MATLANG. In addition, we discuss some design decisions behind the definition of for-MATLANG, noting that our use of canonical vectors results in the availability of an order relation.
- In Section 4 we show that for-MATLANG can compute important linear algebra algorithms in a natural way. We provide expressions in for-MATLANG for LU decomposition (used to solve linear systems of equations), the determinant and matrix inversion.
- More generally, in Section 5 we report our main technical contribution. We show that every uniform arithmetic circuits of polynomial degree correspond to a for-MATLANG expression, and vice versa, when a for-MATLANG expression has polynomial degree, then there is an equivalent uniform family of arithmetic circuits. As a consequence, for-MATLANG inherits all expressiveness properties of such circuits.
- Finally, in Section 6 we generalize the semantics of for-MATLANG to matrices with values in a semiring K, and show that two natural fragment of for-MATLANG, sum-MATLANG, and FO-MATLANG, are equivalent to the (positive) relational algebra and weighted logics on binary K-relations, respectively. We also briefly comment on a minimal fragment of for-MATLANG, based on prod-MATLANG, that is able to compute matrix inversion.

Due to space limitations, most proofs are referred to the appendix.

Related work. We already mentioned LARA [22] and MATLANG [7] whose expressive power was further analyzed in [4, 8, 15, 16]. Extensions of SQL for matrix manipulations are reported in [27]. Most relevant is [23] in which a recursion mechanism is added to SQL which resembles for-loops. The expressive power of this extension is unknown, however. Classical logics with aggregation [20] and fixed-point logics with counting [19] can also be used for linear algebra. More generally, for the descriptive complexity of linear algebra we refer to [12, 21]. Most of these works require to encode real numbers inside relations, whereas we treat real numbers as atomic values. We refer to relevant papers related to arithmetic circuits and logical formalisms on semiring-annotated relations in the corresponding sections later in the paper.

2 MATLANG

We start by recalling the matrix query language MATLANG, introduced in [7], which serves as our starting point.

Syntax. Let $\mathcal{V} = \{V_1, V_2, \ldots\}$ be a countably infinite set of *matrix variables* and $\mathcal{F} = \bigcup_{k>1} \mathcal{F}_k$ with $\mathcal{F}_k$ a set of *functions* of the form $f : \mathbb{R}^k \to \mathbb{R}$, where $\mathbb{R}$ denotes the set of real numbers. The syntax of MATLANG expressions is defined by the following grammar[1]:

e	::=	$V \in \mathcal{V}$	(matrix variable)
	|	e^T	(transpose)
	|	$1(e)$	(one-vector)
	|	$\mathtt{diag}(e)$	(diagonalization of a vector)
	|	$e_1 \cdot e_2$	(matrix multiplication)
	|	$e_1 + e_2$	(matrix addition)
	|	$e_1 \times e_2$	(scalar multiplication)
	|	$f(e_1, \ldots, e_k)$	(pointwise application of $f \in \mathcal{F}_k$).

MATLANG is parametrized by a collection of functions $\mathcal{F}$ but in the remainder of the paper we only make this dependence explicit, and write MATLANG[$\mathcal{F}$], for some set $\mathcal{F}$ of functions, when these functions are crucial for some results to hold. When we simply write MATLANG, we mean that any function can be used (including not using any function at all).

Schemas and typing. To define the semantics of MATLANG expressions we need a notion of schema and well-typedness of expressions. A MATLANG *schema* $\mathcal{S}$ is a pair $\mathcal{S} = (\mathcal{M}, \mathtt{size})$, where $\mathcal{M} \subset \mathcal{V}$ is a finite set of matrix variables, and $\mathtt{size} : \mathcal{M} \mapsto \mathtt{Symb} \times \mathtt{Symb}$ is a function that maps each matrix variable in $\mathcal{M}$ to a pair of *size symbols*. The size function helps us determine whether certain matrix operations, such as matrix multiplication, can be performed for matrices adhering to a schema. We denote size symbols by Greek letters α, β, γ. We also assume that $1 \in \mathtt{Symb}$. To help us determine whether a MATLANG expression can always be evaluated, we define the *type* of an expression e, with respect to a schema $\mathcal{S}$, denoted by $\mathtt{type}_{\mathcal{S}}(e)$, inductively as follows:

- $\mathtt{type}_{\mathcal{S}}(V) := \mathtt{size}(V)$, for a matrix variable $V \in \mathcal{M}$;
- $\mathtt{type}_{\mathcal{S}}(e^T) := (\beta, \alpha)$ if $\mathtt{type}_{\mathcal{S}}(e) = (\alpha, \beta)$;
- $\mathtt{type}_{\mathcal{S}}(1(e)) := (\alpha, 1)$ if $\mathtt{type}_{\mathcal{S}}(e) = (\alpha, \beta)$;
- $\mathtt{type}_{\mathcal{S}}(\mathtt{diag}(e)) := (\alpha, \alpha)$, if $\mathtt{type}_{\mathcal{S}}(e) = (\alpha, 1)$;
- $\mathtt{type}_{\mathcal{S}}(e_1 \cdot e_2) := (\alpha, \gamma)$ if $\mathtt{type}_{\mathcal{S}}(e_1) = (\alpha, \beta)$, and $\mathtt{type}_{\mathcal{S}}(e_2) = (\beta, \gamma)$;
- $\mathtt{type}_{\mathcal{S}}(e_1 + e_2) := (\alpha, \beta)$ if $\mathtt{type}_{\mathcal{S}}(e_1) = \mathtt{type}_{\mathcal{S}}(e_2) = (\alpha, \beta)$;

[1] The original syntax also permits the operator $\mathtt{let}\ V = e_1\ \mathtt{in}\ e_2$, which replaces every occurrence of V in e_2 with the value of e_1. Since this is just syntactic sugar, we omit this operator. We also explicitly include matrix addition and scalar multiplication, although these can be simulated by pointwise function applications. Finally, we use transposition instead of conjugate transposition since we work with matrices over $\mathbb{R}$.

- $\mathsf{type}_{\mathcal{S}}(e_1 \times e_2) := (\alpha, \beta)$ if $\mathsf{type}_{\mathcal{S}}(e_1) = (1,1)$ and $\mathsf{type}_{\mathcal{S}}(e_2) = (\alpha, \beta)$; and
- $\mathsf{type}_{\mathcal{S}}(f(e_1, \ldots, e_k)) := (\alpha, \beta)$, whenever $\mathsf{type}_{\mathcal{S}}(e_1) = \cdots = \mathsf{type}_{\mathcal{S}}(e_k) := (\alpha, \beta)$ and $f \in \mathcal{F}_k$.

We call an expression *well-typed* according to the schema $\mathcal{S}$, if it has a defined type. A well-typed expression can be evaluated regardless of the actual sizes of the matrices assigned to matrix variables, as we describe next.

Semantics. We use $\mathsf{Mat}[\mathbb{R}]$ to denote the set of all real matrices and for $A \in \mathsf{Mat}[\mathbb{R}]$, $\dim(A) \in \mathbb{N}^2$ denotes its dimensions. A (MATLANG) *instance* $\mathcal{I}$ over a schema $\mathcal{S}$ is a pair $\mathcal{I} = (\mathcal{D}, \mathsf{mat})$, where $\mathcal{D} : \mathsf{Symb} \mapsto \mathbb{N}$ assigns a value to each size symbol (and thus in turn dimensions to each matrix variable), and $\mathsf{mat} : \mathcal{M} \mapsto \mathsf{Mat}[\mathbb{R}]$ assigns a concrete matrix to each matrix variable $V \in \mathcal{M}$, such that $\dim(\mathsf{mat}(V)) = \mathcal{D}(\alpha) \times \mathcal{D}(\beta)$ if $\mathsf{size}(V) = (\alpha, \beta)$. That is, an instance tells us the dimensions of each matrix variable, and also the concrete matrices assigned to the variable names in $\mathcal{M}$. We assume that $\mathcal{D}(1) = 1$, for every instance $\mathcal{I}$. If e is a well-typed expression according to $\mathcal{S}$, then we denote by $[\![e]\!](\mathcal{I})$ the matrix obtained by evaluating e over $\mathcal{I}$, and define it as follows:

- $[\![V]\!](\mathcal{I}) := \mathsf{mat}(V)$, for $V \in \mathcal{M}$;
- $[\![e^T]\!](\mathcal{I}) := [\![e]\!](\mathcal{I})^T$, where A^T is the transpose of matrix A;
- $[\![1(e)]\!](\mathcal{I})$ is a $n \times 1$ vector with 1 as all of its entries, where $\dim([\![e]\!](\mathcal{I})) = (n, m)$;
- $[\![\mathsf{diag}(e)]\!](\mathcal{I})$ is a diagonal matrix with the vector $[\![e]\!](\mathcal{I})$ on its main diagonal, and zero in every other position;
- $[\![e_1 \cdot e_2]\!](\mathcal{I}) := [\![e_1]\!](\mathcal{I}) \cdot [\![e_2]\!](\mathcal{I})$;
- $[\![e_1 + e_2]\!](\mathcal{I}) := [\![e_1]\!](\mathcal{I}) + [\![e_2]\!](\mathcal{I})$;
- $[\![e_1 \times e_2]\!](\mathcal{I}) := a \times [\![e_2]\!](\mathcal{I})$ with $[\![e_1]\!](\mathcal{I}) = [a]$; and
- $[\![f(e_1, \ldots, e_k)]\!](\mathcal{I})$ is a matrix A of the same size as $[\![e_1]\!](\mathcal{I})$, and where A_{ij} has the value $f([\![e_1]\!](\mathcal{I})_{ij}, \ldots, [\![e_k]\!](\mathcal{I})_{ij})$.

Although MATLANG forms a solid basis for a matrix query language, it is limited in expressive power. Indeed, MATLANG is subsumed by first order logic with aggregates that uses only three variables [7]. Hence, no MATLANG expression exists that can compute the transitive closure of a graph (represented by its adjacency matrix) or can compute the inverse of a matrix. Rather than extending MATLANG with specific linear algebra operators, such as matrix inversion, we next introduce a limited form of recursion in MATLANG.

3 EXTENDING MATLANG WITH FOR LOOPS

To extend MATLANG with recursion, we take inspiration from classical linear algebra algorithms, such as those described in [28]. Many of these algorithms are based on *for-loops* in which the termination condition for each loop is determined by the matrix dimensions. We have seen how the transitive closure of a matrix can be computed using for-loops in the Introduction. Here we add this ability to MATLANG, and show that the resulting language, called for-MATLANG, can compute properties outside of the scope of MATLANG. We see more advanced examples, such as Gaussian elimination and matrix inversion, later in the paper.

3.1 Syntax and semantics of for-MATLANG

The syntax of for-MATLANG is defined just as for MATLANG but with an extra rule in the grammar:

$$\mathsf{for}\, v, X.\ e \quad \text{(canonical for loop, with } v, X \in \mathcal{V}\text{)}.$$

Intuitively, X is a matrix variable which is iteratively updated according to the expression e. We simulate iterations of the form "for $i \in [1..n]$" by letting v loop over the *canonical vectors* $b_1^n, \ldots, b_n^n$ of dimension n. Here, $b_1^n = [1\, 0 \cdots 0]^T$, $b_2^n = [0\, 1\, 0 \cdots 0]^T$, etc. When n is clear from the context we simply write $b_1, b_2, \ldots$. In addition, the expression e in the rule above may depend on v.

We next make the semantics precise and start by declaring the type of loop expressions. Given a schema $\mathcal{S}$, the type of a for-MATLANG expression e, denoted $\mathsf{type}_{\mathcal{S}}(e)$, is defined inductively as in MATLANG but with following extra rule:

- $\mathsf{type}_{\mathcal{S}}(\mathsf{for}\, v, X.\ e) := (\alpha, \beta)$, if $\mathsf{type}_{\mathcal{S}}(e) = \mathsf{type}_{\mathcal{S}}(X) = (\alpha, \beta)$ and $\mathsf{type}_{\mathcal{S}}(v) = (\gamma, 1)$.

We note that $\mathcal{S}$ now necessarily includes v and X as variables and assigns size symbols to them. We also remark that in the definition of the type of $\mathsf{for}\, v, X.\ e$, we require that $\mathsf{type}_{\mathcal{S}}(X) = \mathsf{type}_{\mathcal{S}}(e)$ as this expression updates the content of the variable X in each iteration using the result of e. We further restrict the type of v to be a vector, i.e., $\mathsf{type}_{\mathcal{S}}(v) = (\gamma, 1)$, since v will be instantiated with canonical vectors. A for-MATLANG expression e is well-typed over a schema $\mathcal{S}$ if its type is defined.

For well-typed expressions we next define their semantics. This is done in an inductive way, just as for MATLANG. To define the semantics of $\mathsf{for}\, v, X.\ e$ over an instance $\mathcal{I}$, we need the following notation. Let $\mathcal{I}$ be an instance and $V \in \mathcal{M}$. Then $\mathcal{I}[V := A]$ denotes an instance that coincides with $\mathcal{I}$, except that the value of the matrix variable V is given by the matrix A. Assume that $\mathsf{type}_{\mathcal{S}}(v) = (\gamma, 1)$, and $\mathsf{type}_{\mathcal{S}}(e) = (\alpha, \beta)$ and $n := \mathcal{D}(\gamma)$. Then, $[\![\mathsf{for}\, v, X.\ e]\!](\mathcal{I})$ is defined iteratively, as follows:

- Let $A_0 := 0$ be the zero matrix of size $\mathcal{D}(\alpha) \times \mathcal{D}(\beta)$.
- For $i = 1, \ldots n$, compute $A_i := [\![e]\!](\mathcal{I}[v := b_i^n, X := A_{i-1}])$.
- Finally, set $[\![\mathsf{for}\, v, X.\ e]\!](\mathcal{I}) := A_n$.

For better understanding how for-MATLANG works, we next provide some examples. We start by showing that the one-vector and diag operators are redundant in for-MATLANG.

Example 3.1. We first show how the one-vector operator $1(e)$ can be expressed using for loops. It suffices to consider the expression

$$e_1 := \mathsf{for}\, v, X.\ X + v,$$

with $\mathsf{type}_{\mathcal{S}}(v) = (\alpha, 1) = \mathsf{type}_{\mathcal{S}}(X)$ if $\mathsf{type}_{\mathcal{S}}(e) = (\alpha, \beta)$. This expression is well-typed and is of type $(\alpha, 1)$. When evaluated over some instance $\mathcal{I}$ with $n = \mathcal{D}(\alpha)$, $[\![e_1]\!](\mathcal{I})$ is defined as follows. Initially, $A_0 := 0$. Then $A_i := A_{i-1} + b_i^n$, i.e., the ith canonical vector is added to A_{i-1}. Finally, $[\![e_1]\!](\mathcal{I}) := A_n$ and this now clearly coincides with $[\![1(e)]\!](\mathcal{I})$. □

Example 3.2. We next show that the diag operator is redundant in for-MATLANG. Indeed, it suffices to consider the expression

$$e_{\mathsf{diag}} := \mathsf{for}\, v, X.\ X + (v^T \cdot e) \times v \cdot v^T,$$

where e is a for-MATLANG expression of type $(\alpha, 1)$. For this expression to be well-typed, v has to be a vector variable of type

$\alpha \times 1$ and X a matrix variable of type (α, α). Then, $[\![e_{\mathsf{diag}}]\!](\mathcal{I})$ is defined as follows. Initially, A_0 is the zero matrix of dimension $n \times n$, where $n = \mathcal{D}(\alpha)$. Then, in each iteration $i \in [1..n]$, $A_i := A_{i-1} + ((b_i^n)^T \cdot [\![e]\!](\mathcal{I})) \times (b_i^n \cdot (b_i^n)^T)$. In other words, A_i is obtained by adding the matrix with value $([\![e]\!](\mathcal{I}))_i$ on position (i, i) to A_{i-1}. Hence, $[\![e_{\mathsf{diag}}]\!](\mathcal{I}) := A_n = [\![\mathtt{diag}(e)]\!](\mathcal{I})$. □

These examples illustrate that we can limit for-MATLANG to consist of the following "core" operators: transposition, matrix multiplication and addition, scalar multiplication, pointwise function application, and for-loops. More specific, for-MATLANG is defined by the following simplified syntax:

$$e ::= V \mid e^T \mid e_1 \cdot e_2 \mid e_1 + e_2 \mid e_1 \times e_2 \mid f(e_1, \ldots, e_k) \mid \mathtt{for}\, v, X.\, e$$

Similarly as for MATLANG, we write for-MATLANG$[\mathcal{F}]$ for some set $\mathcal{F}$ of functions when these are required for the task at hand.

As a final example, we show that we can compute whether a graph contains a 4-clique using for-MATLANG.

Example 3.3. To test for 4-cliques it suffices to consider the following expression with for-loops nested four times:

$$\begin{aligned}
&\mathtt{for}\, u, X_1.\ X_1 + \\
&\quad \mathtt{for}\, v, X_2.\ X_2 + \\
&\qquad \mathtt{for}\, w, X_3.\ X_3 + \\
&\qquad\quad \mathtt{for}\, x, X_4.\ X_4 + \\
&\qquad\qquad u^T \cdot V \cdot v \cdot u^T \cdot V \cdot w \cdot u^T \cdot V \cdot x \cdot \\
&\qquad\qquad v^T \cdot V \cdot w \cdot v^T \cdot V \cdot x \cdot w^T \cdot V \cdot x \cdot g(u, v, w, x)
\end{aligned}$$

with $g(u, v, w, x) = f(u, v) \cdot f(u, w) \cdot f(u, x) \cdot f(v, w) \cdot f(v, x) \cdot f(w, x)$ and $f(u, v) = 1 - u^T \cdot v$. Note that $f(b_i^n, b_j^n) = 1$ if $i \neq j$ and $f(b_i^n, b_j^n) = 0$ otherwise. Hence, $g(b_i^n, b_j^n, b_k^n, b_\ell^n) = 1$ if and only if all i, j, k, l are pairwise different. When evaluating the expression on an instance $\mathcal{I}$ such that V is assigned to the adjacency matrix of a graph, the expression above evaluates to a non-zero value if and only if the graph contains a four-clique. □

Given that MATLANG can not check for 4-cliques [7], we easily obtain the following.

Proposition 3.4. *For any collection of functions $\mathcal{F}$,* MATLANG$[\mathcal{F}]$ *is properly subsumed by* for-MATLANG$[\mathcal{F}]$.

3.2 Design decisions behind for-MATLANG

Loop Initialization. As the reader may have observed, in the semantics of for-loops we always initialize A_0 to the zero matrix 0 (of appropriate dimensions). It is often convenient to start the iteration given some concrete matrix originating from the result of evaluation a for-MATLANG expression e_0. To make this explicit, we write $\mathtt{for}\, v, X = e_0.\, e$ and its semantics is defined as above with the difference that $A_0 := [\![e_0]\!](\mathcal{I})$. We observe, however, that $\mathtt{for}\, v, X = e_0.\, e$ can already be expressed in for-MATLANG. In other words, we do not loose generality by assuming an initialization of A_0 by 0. The key insight is that in for-MATLANG we can check during evaluation whether or not the current canonical vector b_i^n is equal to the b_1^n. This is due to the fact that for-loops iterate over the canonical vectors in a fixed order. We discuss this more in the next paragraph. In particular, we can define a for-MATLANG expression min(), which when evaluated on an instance, returns 1 if its input vector is b_1^n, and returns 0 otherwise. Given min(), consider now the for-MATLANG expression

$$\mathtt{for}\, v, X.\, \mathtt{min}(v) \cdot e(v, X/e_0) + (1 - \mathtt{min}(v)) \cdot e(v, X),$$

where we explicitly list v and X as matrix variables on which e potentially depends on, and where $e(v, X/e_0)$ denotes the expression obtained by replacing every occurrence of X in e with e_0. When evaluating this expression on an instance $\mathcal{I}$, A_0 is initial set to the zero matrix, in the first iteration (when $v = b_1^n$ and thus $\mathtt{min}(v) = 1$) we have $A_1 = [\![e]\!](\mathcal{I}[v := b_1^n, X := [\![e_0]\!](\mathcal{I})])$, and for consecutive iterations (when only the part related to $1 - \mathtt{min}(v)$ applies) A_i is updated as before. Clearly, the result of this evaluation is equal to $[\![\mathtt{for}\, v, X = e_0.\, e]\!](\mathcal{I})$.

As an illustration, we consider the Floyd-Warshall algorithm given in the Introduction.

Example 3.5. Consider the following expression:

$$\begin{aligned}
e_{FW} := &\mathtt{for}\, v_k, X_1 = A.\ X_1 + \\
&\quad \mathtt{for}\, v_i, X_2.\ X_2 + \\
&\qquad \mathtt{for}\, v_j, X_3.\ X_3 + \\
&\qquad\quad (v_i^T \cdot X_1 \cdot v_k \cdot v_k^T \cdot X_1 \cdot v_j) \times v_i \cdot v_j^T
\end{aligned}$$

The expression e_{FW} simulates the Floyd-Warshall algorithm by updating the matrix A, which is stored in the variable X_1. The inner sub-expression here constructs an $n \times n$ matrix that contains one in the position (i, j) if and only if one can reach the vertex j from i by going through k, and zero elsewhere. If an instance $\mathcal{I}$ assigns to A the adjacency matrix of a graph, then $[\![e_{FW}]\!](\mathcal{I})$ will be equal to the matrix produced by the algorithm given in the Introduction. □

Order. By introducing for-loops we not only extend MATLANG with bounded recursion, we also introduce order information. Indeed, the semantics of the for operator assumes that the canonical vectors $b_1, b_2, \ldots$ are accessed in this order. It implies, among other things, that for-MATLANG expressions are not permutation-invariant. We can, for example, return the bottom right-most entry in a matrix. Indeed, consider the expression $e_{\mathsf{max}} := \mathtt{for}\, v, X.\, v$ which, for it to be well-typed, requires both v and X to be of type $(\alpha, 1)$. Then, $[\![e_{\mathsf{max}}]\!](\mathcal{I}) = b_n^n$, for $n = \mathcal{D}(\alpha)$, simply because initially, $X = 0$, but X will be overwritten by $b_1^n, b_2^n, \ldots, b_n^n$, in this order. Hence, at the end of the evaluation b_n^n is returned. To extract the bottom right-most entry from a matrix, we now simply use $e_{\mathsf{max}}^T \cdot V \cdot e_{\mathsf{max}}$.

Although the order is implicit in for-MATLANG, we can explicitly use this order in for-MATLANG expressions. More precisely, the order on canonical vectors is made accessible by using the matrix:

$$S_{\leq} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

We observe that $S_{\leq}$ has the property that $b_i^T \cdot S_{\leq} \cdot b_j = 1$, for two canonical vectors b_i and b_j of the same dimension, if and only if $i \leq j$. Otherwise, $b_i^T \cdot S_{\leq} \cdot b_j = 0$. Interestingly, we can build the matrix $S_{\leq}$ with the following for-MATLANG expression:

$$e_{\leq} = \mathtt{for}\, v, X.\, X + ((X \cdot e_{\mathsf{max}}) + v) \cdot v^T + v \cdot e_{\mathsf{max}}^T,$$

where e_{max} is as defined above. The intuition behind this expression is that by using the last canonical vector b_n, as returned by e_{max},

we have access to the last column of X (via the product $X \cdot e_{\mathsf{max}}$). We use this column such that after the i-th iteration, this column contains the i-th column of $S_{\leq}$. This is done by incrementing X with $v \cdot e_{\mathsf{max}}^T$. To construct $S_{\leq}$, in the i-th iteration we further increment X with (i) the current last column in X (via $X \cdot e_{\mathsf{max}} \cdot v^T$) which holds the $(i-1)$-th column of $S_{\leq}$; and (ii) the current canonical vector (via $v \cdot v^T$). Hence, after iteration i, X contains the first i columns of $S_{\leq}$ and holds the ith column of $S_{\leq}$ in its last column. It is now readily verified that $X = S_{\leq}$ after the nth iteration.

It should be clear that if we can compute $S_{\leq}$ using $e_{\leq}$, then we can easily define the following predicates and vectors related with the order of canonical vectors:

- $\mathsf{succ}(u, v)$ such that $\mathsf{succ}(b_i^n, b_j^n) = 1$ if $i \leq j$ and 0 otherwise. Similarly, we can define $\mathsf{succ}^+(u, v)$ such that $\mathsf{succ}^+(b_i^n, b_j^n) = 1$ if $i < j$ and 0 otherwise;
- $\mathsf{min}(u)$ such that $\mathsf{min}(b_i^n) = 1$ if $i = 1$ and $\mathsf{min}(b_i^n) = 0$ otherwise;
- $\mathsf{max}(u)$ such that $\mathsf{max}(b_i^n) = 1$ if $i = n$ and $\mathsf{min}(b_i^n) = 0$ otherwise; and
- e_{min} and e_{max} such that $[\![e_{\mathsf{min}}]\!](\mathcal{I}) = b_1^n$ and $[\![e_{\mathsf{max}}]\!](\mathcal{I}) = b_n^n$, respectively.

The definitions of these expressions are detailed in the appendix.

Having order information available results in for-MATLANG to be quite expressive. We heavily rely on order information in the next sections to compute the inverse of matrices and more generally to simulate low complexity Turing machines and arithmetic circuits.

4 ALGORITHMS IN LINEAR ALGEBRA

One of our main motivations to introduce for-loops is to be able to express classical linear algebra algorithms in a natural way. We have seen that for-MATLANG is quite expressive as it can check for cliques, compute the transitive closure, and can even leverage a successor relation on canonical vectors. The big question is how expressive for-MATLANG actually is. We will answer this in the next section by connecting for-MATLANG with arithmetic circuits of polynomial degree. Through this connection, one can move back and forth between for-MATLANG and arithmetic circuits, and as a consequence, anything computable by such a circuit can be computed by for-MATLANG as well. When it comes to specific linear algebra algorithms, the detour via circuits can often be avoided. Indeed, in this section we illustrate that for-MATLANG is able to compute LU decompositions of matrices. These decompositions form the basis of many other algorithms, such as solving linear systems of equations. We further show that for-MATLANG is expressive enough to compute matrix inversion and the determinant. We recall that matrix inversion and determinant need to be explicitly added as separate operators in MATLANG [7] and that the LARA language is unable to invert matrices under usual complexity-theoretic assumptions [4].

4.1 LU decomposition

A lower-upper (LU) decomposition factors a matrix A as the product of a lower triangular matrix L and upper triangular matrix U. This decomposition, and more generally LU decomposition with row pivoting (PLU), underlies many linear algebra algorithms and we next show that for-MATLANG can compute these decompositions.

LU decomposition by Gaussian elimination. LU decomposition can be seen as a matrix form of Gaussian elimination in which the columns of A are reduced, one by one, to obtain the matrix U. The reduction of columns of A is achieved as follows. Consider the first column $[A_{11}, \ldots, A_{n1}]^T$ of A and define $c_1 := [0, \alpha_{21}, \ldots, \alpha_{n1}]^T$ with $\alpha_{j1} := -\frac{A_{j1}}{A_{11}}$. Let $T_1 := I + c_1 \cdot b_1^T$ and consider $T_1 \cdot A$. That is, the jth row of $T_1 \cdot A$ is obtained by multiplying the first row of A by α_{j1} and adding it to the jth row of A. As a result, the first column of $T_1 \cdot A$ is equal to $[A_{11}, 0, \ldots, 0]^T$, i.e., all of its entries below the diagonal are zero. One then iteratively performs a similar computation, using a matrix $T_i := I + c_i \cdot b_i^T$, where c_i now depends on the ith column in $T_{i-1} \cdots T_1 \cdot A$. As a consequence, $T_i \cdot T_{i-1} \cdots T_1 \cdot A$ is upper triangular in its first i columns. At the end of this process, $T_n \cdots T_1 \cdot A = U$ where U is the desired upper triangular matrix. Furthermore, it is easily verified that each T_i is invertible and by defining $L := T_1^{-1} \cdot \cdots \cdot T_n^{-1}$ one obtains a lower triangular matrix satisfying $A = L \cdot U$. The above procedure is only successful when the denominators used in the definition of the vectors c_i are non-zero. When this is the case we call a matrix A *LU-factorizable*.

In case when such a denominator is zero in one of the reduction steps, one can remedy this situation by *row pivoting*. That is, when the ith entry of the ith row in $T_{i-1} \cdots T_1 \cdot A$ is zero, one replaces the ith row by jth row in this matrix, with $j > i$, provided that $i$the entry of the jth row is non-zero. If no such row exists, this implies that all elements below the diagonal are zero already in column i and one can proceed with the next column. One can formulate this in matrix terms by stating that there exists a permutation matrix P, which pivots rows, such that $P \cdot A = L \cdot U$. Any matrix A is LU-factorizable *with pivoting*.

Implementing LU decomposition in for-MATLANG. We first assume that the input matrices are LU-factorizable. We deal with general matrices later on. To implement the above procedure, we need to compute the vector c_i for each column i. We do this in two steps. First, we extract from our input matrix its ith column and set all its upper diagonal entries to zero by means of the expression:

$$\mathtt{col}(V, y) := \mathsf{for}\, v, X.\ \mathsf{succ}^+(y, v) \cdot (v^T \cdot V \cdot y) \cdot v + X.$$

Indeed, when V is assigned to a matrix A and y to b_i, we have that X will be initially assigned $A_0 = 0$ and in consecutive iterations, $A_j = A_{j-1} + b_j^T \cdot A \cdot b_i$ if $j > i$ (because $\mathsf{succ}^+(b_i, b_j) = 1$ if $j > i$) and $A_j = A_{j-1}$ otherwise (because $\mathsf{succ}^+(b_i, b_j) = 0$ for $j \leq i$). The result of this evaluation is the desired column vector. Using $\mathtt{col}(V, y)$, we can now compute T_i by the following expression:

$$\mathtt{reduce}(V, y) := e_{\mathsf{Id}} + f_{/}(\mathtt{col}(V, y), -(y^T \cdot V \cdot y) \cdot 1(y)) \cdot y^T,$$

where $f_{/} : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x/y$ is the division function. When V is assigned to A and y to b_i, $f_{/}(\mathtt{col}(A, b_i), -(b_i^T \cdot A \cdot b_i) \cdot 1(b_i))$ is equal to the vector c_i used in the definition of T_i. To perform the reduction steps for all columns, we consider the expression:

$$e_U(V) := (\mathsf{for}\, y, X = e_{\mathsf{Id}}.\ \mathtt{reduce}(X \cdot V, y) \cdot X) \cdot V.$$

That is, when V is assigned A, X will be initially $A_0 = I$, and then $A_i = \mathtt{reduce}(A_{i-1} \cdot A, b_i) = T_i \cdot T_{i-1} \cdots T_1 \cdot A$, as desired. We show in the appendix that, because we can obtain the matrices T_i in for-MATLANG and that these are easily invertible, we can also

construct an expression $e_L(V)$ which evaluates to L when V is assigned to A. We may thus conclude the following.

Proposition 4.1. *There exists* for-MATLANG$[f_/]$ *expressions* $e_L(V)$ *and* $e_U(V)$ *such that* $[\![e_L]\!](\mathcal{I}) = L$ *and* $[\![e_U]\!](\mathcal{I}) = U$ *form an LU-decomposition of* A*, where* $\mathsf{mat}(V) = A$ *and* A *is LU-factorizable.* □

We remark that the proposition holds when division is added as a function in $\mathcal{F}$ in for-MATLANG. When row pivoting is needed, we can also obtain a permutation matrix P such that $P \cdot A = L \cdot U$ holds by means of an expression in for-MATLANG, provided that we additionally allow the function $f_{>0}$, where $f_{>0} : \mathbb{R} \to \mathbb{R}$ is such that $f_{>0}(x) := 1$ if $x > 0$ and $f_{>0}(x) := 0$ otherwise.

Proposition 4.2. *There exist expressions* $e_{L^{-1}P}(M)$ *and* $e_U(M)$ *in* for-MATLANG$[f_/, f_{>0}]$ *such that* $L^{-1} \cdot P = [\![e_{L^{-1}P}]\!](\mathcal{I})$ *and* $U = [\![e_U]\!](\mathcal{I})$*, satisfy* $L^{-1} \cdot P \cdot A = U$. □

Intuitively, by allowing $f_{>0}$ we introduce a limited form of if-then-else in for-MATLANG, which is needed to continue reducing columns only when the right pivot has been found.

4.2 Determinant and inverse

Other key linear algebra operations include the computation of the determinant and the inverse of a matrix (if the matrix is invertible). As a consequence of the expressibility in for-MATLANG$[f_/, f_{>0}]$ of LU-decompositions with pivoting, it can be shown that the determinant and inverse can be expressed as well.

However, the results in the next section (connecting for-MATLANG with arithmetic circuits) imply that the determinant and inverse of a matrix can already be defined in for-MATLANG$[f_/]$. So instead of using LU decomposition with pivoting for matrix inversion and computing the determinant, we provide an alternative solution.

More specifically, we rely on Csanky's algorithm for computing the inverse of a matrix [11]. This algorithm uses the characteristic polynomial $p_A(x) = \mathsf{det}(xI - A)$ of a matrix. When expanded as a polynomial $p_A(x) = \sum_{i=0}^{n} c_i x^i$ and it is known that $A^{-1} = \frac{-1}{c_n} \sum_{i=0}^{n-1} c_i A^{n-1-i}$ if $c_n \neq 0$. Furthermore, $c_0 = 1$, $c_n = (-1)^n \mathsf{det}(A)$ and the coefficients c_i of $p_A(x)$ are known to satisfy the system of equations $S \cdot c = s$ given by:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ S_1 & 2 & 0 & \cdots & 0 & 0 \\ S_2 & S_1 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ S_{n-1} & S_{n-2} & S_{n-3} & \cdots & S_1 & n \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_n \end{pmatrix},$$

with $S_i = \mathsf{tr}(A^i)$. We show, in the appendix, that we can construct all ingredients of this system of equations in for-MATLANG$[f_/]$. By observing that the matrix S is a lower triangular matrix with non-zero elements on its diagonal, we can write it in the form $D_S + (S - D_S) = D_S \cdot (I + D_S^{-1} \cdot (S - D_S))$ with D_S the diagonal matrix consisting of the diagonal entries of S. Hence $S^{-1} = (I + D_S^{-1} \cdot (S - D_S))^{-1} \cdot D_S^{-1}$. We remark D_S^{-1} can simply be obtained by inverting the (non-zero) elements on the diagonal by means of $f_/$ in for-MATLANG$[f_/]$. Furthermore, we observe that $(I + D_S^{-1}(S - D_S))^{-1} = \sum_{i=0}^{n} (D_S^{-1}(S - D_S))^i$ which is something we can compute in for-MATLANG$[f_/]$ as well. Hence, we can invert S and obtain the vector $(c_1, \ldots, c_n)^T$ as $S^{-1} \cdot s$. To compute A^{-1} it now suffices to compute $\frac{-1}{c_n} \sum_{i=0}^{n-1} c_i A^{n-1-i}$. To find the determinant, we compute $(-1)^n c_n$. All this can be done in for-MATLANG$[f_/]$. We may thus conclude:

Proposition 4.3. *There are* for-MATLANG$[f_/]$ *expressions* $e_{\mathsf{det}}(V)$ *and* $e_{\mathsf{inv}}(V)$ *such that* $[\![e_{\mathsf{det}}]\!](\mathcal{I}) = \mathsf{det}(A)$*, and* $[\![e_{\mathsf{inv}}]\!](\mathcal{I}) = A^{-1}$ *when* $\mathcal{I}$ *assigns* V *to* A *and* A *is invertible.* □

5 EXPRESSIVENESS OF FOR LOOPS

In this section we explore the expressive power of for-MATLANG. Given that arithmetic circuits [1] capture most standard linear algebra algorithms [29, 30], they seem as a natural candidate for comparison. Intuitively, an arithmetic circuit is similar to a boolean circuit [3], except that it has gates computing the sum and the product function, and processes elements of $\mathbb{R}$ instead of boolean values. To connect for-MATLANG to arithmetic circuits we need a notion of uniformity of such circuits. After all, a for-MATLANG expression can take matrices of arbitrary dimensions as input and we want to avoid having different circuits for each dimension. To handle inputs of different sizes, we thus consider a notion of uniform families of arithmetic circuits, defined via a Turing machine generating a description of the circuit for each input size n.

What we show in the remainder of this section is that any function f which operates on matrices, and is computed by a uniform family of arithmetic circuits of bounded degree, can also be computed by a for-MATLANG expression, and vice versa. In order to keep the notation light, we will focus on for-MATLANG schemas over "square matrices" where each variable has type (α, α), $(\alpha, 1)$, $(1, \alpha)$, or $(1, 1)$, although all of our results hold without these restrictions as well. In what follows, we will write for-MATLANG to denote for-MATLANG$[\emptyset]$, i.e. the fragment of our language with no additional pointwise functions. We begin by defining circuits and then show how circuit families can be simulated by for-MATLANG.

5.1 From arithmetic circuits to for-MATLANG

Let us first recall the definition of arithmetic circuits. An *arithmetic circuit* Φ over a set $X = \{x_1, \ldots, x_n\}$ of input variables is a directed acyclic labeled graph. The vertices of Φ are called *gates* and denoted by $g_1, \ldots, g_m$; the edges in Φ are called *wires*. The children of a gate g correspond to all gates g' such that (g, g') is an edge. The parents of g correspond to all gates g' such that (g', g) is an edge. The *in-degree*, or a *fan-in*, of a gate g refers to its number of children, and the *out-degree* to its number of parents. We will not assume any restriction on the in-degree of a gate, and will thus consider circuits with unbounded fan-in. Gates with in-degree 0 are called *input gates* and are labeled by either a variable in X or a constant 0 or 1. All other gates are labeled by either + or ×, and are referred to as *sum gates* or *product gates*, respectively. Gates with out-degree 0 are called *output gates*. When talking about arithmetic circuits, one usually focuses on circuits with n input gates and a single output gate.

The *size* of Φ, denoted by $|\Phi|$, is its number of gates and wires. The *depth* of Φ, denoted by $\mathsf{depth}(\Phi)$, is the length of the longest directed path from any of its output gates to any of the input gates. The *degree* of a gate is defined inductively: an input gate has degree 1, a sum gate has a degree equal to the maximum of degrees of its children, and a product gate has a degree equal to the

sum of the degrees of its children. When Φ has a single output gate, the *degree* of Φ, denoted by $\mathsf{degree}(\Phi)$, is defined as the degree of its output gate. If Φ has a single output gate and its input gates take values from $\mathbb{R}$, then Φ corresponds to a polynomial in $\mathbb{R}[X]$ in a natural way. In this case, the degree of Φ equals the degree of the polynomial corresponding to Φ. If $a_1, \ldots, a_n$ are values in $\mathbb{R}$, then the result of the circuit on this input is the value computed by the corresponding polynomial, denoted by $\Phi(a_1, \ldots, a_k)$.

In order to handle inputs of different sizes, we use the notion of uniform circuit families. An *arithmetic circuit family* is a set of arithmetic circuits $\{\Phi_n \mid n = 1, 2, \ldots\}$ where Φ_n has n input variables and a single output gate. An arithmetic circuit family is *uniform* if there exists a LOGSPACE-Turing machine, which on input 1^n, returns an encoding of the arithmetic circuit Φ_n for each n. We observe that uniform arithmetic circuit families are necessarily of polynomial size. Another important parameter is the circuit depth. A circuit family is of logarithmic depth, whenever $\mathsf{depth}(\Phi_n) \in O(\log n)$. We now show that for-MATLANG subsumes uniform arithmetic circuit families that are of logarithmic depth.

Theorem 5.1. *For any uniform arithmetic circuit family* $\{\Phi_n \mid n = 1, 2, \ldots\}$ *of logarithmic depth there is a* for-MATLANG *schema* $\mathcal{S}$ *and an expression* e_Φ *using a matrix variable* v*, with* $\mathsf{type}_{\mathcal{S}}(v) = (\alpha, 1)$ *and* $\mathsf{type}_{\mathcal{S}}(e) = (1, 1)$*, such that for any input values* $a_1, \ldots, a_n$*:*

- *If* $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$ *is a* MATLANG *instance such that* $\mathcal{D}(\alpha) = n$ *and* $\mathtt{mat}(v) = [a_1 \ldots a_n]^T$*.*
- *Then* $[\![e_\Phi]\!](\mathcal{I}) = \Phi_n(a_1, \ldots, a_n)$*.*

It is important to note that the expression e_Φ does not change depending on the input size, meaning that it is uniform in the same sense as the circuit family being generated by a single Turing machine. The different input sizes for a for-MATLANG instance are handled by the typing mechanism of the language.

Proof sketch. The proof of this Theorem, which is the deepest technical result of the paper, depends crucially on two facts: (i) that any polynomial time Turing machine working within linear space and producing linear size output, can be simulated via a for-MATLANG expression; and (ii) that evaluating an arithmetic circuit Φ_n can be done using two stacks of depth n.

Evaluating Φ_n on input $(a_1, \ldots, a_n)$ can be done in a depth-first manner by maintaining two stacks: the gates-stack that tracks the current gate being evaluated, and the values-stack that stores the value that is being computed for this gate. The idea behind having two stacks is that whenever the number of items on the gates-stack is higher by one than the number of items on the values-stack, we know that we are processing a fresh gate, and we have to initialize its current value (to 0 if it is a sum gate, and to 1 if it is a product gate), and push it to the values-stack. We then proceed by processing the children of the head of the gates-stack one by one, and aggregate the results using sum if we are working with a sum gate, and by using product otherwise.

In order to access the information about the gate we are processing (such as whether it is a sum or a product gate, the list of its children, etc.) we use the uniformity of our circuit family. Namely, we know that we can generate the circuit Φ_n with a LOGSPACE-Turing machine M_Φ by running it on the input 1^n. Using this machine, we can in fact compute all the information needed to run the two-stack algorithms described above. For instance, we can construct a LOGSPACE machine that checks, given two gates g_1 and g_2, whether g_2 is a child of g_1. Similarly, we can construct a machine that, given g_1 and g_2 tells us whether g_2 is the final child of g_1, or the one that produces the following child of g_1 (according to the ordering given by the machine M_Φ). Defining these machines based of M_Φ is similar to the algorithm for the composition of two LOGSPACE transducers, and is commonly used to evaluate arithmetic circuits [1].

To simulate the circuit evaluation algorithm that uses two stacks, in for-MATLANG we can use a binary matrix of size $n \times n$, where n is the number of inputs. The idea here is that the gates-stack corresponds to the first $n - 3$ columns of the matrix, with each gate being encoded as a binary number in positions $1, \ldots, n - 3$ of a row. The remaining three columns are reserved for the values-stack, the number of elements on the gates stack, and the number of elements on the values stack, respectively. The number of elements is encoded as a canonical vector of size n. Here we crucially depend on the fact that the circuit is of logarithmic depth, and therefore the size of the two stacks is bounded by n (apart from the portion before the asymptotic bound kicks-in, which can be hard-coded into the expression e_Φ). Similarly, given that the circuits are of polynomial size, we can assume that gate ids can be encoded into $n - 3$ bits.

This matrix is then updated in the same way as the two-stack algorithm. It processes gates one by one, and using the successor relation for canonical vectors determines whether we have more elements on the gates stack. In this case, a new value is added to the values stack (0 if the gate is a sum gate, and 1 otherwise), and the process continues. Information about the next child, last child, or input value, are obtained using the expression which simulates the Turing machine generating this data about the circuit (the machines used never produce an output longer than their input). Given that the size of the circuit is polynomial, say n^k, we can initialize the matrix with the output gate only, and run the simulation of the two-stack algorithm for n^k steps (by iterating k times over size n canonical vectors). After this, the value in position $(1, n - 2)$ (the top of the values stack) holds the final results. □

While Theorem 5.1 gives us an idea on how to simulate arithmetic circuits, it does not tell us which classes of functions over real numbers can be computed by for-MATLANG expressions. In order to answer this question, we note that arithmetic circuits can be used to compute functions over real numbers. Formally, a circuit family $\{\Phi_n \mid n = 1, 2, \ldots\}$ computes a function $f : \bigcup_{n \geq 1} \mathbb{R}^n \mapsto \mathbb{R}$, if for any $a_1, \ldots a_n \in \mathbb{R}$ it holds that $\Phi_n(a_1, \ldots, a_n) = f(a_1, \ldots, a_n)$. To make the connection with for-MATLANG, we need to look at circuit families of bounded degree.

A circuit family $\{\Phi_n \mid n = 1, 2, \ldots\}$ is said to be of *polynomial degree* if $\mathsf{degree}(\Phi_n) \in O(p(n))$, for some polynomial $p(n)$. Note that polynomial size circuit families are not necessarily of polynomial degree. An easy corollary of Theorem 5.1 tells us that all functions computed by uniform family of circuits of polynomial degree and logarithmic depth can be simulated using for-MATLANG expressions. However, we can actually drop the restriction on circuit depth due to the result of Valiant et. al. [32] and Allender et. al. [2] which says that any function computed by a uniform circuit family of polynomial degree (and polynomial depth), can also be computed

by a uniform circuit family of logarithmic depth. Using this fact, we can conclude the following:

Corollary 5.2. *For any function f computed by a uniform family of arithmetic circuits of polynomial degree, there is an equivalent* for-MATLANG *formula e_f.*

Note that there is nothing special about circuits that have a single output, and both Theorem 5.1 and Corollary 5.2 also hold for functions $f : \bigcup_{n\geq 1} \mathbb{R}^n \mapsto \mathbb{R}^{s(n)}$, where s is a polynomial. Namely, in this case, we can assume that circuits for f have multiple output gates, and that the depth reduction procedure of [2] is carried out for each output gate separately. Similarly, the construction underlying the proof of Theorem 5.1 can be performed for each output gate independently, and later composed into a single output vector.

5.2 From for-MATLANG to circuits

Now that we know that arithmetic circuits can be simulated using for-MATLANG expressions, it is natural to ask whether the same holds in the other direction. That is, we are asking whether for each for-MATLANG expression e over some schema $\mathcal{S}$ there is a uniform family of arithmetic circuits computing precisely the same result depending on the input size.

In order to handle the fact that for-MATLANG expressions can produce any matrix, and not just a single value, as their output, we need to consider circuits which have multiple output gates. Similarly, we need to encode matrix inputs of a for-MATLANG expression in our circuits. We will write $\Phi(A_1, \ldots, A_k)$, where Φ is an arithmetic circuit with multiple output gates, and each A_i is a matrix of dimensions $\alpha_i \times \beta_i$, with $\alpha_i, \beta_i \in \{n, 1\}$ to denote the input matrices for a circuit Φ. We will also write $\mathsf{type}(\Phi) = (\alpha, \beta)$, with $\alpha, \beta \in \{n, 1\}$, to denote the size of the output matrix for Φ. We call such circuits *arithmetic circuits over matrices*. When $\{\Phi_n \mid n = 1, 2, \ldots\}$ is a uniform family of arithmetic circuits over matrices, we will assume that the Turing machine for generating Φ_n also gives us the information about how to access a position of each input matrix, and how to access the positions of the output matrix, as is usually done when handling matrices with arithmetic circuits [29]. The notion of degree is extended to be the sum of the degrees of all the output gates. With this at hand, we can now show the following result.

Theorem 5.3. *Let e be a* for-MATLANG *expression over a schema $\mathcal{S}$, and let $V_1, \ldots, V_k$ be the variables of e such that $\mathsf{type}_{\mathcal{S}}(V_i) \in \{(\alpha, \alpha), (\alpha, 1), (1, \alpha), (1, 1)\}$. Then there exists a uniform arithmetic circuit family over matrices $\Phi_n(A_1, \ldots, A_k)$ such that:*

- *For any instance $\mathcal{I} = (\mathcal{D}, \mathsf{mat})$ such that $\mathcal{D}(\alpha) = n$ and $\mathsf{mat}(V_i) = A_i$ it holds that:*
- $[\![e]\!](\mathcal{I}) = \Phi_n(A_1, \ldots, A_k)$.

It is not difficult to see that the proof of Theorem 5.1 can also be extended to support arithmetic circuits over matrices. In order to identify the class of functions computed by for-MATLANG expressions, we need to impose one final restriction: than on the degree of an expression. Formally, the *degree of* for-MATLANG *expression e* over a schema $\mathcal{S}$, is the minimum of the degrees of any circuit family $\{\Phi_n \mid n = 1, 2, \ldots\}$ that is equivalent to e. That is, the expression e is of polynomial degree, whenever there is an equivalent circuit family for e of a polynomial degree. For example, all for-MATLANG expressions seen so far have polynomial degree. With this definition, we can now identify the class of functions for which arithmetic circuits and for-MATLANG formulas are equivalent. This is the main technical contribution of the paper.

Corollary 5.4. *Let f be a function with input matrices $A_1, \ldots, A_k$ of dimensions $\alpha \times \beta$, with $\alpha, \beta \in \{n, 1\}$. Then, f is computed by a uniform circuit family over matrices of polynomial degree if and only if there is a* for-MATLANG *expression of polynomial degree for f.*

Note that this result crucially depends on the fact that expressions in for-MATLANG are of polynomial degree. Some for-MATLANG expression are easily seen to produce results which are not polynomial. An example of such an expression is, for instance, $e_{\mathsf{exp}} = \mathsf{for}\, v, X = A.\ X \cdot X$, over a schema $\mathcal{S}$ with $\mathsf{type}_{\mathcal{S}}(v) = (\gamma, 1)$, and $\mathsf{type}_{\mathcal{S}}(X) = (1, 1)$. Over an instance which assigns n to γ this expression computes the function a^{2^n}, for $A = [a]$. Therefore, a natural question to ask then is whether we can determine the degree of a for-MATLANG expression. Unfortunately, as we show in the following proposition this question is in fact undecidable.

Proposition 5.5. *Given a* for-MATLANG *expression e over a schema $\mathcal{S}$, it is undecidable to check whether e is of polynomial degree.*

Of course, one might wonder whether it is possible to define a syntactic subclass of for-MATLANG expressions that are of polynomial degree and can still express many important linear algebra algorithms. We identify one such class in Section 6.1, called sum-MATLANG, and in fact show that this class is powerful enough to capture relational algebra on (binary) K-relations.

5.3 Supporting additional operators

The equivalence of for-MATLANG and arithmetic circuits we prove above assumes that circuits can only use the sum and product gates (note that even without the sum and the product function, for-MATLANG can simulate these operations via matrix sum/product). However, both arithmetic circuits and expressions in for-MATLANG can be allowed to use a multitude of functions over $\mathbb{R}$. The most natural addition to the set of functions is the division operator, which is crucially needed in many linear algebra algorithms, such as, for instance, Gaussian elimination, or LU decomposition (recall Proposition 4.1). Interestingly, the equivalence in this case still holds, mainly due to a surprising result which shows that (almost all) divisions can in fact be removed for arithmetic circuits which allow sum, product, and division gates [1].

More precisely, in [6, 24, 31] it was shown that for any function of the form $f = g/h$, where g and h are relatively prime polynomials of degree d, if f is computed by an arithmetic circuit of size s, then both g and h can be computed by a circuit whose size is polynomial in $s + d$. Given that we can postpone the division without affecting the final result, this, in essence, tells us that division can be eliminated (pushed to the top of the circuit), and we can work with sum-product circuits instead. The degree of a circuit for f, can then be defined as the maximum of degrees of circuits for g and h. Given this fact, we can again use the depth reduction procedure of [2], and extend Corollary 5.4 to circuits with division.

Corollary 5.6. *Let f be a function taking as its input matrices $A_1, \ldots, A_k$ of dimensions $\alpha \times \beta$, with $\alpha, \beta \in \{n, 1\}$. Then, f is*

computed by a uniform circuit family over matrices of polynomial degree that allows divisions, if and only if there is a for-MATLANG$[f_/]$ *expression of polynomial degree for* f.

An interesting line of future work here is to see which additional functions can be added to arithmetic circuits and for-MATLANG formulas, in order to preserve their equivalence. Note that this will crucially depend on the fact that these functions have to allow the depth reduction of [2] in order to be supported.

6 RESTRICTING THE POWER OF FOR LOOPS

We conclude the paper by zooming in on some special fragments of for-MATLANG and in which matrices can take values from an arbitrary (commutative) semiring K. In particular, we first consider sum-MATLANG, in which iterations can only perform additive updates, and show that it is equivalent in expressive power to the (positive) relational algebra on K-relations. We then extend sum-MATLANG such that also updates involving pointwise-multiplication (Hadamard product) are allowed. The resulting fragment, FO-MATLANG, is shown to be equivalent in expressive power to weighted logics. Finally, we consider the fragment prod-MATLANG in which updates involving sum and matrix multiplication, and possibly order information, is allowed. From the results in Section 4, we infer that the latter fragment suffices to compute matrix inversion. An overview of the fragments and their relationships are depicted in Figure 1.

6.1 Summation matlang and relational algebra

When defining 4-cliques and in several other expressions we have seen so far, we only update X by adding some matrix to it. This restricted form of for-loop proved useful throughout the paper, and we therefore introduce it as a special operator. That is, we define:

$$\Sigma v.e := \mathsf{for}\, v, X\,.\, X + e.$$

We define the subfragment of for-MATLANG, called sum-MATLANG, to consist of the Σ operator plus the "core" operators in MATLANG, namely, transposition, matrix multiplication and addition, scalar multiplication, and pointwise function applications.

One property of sum-MATLANG is that it only allows expressions of polynomial degree. Indeed, one can easily show that sum-MATLANG can only create matrix entries that are polynomial in the dimension n of the expression. More precisely, we can show the following:

Proposition 6.1. *Every expression in* sum-MATLANG *is of polynomial degree.*

Interestingly enough, this restricted version of for-loop already allows us to capture the MATLANG operators that are not present in the syntax of sum-MATLANG. More precisely, we see from Examples 3.1 and 3.2 that the one-vector and diag operator are expressible in sum-MATLANG. Combined with the observation that the 4-clique expression of Example 3.3 is in sum-MATLANG, the following result is immediate.

Corollary 6.2. MATLANG *is strictly subsumed by* sum-MATLANG.

What operations over matrices can be defined with sum-MATLANG that is beyond MATLANG? In [8], it was shown that MATLANG is strictly included in the (positive) relational algebra on K-relations, denoted by RA^+_K [18].[2] It thus seems natural to compare the expressive power of sum-MATLANG with RA^+_K. The main result in this section is that sum-MATLANG and RA^+_K are equally expressive over binary schemas. To make this equivalence precise, we next give the definition of RA^+_K [18] and then show how to connect both formalisms.

Let $\mathbb{D}$ be a data domain and $\mathbb{A}$ a set of attributes. A relational signature is a finite subset of $\mathbb{A}$. A relational schema is a function $\mathcal{R}$ on finite set of symbols $\mathrm{dom}(\mathcal{R})$ such that $\mathcal{R}(R)$ is a relation signature for each $R \in \mathrm{dom}(\mathcal{R})$. To simplify the notation, from now on we write R to denote both the symbol R and the relational signature $\mathcal{R}(R)$. Furthermore, we write $R \in \mathcal{R}$ to say that R is a symbol of $\mathcal{R}$. For $R \in \mathcal{R}$, an R-tuple is a function $t : R \to \mathbb{D}$. We denote by $\mathrm{tuples}(R)$ the set of all R-tuples. Given $X \subseteq R$, we denote by $t[X]$ the restriction of t to the set X.

A semiring $(K, \oplus, \odot, \mathbb{0}, \mathbb{1})$ is an algebraic structure where K is a non-empty set, $\oplus$ and $\odot$ are binary operations over K, and $\mathbb{0}, \mathbb{1} \in K$. Furthermore, $\oplus$ and $\odot$ are associative operations, $\mathbb{0}$ and $\mathbb{1}$ are the identities of $\oplus$ and $\odot$ respectively, $\oplus$ is a commutative operation, $\odot$ distributes over $\oplus$, and $\mathbb{0}$ annihilates K (i.e. $\mathbb{0} \odot k = k \odot \mathbb{0} = \mathbb{0}$). As usual, we assume that all semirings in this paper are commutative, namely, $\odot$ is also commutative. We use $\bigoplus_X$ or $\bigodot_X$ for the $\oplus$- or $\odot$-operation over all elements in X, respectively. Typical examples of semirings are the reals $(\mathbb{R}, +, \times, 0, 1)$, the natural numbers $(\mathbb{N}, +, \times, 0, 1)$, and the boolean semiring $(\{0, 1\}, \vee, \wedge, 0, 1)$.

Fix a semiring $(K, \oplus, \odot, \mathbb{0}, \mathbb{1})$ and a relational schema $\mathcal{R}$. A K-relation of $R \in \mathcal{R}$ is a function $r : \mathrm{tuples}(R) \to K$ such that the support $\mathrm{supp}(r) = \{t \in \mathrm{tuples}(R) \mid r(t) \neq \mathbb{0}\}$ is finite. A K-instance $\mathcal{J}$ of $\mathcal{R}$ is a function that assigns relational signatures of $\mathcal{R}$ to K-relations. Given $R \in \mathcal{R}$, we denote by $R^{\mathcal{J}}$ the K-relation associated to R. Recall that $R^{\mathcal{J}}$ is a function and hence $R^{\mathcal{J}}(t)$ is the value in K assigned to t. Given a K-relation r we denote by $\mathrm{adom}(r)$ the active domain of r defined as $\mathrm{adom}(r) = \{t(a) \mid t \in \mathrm{supp}(r) \wedge a \in R\}$. Then the active domain of an K-instance $\mathcal{J}$ of $\mathcal{R}$ is defined as $\mathrm{adom}(\mathcal{J}) = \bigcup_{R \in \mathcal{R}} \mathrm{adom}(R^{\mathcal{J}})$.

An RA^+_K expression Q over $\mathcal{R}$ is given by the following syntax:

$$Q \quad := \quad R \mid Q \cup Q \mid \pi_X(Q) \mid \sigma_X(Q) \mid \rho_f(Q) \mid Q \bowtie Q$$

where $R \in \mathcal{R}$, $X \subseteq \mathbb{A}$ is finite, and $f : X \to Y$ is a one to one mapping with $Y \subseteq \mathbb{A}$. One can extend the schema $\mathcal{R}$ to any expression over $\mathcal{R}$ recursively as follows: $\mathcal{R}(R) = R$, $\mathcal{R}(Q \cup Q') = \mathcal{R}(Q)$, $\mathcal{R}(\pi_X(Q)) = X$, $\mathcal{R}(\sigma_X(Q)) = \mathcal{R}(Q)$, $\mathcal{R}(\rho_f(Q)) = X$ where $f : X \to Y$, and $\mathcal{R}(Q \bowtie Q') = \mathcal{R}(Q) \cup \mathcal{R}(Q')$ for every expressions Q and Q'. We further assume that any expression Q satisfies the following syntactic restrictions: $\mathcal{R}(Q') = \mathcal{R}(Q'')$ whenever $Q = Q' \cup Q''$, $X \subseteq \mathcal{R}(Q')$ whenever $Q = \pi_X(Q')$ or $Q = \sigma_X(Q')$, and $Y = \mathcal{R}(Q')$ whenever $Q = \rho_f(Q')$ with $f : X \to Y$.

Given an RA^+_K expression Q and a K-instance $\mathcal{J}$ of $\mathcal{R}$, we define the semantics $[\![Q]\!]_{\mathcal{J}}$ as a K-relation of $\mathcal{R}(Q)$ as follows. For $X \subseteq \mathbb{A}$, let $\mathrm{Eq}_X(t) = \mathbb{1}$ when $t(a) = t(b)$ for every $a, b \in X$, and $\mathrm{Eq}_X(t) =$

[2] The algebra used in [8] differs slightly from the one given in [18]. In this paper we work with the original algebra RA^+_K as defined in [18].

$\mathbb{0}$ otherwise. For every tuple $t \in \mathcal{R}(Q)$:

$$\begin{array}{ll}
\text{if } Q = R, \text{ then} & [\![Q]\!]_{\mathcal{J}}(t) = R^{\mathcal{J}}(t) \\
\text{if } Q = Q_1 \cup Q_2, \text{ then} & [\![Q]\!]_{\mathcal{J}}(t) = [\![Q_1]\!]_{\mathcal{J}}(t) \oplus [\![Q_2]\!]_{\mathcal{J}}(t) \\
\text{if } Q = \pi_X(Q'), \text{ then} & [\![Q]\!]_{\mathcal{J}}(t) = \bigoplus_{t':t'[X]=t} [\![Q']\!]_{\mathcal{J}}(t') \\
\text{if } Q = \sigma_X(Q'), \text{ then} & [\![Q]\!]_{\mathcal{J}}(t) = [\![Q']\!]_{\mathcal{J}}(t) \odot \mathrm{Eq}_X(t) \\
\text{if } Q = \rho_f(Q'), \text{ then} & [\![Q]\!]_{\mathcal{J}}(t) = [\![Q']\!]_{\mathcal{J}}(t \circ f) \\
\text{if } Q = Q_1 \bowtie Q_2, \text{ then} & [\![Q]\!]_{\mathcal{J}}(t) = [\![Q_1]\!]_{\mathcal{J}}(t[Y]) \odot [\![Q_2]\!]_{\mathcal{J}}(t[Z]),
\end{array}$$

where $Y = \mathcal{R}(Q_1)$ and $Z = \mathcal{R}(Q_2)$. It is important to note that the $\bigoplus$-operation in the semantics of $\pi_X(Q')$ is well-defined given that the support of $[\![Q']\!]_{\mathcal{J}}$ is always finite.

We are now ready for comparing sum-MATLANG with RA^+_K . First of all, we need to extend sum-MATLANG from $\mathbb{R}$ to any semiring. Let $\mathsf{Mat}[K]$ denote the set of all K-matrices. Similarly as for MATLANG over $\mathbb{R}$, given a MATLANG schema $\mathcal{S}$, a K-instance $\mathcal{I}$ over $\mathcal{S}$ is a pair $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$, where $\mathcal{D} : \mathtt{Symb} \mapsto \mathbb{N}$ assigns a value to each size symbol, and $\mathtt{mat} : \mathcal{M} \mapsto \mathsf{Mat}[K]$ assigns a concrete K-matrix to each matrix variable. Then it is straightforward to extend the semantics of MATLANG, for-MATLANG, and sum-MATLANG from $(\mathbb{R}, +, \times, 0, 1)$ to $(K, \oplus, \odot, \mathbb{0}, \mathbb{1})$ by switching + with $\oplus$ and $\times$ with $\odot$.

The next step to compare sum-MATLANG with RA^+_K is to represent K-matrices as K-relations. Let $\mathcal{S} = (\mathcal{M}, \mathtt{size})$ be a MATLANG schema. On the relational side we have for each size symbol $\alpha \in \mathtt{Symb} \setminus \{1\}$, attributes α, row_α, and col_α in $\mathbb{A}$. Furthermore, for each $V \in \mathcal{M}$ and $\alpha \in \mathtt{Symb}$ we denote by R_V and R_α its corresponding relation name, respectively. Then, given $\mathcal{S}$ we define the relational schema $\mathrm{Rel}(\mathcal{S})$ such that $\mathrm{dom}(\mathrm{Rel}(\mathcal{S})) = \{R_\alpha \mid \alpha \in \mathtt{Symb}\} \cup \{R_V \mid V \in \mathcal{M}\}$ where $\mathrm{Rel}(\mathcal{S})(R_\alpha) = \{\alpha\}$ and:

$$\mathrm{Rel}(\mathcal{S})(R_V) = \begin{cases} \{\mathsf{row}_\alpha, \mathsf{col}_\beta\} & \text{if } \mathtt{size}(V) = (\alpha, \beta) \\ \{\mathsf{row}_\alpha\} & \text{if } \mathtt{size}(V) = (\alpha, 1) \\ \{\mathsf{col}_\beta\} & \text{if } \mathtt{size}(V) = (1, \beta) \\ \{\} & \text{if } \mathtt{size}(V) = (1, 1). \end{cases}$$

Consider now a matrix instance $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$ over $\mathcal{S}$. Let $V \in \mathcal{M}$ with $\mathtt{size}(V) = (\alpha, \beta)$ and let $\mathtt{mat}(V)$ be its corresponding K-matrix of dimension $\mathcal{D}(\alpha) \times \mathcal{D}(\beta)$. To encode $\mathcal{I}$ as a K-instance in RA^+_K, we use as data domain $\mathbb{D} = \mathbb{N} \setminus \{0\}$. Then we construct the K-instance $\mathrm{Rel}(\mathcal{I})$ such that for each $V \in \mathcal{M}$ we define $R_V^{\mathrm{Rel}(\mathcal{I})}(t) := \mathtt{mat}(V)_{ij}$ whenever $t(\mathsf{row}_\alpha) = i \leq \mathcal{D}(\alpha)$ and $t(\mathsf{col}_\beta) = j \leq \mathcal{D}(\beta)$, and $\mathbb{0}$ otherwise. Furthermore, for each $\alpha \in \mathtt{Symb}$ we define $R_\alpha^{\mathrm{Rel}(\mathcal{I})}(t) := \mathbb{1}$ whenever $t(\alpha) \leq \mathcal{D}(\alpha)$, and $\mathbb{0}$ otherwise. In other words, R_α and R_β encodes the active domain of a matrix variable V with $\mathtt{size}(V) = (\alpha, \beta)$. Given that the RA^+_K framework of [18] represents the "absence" of a tuple in the relation with $\mathbb{0}$, we need to separately encode the indexes in a matrix. This is where $R_\alpha^{\mathrm{Rel}(\mathcal{I})}$ and $R_\beta^{\mathrm{Rel}(\mathcal{I})}$ are used for. We are now ready to state the first connection between sum-MATLANG and RA^+_K by using the previous encoding. The proof of the proposition below is by induction on the structure of expressions.

Proposition 6.3. *For each* sum-MATLANG *expression e over schema $\mathcal{S}$ such that $\mathcal{S}(e) = (\alpha, \beta)$ with $\alpha \neq 1 \neq \beta$, there exists an RA^+_K expression $\Phi(e)$ over relational schema $Rel(\mathcal{S})$ such that $Rel(\mathcal{S})(\Phi(e)) = \{\mathsf{row}_\alpha, \mathsf{row}_\beta\}$ and such that for any instance $\mathcal{I}$ over $\mathcal{S}$,*

$$[\![e]\!](\mathcal{I})_{i,j} = [\![\Phi(e)]\!]_{Rel(\mathcal{I})}(t)$$

for tuple $t(\mathrm{row}_\alpha) = i$ and $t(\mathrm{col}_\beta) = j$. Similarly for when e has schema $\mathcal{S}(e) = (\alpha, 1)$, $\mathcal{S}(e) = (1, \beta)$ or $\mathcal{S}(e) = (1, 1)$, then $\Phi(e)$ has schema $Rel(\mathcal{S})(\Phi(e)) = \{\mathrm{row}_\alpha\}$, $Rel(\mathcal{S})(\Phi(e)) = \{\mathrm{col}_\alpha\}$, or $Rel(\mathcal{S})(\Phi(e)) = \{\}$, respectively.

We now move to the other direction. To translate RA^+_K into sum-MATLANG, we must restrict our comparison to RA^+_K over K-relations with at most two attributes. Given that linear algebra works over vector and matrices, it is reasonable to restrict to unary or binary relations as input. Note that this is only a restriction on the input relations and not on intermediate relations, namely, expressions can create relation signatures of arbitrary size from the binary input relations. Thus, from now we say that a relational schema $\mathcal{R}$ is binary if $|R| \leq 2$ for every $R \in \mathcal{R}$. We also make the assumption that there is an (arbitrary) order, denoted by $<$, on the attributes in $\mathbb{A}$. This is to identify which attributes correspond to rows and columns when moving to matrices. Then, given that relations will be either unary or binary and there is an order on the attributes, we write $t = (v)$ or $t = (u, v)$ to denote a tuple over a unary or binary relation R, respectively, where u and v is the value of the first and second attribute with respect to $<$.

Consider a binary relational schema $\mathcal{R}$. With each $R \in \mathcal{R}$ we associate a matrix variable V_R such that, if R is a binary relational signature, then V_R represents a (square) matrix, and, if not (i.e. R is unary), then V_R represents a vector. Formally, fix a symbol $\alpha \in \mathtt{Symb} \setminus \{1\}$. Let $\mathrm{Mat}(\mathcal{R})$ denote the MATLANG schema $(\mathcal{M}_\mathcal{R}, \mathtt{size}_\mathcal{R})$ such that $\mathcal{M}_\mathcal{R} = \{V_R \mid R \in \mathcal{R}\}$ and $\mathtt{size}_\mathcal{R}(V_R) = (\alpha, \alpha)$ whenever $|R| = 2$, and $\mathtt{size}_\mathcal{R}(V_R) = (\alpha, 1)$ whenever $|R| = 1$. Take now a K-instance $\mathcal{J}$ of $\mathcal{R}$ and suppose that $\mathrm{adom}(\mathcal{J}) = \{d_1, \ldots, d_n\}$ is the active domain of $\mathcal{J}$ (the order over $\mathrm{adom}(\mathcal{J})$ is arbitrary). Then we define the matrix instance $\mathrm{Mat}(\mathcal{J}) = (\mathcal{D}_\mathcal{J}, \mathtt{mat}_\mathcal{J})$ such that $\mathcal{D}_\mathcal{J}(\alpha) = n$, $\mathtt{mat}_\mathcal{J}(V_R)_{i,j} = R^{\mathcal{J}}((d_i, d_j))$ whenever $|R| = 2$, and $\mathtt{mat}_\mathcal{J}(V_R)_i = R^{\mathcal{J}}((d_i))$ whenever $|R| = 1$. Note that, although each K-relation can have a different active domain, we encode them as square matrices by considering the active domain of the K-instance. By again using an inductive proof on the structure of RA^+_K expressions, we obtain the following result.

Proposition 6.4. *Let $\mathcal{R}$ be a binary relational schema. For each RA^+_K expression Q over $\mathcal{R}$ such that $|\mathcal{R}(Q)| = 2$, there exists a* sum-MATLANG *expression $\Psi(Q)$ over* MATLANG *schema $Mat(\mathcal{R})$ such that for any K-instance $\mathcal{J}$ with $\mathrm{adom}(\mathcal{J}) = \{d_1, \ldots, d_n\}$ over $\mathcal{R}$,*

$$[\![Q]\!]_{\mathcal{J}}((d_i, d_j)) = [\![\Psi(Q)]\!](Mat(\mathcal{J}))_{i,j}.$$

Similarly for when $|\mathcal{R}(Q)| = 1$, or $|\mathcal{R}(Q)| = 0$ respectively.

It is important to remark that the expression Q of the previous result can have intermediate expressions that are not necessary binary, given that the proposition only restricts that the input relation and the schema of Q must have arity at most two. We recall from [8] that MATLANG corresponds to RA^+_K where intermediate expressions are at most ternary, and this underlies, e.g., the inability of MATLANG to check for 4-cliques. In sum-MATLANG, we can deal with intermediate relations of arbitrary arity. In fact, each new attribute can be seen to correspond to an application of the Σ operator. For example, in the 4-clique expression, four Σ operators are needed, in analogy to how 4-clique is expressed in RA^+_K.

Given the previous two propositions we derive the following conclusion which is the first characterization of relational algebra with a (sub)-fragment of linear algebra.

Corollary 6.5. sum-MATLANG *and* RA_K^+ *over binary relational schemas are equally expressive.*

As a direct consequence, we have that sum-MATLANG cannot compute matrix inversion. Indeed, using similar arguments as in [7], i.e., by embedding RA_K^+ in (infinitary) first-order logic with counting and by leveraging its locality, one can show that sum-MATLANG cannot compute the transitive closure of an adjacency matrix. By contrast, the transitive closure can be expressed by means of matrix inversion [7]. We also note that the evaluation of the Σ operator is independent of the order in which the canonical vectors are considered. This is because $\oplus$ is commutative. Hence, sum-MATLANG cannot express the order predicates mentioned in Section 3.

6.2 Hadamard product and weighted logics

Similarly to using sum, we can use other operations to update X in the for-loop. The next natural choice is to consider products of matrices. In contrast to matrix sum, we have two options: either we can choose to use matrix product or to use the pointwise matrix product, also called the Hadamard product. We treat matrix product in the next subsection and first explain here the connection of sum and Hadamard product operators to weighted logics.

For the rest of this section, fix a semiring $(K, \oplus, \odot, \mathbb{0}, \mathbb{1})$. The Hadamard product over K-matrices can be defined as the pointwise application of $\odot$ between two matrices of the same size. Formally, we define the expression $e \circ e'$ where e, e' are expressions with respect to $\mathcal{S}$ and $\mathtt{type}_{\mathcal{S}}(e) = \mathtt{type}_{\mathcal{S}}(e')$ for some schema $\mathcal{S} = (\mathcal{M}, \mathtt{size})$. Then the semantics of $e \circ e'$ is the pointwise application of $\odot$, namely, $[\![e \circ e']\!](\mathcal{I})_{ij} = [\![e]\!](\mathcal{I})_{ij} \odot [\![e']\!](\mathcal{I})_{ij}$ for any instance $\mathcal{I}$ of $\mathcal{S}$. This enables us to define, similar as for Σv, the pointwise-product quantifier $\Pi^\circ v$ as follows:

$$\Pi^\circ v.\ e := \mathtt{for}\ v, X = \mathbb{1}\ .\ X \circ e.$$

where $\mathbb{1}$ is a matrix with the same type as X and all entries equal to the $\mathbb{1}$-element of K (i.e., we need to initialize X accordingly with the $\odot$-operator). We cal FO-MATLANG the subfragment of for-MATLANG that consists of sum-MATLANG extended with $\Pi^\circ v$.

Example 6.6. Similar to the trace of a matrix, a useful function in linear algebra is to compute the product of the values on the diagonal. Using the $\Pi^\circ v$ operator, this can be easily expressed:

$$e_{\mathsf{dp}}(V) := \Pi^\circ v.\ v^T \cdot V \cdot v.$$

Clearly, the inclusion of this new operator extends the expressive power to sum-MATLANG. For example, $[\![e_{\mathsf{dp}}]\!](\mathcal{I})$ can be an exponentially large number in the dimension n of the input. By contrast, one can easily show that all expressions in sum-MATLANG can only return numbers polynomial in n. That is, FO-MATLANG is more expressive than sum-MATLANG and RA_K^+.

To measure the expressive power of FO-MATLANG, we use weighted logics [13] (WL) as a yardstick. Weighted logics extend monadic second-order logic from the boolean semiring to any semiring K. Furthermore, it has been used extensively to characterize the expressive power of weighted automata in terms of logic [14]. We use here the first-order subfragment of weighted logics to suit our purpose and, moreover, we extend its semantics over weighted structures (similar as in [17]).

A relational vocabulary Γ is a finite collection of relation symbols such that each $R \in \Gamma$ has an associated arity, denoted by $\mathrm{arity}(R)$. A K-weighted structure over Γ (or just structure) is a pair $\mathcal{A} = (A, \{R^{\mathcal{A}}\}_{R \in \Gamma})$ such that A is a non-empty finite set (i.e. the domain) and, for each $R \in \Gamma$, $R^{\mathcal{A}} : A^{\mathrm{arity}(R)} \to K$ is a function that associates to each tuple in $A^{\mathrm{arity}(R)}$ a weight in K.

Let X be a set of first-order variables. A K-weighted logic (WL) formula φ over Γ is defined by the following syntax:

$$\varphi \quad := \quad x = y \mid R(\bar{x}) \mid \varphi \oplus \varphi \mid \varphi \odot \varphi \mid \Sigma x.\varphi \mid \Pi x.\varphi$$

where $x, y \in X$, $R \in \Gamma$, and $\bar{x} = x_1, \ldots, x_k$ is a sequence of variables in X such that $k = \mathrm{arity}(R)$. As usual, we say that x is a free variable of φ, if x is not below Σx or Πx quantifiers (e.g. x is free in $\Sigma y.R(x, y)$ but y is not). Given that K is fixed, from now on we talk about structures and formulas without mentioning K explicitly.

An assignment σ over a structure $\mathcal{A} = (A, \{R^{\mathcal{A}}\}_{R \in \Gamma})$ is a function $\sigma : X \to A$. Given $x \in X$ and $a \in A$, we denote by $\sigma[x \mapsto a]$ a new assignment such that $\sigma[x \mapsto a](y) = a$ whenever $x = y$ and $\sigma[x \mapsto a](y) = \sigma(y)$ otherwise. For $\bar{x} = x_1, \ldots, x_k$, we write $\sigma(\bar{x})$ to say $\sigma(x_1), \ldots, \sigma(x_k)$. Given a structure $\mathcal{A} = (A, \{R^{\mathcal{A}}\}_{R \in \Gamma})$ and an assignment σ, we define the semantics $[\![\varphi]\!]_{\mathcal{A}}(\sigma)$ of φ as follows:

if $\varphi := x = y$, then	$[\![\varphi]\!]_{\mathcal{A}}(\sigma) = \begin{cases} \mathbb{1} & \text{if } \sigma(x) = \sigma(y) \\ \mathbb{0} & \text{otherwise} \end{cases}$
if $\varphi := R(\bar{x})$, then	$[\![\varphi]\!]_{\mathcal{A}}(\sigma) = R^{\mathcal{A}}(\sigma(\bar{x}))$
if $\varphi := \varphi_1 \oplus \varphi_2$, then	$[\![\varphi]\!]_{\mathcal{A}}(\sigma) = [\![\varphi_1]\!]_{\mathcal{A}}(\sigma) \oplus [\![\varphi_2]\!]_{\mathcal{A}}(\sigma)$
if $\varphi := \varphi_1 \odot \varphi_2$, then	$[\![\varphi]\!]_{\mathcal{A}}(\sigma) = [\![\varphi_1]\!]_{\mathcal{A}}(\sigma) \odot [\![\varphi_2]\!]_{\mathcal{A}}(\sigma)$
if $\varphi := \Sigma x.\ \varphi'$, then	$[\![\varphi]\!]_{\mathcal{A}}(\sigma) = \bigoplus_{a \in A} [\![\varphi']\!]_{\mathcal{A}}(\sigma[x \mapsto a])$
if $\varphi := \Pi x.\ \varphi'$, then	$[\![\varphi]\!]_{\mathcal{A}}(\sigma) = \bigodot_{a \in A} [\![\varphi']\!]_{\mathcal{A}}(\sigma[x \mapsto a])$

When φ contains no free variables, we omit σ and write $[\![\varphi]\!]_{\mathcal{A}}$ instead of $[\![\varphi]\!]_{\mathcal{A}}(\sigma)$.

For comparing the expressive power of FO-MATLANG with WL, we have to show how to encode MATLANG instances into structures and vice versa. For this, we make two assumptions to put both languages at the same level: (1) we restrict structures to relation symbols of arity at most two and (2) we restrict instances to square matrices. The first assumption is for the same reasons as when comparing sum-MATLANG with RA_K^+, and the second assumption is to have a crisp translation between both languages. Indeed, understanding the relation of FO-MATLANG with WL for non-square matrices is slightly more complicated and we leave this for future work.

Let $\mathcal{S} = (\mathcal{M}, \mathtt{size})$ be a schema of square matrices, that is, there exists an α such that $\mathtt{size}(V) \in \{1, \alpha\} \times \{1, \alpha\}$ for every $V \in \mathcal{M}$. Define the relational vocabulary $\mathrm{WL}(\mathcal{S}) = \{R_V \mid V \in \mathcal{M}\}$ such that $\mathrm{arity}(R_V) = 2$ if $\mathtt{size}(V) = (\alpha, \alpha)$, $\mathrm{arity}(R_V) = 1$ if $\mathtt{size}(V) \in \{(\alpha, 1), (1, \alpha)\}$, and $\mathrm{arity}(R_V) = 0$ otherwise. Then given a matrix instance $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$ over $\mathcal{S}$ define the structure $\mathrm{WL}(\mathcal{I}) = (\{1, \ldots, n\}, \{R_V^{\mathcal{I}}\})$ such that $\mathcal{D}(\alpha) = n$ and $R_V^{\mathcal{I}}(i, j) = \mathtt{mat}(V)_{i,j}$ if $\mathtt{size}(V) = (\alpha, \alpha)$, $R_V^{\mathcal{I}}(i) = \mathtt{mat}(V)_i$ if $\mathtt{size}(V) \in \{(\alpha, 1), (1, \alpha)\}$, and $R_V^{\mathcal{I}} = \mathtt{mat}(V)$ if $\mathtt{size}(V) = (1, 1)$.

To encode weighted structures into matrices and vectors, the story is similar as for RA_K^+. Let Γ be a relational vocabulary where $\mathrm{arity}(R) \leq 2$. Define $\mathrm{Mat}(\Gamma) = (\mathcal{M}_\Gamma, \mathtt{size}_\Gamma)$ such that $\mathcal{M}_\Gamma = \{V_R \mid R \in \Gamma\}$ and $\mathtt{size}_\Gamma(V_R)$ is equal to (α, α), $(\alpha, 1)$, or $(1, 1)$ if

$\text{arity}(R) = 2$, $\text{arity}(R) = 1$, or $\text{arity}(R) = 0$, respectively, for some $\alpha \in \text{Symb}$. Similarly, let $\mathcal{A} = (A, \{R^{\mathcal{A}}\}_{R \in \Gamma})$ be a structure with $A = \{a_1, \ldots, a_n\}$, ordered arbitrarily. Then we define the matrix instance $\text{Mat}(\mathcal{A}) = (\mathcal{D}, \texttt{mat})$ such that $\mathcal{D}(\alpha) = n$, $\texttt{mat}(V_R)_{i,j} = R^{\mathcal{A}}(a_i, a_j)$ if $\text{arity}(R) = 2$, $\texttt{mat}(V_R)_i = R^{\mathcal{A}}(a_i)$ if $\text{arity}(R) = 1$, and $\texttt{mat}(V_R) = R^{\mathcal{A}}$ otherwise.

Let $\mathcal{S}$ be a MATLANG schema of square matrices and Γ a relational vocabulary of relational symbols of arity at most 2. We can then show the equivalence of FO-MATLANG and WL as follows.

Proposition 6.7. *Weighted logics over Γ and* FO-MATLANG *over $\mathcal{S}$ have the same expressive power. More specifically,*

- *for each* FO-MATLANG *expression e over $\mathcal{S}$ such that $\mathcal{S}(e) = (1, 1)$, there exists a WL-formula $\Phi(e)$ over $WL(\mathcal{S})$ such that for every instance $\mathcal{I}$ of $\mathcal{S}$, $[\![e]\!](\mathcal{I}) = [\![\Phi(e)]\!]_{WL(\mathcal{I})}$.*
- *for each WL-formula φ over Γ without free variables, there exists a* FO-MATLANG *expression $\Psi(\varphi)$ such that for any structure $\mathcal{A}$ over $Mat(\Gamma)$, $[\![\varphi]\!]_{\mathcal{A}} = [\![\Psi(\varphi)]\!](Mat(\mathcal{A}))$.*

6.3 Matrix multiplication as a quantifier

In a similar way, we can consider a fragment in which sum and the usual product of matrices can be used in for-loops. Formally, for an expression e we define the operator:

$$\Pi v.\, e = \texttt{for}\, v, X = I.\, X \cdot e.$$

where I is the identity matrix. We call prod-MATLANG the subfragment of for-MATLANG that consists of sum-MATLANG extended with Πv. It is readily verified that $\Pi^{\circ} v$ can expressed in terms of Πv. Furthermore, by contrast to the Hadamard product, matrix multiplication is a non-commutative operator. As a consequence, one can formulate expressions that are not invariant under the order in which the canonical vectors are processed.

Proposition 6.8. *Every expression in* FO-MATLANG *can be defined in* prod-MATLANG. *Moreover, there exists an expression that uses the Πv quantifier that cannot be defined in* FO-MATLANG.

What is interesting is that sum-MATLANG extended with Πv suffices to compute the transitive closure, provided that we allow for the $f_{>0}$ function. Indeed, one can use the expression $e_{\mathsf{TC}}(V) := f_{>0}\big(\Pi v.\, (e_{\mathsf{Id}} + V)\big)$ for this purpose because $[\![e_{\mathsf{TC}}]\!](\mathcal{I}) = f_{>0}\big((I + A)^n\big)$ when $\mathcal{I}$ assigns an $n \times n$ adjacency matrix A to V, and non-zero entries in $(I + A)^n$ coincide with non-zero entries in the transitive closure of A. Furthermore, if we extend this fragment with access to the matrix $S_<$, defining the (strict) order on canonical vectors, then Csanky's matrix inversion algorithm becomes expressible (if $f_/$ is allowed). We leave the study of this fragment and, in particular, the relationship to full for-MATLANG, for future work.

Finally, in Figure 1 we show a diagram of all the fragments of for-MATLANG introduced in this section and their corresponding equivalent formalisms.

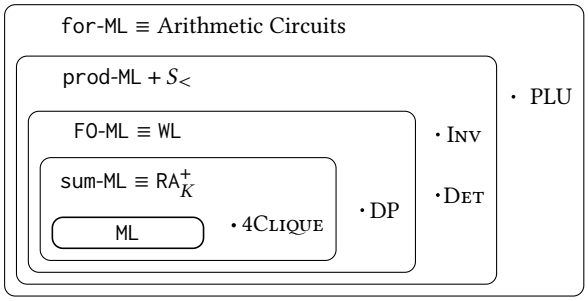


Figure 1: Fragments of for-MATLANG and their equivalences. The functions 4Clique, DP (diagonal product), Inv, Det, and PLU decomposition are placed in their fragments.

7 CONCLUSIONS

We proposed for-MATLANG, an extension of MATLANG with limited recursion, and showed that it is able to capture most of linear algebra due to its connection to arithmetic circuits. We further revealed interesting connections to logics on annotated relations. Our focus was on language design and expressivity. An interesting direction for future work relates to efficient evaluation of (fragments) of for-MATLANG. A possible starting point is [9] in which a general methodology for communication-optimal algorithms for for-loop linear algebra programs is proposed.

REFERENCES

[1] Eric Allender. 2004. Arithmetic circuits and counting complexity classes. *Complexity of Computations and Proofs, Quaderni di Matematica* 13 (2004), 33–72.
[2] Eric Allender, Jia Jiao, Meena Mahajan, and V. Vinay. 1998. Non-Commutative Arithmetic Circuits: Depth Reduction and Size Lower Bounds. *Theor. Comput. Sci.* 209, 1-2 (1998), 47–86.
[3] Sanjeev Arora and Boaz Barak. 2009. Complexity theory: A modern approach.
[4] Pablo Barceló, Nelson Higuera, Jorge Pérez, and Bernardo Subercaseaux. 2020. On the Expressiveness of LARA: A Unified Language for Linear and Relational Algebra. In *ICDT*. 6:1–6:20.
[5] Matthias Boehm, Arun Kumar, and Jun Yang. 2019. *Data Management in Machine Learning Systems.* Morgan & Claypool Publishers.
[6] Allan Borodin, Joachim von zur Gathem, and John Hopcroft. 1982. Fast parallel matrix and GCD computations. In *FOCS*. 65–71.
[7] Robert Brijder, Floris Geerts, Jan Van den Bussche, and Timmy Weerwag. 2019. On the Expressive Power of Query Languages for Matrices. *ACM Trans. Database Syst.* 44, 4 (2019), 15:1–15:31.
[8] Robert Brijder, Marc Gyssens, and Jan Van den Bussche. 2020. On Matrices and K-Relations. In *FoIKS*. 42–57.
[9] Michael Christ, James Demmel, Nicholas Knight, Thomas Scanlon, and Katherine A. Yelick. 2013. Communication Lower Bounds and Optimal Algorithms for Programs that Reference Arrays - Part 1. (2013).
[10] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. 2001. *Introduction to Algorithms* (2nd ed.). The MIT Press.
[11] L. Csanky. 1976. Fast Parallel Matrix Inversion Algorithms. *SIAM J. Comput.* 5, 4 (1976), 618–623.
[12] Anuj Dawar, Martin Grohe, Bjarki Holm, and Bastian Laubner. 2009. Logics with Rank Operators. In *LICS*. 113–122.
[13] Manfred Droste and Paul Gastin. 2005. Weighted Automata and Weighted Logics. In *ICALP*, Vol. 3580. 513–525.
[14] Manfred Droste, Werner Kuich, and Heiko Vogler. 2009. *Handbook of weighted automata.* Springer Science & Business Media.
[15] Floris Geerts. 2019. On the Expressive Power of Linear Algebra on Graphs. In *ICDT*. 7:1–7:19.
[16] Floris Geerts. 2020. When Can Matrix Query Languages Discern Matrices?. In *ICDT*. 12:1–12:18.
[17] Erich Grädel and Val Tannen. 2017. Semiring Provenance for First-Order Model Checking. *CoRR* abs/1712.01980 (2017).
[18] Todd J. Green, Gregory Karvounarakis, and Val Tannen. 2007. Provenance semirings. In *PODS*. 31–40.
[19] Martin Grohe and Wied Pakusa. 2017. Descriptive complexity of linear equation systems and applications to propositional proof complexity. In *LICS*. 1–12.
[20] Lauri Hella, Leonid Libkin, Juha Nurmonen, and Limsoon Wong. 2001. Logics with Aggregate Operators. *J. ACM* 48, 4 (2001), 880–907.
[21] Bjarki Holm. 2010. *Descriptive Complexity of Linear Algebra.* Ph.D. Dissertation. University of Cambridge.
[22] Dylan Hutchison, Bill Howe, and Dan Suciu. 2017. LaraDB: A Minimalist Kernel for Linear and Relational Algebra Computation. In *BeyondMR*. 2:1–2:10.
[23] Dimitrije Jankov, Shangyu Luo, Binhang Yuan, Zhuhua Cai, Jia Zou, Chris Jermaine, and Zekai J. Gao. 2019. Declarative Recursive Computation on an RDBMS.

Proc. VLDB Endow. 12, 7 (2019), 822–835.

[24] Erich Kaltofen. 1988. Greatest common divisors of polynomials given by straight-line programs. *Journal of the ACM (JACM)* 35, 1 (1988), 231–264.

[25] Mahmoud Abo Khamis, Hung Q. Ngo, XuanLong Nguyen, Dan Olteanu, and Maximilian Schleich. 2018. AC/DC: In-Database Learning Thunderstruck. In *DEEM*, Sebastian Schelter, Stephan Seufert, and Arun Kumar (Eds.). 8:1–8:10.

[26] Andreas Kunft, Alexander Alexandrov, Asterios Katsifodimos, and Volker Markl. 2016. Bridging the Gap: Towards Optimization Across Linear and Relational Algebra. In *BeyondMR*. 1:1–1:4.

[27] Shangyu Luo, Zekai J. Gao, Michael Gubanov, Luis L. Perez, and Christopher Jermaine. 2018. Scalable Linear Algebra on a Relational Database System. *SIGMOD Rec.* 47, 1 (2018), 24–31.

[28] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. 1992. *Numerical Recipes in C, 2nd Edition.* Cambridge University Press.

[29] Ran Raz. 2003. On the Complexity of Matrix Product. *SIAM J. Comput.* 32, 5 (2003), 1356–1369.

[30] Amir Shpilka and Amir Yehudayoff. 2010. Arithmetic Circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science* 5, 3-4 (2010), 207–388.

[31] Volker Strassen. 1973. Vermeidung von Divisionen. *Journal für die reine und angewandte Mathematik* 264 (1973), 184–202.

[32] Leslie G Valiant and Sven Skyum. 1981. Fast parallel computation of polynomials using few processors. In *International Symposium on Mathematical Foundations of Computer Science*. Springer, 132–139.

APPENDIX

A PRELIMINARIES

We first introduce some additional notations and describe simplifications that will be used later in the appendix.

A.1 Definitions

We sometimes want to iterate over k canonical vectors. We define the following shorthand notation:

$$\begin{aligned}\texttt{for}\, v_1,\ldots,v_k,X\,.\; e(X,v_1,\ldots,v_n) :=& \texttt{for}\, v_1, X_1\,.\; X_1 + \\ & \texttt{for}\, v_2, X_2 = X_1\,.\; X_2 + \\ & \quad \texttt{for}\, v_3, X_3 = X_2\,.\; X_3 + \\ & \qquad \ddots \\ & \qquad \texttt{for}\, v_k, X_k = X_{k-1}\,.\; e(X_k, v_1,\ldots,v_k).\end{aligned}$$

To reference ℓ different vector variables $X_1,\ldots,X_\ell$ in every iteration and update them in different ways we define:

$$\begin{aligned}\texttt{for}\, v, X_1,\ldots,X_\ell\,.\; (e_1(X_1,v), e_2(X_2,v),\ldots,e_l(X_\ell,v)) :=\; & \texttt{for}\, v, X\,.\; e_1(X\cdot e_{\mathsf{min}}, v)\cdot (e_{\texttt{diag}}(e_1(X^T))\cdot e_{\mathsf{min}})^T + \\ & e_2(X\cdot e_{\mathsf{min+1}}, v)\cdot (e_{\texttt{diag}}(e_1(X^T))\cdot e_{\mathsf{min+1}})^T + \ldots + e_\ell(X\cdot e_{\mathsf{max}}, v)\cdot (e_{\texttt{diag}}(e_1(X^T))\cdot e_{\mathsf{max}})^T\end{aligned}$$

We note that for the latter expression to be semantically correct v has to be of type $\gamma\times 1$, both X_i and e_i for $i=1,\ldots,\ell$ have to be of type $\alpha\times 1$, and X has to be of type $\alpha\times\beta$, where $\mathcal{D}(\beta)=\ell$. Here we use $e_{\texttt{diag}}(e_1(X^T))$ to compute the $\beta\times\beta$ identity and ensure the typing of the $e_{\mathsf{min}+i}$. When evaluated on an instance $\mathcal{I}$, $e_{\mathsf{min}}, e_{\mathsf{min}+i}$ evaluate to $b_1^{\mathcal{D}(\beta)}$ and $b_{1+i}^{\mathcal{D}(\beta)}$, respectively, and we show their defining expressions in section B.1. Similarly for $e_{\mathsf{max}} = b_n^{\mathcal{D}(\beta)}$. The combinations of both previous operators results in:

$$\texttt{for}\, v_1,\ldots,v_k,X_1,\ldots,X_\ell\,.\; (e_1(X_1,v_1,\ldots,v_k), e_2(X_2,v_1,\ldots,v_k),\ldots,e_\ell(X_\ell,v_1,\ldots,v_k)) := \texttt{for}\, v_1,\ldots,v_k,X\,.\; e'(X,v_1,\ldots,v_k)$$

where

$$\begin{aligned}e'(X,v_1,\ldots,v_k) :=& e_1(X\cdot e_{\mathsf{min}}, v_1,\ldots,v_k)\cdot (e_{\texttt{diag}}(e_1(X^T))\cdot e_{\mathsf{min}})^T && (1)\\ & + e_2(X\cdot e_{\mathsf{min+1}}, v_1,\ldots,v_k)\cdot (e_{\texttt{diag}}(e_1(X^T))\cdot e_{\mathsf{min+1}})^T && (2)\\ & + \ldots + e_\ell(X\cdot e_{\mathsf{max}}, v_1,\ldots,v_k)\cdot (e_{\texttt{diag}}(e_1(X^T))\cdot e_{\mathsf{max}})^T && (3)\end{aligned}$$

It is clear that this expression iterates over k canonical vectors and references ℓ independent vectors updating each of them in their particular way.

A.2 Simplifications

When showing results based on induction of expressions in for-MATLANG, it is often convenient to assume that function applications $f(e_1,\ldots,e_k)$ for $f\in\mathcal{F}_k$ are restricted to the case when all expressions $e_1,\ldots,e_k$ have type 1×1. This does not loose generality. Indeed, for general function applications $f(e_1,\ldots,e_k)$, if we have Σ, scalar product and function application on scalars (here denoted by $f_{1\times 1}$), we can simulate full function application, as follows:

$$f(e_1,\ldots,e_k) := \Sigma v_i \Sigma v_j . f_{1\times 1}(v_i^T\cdot e_1\cdot v_j,\ldots,v_i^T\cdot e_k\cdot v_j)\times v_i\cdot v_j^T.$$

Furthermore, it also convenient at times to use the pointwise functions $f_{\odot}^k:\mathbb{R}^k\mapsto\mathbb{R}:(x_1,\ldots,x_k)\mapsto x_1\times\cdots\times x_k$ and $f_{\oplus}^k:\mathbb{R}^k\mapsto\mathbb{R}:(x_1,\ldots,x_k)\mapsto x_1+\cdots+x_k$. In fact, it is readily observed that adding these functions does not extend the expressive power of for-MATLANG:

Lemma A.1. *We have that* for-MATLANG$[\emptyset]\equiv$ for-MATLANG$[\{f_{\odot}^k, f_{\oplus}^k \mid k\in\mathrm{N}\}]$.

In fact, this lemma also holds for the smaller fragments we consider. We also observe that having $f_{\odot}^2:\mathbb{R}^2\to\mathbb{R}$ allows us to define scalar multiplication:

$$e_1\times e_2 := f_{\odot}(1(e_2)^T\cdot e_1\cdot 1(e_2)^T, e_2).$$

Conversely, $f_{\odot}^k$ can be expressed using scalar multiplication, as can be seen from our simulation of general function applications by pointwise function application on scalars. Finally, a notational simplification is that when using scalars $a\in\mathbb{R}$ in our expressions, we write sometimes a instead of $[a]$. For example, $(1-e_1(v)^T\cdot v)$ stands for $([1]-e_1(v)^T\cdot v)$.

B PROOFS OF SECTION 3

B.1 Order predicates

We detail how order information on canonical vectors can be obtained in for-MATLANG. We provide explicit expressions for the operators mentioned in Section 3 and furthermore, we also define expressions for operators that will be used in our proofs.

To begin with, we can easily obtain the last canonical vector using the expression

$$e_{\mathsf{max}} := \mathtt{for}\, v, X\,.\, v.$$

In other words, we simply overwrite X with the current canonical vector in each iteration. Hence, at the end, X is assigned to the last canonical vector.

As already mentioned in the main body of the paper, to define an order relation for canonical vectors, we notice that the following matrix:

$$S_{\leq} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & \ddots & \ddots & \vdots \\ \cdots & \cdots & \cdots & 1 \\ 0 & \cdots & \cdots & 1 \end{bmatrix}.$$

has the property that for two canonical vectors b_i and b_j of the same dimension,

$$b_i^T \cdot S_{\leq} \cdot b_j = \begin{cases} 1 & \text{if } i \leq j \\ 0 & \text{otherwise.} \end{cases}$$

We observe that $S_{\leq}$ can be expressed in for-MATLANG as follows:

$$S_{\leq} := \mathtt{for}\, v, X\,.\, X + \big((X \cdot e_{\mathsf{max}}) + v\big) \cdot v^T + v \cdot e_{\mathsf{max}}^T,$$

where e_{max} is as defined above. The intuition behind this expression is that by using the last canonical vector b_n, as returned by e_{max}, we have access to the last column of X (via the product $X \cdot e_{\mathsf{max}}$). We use this column such that after the i-th iteration, this column contains the i-th column of $S_{\leq}$. This is done by incrementing X with $v \cdot e_{\mathsf{max}}^T$. To construct $S_{\leq}$, in the i-th iteration we further increment X with (i) the current last column in X (via $X \cdot e_{\mathsf{max}} \cdot v^T$) which holds the $(i-1)$-th column of $S_{\leq}$; and (ii) the current canonical vector (via $v \cdot v^T$). Hence, after iteration i, X contains the first i columns of $S_{\leq}$ and holds the ith column of $S_{\leq}$ in its last column. It is now readily verified that $X = S_{\leq}$ after the nth iteration.

By defining

$$\mathsf{succ}(u, v) := u^T \cdot S_{\leq} \cdot v,$$

we obtain an order relation that allows us to discern whether one canonical vector comes before the other in the order given by $S_{\leq}$. If we want a strict order, we can just use the matrix $S_{<} := S_{\leq} - e_{\mathsf{Id}}$, where e_{Id} is an expression in for-MATLANG which returns the identity matrix (of appropriate dimension). Given this, we define

$$\mathsf{succ}^+(u, v) := u^T \cdot S_{<} \cdot v.$$

from which we can also derive

$$\mathsf{max}(u) := u^T \cdot e_{\mathsf{max}}.$$

which is an expression that returns the last canonical vector.

Interestingly, we can also define the *previous* relation between canonical vectors. For this, we require the following matrix:

$$\mathsf{Prev} = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \cdots & \cdots & \cdots & 1 \\ 0 & \cdots & \cdots & 0 \end{bmatrix},$$

Using this matrix, we have that for a canonical vector b_i:

$$\mathsf{Prev} \cdot b_i = \begin{cases} b_{i-1}, & \text{if } i > 1. \\ 0, & \text{if } i = 1. \end{cases}$$

where 0 is a vector of zeros of the same type as b_i. Notice also that $1(u)^T \cdot \mathsf{Prev} \cdot u$ is equal to zero, for a canonical vector u, if and only if $u = b_1$ is the first canonical vector, and zero otherwise. Therefore the expression $\mathsf{min}(u)$ is defined as

$$\mathsf{min}(u) := 1 - 1(u)^T \cdot \mathsf{Prev} \cdot u,$$

and, when evaluated over canonical vectors, will result in 1 if and only if $u = b_1$ is the first canonical vector. To define the first canonical vector in the order given by for, we can then write:

$$e_{\mathsf{min}} := \mathtt{for}\, v, X\,.\, X + \mathsf{min}(v) \times v,$$

Finally, we show that Prev can be defined using the following for-MATLANG expression:

$$e_{\mathsf{Prev}} := \mathsf{for}\, v, X.\ X + \left((1 - \mathsf{max}(v)) \times v \cdot e_{\mathsf{max}}^T - (X \cdot e_{\mathsf{max}}) \cdot e_{\mathsf{max}}^T + (X \cdot e_{\mathsf{max}}) \cdot v^T\right).$$

Here, X is initialized as 0 and thus in the first iteration we put b_1 in the last column of X (note that $X \cdot e_{\mathsf{max}}$ is also zero in the first iteration). Next, in iteration two, we add a matrix that has the stored vector $X \cdot e_{\mathsf{max}}$ (the previous canonical vector) in the column indicated by v (the current canonical vector) and $v - X \cdot e_{\mathsf{max}}$ in the last column, to replace the vector stored. As a consequence, b_2 is now stored in the last column. In the last iteration, we have b_{n-1} already in the last column, so no further update of X is required.

To get the *next* relation we simply do $e_{\mathsf{Next}} = e_{\mathsf{Prev}}^T$. We have that for a canonical vector b_i:

$$\mathsf{Next} \cdot b_i = \begin{cases} b_{i+1}, \text{ if } i < n. \\ 0, \text{ if } i = n. \end{cases}$$

In this way, we also can obtain the following operators for a canonical vector v:

$$\mathsf{prev}(v) := e_{\mathsf{Prev}} \cdot v.$$

$$\mathsf{next}(v) := e_{\mathsf{Next}} \cdot v.$$

More generally, we define

$$e_{\mathsf{getPrevMatrix}}(v) := \Pi w.\mathsf{succ}(w, v) \times e_{\mathsf{Prev}} + (1 - \mathsf{succ}(w, v)) \times e_{\mathsf{Id}}$$

$$e_{\mathsf{getNextMatrix}}(v) := \Pi w.\mathsf{succ}(w, v) \times e_{\mathsf{Next}} + (1 - \mathsf{succ}(w, v)) \times e_{\mathsf{Id}}$$

expressions that, when v is interpreted as canonical vector b_i, output Prev^i and Next^i respectively. Note that

$$\mathsf{Prev}^j \cdot b_i = \begin{cases} b_{i-j}, \text{ if } i > j. \\ 0, \text{ if } i \leq j. \end{cases}$$

and

$$\mathsf{Next}^j \cdot b_i = \begin{cases} b_{i+j}, \text{ if } i + j \leq n. \\ 0, \text{ if } i + j > n. \end{cases}$$

Finally, define

$$e_{\mathsf{min}+i} := \underbrace{e_{\mathsf{getNextMatrix}}(\ldots e_{\mathsf{getNextMatrix}}}_{i \text{ times}}(e_{\mathsf{min}}))$$

and

$$e_{\mathsf{max}-i} := \underbrace{e_{\mathsf{getPrevMatrix}}(\ldots e_{\mathsf{getPrevMatrix}}}_{i \text{ times}}(e_{\mathsf{max}}))$$

We note that some these expressions were already used in Section A.1.

C PROOFS OF SECTION 4

We next provide more details about how to perform LU-decomposition (without and with pivoting) and to compute the determinant and inverse of a matrix.

C.1 LU-decomposition

We start with LU-decomposition without pivoting. We recall proposition 4.1:

Proposition 4.1. *There exists* for-MATLANG$[f_/]$ *expressions* $e_L(V)$ *and* $e_U(V)$ *such that* $[\![e_L]\!](\mathcal{I}) = L$ *and* $[\![e_U]\!](\mathcal{I}) = U$ *form an LU-decomposition of* A*, where* $\mathtt{mat}(V) = A$ *and* A *is LU-factorizable.*

Proof. Let A be an LU-factorizable matrix. We already explained how the expression $e_U(V)$ is obtained in the main body of the paper, i.e.,

$$e_U(V) := (\mathsf{for}\, y, X{=}e_{\mathsf{Id}}.\ \mathtt{reduce}(X \cdot V, y) \cdot X) \cdot V.$$

We recall that $e_U(A) = T_n \cdot \cdots \cdot T_1 \cdot A$ with $L^{-1} = T_n \cdot \cdots \cdot T_1$. Let

$$e_{L^{-1}}(V) := \mathsf{for}\, y, X{=}e_{\mathsf{Id}}.\ \mathtt{reduce}(X \cdot V, y) \cdot X.$$

such that

$$e_{\mathsf{U}}(V) := e_{L^{-1}}(V) \cdot V.$$

It now suffices to observe that, since $T_n = I$,

$$\begin{aligned} L^{-1} &= (I - c_1 \cdot b_1^T) \cdots (I - c_{n-1} \cdot b_{n-1}^T) \\ &= I - c_1 \cdot b_1^T - \cdots - c_{n-1} \cdot b_{n-1}^T \end{aligned}$$

and hence,

$$L = (I + c_1 \cdot b_1^T) \cdots (I + c_{n-1} \cdot b_{n-1}^T)$$
$$= I + c_1 \cdot b_1^T + \cdots + c_{n-1} \cdot b_{n-1}^T.$$

As a consequence, to obtain L from L^{-1} we just need to multiply every entry below the diagonal by -1. Since both L and L^{-1} are lower triangular, this can done by computing $L = -1 \times L^{-1} + 2 \times I$. Translated into for-MATLANG, this means that we can define

$$e_L(V) := -1 \times e_{L^{-1}}(V) + 2 \times e_{\mathsf{Id}},$$

which concludes the proof of the proposition. □

C.2 LU-decomposition with pivoting

We next consider LU-decomposition with pivoting. We recall proposition 4.2:

Proposition 4.2. *There exist expressions $e_{L^{-1}P}(V)$ and $e_U(V)$ in* for-MATLANG$[f_/, f_{>0}]$ *such that $L^{-1} \cdot P = [\![e_{L^{-1}P}]\!](\mathcal{I})$ and $U = [\![e_U]\!](\mathcal{I})$, satisfy $L^{-1} \cdot P \cdot A = U$, where $\mathcal{I}$ is an instance such that* $\mathtt{mat}(V) = A$.

Proof. We assume that $f_/$ and $f_{>0}$ are in $\mathcal{F}$. Let A be an arbitrary matrix. By contrast to when A is LU-factorizable, during the LU-decomposition process we may need row interchange (pivoting) in each step of the iteration. Let us assume that row interchange is needed immediately before step k, $1 \leq k \leq n$. In other words, we now aim to reduce the k-th column of $A_k = T_{k-1} \cdots T_1 \cdot A$, or $A_k = A$ if $k = 1$, but now A_k has a zero pivot, i.e., $(A_k)_{kk} = 0$. Let P be the matrix that denotes the necessary row interchange. If we know P, then to compute T_k we need to perform $\mathtt{reduce}(P \cdot X \cdot A, v)$ in this iteration, where $\mathtt{reduce}(\cdot,)$ is the expression in for-MATLANG reducing a column, as explained in the main body of the paper. Furthermore, we need to apply the permutation P to the current result, resulting in the expression $\mathtt{for}\, v, X = I.\ \mathtt{reduce}(P \cdot X \cdot A, v) \cdot P \cdot X$. We now remark that P is a permutation matrix of the form $P = I - u \cdot u^T$ and it denotes an interchange (if multiplied by left) of rows i and j if $u = (b_i - b_j)$. Note that we are performing a row interchange for column k and thus $i = k$ and $j > k - 1$. If no interchange is needed, $i = j = k$ and $P = I$. Also note that when $k = n$ no interchange takes place. Furthermore, if no suitable b_j can be found, this implies that no interchange is required as well and we can move on to next column.

To find the vector u in P, we can, for example, find the first entry $j \geq k$ in column k of A_k that holds a non-zero value. More generally, we can find the first entry in a vector a that holds a non-zero value by using the function $f_{>0}$. Indeed, consider the following expression:

$$\mathtt{neq}(a, u) := \mathtt{for}\, v, X.\ \left(1 - e_1(v)^T \cdot X\right) \times f_{>0}\left((v^T \cdot a)^2\right) \times v + \mathsf{max}(v) \times \left(1 - e_1(v)^T \cdot X\right) \times \left(1 - f_{>0}\left((v^T \cdot a)^2\right)\right) \times u$$

Here, $\mathtt{neq}(a, u)$ receives two n dimensional vectors a and u and outputs a canonical vector b_j such that a_j is the first non-zero entry of a, or u if such non-zero value does not exist in a. We check for $f_{>0}((\cdot)^2)$ in case a negative number is tested. The above expression simply checks in each iteration whether X already holds a canonical vector. If so, then X is not updated. Otherwise, X is replaced by the current canonical vector b_j if and only if $b_j^T \cdot a$ is non-zero. Furthermore, when the final canonical vector is considered and X does not hold a canonical vector yet and $b_n^T \cdot a$ is zero, the vector u is returned.

We use $\mathtt{neq}(a, u)$ to find a pivot for a specific column. Let us assume again that we want to find a pivot in column k of A_k. We can then first make all entries in that column, with indexes smaller or equal to k, zero, just as we did by means of $\mathtt{col}(\cdot, \cdot)$ in the definition of $\mathtt{reduce}(\cdot,)$. Except, now we also need to make the $k$the entry zero as well. Let us denote by $\mathtt{col}_{\mathtt{eq}}(\cdot, \cdot)$ the operation $\mathtt{col}(\cdot, \cdot)$, as defined in the main body of the paper, but using succ instead of succ^+ (to include the k entry). Given this, we can construct $P = I - u \cdot u^T$ as follows:

$$e_{P_u}(A, u) := e_{\mathsf{Id}} - \left[u - \mathtt{neq}(\mathtt{col}_{\mathtt{eq}}(A, u), u)\right] \cdot \left[u - \mathtt{neq}(\mathtt{col}_{\mathtt{eq}}(A, u), u)\right]^T.$$

From the explanations given above, it should be clear that $e_{P_u}(A, u)$ computes the necessary permutation matrix of A_k for the column indicated by u, or I if no permutation is needed, or if such permutation does not exist (so we skip the current column). Also, we have to modify the $\mathtt{reduce}(V, y)$ operators, as follows:

$$\mathtt{reduce}(V, y) := e_{\mathsf{Id}} + f_{>0}\left((y^T \cdot V \cdot y)^2\right) \times f_/(\mathtt{col}(V, y), \left[-(y^T \cdot V \cdot y) \times e_1(y) + \left(1 - f_{>0}\left((y^T \cdot V \cdot y)^2\right)\right) \times e_1(y)\right]) \cdot y^T,$$

so that when V is interpreted by a matrix B and $y = b_i$, it returns $I + c_i b_i^T$ if B_{ii} is not zero. If $B_{ii} = 0$ then we divide $\mathtt{col}(B, b_i)$ by $e_1(b_i)$ (so we don't get *undefined*), but we don't add $c_i b_i^T$ precisely because $B_{ii} = 0$, and return the identity so nothing happens. We check for $f_{>0}((\cdot)^2)$ in case a negative number is tested.

Finally, we define

$$e_{L^{-1}P}(V) := \mathtt{for}\, v, X = e_{\mathsf{Id}}.\ \mathtt{reduce}(e_{P_v}(X \cdot V, v) \cdot X \cdot V, v) \cdot e_{P_v}(X \cdot V, v) \cdot X$$

and $e_{\mathsf{U}}(V) := e_{L^{-1}P}(V) \cdot V$ as the desired expressions.

As a final observation, in the definition of $e_{L^{-1}P}(V)$ we interlaced permutation matrices with the T_i's. More specifically, $A_k = T_k \cdot P \cdot T_{k-1} \cdots T_1 \cdot A$. We observe, however, that for $\ell \leq k - 1$ and $T_\ell = I - c_\ell \cdot b_\ell^T$, we have that $b_\ell^T \cdot P = b_\ell$ because b_ℓ has zeroes in positions in the rows involved in the row exchange P. Also, note that $P^2 = I$ and thus

$$P \cdot T_\ell \cdot P = P^2 - P \cdot c_\ell \cdot b_\ell^T \cdot P = I - \widehat{c}_\ell \cdot b_l^T = \widehat{T}_\ell.$$

As a consequence,

$$T_k \cdot P \cdot T_{k-1} \cdots T_1 = T_k \cdot P \cdot T_{k-1} \cdot P^2 \cdot T_{k-2} \cdot P^2 \cdots P^2 \cdot T_1 \cdot P^2 = T_k \cdot (P \cdot T_{k-1} \cdot P) \cdots (P \cdot T_1 \cdot P) \cdot P = \widehat{T}_{k-1} \cdots \widehat{T}_1 \cdot P,$$

and thus we may assume that P occurs at the end. Hence, we obtain $L^{-1} \cdot P \cdot A = U$. □

C.3 Determinant and inverse

We next turn our attention to computing the inverse and determinant of a matrix. To show Proposition 4.3 we first show that it holds when considering non-singular lower or upper triangular matrices.

Lemma C.1. *There are* for-MATLANG$[f_/]$ *expressions* $e_{\mathsf{upperDiagInv}}(V)$ *and* $e_{\mathsf{lowerDiagInv}}(V)$ *such that* $[\![e_{\mathsf{upperDiagInv}}]\!](\mathcal{I}) = A^{-1}$ *when* $\mathcal{I}$ *assigns* V *to an invertible upper triangular matrix* A *and* $[\![e_{\mathsf{lowerDiagInv}}]\!](\mathcal{I}) = A^{-1}$ *when* $\mathcal{I}$ *assigns* V *to an invertible lower triangular matrix* A.

Proof. We start by considering the expression:

$$e_{\mathsf{ps}}(V) := e_{\mathsf{Id}} + \Sigma v.\Pi w.\, [\mathsf{succ}(w, v) \times V + (1 - \mathsf{succ}(w, v)) \times e_{\mathsf{Id}}]\,.$$

Here, $e_{\mathsf{ps}}(A)$ results in $I + A + A^2 + \cdots + A^n$ for any matrix A. In the expression, the outer loop defines which power we compute. That is, when v is the ith canonical vector, we compute A^i. Computing A^i is achieved via the inner product loop, which uses $\mathsf{succ}(w, v)$ to determine whether w comes before or is v in the ordering of canonical vectors. When this is the case, we multiply the current result by A, and when w is greater than v, we use the identity as not to affect the already computed result. We add the identity at the end.

Now, let A be an $n \times n$ matrix that is upper triangular and let D_A be the matrix consisting of the diagonal elements of A, i.e.,

$$D_A = \begin{bmatrix} a_{11} & \cdots & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \cdots & \cdots & a_{nn} \end{bmatrix}.$$

We can compute D_A by the expression:

$$e_{\mathsf{getDiag}}(V) := \Sigma v.(v^T \cdot V \cdot v) \times v \cdot v^T.$$

Let $T = A - D_A$, then

$$A^{-1} = [D_A + T]^{-1} = \left[D_A \left(I + D_A^{-1} T\right)\right]^{-1} = \left(I + D_A^{-1} T\right)^{-1} D_A^{-1}.$$

We now observe that D_A^{-1} simply consists of the inverses of the elements on the diagonal. This can be expressed, as follows:

$$e_{\mathsf{diagInverse}}(V) := \Sigma v.f_/(1, v^T \cdot V \cdot v) \times v \cdot v^T = \Sigma v.f_/(1, v^T \cdot V \cdot v) \times v \cdot v^T,$$

Where $f_/$ is the division function. In the last equality we take advantage of the fact that the diagonals of A and D_A are the same.

We now focus on the computation of $\left(I + D_A^{-1} \cdot T\right)^{-1}$. First, by construction, $D_A^{-1} \cdot T$ is strictly upper triangular and thus nilpotent, such that $\left(D_A^{-1} \cdot T\right)^n = 0$, where n is the dimension of A. Recall the following algebraic identity

$$(1 + x)\left(\sum_{i=0}^{m} (-x)^i\right) = 1 - (-x)^{m+1}.$$

By choosing $m = n - 1$ and applying it to $x = D_A^{-1} \cdot T$, we have

$$\left(I + D_A^{-1} \cdot T\right)\left(\sum_{i=0}^{n-1} (-D_A^{-1} \cdot T)^i\right) = I - \left(-D_A^{-1} \cdot T\right)^n = I.$$

Hence,

$$\left(I + D_A^{-1} \cdot T\right)^{-1} = \sum_{i=0}^{n-1} (-D_A^{-1} \cdot T)^i = \sum_{i=0}^{n} (-D_A^{-1} \cdot T)^i.$$

We now observe that

$$e_{\mathsf{ps}}(-1 \times D_A^{-1} \cdot T) = \sum_{i=0}^{n} (-D_A^{-1} \cdot T)^i = \left(I + D_A^{-1} \cdot T\right)^{-1},$$

and thus

$$A^{-1} = e_{\mathsf{ps}}\left(-1 \times \left[e_{\mathsf{diagInverse}}(A)(A - e_{\mathsf{getDiag}}(A))\right]\right) e_{\mathsf{diagInverse}}(A).$$

Seeing this as an expression:

$$e_{\mathsf{upperDiagInv}}(V) := e_{\mathsf{ps}}\left(-1 \times \left[e_{\mathsf{diagInverse}}(V)(V - e_{\mathsf{getDiag}}(V))\right]\right) e_{\mathsf{diagInverse}}(V),$$

we see that when interpreting V as an upper triangular invertible matrix, $e_{\mathsf{upperDiagInv}}(A)$ evaluates to A^{-1}.

To deal with invertible lower triangular matrices A, we observe that $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$ and A^T is upper triangular. Hence, it suffices to define

$$e_{\mathsf{lowerDiagInv}}(V) := e_{\mathsf{upperDiagInv}}(V^T)^T.$$

This concludes the proof of the lemma. □

We are now ready to prove proposition 4.3. We recall:

Proposition 4.3. *There are* for-MATLANG$[f_/]$ *expressions* $e_{\mathsf{det}}(V)$ *and* $e_{\mathsf{inv}}(V)$ *such that* $[\![e_{\mathsf{det}}]\!](\mathcal{I}) = \mathsf{det}(A)$, *and* $[\![e_{\mathsf{inv}}]\!](\mathcal{I}) = A^{-1}$ *when* $\mathcal{I}$ *assigns* V *to* A *and* A *is invertible.*

Proof. Let A be an $n \times n$ matrix. As mentioned in the main body of the paper, we will implement Csanky's algorithm. Let $p_A(x) := \mathsf{det}(xI - A)$ denote characteristic polynomial of A. We write $p_A(x) = 1 + \sum_{i=1}^{n} c_i x^i$ and let $S_i := \frac{1}{i+1}\mathsf{tr}(A^i)$ with $\mathsf{tr}(\cdot)$ the trace operator which sums up the diagonal elements of a matrix. Then, the coefficients $c_1, \ldots, c_n$ are known to satisfy[3]

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ S_1 & 1 & 0 & \cdots & 0 & 0 \\ S_2 & S_1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ S_{n-1} & S_{n-2} & S_{n-3} & \cdots & S_1 & 1 \end{pmatrix}}_{S} \cdot \underbrace{\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix}}_{\bar{c}} = \underbrace{\begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_n \end{pmatrix}}_{\bar{b}}$$

and furthermore, $c_n = (-1)^n \mathsf{det}(A)$ and if $c_n \neq 0$, then

$$A^{-1} = \frac{1}{c_n} \sum_{i=0}^{n-1} c_i A^{n-1-i},$$

with $c_0 = 1$. It is now easy to see that we can compute the $S_i's$ in for-MATLANG. Indeed, for $i = 1, \ldots, n$ we can consider

$$e_{\mathsf{powTr}}(V, v) := \Sigma w. w^T \cdot \left(e_{\mathsf{pow}}(V, v) \cdot V\right) \cdot w$$

with

$$e_{\mathsf{pow}}(V, v) := \Pi w. (\mathsf{succ}(w, v) \times V + (1 - \mathsf{succ}(w, v)) \times e_{\mathsf{Id}}).$$

We have that $e_{\mathsf{pow}}(A, b_j) = A^j$ and thus $e_{\mathsf{powTr}}(A, b_j) = \mathsf{tr}(A^j)$. Define:

$$e_S(V, v) := f_/(1, 1 + \Sigma w. \mathsf{succ}(w, v)) \times e_{\mathsf{powTr}}(V, v).$$

Here $e_S(A, b_i) = S_i$. Note that $i + 1$ is computed summing up to the dimension indicated by v, and adding 1. We can now easily construct the vector $\bar{b}$ used in the system of equations by means of the expression:

$$e_{\bar{b}}(V) := \Sigma w. e_S(V, w) \times w.$$

We next construct the matrix S. We need to be able to *shift* a vector a in k positions, i.e., such that $(a_1, \ldots, a_n) \mapsto (0, \ldots, a_1, \ldots, a_{n-k})$. We use $e_{\mathsf{getNextMatrix}}$ defined in section B.1, i.e., we define:

$$e_{\mathsf{shift}}(a, v) := \Sigma w. (w^T \cdot a) \times (e_{\mathsf{getNextMatrix}}(v) \cdot w)$$

performs the desired shift when u is assigned a vector a and v is b_k. The matrix S is now obtained as follows:

$$S(V) := e_{\mathsf{Id}} + \Sigma v. e_{\mathsf{shift}}(e_{\bar{b}}(V), v) \cdot v^T$$

We now observe that S is lower triangular with nonzero diagonal entries. So, Lemma C.1 tells us that we can invert it, i.e., $e_{\mathsf{lowerDiagInv}}(S) = S^{-1}$. As a consequence,

$$e_{\bar{c}}(V) := e_{\mathsf{lowerDiagInv}}(S(V)) \cdot e_{\bar{b}}(V).$$

outputs $\bar{c}$ when V is interpreted as matrix A. Observe that we only use the division operator. We now have all coefficients of the characteristic polynomial of A.

We can now define

$$e_{\mathsf{det}}(V) := \left(\left((\Pi w.(-1) \times e_1(V))^T \cdot e_{\mathsf{max}}\right) \times e_{\bar{c}}(V)\right)^T \cdot e_{\mathsf{max}},$$

an expression that, when V is interpreted as any matrix A, outputs $\mathsf{det}(A)$. Here, $(\Pi w.(-1) \times e_1(V))$ is the n dimensional vector with $(-1)^n$ in all of its entries. Since $c_n = (-1)^n \mathsf{det}(A)$, we extract $(-1)^n(-1)^n \mathsf{det}(A) = \mathsf{det}(A)$ with e_{max}.

[3] We use a slightly different, but equivalent, system of equations than the one mentioned in the paper.

For the inverse, we have that

$$A^{-1} = \frac{1}{c_n}\sum_{i=0}^{n-1} c_i A^{n-1-i} = \frac{1}{c_n}A^{n-1} + \sum_{i=1}^{n-1}\frac{c_i}{c_n}A^{n-1-i}.$$

We compute $\frac{1}{c_n}A^{n-1}$ as

$$f_/(1, e_{\bar{c}}(A)^T \cdot e_{\mathsf{max}}) \times e_{\mathsf{pow}}(A, e_{\mathsf{max}})$$

and $\sum_{i=1}^{n-1}\frac{c_i}{c_n}A^{n-1-i}$ as

$$\Sigma v. f_/\left(e_{\bar{c}}(A)^T \cdot v, e_{\bar{c}}(A)^T \cdot e_{\mathsf{max}}\right) \times e_{\mathsf{invPow}}(A, v),$$

where

$$e_{\mathsf{invPow}}(V, v) := \Pi w.(1 - \mathsf{max}(w)) \times [(1 - \mathsf{succ}(w, v)) \times V + \mathsf{succ}(w, v) \times e_{\mathsf{Id}}] + \mathsf{max}(w) \times e_{\mathsf{Id}}.$$

Here, $e_{\mathsf{invPow}}(A, b_i) = A^{n-1-i}$ and $e_{\mathsf{invPow}}(A, b_n) = I$. Note that we always multiply by e_{Id} in the last step. To conclude, we define:

$$e_{\mathsf{inv}}(V) := f_/(1, e_{\bar{c}}(V)^T \cdot e_{\mathsf{max}}) \times e_{\mathsf{pow}}(V, e_{\mathsf{max}}) + \left[\Sigma v. f_/\left(e_{\bar{c}}(V)^T \cdot v, e_{\bar{c}}(V)^T \cdot e_{\mathsf{max}}\right) \times e_{\mathsf{invPow}}(V, v)\right],$$

an expression that, when V is interpreted as any invertible matrix A, computes A^{-1}. □

As an observation, here we only use operators Σ and Π defined in section 6. We also assume access to order.

D PROOFS OF SECTION 5

D.1 Linear space functions

We start by showing a crucial ingredient for making the correspondence between for-MATLANG and arithmetic circuits. More specifically, we show that any polynomial time Turing machine, working within linear space and producing linear space output, can be simulated in for-MATLANG. For this proof and section only, we will denote the canonical vectors as $\mathrm{e}_1, \ldots, \mathrm{e}_n$, since b will be used to represent a value on a position of a tape.

We consider deterministic Turing Machines (TM) T consisting of ℓ read-only input tapes, denoted by $R_1, \ldots, R_\ell$, a work tape, denoted by W, and a write-only output tape, denoted by O. The TM T has a set Q of m states, denoted by $q_0, \ldots, q_m$. We assume that q_0 is the initial state and q_m is the accepting state. The input and tape alphabet are $\Sigma = \{0, 1\}$ and $\Gamma = \Sigma \cup \{\triangleright, \triangleleft\}$, respectively. The special symbol $\triangleright$ denotes the beginning of each of the tapes, the symbol $\triangleleft$ denotes the end of the ℓ input tapes. The transition function Δ is defined as usual, i.e., $\Delta : Q \times \Gamma^{\ell+2} \to Q \times \Gamma^2 \times \{\leftarrow, \sqcup, \rightarrow\}^{\ell+2}$ such that $\Delta(q, (a_1, \ldots, a_\ell, b, c)) = \left(q', (b', c'), (\mathsf{d}_1, \ldots, \mathsf{d}_{\ell+2})\right)$ with $\mathsf{d}_i \in \{\leftarrow, \sqcup, \rightarrow\}$, means that when T is in state q and the $\ell + 2$ heads on the tapes read symbols $a_1, \ldots, a_\ell, b, c$, respectively, then T transitions to state q', writes b', c' on the work and output tapes, respectively, at the position to which the work and output tapes' heads points at, and finally moves the heads on the tapes according $\mathsf{d}_1, \ldots, \mathsf{d}_{\ell+2}$. More specifically, $\leftarrow$ indicates a move to the left, $\rightarrow$ a move to the right, and finally, $\sqcup$ indicates that the head does not move.

We assume that Δ is defined such that it ensures that on none of the tapes, heads can move beyond the leftmost marker $\triangleright$. Furthermore, the tapes $R_1, \ldots, R_\ell$ are treated as read-only and the heads on these tapes cannot move beyond the end markers $\triangleleft$. Similarly, Δ ensures that the output tape O is write only, i.e., its head cannot move to the left. We also assume that Δ does not change the occurrences of $\triangleright$ or writes $\triangleleft$ on the work and output tape.

A configuration of T is defined in the usual way. That is, a configuration of the input tapes is of the form $\triangleright w_1 q w_2 \triangleleft$ with $w_1, w_2 \in \Sigma^*$ and represents that the current tape content is $\triangleright w_1 w_2 \triangleleft$, T is in state q and the head is positioned on the first symbol of w_2. Similarly, configurations of the work and output tape are represented by $\triangleright w_1 q w_2$. A configuration of T is consists of configurations for all tapes. Given two configurations c_1 and c_2, we say that c_1 yields c_2 if c_2 is the result of applying the transition function Δ of T based on the information in c_1. As usual, we close this "yields" relation transitively.

Given ℓ input words $w_1, \ldots, w_\ell \in \Sigma^*$, we assume that the initial configuration of T is given by $\left(q_0 \triangleright w_1 \triangleleft, q_0 \triangleright w_2 \triangleleft, \ldots, q_0 \triangleright w_\ell \triangleleft, q_0 \triangleright, q_0 \triangleright\right)$ and an accepting configuration is assumed to be of the form $\left(\triangleright q_m w_1 \triangleleft, \triangleright q_m w_2 \triangleleft, \ldots, \triangleright q_m w_\ell \triangleleft, \triangleright q_m, \triangleright q_m w\right)$ for some $w \in \Sigma^*$. We say that T computes the function $f : (\Sigma^*)^\ell \to \Sigma^*$ if for every $w_1, \ldots, w_\ell \in \Sigma^*$, the initial configuration yields (transitively) an accepting configuration such that the configuration on the output tape is given by $\triangleright q_m f(w_1, \ldots, w_\ell)$.

We assume that once T reaches an accepting configuration it stays indefinitely in that configuration (i.e., it loops). We further assume that T only reaches an accepting configuration when all its input words have the same size. Furthermore, when all inputs have the same size, T will reach an accepting configuration.

We say that T is a *linear space machine* when it reaches an accepting configuration on inputs of size n by using $O(n)$ space on its work tape and additionally needs $O(n^k)$ steps to do so. A *linear input-output function* is a function of the form $f = \bigcup_{n \geq 0} f_n : (\Sigma^n)^\ell \to \Sigma^n$. In other words, for every ℓ words of the same size n, f returns a word of size n. We say that a linear input-output function is a *linear space input-output function* if there exists a linear space machine T that for every $n \geq 0$, on input $w_1, \ldots, w_\ell \in \Sigma^n$ the TM T has $f_n(w_1, \ldots, w_\ell)$ on its the output tape when (necessarily) reaching an accepting configuration.

Proposition D.1. *Let $f = \bigcup_{n\geq 0} f_n : (\Sigma^n)^\ell \to \Sigma^n$ be a linear space input-ouput function computed by a linear space machine T with m states, ℓ input tapes, which consumes $O(n)$ space and runs in $O(n^{k-1})$ time on inputs of size n. There exists (i) a* MATLANG *schema $\mathcal{S} = (\mathcal{M}, size)$ where $\mathcal{M}$ consists matrix variables[4] $Q_1, \ldots, Q_m, R_1, \ldots, R_\ell, H_1, \ldots, H_\ell, W_1, \ldots, W_s, H_{W_1}, \ldots, H_{W_s}, O, H_O, v_1, \ldots, v_k$ with $\mathsf{size}(V) = \alpha \times 1$ for all $V \in \mathcal{M}$; and (ii) a* MATLANG *expression e_f over $\mathcal{S}$ such that for the instance $\mathcal{I} = (\mathcal{D}, mat)$ over $\mathcal{S}$ with $\mathcal{D}(\alpha) = n$ and*

$$\mathsf{mat}(R_i) = \mathsf{vec}(w_i) \in \mathbb{R}^n, \text{ for } i \in [\ell] \text{ and all other matrix variables instantiated with the zero vector in } \mathbb{R}^n$$

for words $w_1, \ldots, w_\ell \in \Sigma^n$ and such that $\mathsf{vec}(w_i)$ is the $n \times 1$-vector encoding the word w_i, we have that $\mathsf{mat}(O) = \mathsf{vec}(f_n(w_1, \ldots, w_n)) \in \mathbb{R}^n$ after evaluating e_f on $\mathcal{I}$.

Proof. The expression e_f we construct will simulate the TM T. To have some more control on the space and time consumption of T, let us first assume that n is large enough, say larger than $n \geq N$, such that T runs in sn space and $cn^{k-1} \leq n^k$ time for constants s and c. We deal with $n < N$ later on.

To simulate T we need to encode states, tapes and head positions. The matrix variables in $\mathcal{M}$ mentioned in the proposition will take these roles. More specifically, the variables $R_1, \ldots, R_\ell$ will hold the input vectors, $W_1, \ldots, W_s$ will hold the contents of the work tape, where s is the constant mentioned earlier, and O will hold the contents of the output tape. The vectors corresponding to the work and output tape are initially set to the zero vector. The vector for the input tape R_i is set to $\mathsf{vec}(w_i)$, for $i \in [\ell]$.

With each tape we associate a matrix variable encoding the position of the head. More specifically, $H_1, \ldots, H_\ell$ correspond to the input tape heads, $H_{W_1}, \ldots, H_{W_s}$ are the heads for the work tape, and H_O is the head of the output tape. All these vectors are initialised with the zero vector. Later on, these vectors will be zero except for a single position, indicating the positions in the corresponding tapes the heads point to. For those positions j, $1 < j < n$, the head vectors will carry value 1. When $j = 1$ or n and when it concerns positions for the input tape, the head vectors can carry value 1 or 2. We need to treat these border cases separately because we only have n positions available to store the input words, whereas the actual input tapes consist of $n + 2$ symbols because of $\triangleright$ and $\triangleleft$. So when, for example, H_1 has a 1 in its first entry, we interpret it as the head is pointing to the first symbol of the input word w_1. When H_1 has a 2 in its first position, we interpret it as the head pointing to $\triangleright$. The end marker $\triangleleft$ is dealt with in the same way, by using value 1 or 2 in the last position of H_1. We use this encoding for all input tapes, and also for the work tape W_1 and output tape O with the exception that no end marker $\triangleleft$ is present.

To encode the states, we use the variables $Q_1, \ldots, Q_m$. We will ensure that when T is in state q_i when $\mathsf{mat}(Q_i) = [1, 0, \ldots, 0]^T \in \mathbb{R}^n$, otherwise $\mathsf{mat}(Q_i)$ is the zero vector in $\mathbb{R}^n$.

Finally, the variables $v_1, \ldots, v_k$ represent k canonical vectors which are used to iterate in for-loops. By iterating over then, we can perform n^k iterations, which suffices for simulating the $O(n^{k-1})$ steps used by T to reach an accepting configuration.

With these matrix variables in place, we start by defining e_f. It will consists of two subexpressions $e_f^{\geq N}$, for dealing with $n \geq N$, and $e_f^{<N}$, for dealing with $n < N$. We explain the expression $e_f^{\geq N}$ first.

In our expressions we use subexpressions which we defined before in section B.1. These subexpression require some auxiliary variables, as detailed below. As a consequence, e_f will be an expressions defined over an extended schema $\mathcal{S}'$. Hence, the instance $\mathcal{I}$ in the statement of the Proposition is an instance $\mathcal{I}'$ of $\mathcal{S}'$ which coincides with $\mathcal{I}$ on $\mathcal{S}$ and in which the auxiliary matrix variables are all instantiated with zero vectors or matrices, depending on their size.

Now, we specify the finite auxiliary variables involved in the `for-MATLANG` expression. These arise when computing the following `for-MATLANG` expressions defined

- $e_{\mathsf{Prev}}(z, Z, z', Z')$, and expression over auxiliary variables z, z', Z and Z' with $\mathsf{size}(z) = \mathsf{size}(z') = \mathsf{size}(Z) = \alpha \times 1$ and $\mathsf{size}(Z') = \alpha \times \alpha$. On input $\mathcal{I}'$ with $\mathsf{mat}(z) = \mathsf{mat}(z') = \mathsf{mat}(Z)$ the zero column vector of dimension n, and $\mathsf{mat}(Z')$ the zero $n \times n$ matrix, $[\![e_{\mathsf{Prev}}]\!](\mathcal{I}')$ returns the $n \times n$ matrix Prev such that
$$\mathsf{Prev} \cdot \mathrm{e}_i := \begin{cases} \mathrm{e}_{i-1} & \text{if } i > 1 \\ 0 & \text{if } i = 1. \end{cases}$$
- $e_{\mathsf{Next}}(z, Z, z', Z')$, and expression over auxiliary variables z, z', Z and Z' with $\mathsf{size}(z) = \mathsf{size}(z') = \mathsf{size}(Z) = \alpha \times 1$ and $\mathsf{size}(Z') = \alpha \times \alpha$. On input $\mathcal{I}'$ with $\mathsf{mat}(z) = \mathsf{mat}(z') = \mathsf{mat}(Z)$ the zero column vector of dimension n, and $\mathsf{mat}(Z')$ the zero $n \times n$ matrix, $[\![e_{\mathsf{Next}}]\!](\mathcal{I}')$ returns the $n \times n$ matrix Next such that
$$\mathsf{Next} \cdot \mathrm{e}_i := \begin{cases} \mathrm{e}_{i+1} & \text{if } i < n \\ 0 & \text{if } i = n. \end{cases}$$
- $\mathsf{min}(v, z, Z, z', Z)$ with auxiliary variables z, z', Z and Z' as before, and v is one of the (vector) variables in $\mathcal{M}$. For an $n \times 1$ vector v, on input $\mathcal{I}'[v \leftarrow \mathrm{v}]$
$$[\![\mathsf{min}]\!](\mathcal{I}'[v \leftarrow \mathrm{v}]) := \begin{cases} 1 & \text{if } \mathrm{v} = \mathrm{e}_1 \\ 0 & \text{otherwise.} \end{cases}$$

[4] We also need a finite number of auxiliary variables, these will be specified in the proof.

- $\mathsf{max}(v, z, Z, z', Z)$ with auxiliary variables z, z', Z and Z' as before, and and v is one of the (vector) variables in $\mathcal{M}$. For an $n \times 1$ vector v, on input $\mathcal{I}'[v \leftarrow \mathrm{v}]$

$$[\![\mathsf{max}]\!](\mathcal{I}'[v \leftarrow \mathrm{v}]) := \begin{cases} 1 & \text{if } \mathrm{v} = \mathrm{e}_n \\ 0 & \text{otherwise.} \end{cases}$$

- $e_{\mathsf{min}}(z, Z, z', Z', z'', Z'')$, an expressions with auxiliary variables z, z', z'', Z, Z' and Z'' with $\mathsf{size}(z) = \mathsf{size}(z') = \mathsf{size}(z'') = \mathsf{size}(Z) = \mathsf{size}(Z'') = \alpha \times 1$ and $\mathsf{size}(Z') = \alpha \times \alpha$. On input $\mathcal{I}'$ with matrix variables instantiated with zero vectors (or matrix for Z'), $[\![e_{\mathsf{min}}]\!](\mathcal{I}') = \mathrm{e}_1$.
- $e_{\mathsf{max}}(z, Z, z', Z', z'', Z'')$, an expressions with auxiliary variables z, z', z'', Z, Z' and Z'' with $\mathsf{size}(z) = \mathsf{size}(z') = \mathsf{size}(z'') = \mathsf{size}(Z) = \mathsf{size}(Z'') = \alpha \times 1$ and $\mathsf{size}(Z') = \alpha \times \alpha$. On input $\mathcal{I}'$ with matrix variables instantiated with zero vectors (or matrix for Z'), $[\![e_{\mathsf{max}}]\!](\mathcal{I}') = \mathrm{e}_n$.

We thus see that we only need z, z', z'', Z, Z', Z'' as auxiliary variables and these can be re-used whenever e_f calls these functions. From now one, we omit the auxiliary variables from the description of e_f.

Let us first define $e_f^{\geq N}$. Since we want to simulate T we need to be able to check which transitions of T can be applied based on a current configuration. More precisely, suppose that we want to check whether $\delta(q_i, (a_1, \ldots, a_\ell, b, c))$ is applicable, then we need to check whether T is in state q_i, we can do this by checking $\mathsf{min}(Q_i)$, and whether the heads on the tapes read symbols $a_1, \ldots, a_\ell, b, c$. We check the latter by the following expressions. For the input tape R_i we define

$$\mathsf{test\_inp}_b^i := \begin{cases} (1 - \mathsf{min}(1/2 \cdot H_i)) \cdot (1 - \mathsf{max}(1/2 \cdot H_i)) \cdot (1 - R_i^T \cdot H_i) & \text{if } b = 0 \\ (1 - \mathsf{min}(1/2 \cdot H_i)) \cdot (1 - \mathsf{max}(1/2 \cdot H_i)) \cdot (R_i^T \cdot H_i) & \text{if } b = 1 \\ \mathsf{min}(1/2 \cdot H_i) & \text{if } b = \triangleright \\ \mathsf{max}(1/2 \cdot H_i) & \text{if } b = \triangleleft, \end{cases}$$

which returns 1 if and only if either $b \in \{0, 1\}$ is the value in $\mathsf{mat}(R_i)$ at the position encoded by $\mathsf{mat}(H_i)$, or when $b = \triangleright$ and $\mathsf{mat}(H_i)$ is the vector $(2, 0, \ldots, 0) \in \mathbb{R}^n$, or when $b = \triangleleft$ and $\mathsf{mat}(H_i)$ is the vector $(0, 0, \ldots, 2) \in \mathbb{R}^n$. Similarly, for the output tape we define

$$\mathsf{test\_out}_b := \begin{cases} (1 - \mathsf{min}(1/2 \cdot H_O)) \cdot (1 - O^T \cdot H_O) & \text{if } b = 0 \\ (1 - \mathsf{min}(1/2 \cdot H_O)) \cdot (O^T \cdot H_O) & \text{if } b = 1 \\ \mathsf{min}(1/2 \cdot H_O) & \text{if } b = \triangleright, \end{cases}$$

and for the work tapes $W_1, \ldots, W_s$ we define

$$\mathsf{test\_work}_b^i := \begin{cases} (1 - \mathsf{min}(1/2 \cdot H_{W_i})) \cdot (1 - W_i^T \cdot H_{W_i})) & \text{if } b = 0 \\ (1 - \mathsf{min}(1/2 \cdot H_{W_i})) \cdot (W_i^T \cdot H_{W_i}) & \text{if } b = 1 \\ \mathsf{min}(1/2 \cdot H_{W_i}) & \text{if } b = \triangleright \text{ and } i = 1. \end{cases}$$

We then combine all these expressions into a single expression for $q_i \in Q$, $a_1, \ldots, a_\ell, b, c \in \Gamma$:

$$\mathsf{isconf}_{q_i, a_1, \ldots, a_\ell, b, c} := \mathsf{min}(Q_i) \cdot \left(\prod_{j=1}^{\ell} \mathsf{test\_inp}_{a_j}^j \right) \cdot \left(\sum_{j=1}^{s} \mathsf{test\_work}_b^j \right) \cdot \mathsf{test\_out}_c.$$

This expression will return 1 if and only if the vectors representing the tapes, head positions and states are such that $\mathsf{Q_i}$ is the first canonical vector (and thus T is in state q_i), the heads point to entries in the tape vectors storing the symbols $a_1, \ldots, a_\ell, b, c$ or they point to the first (or last for input tapes) positions but have value 2 (when the symbols are $\triangleright$ or $\triangleleft$).

To ensure that at the beginning of the simulation of T by $e_f^{\geq N}$ we correctly encode that we are in the initial configuration, we thus need to initialise all vectors $\mathsf{mat}(H_1), \mathsf{mat}(H_2), \ldots, \mathsf{mat}(H_\ell), \mathsf{mat}(H_{W_1}), \mathsf{mat}(H_O)$ with the vector $(2, 0, 0, \ldots, 0) \in \mathbb{R}$ since all heads read the symbol $\triangleright$. Similarly, we have to initialise $\mathsf{Q_1}$ with the first canonical vector since T is in state q_0.

We furthermore need to be able to correctly adjust head positions. We do this by means of the predecessor and successor expressions described above. A consequence of our encoding is that we need to treat the border cases (corresponding to $\triangleright$ and $\triangleleft$) differently. More specifically, for the input tapes R_i and heads H_i we define

$$\mathsf{move\_inp}_{\mathsf{d}}^i := \begin{cases} 2 \times \mathsf{min}(H_i) \times H_i + 1/2 \times \mathsf{max}(1/2 \times H_i) \times H_i + (1 - \mathsf{min}(H_i))(1 - \mathsf{max}(1/2 \times H_i)) \times e_{\mathsf{Prev}} \cdot H_i & \text{if } \mathsf{d} = \leftarrow \\ 2 \times \mathsf{max}(H_i) \times H_i + 1/2 \times \mathsf{min}(1/2 \times H_i) \times H_i + (1 - \mathsf{min}(1/2 \times H_i))(1 - \mathsf{max}(H_i)) \times e_{\mathsf{Next}} \cdot H_i & \text{if } \mathsf{d} = \rightarrow \\ H_i & \text{if } \mathsf{d} = \sqcup. \end{cases}$$

In other words, we shift to the previous (or next) canonical vector when d is $\leftarrow$ or $\rightarrow$, respectively, unless we need to move to or from the position that will hold $\triangleright$ or $\triangleleft$. In those case we readjust $\mathsf{mat}(H_i)$ (which will either $(1, 0, \ldots, 0)$, $(2, 0, \ldots, 0)$, $(0, \ldots, 0, 1)$ or $(0, \ldots, 0, 2)$) by

either dividing or multiplying with 2. In this way we can correctly infer whether or not the head points to the begin and end markers. For the output tape we proceed in a similar way, but only taking into account the begin marker and recall that we do not have moves to the left:

$$\mathsf{move\_outp}_{\mathsf{d}} := \begin{cases} 1/2 \times \mathsf{min}(1/2 \times H_O) \times H_O + (1 - \mathsf{min}(1/2 \times H_O)) \times e_{\mathsf{Next}} \times H_O & \text{if } \mathsf{d} = \rightarrow \\ H_O & \text{if } \mathsf{d} = \sqcup. \end{cases}$$

Since we represent the work tape by s vectors $W_1, \ldots, W_s$ we need to ensure that only one of the head vectors H_{W_i} has a non-zero value and that by moving left or right, we need to appropriately update the right head vector. We do this as follows. We first consider the work tapes W_i for $i \neq 1, s$ and define

$$\mathsf{move\_work}^i_{\mathsf{d}} := \begin{cases} -\mathsf{min}(H_{W_i}) \times H_{W_i} + (1 - \mathsf{min}(H_{W_i})) \times e_{\mathsf{Prev}} \cdot H_{W_i} + \mathsf{min}(H_{W_{i+1}}) \times e_{\mathsf{max}} & \text{if } \mathsf{d} = \leftarrow \\ -\mathsf{max}(H_{W_i}) \times H_{W_i} + (1 - \mathsf{max}(H_{W_i})) \times e_{\mathsf{Next}} \cdot H_{W_i} + \mathsf{max}(H_{W_{i-1}}) \times e_{\mathsf{min}} & \text{if } \mathsf{d} = \rightarrow \\ H_{W_i} & \text{if } \mathsf{d} = \sqcup. \end{cases}$$

In other words, we set the H_{W_i} to zero when a move brings us to either W_{i-1} or W_{i+1}, we move the successor or predecessor when staying within W_i, or initialise H_{W_i} with the first or last canonical vector when moving from W_{i-1} to W_i (right move) or from W_{i+1} to W_i (left move). For $i = s$ we can ignore the parts in the previous expression that involve W_{s+1} (which does not exist):

$$\mathsf{move\_work}^s_{\mathsf{d}} := \begin{cases} -\mathsf{min}(H_{W_s}) \times H_{W_i} + (1 - \mathsf{min}(H_{W_s})) \times e_{\mathsf{Prev}} \cdot H_{W_s} & \text{if } \mathsf{d} = \leftarrow \\ -\mathsf{max}(H_{W_s}) \times H_{W_s} + (1 - \mathsf{max}(H_{W_s})) \times e_{\mathsf{Next}} \cdot H_{W_s} + \mathsf{max}(H_{W_{s-1}}) \times e_{\mathsf{min}} & \text{if } \mathsf{d} = \rightarrow \\ H_{W_s} & \text{if } \mathsf{d} = \sqcup. \end{cases}$$

For $i = 1$, we can ignore the part involving W_0 (which does not exist) but have to take $\triangleright$ into account:

$$\mathsf{move\_work}^1_{\mathsf{d}} := \begin{cases} 2 \times \mathsf{min}(H_{W_1}) \times H_{W_i} + (1 - \mathsf{min}(H_{W_1})) \times e_{\mathsf{Prev}} \cdot H_{W_1} + \mathsf{min}(H_{W_2}) \times e_{\mathsf{max}} & \text{if } \mathsf{d} = \leftarrow \\ 1/2 \times \mathsf{min}(1/2 \times H_{W_1}) \times H_{W_1} + (1 - \mathsf{max}(1/2 \times H_{W_1})) \times e_{\mathsf{Next}} \cdot H_{W_1} & \text{if } \mathsf{d} = \rightarrow \\ H_{W_1} & \text{if } \mathsf{d} = \sqcup. \end{cases}$$

A final ingredient for defining $e_f^{\geq N}$ are expressions which update the work and output tape. To define these expression, we need the position and symbol to put on the tape. For the output tape we define

$$\mathsf{write\_outp}_b := \begin{cases} \mathsf{min}(1/2 \times H_O) \times O & \text{if } b = \triangleright \\ (1 - \mathsf{min}(1/2 \times H_O)) \times \left((1 - O^T \cdot H_O) \times O + (O^T \cdot H_O) \times (O - H_O) \right) & \text{if } b = 0 \\ (1 - \mathsf{min}(1/2 \times H_O)) \times \left((1 - O^T \cdot H_O) \times (O + H_O) + (O^T \cdot H_O) \times O \right) & \text{if } b = 1 \end{cases}$$

and similarly for the work tapes $i \neq 1$:

$$\mathsf{write\_work}^i_b := \begin{cases} W_i & \text{if } b = \triangleright \\ (1 - W_i^T \cdot H_{W_i}) \times W_i + (W_i^T \cdot H_{W_i}) \times (W_i - H_{W_i}) & \text{if } b = 0 \\ (1 - W_i^T \cdot H_{W_i}) \times (W_i + H_{W_i}) + (W_i^T \cdot H_{W_i}) \times W_i & \text{if } b = 1, \end{cases}$$

and for W_1 we have to take care again of the begin marker:

$$\mathsf{write\_work}^1_b := \begin{cases} \mathsf{min}(1/2 \times H_{W_1}) \times W_1 & \text{if } b = \triangleright \\ (1 - \mathsf{min}(1/2 \times H_{W_1}) \times \left((1 - W_1^T \cdot H_{W_1}) \times W_1 + (W_1^T \cdot H_{W_1}) \times (W_1 - H_{W_1}) \right) & \text{if } b = 0 \\ (1 - \mathsf{min}(1/2 \times H_{W_1}) \times \left((1 - W_1^T \cdot H_{W_1}) \times (W_1 + H_{W_1}) + (W_1^T \cdot H_{W_1}) \times W_1 \right) & \text{if } b = 1. \end{cases}$$

We are now finally ready to define $e_f^{\geq N}$:

$$e_f^{\geq N} := \mathsf{for}\, v_1, \ldots, v_k, Q_1, \ldots, Q_m, H_1, \ldots, H_\ell, W_1, \ldots, W_s, H_{W_1}, \ldots, H_{W_s}, O, H_O. \\ (e_{Q_1}, \ldots, e_{Q_m}, e_{H_1}, \ldots, e_{H_\ell}, e_{W_1}, \ldots, e_{W_s}, e_{H_{W_1}}, \ldots, e_{H_{W_s}}, e_O, e_{H_O}).$$

We refer to section A.1 for the definition of this form of the for-loop. The expressions used are (we use $\star$ below to mark irrelevant information in the transitions):

$$e_{Q_1} := \left(\prod_{j=1}^{k} \mathsf{min}(v_i)\right) \times e_{\mathsf{min}} + \sum_{\substack{(q_i,a_1,\ldots,a_\ell,b,c) \\ \Delta(q_i,a_1,\ldots,a_\ell,b,c)=(q_1,\star)}} \mathsf{isconf}_{q_i,a_1,\ldots,a_\ell,b,c} \times e_{\mathsf{min}}$$

$$e_{Q_j} := \sum_{\substack{(q_i,a_1,\ldots,a_\ell,b,c) \\ \Delta(q_i,a_1,\ldots,a_\ell,b,c)=(q_j,\star)}} \mathsf{isconf}_{q_i,a_1,\ldots,a_\ell,b,c} \times e_{\mathsf{min}} \quad \text{for } j \neq 1$$

$$e_{H_i} := 2\left(\prod_{j=1}^{k} \mathsf{min}(v_i)\right) \times e_{\mathsf{min}} + \sum_{\substack{(q,a_1,\ldots,a_\ell,b,d) \\ \Delta(q,a_1,\ldots,a_\ell,b,c)=(\star,\mathsf{d_i},\star)}} \mathsf{isconf}_{q,a_1,\ldots,a_\ell,b,c} \times \mathsf{move\_inp}^i_{\mathsf{d}_i}$$

$$e_{H_{W_i}} := 2\left(\prod_{j=1}^{k} \mathsf{min}(v_i)\right) \times e_{\mathsf{min}} + \sum_{\substack{(q,a_1,\ldots,a_\ell,b,d) \\ \Delta(q,a_1,\ldots,a_\ell,b,c)=(\star,\mathsf{d}_{\ell+1},\star)}} \mathsf{isconf}_{q,a_1,\ldots,a_\ell,b,c} \times \mathsf{move\_work}^i_{\mathsf{d}_{\ell+1}}$$

$$e_{H_O} := 2\left(\prod_{j=1}^{k} \mathsf{min}(v_i)\right) \times e_{\mathsf{min}} + \sum_{\substack{(q,a_1,\ldots,a_\ell,b,d) \\ \Delta(q,a_1,\ldots,a_\ell,b,c)=(\star,\mathsf{d}_{\ell+2})}} \mathsf{isconf}_{q,a_1,\ldots,a_\ell,b,c} \times \mathsf{move\_outp}_{\mathsf{d}_{\ell+2}}$$

$$e_{W_i} := \sum_{\substack{(q,a_1,\ldots,a_\ell,b,d) \\ \Delta(q,a_1,\ldots,a_\ell,b,c)=(\star,b',c',\star)}} \mathsf{isconf}_{q,a_1,\ldots,a_\ell,b,c} \times \mathsf{write\_work}^i_{b'}$$

$$e_O := \sum_{\substack{(q,a_1,\ldots,a_\ell,b,d) \\ \Delta(q,a_1,\ldots,a_\ell,b,c)=(\star,b',c',\star)}} \mathsf{isconf}_{q,a_1,\ldots,a_\ell,b,c} \times \mathsf{write\_outp}_{c'}.$$

The correctness of $e_f^{\geq N}$ should be clear from the construction (one can formally verify this by induction on the number of iterations). We next explain how the border cases $n < N$ can be dealt with. For each $n < N$ and every possible input words $w_1, \ldots, w_\ell$ of size n, we define a for-MATLANG expression which checks whether $\mathsf{mat}(R_i) = \mathsf{vec}(w_i)$ for each $i \in [\ell]$. This can be easily done since n can be regarded as a constant. For example, to check whether $\mathsf{mat}(R_i) = [0,1,1]^T$ we simply write

$$(1 - R_i^T \cdot e_{\mathsf{min}}) \times (R_i^T \cdot e_{\mathsf{Next}} \cdot e_{\mathsf{min}}) \times (1 - R_i^T \cdot e_{\mathsf{Next}} \cdot e_{\mathsf{Next}} \cdot e_{\mathsf{min}}) \times (1 - e_1(R_i)^T \cdot e_{\mathsf{Next}} \cdot e_{\mathsf{Next}} \cdot e_{\mathsf{Next}} \cdot e_{\mathsf{min}})$$

which will evaluate to 1 if and only if $\mathsf{mat}(R_i) = [0,1,1]^T$. We note that the final factor is in place to check that the dimension of $\mathsf{mat}(R_i)$ is three. We denote by $e^i_{n,w}$ the expression which evaluates to 1 if and only if $\mathsf{mat}(R_i) = \mathsf{vec}(w)$ for $|w| = n$. We can similarly write any word w of fixed size in the matrix variable O. For example, suppose that $w = 101$ then we write

$$O + e_{\mathsf{min}} + e_{\mathsf{Next}} \cdot e_{\mathsf{Next}} \cdot e_{\mathsf{min}}.$$

We write $e_{n,w}$ be the expression which write w of size $|w| = n$ in matrix variable O. Then, the expressions

$$e_{n,w_1,\ldots,w_n,w} := e^1_{n,w_1} \cdot \cdots \cdot e^\ell_{n w_\ell} \cdot e_{n,w}$$

will write w in O if and only if $\mathsf{mat}(R_i) = \mathsf{vec}(w_i)$ for $i \in [\ell]$. We now simply take the disjunction over all words $w_1, \ldots, w_\ell \in \Sigma^n$ and $w = f_n(w_1, \ldots, w_\ell) \in \Sigma^n$:

$$e_n := \sum_{w_1,\ldots,w_\ell \in \Sigma^n} e_{n,w_1,\ldots,w_\ell,f_n(w_1,\ldots,w_\ell)},$$

which correctly evaluates f_n. We next take a further disjunction by letting ranging from $n = 0, \ldots, N-1$:

$$e_f^{<N} := \sum_{n=0}^{N-1} e_n$$

Since every possible input is covered and only a unique expression $e_{n,w_1,\ldots,w_\ell,f_n(w_1,\ldots,w_\ell)}$ will be triggered $e_f^{<N}$ will correctly evaluate f on inputs smaller than N.

Our final expression e_f is now given by

$$e_f := e_f^{<N} + \mathsf{dim\_is\_greater\_than}_{\mathsf{N}} \times e_f^{\geq N}$$

where $\mathsf{dim\_is\_greater\_than_N}$ is the expression $e_1(R_i)^T \cdot \underbrace{e_{\mathsf{Next}} \cdot \cdots \cdot e_{\mathsf{Next}}}_{N \text{ times}}$ which will evaluate to 1 if an only if the input dimension is larger or equal than N. □

D.2 Circuit evaluation

We prove theorem 5.1:

Theorem 5.1. *For any uniform arithmetic circuit family $\{\Phi_n \mid n = 1, 2, \ldots\}$ of logarithmic depth there is a* for-MATLANG *schema $\mathcal{S}$ and an expression e_Φ using a matrix variable v, with $\mathtt{type}_\mathcal{S}(v) = (\alpha, 1)$ and $\mathtt{type}_\mathcal{S}(e) = (1, 1)$, such that for any input $a_1, \ldots, a_n$ to the circuit Φ_n:*

- *If $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$ is a* MATLANG *instance such that $\mathcal{D}(\alpha) = n$ and $\mathtt{mat}(v) = [a_1 \ldots a_n]^T$*
- *Then $[\![e]\!](\mathcal{I}) = \Phi_n(a_1, \ldots, a_n)$.*

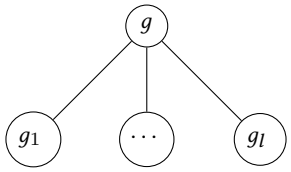
Proof. Let Φ_n be a circuit with n input gates and such that it can be computed by a $L-$uniform arithmetic circuit of log-depth. Each gate of the circuit that encodes f has an $\mathtt{id} \in \{0,1\}^n$. From now on, when we write g for a gate of the circuit, we mean the $\mathtt{id}$ encoding g. Let n^k be a polynomial such that the number of wires $W(n) \leq n^k$ for n big enough. Further, we assume that $2W(n) \leq n^k$. We need this because the for-matlang simulation of the circuit is in a depth first search way, so $2W(n)$ wires will be traversed. Then we have that:

- the number of gates is bounded by n^k.
- we need at most $k \log(n)$ bits to store the *id* of a gate.
- the depth of the circuit is at most $k' \log(n)$ for some k'.

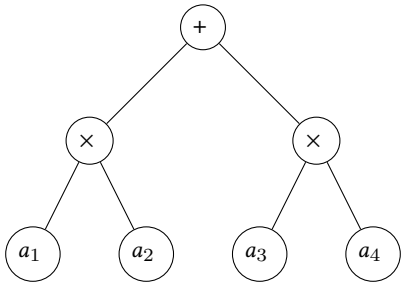
So, let n_0 and k such that $\forall n \geq n_0$:

$$\left.\begin{aligned} 2W(n) &\leq n^k \\ k\lceil \log(n) \rceil &\leq n - 3 \\ k'\lceil \log(n) \rceil &\leq n \end{aligned}\right\} (\star)$$

We know n_0 and k exist. Let $n \geq n_0$. Towards the end, we will deal with the case when $n < n_0$.
Let g be a gate. The children of g are denoted by $g_1, \ldots, g_l$.



For example, a circuit that encodes the function $f(a_1, a_2, a_3, a_4) = a_1a_2 + a_3a_4$ is



We can simulate the polynomial $x^2 + xy$ by doing $f(A)$ where $A = [x\ x\ x\ y]^T$. The main idea is to traverse the circuit top down in a depth first search way and store visited gates in a stack and its corresponding current values in another stack, and aggregate in the iterations according to the gate type.

For a stack S, the operations are standard:

- $S.\mathtt{push}(s)$: pushes s into S.
- $S.\mathtt{pop}$: pops the top element.
- $S.\mathtt{size}$: the length of the stack.
- $S.\mathtt{top}$: the top element in the stack.

For the pseudo-code, $\mathcal{G}$ and $\mathcal{V}$ denote stacks of gates and values, respectively. The property that holds during the simulation is that the value in $\mathcal{V}[i]$ is the value that $\mathcal{G}[i]$ currently outputs. The algorithm ends with $\mathcal{G} = [g_{\texttt{root}}]$ and $\mathcal{V} = [v_{\texttt{root}}]$ after traversing the circuit, and returns $v_{\texttt{root}}$

During the evaluation algorithm there will be two possible configurations of $\mathcal{G}$ and $\mathcal{V}$.

(1) $\mathcal{G}.\texttt{size} = \mathcal{V}.\texttt{size} + 1$: this means that $\mathcal{G}.\texttt{top}$ is a gate that we visit for the first time and we need to initialize its value.
(2) $\mathcal{G}.\texttt{size} = \mathcal{V}.\texttt{size}$: here $\mathcal{V}.\texttt{top}$ is the value of evaluating the circuit in gate $\mathcal{G}.\texttt{top}$. Therefore, we need to aggregate the value $\mathcal{V}.\texttt{top}$ to the parent gate of g.

We assume the circuit has input gates, $+, \times$-gates and allow constant 1-gate.

The idea is to traverse the circuit top down in a depth first search way. For example, in the circuit $f(a_1, a_2, a_3, a_4) = a_1a_2 + a_3a_4$ above, we would initialize the output gate value as 0 because it is a + gate, so $\mathcal{G} = \{+\}$, $\mathcal{V} = \{0\}$. Then stack the left $\times$ gate to $\mathcal{G}$, stack its initial value (i.e. 1) to $\mathcal{V}$. Now stack a_1 to $\mathcal{G}$ and its value (i.e. a_1) to $\mathcal{V}$. Since we are on an input gate we pop the gate and value pair off of $\mathcal{G}$ and $\mathcal{V}$ respectively, aggregate a_1 to $\mathcal{V}.\texttt{top}$ and continue by stacking the a_2 gate to $\mathcal{G}$. We pop a_2 off of $\mathcal{V}$ (and its gate off of $\mathcal{G}$) and aggregate its value to $\mathcal{V}.\texttt{top}$. We pop and aggregate the value of the left $\times$ gate to $\mathcal{V}.\texttt{top}$ (the root value). Then continue with the right $\times$ gate branch similarly.

For the pseudo-code, we supply ourselves with the following functions:

- $\texttt{isplus}(g)$: true if and only if g is a +-gate.
- $\texttt{isprod}(g)$: true if and only if g is a $\times$-gate.
- $\texttt{isone}(g)$: true if and only if g is a 1-gate.
- $\texttt{isinput}(g)$: true if and only if g is an input gate.
- $\texttt{getfirst}(g)$: outputs the first child of g.
- $\texttt{getinput}(g)$: outputs $A[i]$ when g is the i-th input.
- $\texttt{not\_last}(g_1, g_2)$: true if and only if g_2 is not the last child gate of g_1.
- $\texttt{next\_gate}(g_1, g_2)$: outputs the next child gate of g_1 after g_2.
- $\texttt{getroot}()$: outputs the root gate of the circuit.

The corresponding $\{0,1\}^n \to \{0,1\}^n$ functions are:

- $\texttt{isplus}(g)$: 1 if and only if g is a +-gate.
- $\texttt{isprod}(g)$: 1 if and only if g is a $\times$-gate.
- $\texttt{isone}(g)$: 1 if and only if g is a 1-gate.
- $\texttt{isinput}(g)$: 1 if and only if g is an input gate.
- $\texttt{getfirst}(g)$: outputs the $\texttt{id}$ of the first child of g.
- $\texttt{getinput}(g)$: outputs canonical vector b_i, where the i-th input gate of Φ_n is encoded by g.
- $\texttt{not\_last}(g_1, g_2)$: 1 if and only if g_2 is not the last child gate of g_1.
- $\texttt{next\_gate}(g_1, g_2)$: outputs the $\texttt{id}$ of the next child gate of g_1 after g_2.
- $\texttt{getroot}()$: outputs the $\texttt{id}$ of the root gate of the circuit.

The previous functions are all definable by an L-transducer and can be defined from the L-transducer of f. Then, by proposition D.1, for each of these functions there is a for-MATLANG expression that simulates them.

Now, we give the pseudo-code of the top-down evaluation. We define the functions *Initialize* (algorithm 1), *Aggregate* (algorithm 2) and *Evaluate* (algorithm 3). The main algorithm is *Evaluate*.

Algorithm 1 Initialize (pseudo-code)

```
1:  function Initialize(G, V, A)          ▷ The stacks and input. Here, G.size = V.size + 1
2:      if isplus(G.top) then
3:          V.push(0)
4:          G.push(getfirst(G.top))
5:      else if isprod(G.top) then
6:          V.push(1)
7:          G.push(getfirst(G.top))
8:      else if isone(G.top) then
9:          V.push(1)
10:     else if isinput(G.top) then
11:         V.push(A[getinput(G.top)])
12:     return G, V
```

Algorithm 2 Aggregate (pseudo-code)

```
1:  function AGGREGATE(G, V)                          ▷ Here, G.size = V.size
2:      g = G.pop
3:      v = V.pop
4:      if isplus (G.top) then
5:          V.top = V.top + v
6:      else if isprod (G.top) then
7:          V.top = V.top · v
8:      if not_last (G.top, g) then
9:          G.push(next_gate (G.top, g))
10:     return G, V
```

Algorithm 3 Evaluate (pseudo-code)

```
1:  function EVALUATE(A)          ▷ Input n × 1 vector A. Here, G and V are empty
2:      G.push(getroot())
3:      while G.size ≠ 1 or V.size ≠ 1 do
4:          if G.size ≠ V.size then
5:              (G, V) := Initialize(G, V, A)
6:          else
7:              (G, V) := Aggregate(G, V)
8:      return V.top
```

The *Evaluate* algorithm gives us the output of the circuit. Note that after each iteration it either holds that $\mathcal{G}.\texttt{size} = \mathcal{V}.\texttt{size} + 1$ or $\mathcal{G}.\texttt{size} = \mathcal{V}.\texttt{size}$. Furthermore, when we start we have $\mathcal{G}.\texttt{size} = 1$ and $\mathcal{V}.\texttt{size} = 0$. The condition $\mathcal{G}.\texttt{size} = 1$ and $\mathcal{V}.\texttt{size} = 1$ holds only when we have traversed all the circuit, and the value in $\mathcal{V}.\texttt{top}$ is the value that the root of the circuit outputs after its computation.

Next, we show how to encode this algorithm in for-MATLANG.

Let $n_0 \in \mathbb{N}$ be big enough for ($\star$) to hold and let $n \geq k$. Hence, the number of gates (values) is bounded by n^k and we need $k \log(n)$ bits to encode the id of each gate.

To simulate the two stacks $\mathcal{G}$ and $\mathcal{V}$ we keep a matrix X of dimensions $n \times n$.

- Column n will store a canonical vector that marks the top of stack V (values).
- Column $n-1$ will store a canonical vector that marks the top of stack G (gates).
- Column $n-2$ is the stack of values where $X[1, n-2]$ is the bottom of the stack.
- Columns 1 to $n-3$ are the stack of gates.

If we have j gates in the stack and currently $\mathcal{G}.\texttt{size} = \mathcal{V}.\texttt{size}$ then X would look like:

$$X = \begin{bmatrix} \texttt{id}_1 & v_1 & 0 & 0 \\ \texttt{id}_2 & v_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \texttt{id}_j & v_j & 1 & 1 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since $n \geq n_0$, ($\star$) holds and thus we never use more than $n-3$ bits to encode an id. Also, $j \leq n$ given that we never keep more gates than the depth of the tree. As a consequence, we never keep more than n values either.

An important detail is that the ids of the gates are encoded as $\texttt{id}_r 000$ for it to have dimension n, where $\texttt{id}_r$ is the corresponding binary number in reverse.

We make a series of definitions to make the notation more clear. Refer to section B.1 for more information about these expressions.

Let b_i be the i-th canonical vector. Next and Prev denote the successor and predecessor matrices respectively, such that

$$\mathsf{Next} \cdot b_i = \begin{cases} b_{i+1} \text{ if } i \leq n \\ 0 \text{ otherwise} \end{cases}$$

$$\mathsf{Prev} \cdot b_i = \begin{cases} b_{i-1} \text{ if } i \geq n \\ 0 \text{ otherwise} \end{cases}$$

We write expressions e_{min} for the first canonical vector and e_{max} for the last canonical vector. For any i we write

$$e_{\mathsf{min+i}} = \mathsf{Next}^i \cdot e_{\mathsf{min}}$$
$$e_{\mathsf{max+i}} = \mathsf{Prev}^i \cdot e_{\mathsf{max}}$$

We use the extra $\{0,1\}^n \to \{0,1\}^n$ functions that have a for-MATLANG translation:

$$\mathsf{min}(e) = \begin{cases} 1 \text{ if } e = e_{\mathsf{min}} \\ 0 \text{ otherwise} \end{cases}$$

$$\mathsf{max}(e) = \begin{cases} 1 \text{ if } e = e_{\mathsf{max}} \\ 0 \text{ otherwise} \end{cases}$$

$$\mathsf{succ}(b_i, b_j) = \begin{cases} 1 \text{ if } i \leq j \\ 0 \text{ otherwise} \end{cases}$$

When used in for-MATLANG these functions output $[0]$ and $[1]$.
Now

$$e_V := e_{\mathsf{max-2}}$$
$$e_{G_{top}} := e_{\mathsf{max-1}}$$
$$e_{V_{top}} := e_{\mathsf{max}}$$

For a canonical vector, let

$$\mathtt{Iden}(b_i) := \Sigma v.\mathsf{succ}(v, b_i) \cdot (v \cdot v^T).$$

This matrix has ones in the diagonal up to position i marked by e_i. We define the following sub-matrices of X:

$$V_{top} := X \cdot e_{V_{top}}$$
$$V := \mathtt{Iden}(V_{top}) \cdot X \cdot e_V$$
$$G_{top} := X \cdot e_{G_{top}}$$
$$G := \mathtt{Iden}(G_{top}) \cdot X \cdot \mathtt{Iden}(e_{\mathsf{max-3}})$$

For example, if we are in a step where $\mathcal{G}.\mathtt{size} = \mathcal{V}.\mathtt{size} + 1$ then

$$X = \begin{bmatrix} \mathtt{id}_1 & v_1 & 0 & 0 \\ \mathtt{id}_2 & v_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathtt{id}_{j-1} & v_{j-1} & 0 & 1 \\ \mathtt{id}_j & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}, G = \begin{bmatrix} \mathtt{id}_1 & 0 & 0 & 0 \\ \mathtt{id}_2 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathtt{id}_{j-1} & 0 & 0 & 0 \\ \mathtt{id}_j & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}, V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{j-1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, G_{top} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, V_{top} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Here, V is a vector encoding the stack of values in X and G is a matrix encoding the stack of gates in X. Note that what is *over* the top of the stacks is always set to zero due to $\mathtt{Iden}(G_{top})$ and $\mathtt{Iden}(V_{top})$. Also, note that G is of the same size as X. We sometimes omit the zeroes due to simplicity.

To set the initial state (algorithm 3 line 2) we define the for-MATLANG expression:

$$\text{START} := e_{\mathsf{min}} \cdot \mathtt{getroot}()^T + e_{\mathsf{min}} \cdot e_{G_{top}}^T.$$

For the initialize step, we define the for-MATLANG expressions: INIT_PLUS (algorithm 1, lines 2, 3, 4), INIT_PROD (algorithm 1, lines 5, 6, 7), CONST (algorithm 1, lines 8, 9) and INPUT (algorithm 1, lines 10, 11):

$$\text{INIT\_PLUS} := \texttt{isplus}\left(G^T \cdot G_{top}\right) \times \left[G + (\mathsf{Next} \cdot G_{top}) \cdot \texttt{getfirst}\left(G^T \cdot G_{top}\right)^T + \mathsf{Next} \cdot G_{top} \cdot e_{G_{top}}^T + V \cdot e_V^T + \mathsf{Next} \cdot V_{top} \cdot e_{V_{top}}^T\right]$$

$$\text{INIT\_PROD} := \texttt{isprod}\left(G^T \cdot G_{top}\right) \times \left[G + (\mathsf{Next} \cdot G_{top}) \cdot \texttt{getfirst}\left(G^T \cdot G_{top}\right)^T + \mathsf{Next} \cdot G_{top} \cdot e_{G_{top}}^T + (V + \mathsf{Next} \cdot v_{top}) \cdot e_V^T + \mathsf{Next} \cdot V_{top} \cdot e_{V_{top}}^T\right]$$

$$\text{CONST} := \texttt{isone}\left(G^T \cdot G_{top}\right) \times \left[G + (V + \mathsf{Next} \cdot V_{top}) \cdot e_V^T + \mathsf{Next} \cdot V_{top} \cdot e_{V_{top}}^T\right]$$

$$\text{INPUT} := \texttt{isinput}\left(G^T \cdot G_{top}\right) \times \left[G + \left(V + \left(v^T \cdot \mathsf{Next} \cdot V_{top} \cdot \texttt{getinput}\left(G^T \cdot G_{top}\right)^T\right)\right) \cdot e_V^T + \mathsf{Next} \cdot V_{top} \cdot e_{V_{top}}^T\right]$$

Where v is the matrix variable stated in the theorem, the one associated with the input A of the circuit. Here, $G^T \cdot G_{top}$ is to get the current id in the top of the stack. In INIT_PLUS we get the current stack G, we add $\mathsf{Next} \cdot G_{top} \cdot \texttt{getfirst}\left(G^T \cdot G_{top}\right)^T$ which is an $n \times n$ matrix with the first child of $G^T \cdot G_{top}$ in the next row. Then $\mathsf{Next} \cdot G_{top} \cdot e_{G_{top}}^T$ adds $\mathsf{Next} \cdot G_{top}$ to the $n-1$ column to mark the gate we added as the top. Next, we do the same with the values by adding $V \cdot e_V + \mathsf{Next} \cdot V_{top} \cdot e_{V_{top}}^T$.

The for-MATLANG expression equivalent to algorithm 1 is

$$\text{INIT} := \text{INIT\_PLUS} + \text{INIT\_PROD} + \text{CONST} + \text{INPUT}.$$

The idea is to return the matrix for the next iteration. Recall that here $\mathcal{G}.\texttt{size} = \mathcal{V}.\texttt{size} + 1$. So, when the operation is INPUT or CONST, if we start with

$$\begin{bmatrix} \mathtt{id}_1 & v_1 & 0 & 0 \\ \mathtt{id}_2 & v_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathtt{id}_{j-1} & v_{j-1} & 0 & 1 \\ \mathtt{id}_j & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}, \text{ then we return } \begin{bmatrix} \mathtt{id}_1 & v_1 & 0 & 0 \\ \mathtt{id}_2 & v_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathtt{id}_{j-1} & v_{j-1} & 0 & 0 \\ \mathtt{id}_j & v_j & 1 & 1 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

When the operation is INIT_PLUS or INIT_PROD, if we start with

$$\begin{bmatrix} \mathtt{id}_1 & v_1 & 0 & 0 \\ \mathtt{id}_2 & v_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathtt{id}_{j-1} & v_{j-1} & 0 & 1 \\ \mathtt{id}_j & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}, \text{ then we return } \begin{bmatrix} \mathtt{id}_1 & v_1 & 0 & 0 \\ \mathtt{id}_2 & v_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathtt{id}_{j-1} & v_{j-1} & 0 & 0 \\ \mathtt{id}_j & v_j & 0 & 1 \\ \mathtt{id}_{j+1} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

For the aggregate expression (algorithm 2) we do the following. Let

$$\mathsf{E}\,[b_i, c] = \Sigma v.(v^T \cdot b_i) \cdot c \cdot v \cdot v^T + (1 - v^T \cdot b_i) \cdot v \cdot v^T,$$

namely, it is the identity with c in position (i, i).

We define the expressions: AGG_PLUS (algorithm 2, lines 4, 5), AGG_PROD (algorithm 2, lines 6, 7), IS_NOT_LAST (algorithm 2, lines 8, 9), IS_LAST and POP:

$$\text{POP} := \texttt{Iden}(\mathsf{Prev} \cdot G_{top}) \cdot G + \mathsf{Prev} \cdot V_{top} \cdot e_{V_{top}}^T$$

$$\text{AGG\_PLUS} := \texttt{isplus}\left(G^T \cdot (P \cdot G_{top})\right) \times \left[\left(\texttt{Iden}(\mathsf{Prev} \cdot V_{top}) \cdot V + \left(V^T \cdot V_{top}\right)(\mathsf{Prev} \cdot V_{top})\right) \cdot e_V^T\right]$$

$$\text{AGG\_PROD} := \texttt{isprod}\left(G^T \cdot (P \cdot G_{top})\right) \times \left[\left(\texttt{E}\left[\mathsf{Prev} \cdot V_{top}, V^T \cdot V_{top}\right] \cdot \texttt{Iden}(\mathsf{Prev} \cdot V_{top}) \cdot V\right) \cdot e_V^T\right]$$

$$\text{IS\_NOT\_LAST} := \texttt{not\_last}\left(G^T \cdot (P \cdot G_{top}), G^T \cdot G_{top}\right) \times \left[G_{top} \cdot \texttt{next\_gate}\left(G^T \cdot (\mathsf{Prev} \cdot G_{top}), G^T \cdot G_{top}\right)^T + G_{top} \cdot e_{G_{top}}^T\right]$$

$$\text{IS\_LAST} := \left(1 - \texttt{not\_last}\left(G^T \cdot (P \cdot G_{top}), G^T \cdot G_{top}\right)\right) \times \left[(\mathsf{Prev} \cdot G_{top}) \cdot e_{G_{top}}^T\right]$$

The for-MATLANG expression equivalent to algorithm 2 is

$$\text{AGG} := \text{POP} + \text{AGG\_PLUS} + \text{AGG\_PROD} + \text{IS\_NOT\_LAST} + \text{IS\_LAST}.$$

The *Evaluate* method (algorithm 3) is defined as follows:

$$\begin{aligned}
&\text{EVAL}[v] = \\
&\quad e_{\mathsf{min}}^T \cdot \{\texttt{for}\, X, v_1, \ldots, v_k\,.\ : \\
&\quad \left(\Pi_{i=1}^k \mathsf{min}(v_i)\right) \times \mathit{START} + \\
&\quad \left(1 - \Pi_{i=1}^k \mathsf{min}(v_i)\right) \times \left((1 - \mathsf{min}(G_{top}) \cdot \mathsf{min}(V_{top})) \times \left[\left(1 - G_{top}^T \cdot V_{top}\right) \times \text{INIT} + \left(G_{top}^T \cdot V_{top}\right) \times \text{AGG}\right] + \mathsf{min}(G_{top}) \times \mathsf{min}(V_{top}) \times X\right) \\
&\quad \} \cdot e_V
\end{aligned}$$

Note that the for-expression does the evaluation. The final output is in $X[1, max - 2]$, we extract this value by multiplying the final result as $e_{\mathsf{min}}^T \cdot [\texttt{for}(\ldots)] \cdot e_V$.

Finally, we need to take care of all $n < n_0$, where ($\star$) does not necessarily hold. For any i, let:

$$\text{Eval}[i, A] := \text{ the } 1 \times 1 \text{ matrix with the value of the polynomial } \Phi_n(A) \text{ when } n = i.$$

Then we define:

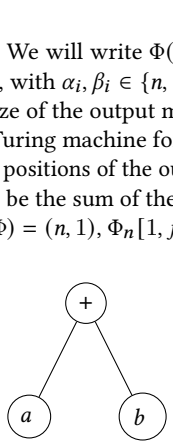
$$\Phi_n(a_1, \ldots, a_i) = \Sigma_{i=0}^{n_0 - 1}\left(e_{\mathsf{min+i}}^T \cdot (e_{\mathsf{diag}}(e_1(v)) \cdot e_{\mathsf{max}})\right) \times \text{EVAL}[i, v] + \left((\mathsf{Next}^{n_0} \cdot e_{\mathsf{min}})^T \cdot e_1(v)\right) \times \text{EVAL}[v].$$

Above, $(e_{\mathsf{min+i}}^T \cdot e_{\mathsf{max}})$ checks if the dimension is equal to i (we multiply by the $n \times n$ identity $e_{\mathsf{diag}}(e_1(v))$ to ensure typing), and $(\mathsf{Next}^{n_0} \cdot e_{\mathsf{min}})^T \cdot e_1(e_{\mathsf{min}})$ checks if the dimension is greater or equal than n_0.

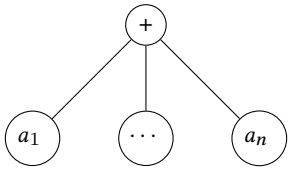
□

D.3 From MATLANG to uniform ACs

We consider circuits over matrices (multiple output gates). We will write $\Phi(A_1, \ldots, A_k)$, where Φ is an arithmetic circuit with multiple output gates, and each A_i is a matrix of dimensions $\alpha_i \times \beta_i$, with $\alpha_i, \beta_i \in \{n, 1\}$ to denote the input matrices for a circuit Φ. We will also write $\texttt{type}_{\mathcal{S}}(\Phi) = (\alpha, \beta)$, with $\alpha, \beta \in \{n, 1\}$, to denote the size of the output matrix for Φ. When $\{\Phi_n \mid n = 1, 2, \ldots\}$ is a uniform family of arithmetic circuits over matrices, we will assume that the Turing machine for generating Φ_n also gives us the information about how to access a position of each input matrix, and how to access the positions of the output matrix, as is usually done when handling matrices with arithmetic circuits [29]. The notion of degree is extended to be the sum of the degrees of all the output gates. The former will be denoted as $\Phi_n[i, j]$ when $\texttt{type}_{\mathcal{S}}(\Phi) = (n, n)$, $\Phi_n[i, 1]$ when $\texttt{type}_{\mathcal{S}}(\Phi) = (n, 1)$, $\Phi_n[1, j]$ when $\texttt{type}_{\mathcal{S}}(\Phi) = (1, n)$ and Φ_n when $\texttt{type}_{\mathcal{S}}(\Phi) = (1, 1)$. Also, when we write $a \oplus b$ we mean



When we write $\bigoplus_{l=1}^n a_l$ we mean



Same with $\otimes$. Now we prove the statement.

Theorem 5.3. *Let e be a* for-MATLANG *expression over a schema $\mathcal{S}$, and let $V_1, \ldots, V_k$ be the variables of e such that* $\mathtt{type}_{\mathcal{S}}(V_i) \in \{(\alpha,\alpha),(\alpha,1),(1,\alpha),(1,1)\}$. *Then there exists a uniform arithmetic circuit family over matrices $\Phi_n(A_1,\ldots,A_k)$ such that:*

- *For any instance $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$ such that $\mathcal{D}(\alpha) = n$ and $\mathtt{mat}(V_i) = A_i$ it holds that:*
- $[\![e]\!](\mathcal{I}) = \Phi_n(A_1,\ldots,A_k)$.

Proof. Let e be a for-MATLANG expression.
If $e = V$ then $\Phi^e_n := \Phi(A)$, and we have that

- If $\mathtt{type}_{\mathcal{S}}(V) = (1,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (1,1)$ and Φ^e_n has the one input/output gate.
- If $\mathtt{type}_{\mathcal{S}}(V) = (1,\alpha)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (1,n)$ and Φ^e_n has n input/output gates.
- If $\mathtt{type}_{\mathcal{S}}(V) = (\alpha,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (n,1)$ and Φ^e_n has n input/output gates.
- If $\mathtt{type}_{\mathcal{S}}(V) = (\alpha,\alpha)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (n,n)$ and Φ^e_n has n^2 input/output gates.

If $e = e'^T$ then $\Phi^e_n = \Phi^{e'}_n$.

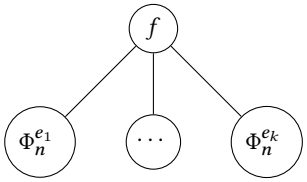
- If $\mathtt{type}_{\mathcal{S}}(\Phi^{e'}_n) = (1,1)$ then $\Phi^e_n = \Phi^{e'}_n$ and $\mathtt{type}(\Phi^e_n) = (1,1)$.
- If $\mathtt{type}_{\mathcal{S}}(\Phi^{e'}_n) = (1,n)$ then $\mathtt{type}(\Phi^e_n) = (n,1)$ and $\Phi^e_n[i,1] := \Phi^{e'}_n[1,i]$.
- If $\mathtt{type}_{\mathcal{S}}(\Phi^{e'}_n) = (n,1)$ then $\mathtt{type}(\Phi^e_n) = (1,n)$ and $\Phi^e_n[1,j] := \Phi^{e'}_n[j,1]$.
- If $\mathtt{type}_{\mathcal{S}}(\Phi^{e'}_n) = (n,n)$ then $\mathtt{type}(\Phi^e_n) = (n,n)$ and $\Phi^e_n[i,j] := \Phi^{e'}_n[j,i]$.

If $e = 1(e')$ where $\mathtt{type}_{\mathcal{S}}(\Phi^{e'}_n) = (\alpha,\beta)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (\alpha,1)$ and $\Phi^e_n[i,1] := 1$.
If $e = e_1 + e_2$ we have

- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (1,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (1,1)$ and $\Phi^e_n := \Phi^{e_1}_n \oplus \Phi^{e_2}_n$.
- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (1,n)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (1,n)$ and $\Phi^e_n[1,j] := \Phi^{e_1}_n[1,j] \oplus \Phi^{e_2}_n[1,j]$.
- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (n,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (n,1)$ and $\Phi^e_n[i,1] := \Phi^{e_1}_n[i,1] \oplus \Phi^{e_2}_n[i,1]$.
- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (n,n)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (n,n)$ and $\Phi^e_n[i,j] := \Phi^{e_1}_n[i,j] \oplus \Phi^{e_2}_n[i,j]$.

If $e = f(e_1,\ldots,e_k)$ we have two cases

- When f is the function $f_{\odot}$ (recall that this function is definable in MATLANG$[\emptyset]$ by Lemma A.1) then
  - If $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \ldots = \mathtt{type}_{\mathcal{S}}(\Phi^{e_k}_n) = (1,1)$ then $\Phi^e_n := \bigotimes_{l=1}^{k} \Phi^{e_l}_n$.
  - If $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \ldots = \mathtt{type}_{\mathcal{S}}(\Phi^{e_k}_n) = (1,n)$ then $\Phi^e_n[1,j] := \bigotimes_{l=1}^{k} \Phi^{e_l}_n[1,j]$.
  - If $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \ldots = \mathtt{type}_{\mathcal{S}}(\Phi^{e_k}_n) = (n,1)$ then $\Phi^e_n[i,1] := \bigotimes_{l=1}^{k} \Phi^{e_l}_n[i,1]$.
  - If $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \ldots = \mathtt{type}_{\mathcal{S}}(\Phi^{e_k}_n) = (n,n)$ then $\Phi^e_n[i,j] := \bigotimes_{l=1}^{k} \Phi^{e_l}_n[i,j]$.
- When f is any function, we prove the case when $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = \ldots = \mathtt{type}_{\mathcal{S}}(\Phi^{e_k}_n) = (1,1)$ (only case necessary, as discussed in Appendix A.2). Here Φ^e_n is



  Note that since for the context of this result we only consider for-MATLANG = MATLANG$[\emptyset]$, this case is not strictly necessary, modulo for $f_{\odot}, f_{\oplus}$ due to Lemma A.1. However, if we extend the circuits with the same functions allowed in for-MATLANG, then our inductive construction still goes through, as just illustrated.

If $e = e_1 \cdot e_2$ we have

- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = (1,1)$ and $\mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (1,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (1,1)$ and $\Phi^e_n := \Phi^{e_1}_n \otimes \Phi^{e_2}_n$.
- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = (1,1)$ and $\mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (1,n)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (1,n)$ and $\Phi^e_n[1,j] := \Phi^{e_1}_n \otimes \Phi^{e_2}_n[1,j]$.
- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = (n,1)$ and $\mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (1,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (n,1)$ and $\Phi^e_n[i,1] := \Phi^{e_1}_n[i,1] \otimes \Phi^{e_2}_n$.
- When $\mathtt{type}_{\mathcal{S}}(\Phi^{e_1}_n) = (n,1)$ and $\mathtt{type}_{\mathcal{S}}(\Phi^{e_2}_n) = (1,n)$ then $\mathtt{type}_{\mathcal{S}}(\Phi^e_n) = (n,n)$ and $\Phi^e_n[i,j] := \Phi^{e_1}_n[i,1] \otimes \Phi^{e_2}_n[1,j]$.

- When $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_1}) = (1,n)$ and $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_2}) = (n,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e}) = (1,1)$ and
$$\Phi_n^e := \bigoplus_{k=1}^{n} \left(\Phi_n^{e_1}[1,k] \otimes \Phi_n^{e_2}[k,1]\right).$$
- When $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_1}) = (1,n)$ and $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_2}) = (n,n)$ then $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e}) = (1,n)$ and
$$\Phi_n^e[1,j] := \bigoplus_{k=1}^{n} \left(\Phi_n^{e_1}[1,k] \otimes \Phi_n^{e_2}[k,j]\right).$$
- When $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_1}) = (n,n)$ and $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_2}) = (n,1)$ then $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e}) = (n,1)$ and
$$\Phi_n^e[i,1] := \bigoplus_{k=1}^{n} \left(\Phi_n^{e_1}[i,k] \otimes \Phi_n^{e_2}[k,1]\right).$$
- When $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_1}) = (n,n)$ and $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e_2}) = (n,n)$ then $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e}) = (n,n)$ and
$$\Phi_n^e[i,j] := \bigoplus_{k=1}^{n} \left(\Phi_n^{e_1}[i,k] \otimes \Phi_n^{e_2}[k,j]\right).$$

If $e = \mathtt{for}\, X, v.\ e'(X,v)$, then define Φ^0 as the zero matrix circuit $\mathtt{type}_{\mathcal{S}}(\Phi^0) = (1,1)$ if $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e'}) = (1,1)$ and $\mathtt{type}_{\mathcal{S}}(\Phi^0) = (n,n)$ if $\mathtt{type}_{\mathcal{S}}(\Phi_n^{e'}) = (n,n)$. Also, $\Phi^0 = 0$ and $\Phi^0[i,j] = 0\ \forall i,j$ for each case respectively. Now for $i = 1,\ldots,n$, define Φ^{v_i} as the circuit such that $\mathtt{type}_{\mathcal{S}}(\Phi^{v_i}) = (n,1)$ and $\Phi^{v_i}[i,1] := 1$ and zero otherwise. Finally, define

$$\Phi_n^e = \Phi_n^{e'}\left(\Phi_n^{e'}\left(\cdots\left(\Phi_n^{e'}\left(\Phi^0, \Phi^{v_1}\right), \Phi^{v_2}\right)\cdots, \Phi^{v_{n-1}}\right), \Phi^{v_n}\right).$$

Note that every circuit adds a constant number of layers except when $e = \mathtt{for}\, X, v.\ e'(X,v)$. This means that the depth still is polynomial. When $e = \mathtt{for}\, X, v.\ e'(X,v)$ we have that the depth of the circuit is $n \cdot p(n)$, where the depth of $e'(X,v)$ is $p(n)$, so it also remains polynomial.

Here, we do not need to translate scalar multiplication because it can be simulated using the ones operator and $f_{\odot}$ (see section A.2).

Finally, we remark that when composing the circuits the fact that uniformity is preserved (i.e. the resulting circuit can be generated by a LOGSPACE machine) is proved analogously as when composing two LOGSPACE transducers [3]. The only more involved case is treating for-loop construction, however, notice here that we only need to keep track of where we are in the evaluation (i.e. which v_i we are processing), and not of all the previous results, given that they update the resulting matrix in a fixed order.

□

D.4 Undecidability

Let e be a for-MATLANG expression over a matrix schema $\mathcal{S} = (\mathcal{M}, \mathrm{size})$ and let $V_1,\ldots,V_k$ be the variables of e, each of type (α,α), $(1,\alpha)$, $(\alpha,1)$ or $(1,1)$. We know from Theorem 5.3 that there exists a uniform arithmetic circuit family $\{\Phi_n \mid n = 1,2,\ldots\}$ such that $[\![e]\!](\mathcal{I}) = \Phi_n(A_1,\ldots,A_k)$ for any instance $\mathcal{I}$ such that $\mathcal{D}(\alpha) = n$ and $\mathtt{mat}(V_i) = A_i$ for $i = 1,\ldots,k$. We are interested in deciding whether there exists such a uniform arithmetic circuit family $\{\Phi_n \mid n = 1,2,\ldots\}$ of polynomial degree, i.e., such that $\mathsf{degree}(\Phi_n) = O(p(n))$ for some polynomial $p(x)$. If such a circuit family exists, we call e of polynomial degree.

Proposition 5.5. *Given a* for-MATLANG *expression e over a schema $\mathcal{S}$, it is undecidable to check whether e is of polynomial degree.*

Proof. We show undecidability based on the following undecidable language:

$$\{\langle M\rangle \mid M \text{ is a deterministic TM which halts on the empty input}\},$$

where $\langle M\rangle$ is some string encoding of M. Consider a TM M described by $(Q, \Gamma = \{0,1\}, q_0, q_m, \Delta)$ with $Q = \{q_1,\ldots,q_m\}$ its states, q_1 being the initial state and q_m being the halting state, Γ is the tape alphabet, and Δ is a transition function from $Q \times \Gamma \to Q \times \Gamma \times \{\leftarrow, \sqcup, \rightarrow\}$. The simulation of linear space TM, as given in the proof of Proposition D.1 can be easily modified to any TM M provided that we limit the execution of M to exactly n steps. Let e_M denote this expression. Similarly as in the linear space TM simulation, we have vector variables $Q_1,\ldots,Q_m$ encoding the states, a single relation T encoding the tape and relation H_T encoding the position of the tape. When an instance $\mathcal{I}$ assigns n to α, we have a tape of length n at our disposal. This suffices if we let M run for n steps. We further observe that all vector variables can be assumed to be zero, initially. This is because we do not have input. So, let $\mathcal{I}_n^0$ denote the instance which assigns vector variables to the n-dimensional zero vector. Furthermore, by contrast to the linear space TM simulation, we use a single vector v (instead of k such vectors) to simulate n steps of M. Finally, we modify the expression given in the proof of Proposition D.1 such $[\![e_M]\!](\mathcal{I}_n^0)$ returns 1 if M halts in at most n steps, and 0 if M did not halt yet after n steps.

As a consequence, when M does not halt, $[\![e_M]\!](\mathcal{I}_n^0) = 0$ for all $n \geq 0$. When M halts, there will be an n such that $[\![e_M]\!](\mathcal{I}_n^0) = 1$ It now suffices to consider the for-MATLANG expression

$$d_M := e_M \cdot e_{\mathsf{exp}}$$

where $e_{\mathsf{exp}} = \mathsf{for}\, v, X = 1(X)^T \cdot 1(X).\ X \cdot X$ such that $e_{\mathsf{exp}}(\mathcal{I}_n^0) = n^{2^n}$. Then, when M does not halt we can clearly compute d_M with a constant degree circuit "0" for any n, otherwise, the circuit needed will be of exponential degree for at least one n, simply because no polynomial degree uniform circuit family can compute n^{2^n}. In other words, deciding whether d_M has polynomial degree coincides with deciding whether M halts. □

E PROOFS OF SECTION 6

E.1 From sum-MATLANG to RA_K^+

We prove proposition 6.3.

Proposition 6.3. *For each* sum-MATLANG *expression e over schema $\mathcal{S}$ such that $\mathcal{S}(e) = (\alpha, \beta)$ with $\alpha \neq 1 \neq \beta$, there exists a RA_K^+ expression $Q(e)$ over relational schema $Rel(\mathcal{S})$ such that $Rel(\mathcal{S})(Q(e)) = \{\mathsf{row}_\alpha, \mathsf{row}_\beta\}$ and such that for any instance $\mathcal{I}$ over $\mathcal{S}$,*

$$[\![e]\!](\mathcal{I})_{i,j} = [\![Q(e)]\!]_{Rel(\mathcal{I})}(t)$$

for tuple $t(\mathrm{row}_\alpha) = i$ and $t(\mathrm{col}_\beta) = j$. Similarly for when e has schema $\mathcal{S}(e) = (\alpha, 1)$, $\mathcal{S}(e) = (1, \beta)$ or $\mathcal{S}(e) = (1, 1)$, then $Q(e)$ has schema $Rel(\mathcal{S})(Q(e)) = \{\mathrm{row}_\alpha\}$, $Rel(\mathcal{S})(Q(e)) = \{\mathrm{col}_\alpha\}$, or $Rel(\mathcal{S})(Q(e)) = \{\}$, respectively.

Proof. We start from a matrix schema $\mathcal{S} = (\mathcal{M}, \mathtt{size})$, where $\mathcal{M} \subset \mathcal{V}$ is a finite set of matrix variables, and $\mathtt{size} : \mathcal{V} \mapsto \mathtt{Symb} \times \mathtt{Symb}$ is a function that maps each matrix variable to a pair of size symbols. On the relational side we have for each size symbol $\alpha \in \mathtt{Symb} \setminus \{1\}$, attributes α, row_α, and col_α in $\mathbb{A}$. We also reserve some special attributes $\gamma_1, \gamma_2, \ldots$ whose role will become clear shortly. For each $V \in \mathcal{M}$ and $\alpha \in \mathtt{Symb}$ we denote by R_V and R_α its corresponding relation name, respectively.

Then, given $\mathcal{S}$ we define the relational schema $\mathrm{Rel}(\mathcal{S})$ such that $\mathrm{dom}(\mathrm{Rel}(\mathcal{S})) = \{R_\alpha \mid \alpha \in \mathtt{Symb}\} \cup \{R_V \mid V \in \mathcal{M}\}$ where $\mathrm{Rel}(\mathcal{S})(R_\alpha) = \{\alpha\}$ and for all $V \in \mathcal{M}$:

$$\mathrm{Rel}(\mathcal{S})(R_V) = \begin{cases} \{\mathsf{row}_\alpha, \mathsf{col}_\beta\} & \text{if } \mathtt{size}(V) = (\alpha, \beta) \\ \{\mathsf{row}_\alpha\} & \text{if } \mathtt{size}(V) = (\alpha, 1) \\ \{\mathsf{col}_\beta\} & \text{if } \mathtt{size}(V) = (1, \beta) \\ \{\} & \text{if } \mathtt{size}(V) = (1, 1). \end{cases}$$

Next, for a matrix instance $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$ over $\mathcal{S}$, let $V \in \mathcal{M}$ with $\mathtt{size}(V) = (\alpha, \beta)$ and let $\mathtt{mat}(V)$ be its corresponding K-matrix of dimension $\mathcal{D}(\alpha) \times \mathcal{D}(\beta)$. The K-instance in RA_K^+ according to $\mathcal{I}$ is $\mathrm{Rel}(\mathcal{I})$ with data domain $\mathbb{D} = \mathbb{N} \setminus \{0\}$. For each $V \in \mathcal{M}$ we define $R_V^{\mathrm{Rel}(\mathcal{I})}(t) := \mathtt{mat}(V)_{ij}$ whenever $t(\mathsf{row}_\alpha) = i \leq \mathcal{D}(\alpha)$ and $t(\mathsf{col}_\beta) = j \leq \mathcal{D}(\beta)$, and $\mathbb{0}$ otherwise. Also, for each $\alpha \in \mathtt{Symb}$ we define $R_\alpha^{\mathrm{Rel}(\mathcal{I})}(t) := \mathbb{1}$ whenever $t(\alpha) \leq \mathcal{D}(\alpha)$, and $\mathbb{0}$ otherwise. If $\mathtt{size}(V) = (\alpha, 1)$ then $R_V^{\mathrm{Rel}(\mathcal{I})}(t) := \mathtt{mat}(V)_{i1}$ whenever $t(\mathsf{row}_\alpha) = i \leq \mathcal{D}(\alpha)$ and $\mathbb{0}$ otherwise. Similarly, if $\mathtt{size}(V) = (1, \beta)$ then $R_V^{\mathrm{Rel}(\mathcal{I})}(t) := \mathtt{mat}(V)_{1j}$ whenever $t(\mathsf{col}_\beta) = j \leq \mathcal{D}(\beta)$ and $\mathbb{0}$ otherwise. If $\mathtt{size}(V) = (1, 1)$ then $R_V^{\mathrm{Rel}(\mathcal{I})}(()) := \mathtt{mat}(V)_{11}$.

Let e be a sum-MATLANG expression. In the following we need to distinguish between matrix variables v that occur in e as part of a sub-expression $\Sigma v.(\cdot)$, i.e., those variables that will be used to iterate over by means of canonical vectors, and those that are not. To make this distinction clear, we use $v_1, v_2, \ldots$ for those "iterator" variables, and capital V for the other variables occurring in e. For simplicity, we assume that each occurrence of Σ has its own iterator variable associated with it.

We define free (iterator) variables, as follows. $\mathsf{free}(V) := \emptyset$, $\mathsf{free}(v) := \{v\}$, $\mathsf{free}(e^T) := \mathsf{free}(e)$, $\mathsf{free}(e_1 + e_2) := \mathsf{free}(e_1) \cup \mathsf{free}(e_2)$, $\mathsf{free}(e_1 \cdot e_2) := \mathsf{free}(e_1) \cup \mathsf{free}(e_2)$, $\mathsf{free}(f_\odot(e_1, \ldots, e_k)) := \mathsf{free}(e_1) \cup \cdots \cup \mathsf{free}(e_k)$, and $\mathsf{free}(e = \Sigma V.e_1) = \mathsf{free}(e_1) \setminus \{v\}$. We will explicitly denote the free variables in an expression e by writing $e(v_1, \ldots, v_k)$.

We now use the following induction hypotheses:

- If $e(v_1, \ldots, v_k)$ is of type (α, β) then there exists a RA_K^+ expression Q such that $\mathrm{Rel}(\mathcal{S})(Q(e)) = \{\mathsf{row}_\alpha, \mathsf{col}_\beta, \gamma_1, \ldots, \gamma_k\}$ and such that
$$[\![Q(e)]\!]_{\mathrm{Rel}(\mathcal{I})}(t) = [\![e]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])_{i,j}$$
for tuple $t(\mathrm{row}_\alpha) = i$, $t(\mathrm{col}_\beta) = j$ and $t(\gamma_s) = i_s$ for $s = 1, \ldots, k$.
- If $e(v_1, \ldots, v_k)$ is of type $(\alpha, 1)$ then there exists a RA_K^+ expression Q such that $\mathrm{Rel}(\mathcal{S})(Q(e)) = \{\mathsf{row}_\alpha, \gamma_1, \ldots, \gamma_k\}$ and such that
$$[\![Q(e)]\!]_{\mathrm{Rel}(\mathcal{I})}(t) = [\![e]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])_{i,1}$$
for tuple $t(\mathrm{row}_\alpha) = i$, and $t(\gamma_s) = i_s$ for $s = 1, \ldots, k$. And similarly for when e is type $(1, \beta)$.
- If $e(v_1, \ldots, v_k)$ is of type $(1, 1)$ then there exists a RA_K^+ expression Q such that $\mathrm{Rel}(\mathcal{S})(Q(e)) = \{\gamma_1, \ldots, \gamma_k\}$ and such that
$$[\![Q(e)]\!]_{\mathrm{Rel}(\mathcal{I})}(t) = [\![e]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])_{1,1}$$
for tuple $t(\gamma_s) = i_s$ for $s = 1, \ldots, k$.

Clearly, this suffices to show the proposition since we there consider expressions e for which $\mathsf{free}(e) = \emptyset$, in which case the above statements reduce to the one given in the proposition.

The proof is by induction on the structure of sum-MATLANG expressions. In line with the simplifications in Section A.2, it suffices to consider pointwise function application with $f_\odot$ instead of scalar multiplication. (We also note that we can express the one-vector operator in sum-MATLANG, so scalar multiplication can be expressed using $f_\odot$ in sum-MATLANG).

Let e be a sum-MATLANG expression.

- If $e = V$ then $Q(e) := R_V$.
- If $e = v_p$ then $Q(e) := \sigma_{\{\mathsf{row}_\alpha, \gamma_p\}}\big(\rho_{\mathsf{row}_\alpha \to \alpha}(R_\alpha) \bowtie \rho_{\gamma_p \to \alpha}(R_\alpha)\big)$ when v_p is of type $(\alpha, 1)$. It is here that we introduce the attribute γ_p associated with iterator variable v_p. We note that
$$[\![Q(v_p)]\!]_{\mathrm{Rel}(\mathcal{I})}(t) = [\![v_p]\!](\mathcal{I}[v_p \leftarrow b_j])_{i,1} = (b_j)_{i,1}$$
for $t(\mathrm{row}_\alpha) = i$ and $t[\gamma_p] = j$. Indeed, $(b_j)_{i,1} = \mathbb{1}$ if $j = i$ and this holds when $t(\mathrm{row}_\alpha) = t[\gamma_p] = j$, and $(b_j)_{i,1} = \mathbb{0}$ if $j \neq i$ and this also holds when $t(\mathrm{row}_\alpha) \neq t[\gamma_p] = j$.
- If $e(v_1, \ldots, v_k) = (e_1(v_1, \ldots, v_k))^T$ with $\mathcal{S}(e_1) = (\alpha, \beta)$ then
$$Q(e) := \begin{cases} \rho_{\mathrm{row}_\alpha \to \mathrm{col}_\alpha, \mathrm{col}_\beta \to \mathrm{row}_\beta}\big(Q(e_1)\big) & \text{if } \alpha \neq 1 \neq \beta; \\ \rho_{\mathrm{row}_\alpha \to \mathrm{col}_\alpha}\big(Q(e_1)\big) & \text{if } \alpha \neq 1 = \beta; \\ \rho_{\mathrm{col}_\beta \to \mathrm{row}_\beta}\big(Q(e_1)\big) & \text{if } \alpha = 1 \neq \beta; \\ Q(e_1) & \text{if } \alpha = 1 = \beta. \end{cases}$$
- If $e = e_1(v_1, \ldots, v_k) + e_2(v_1, \ldots, v_k)$ with $\mathcal{S}(e_1) = \mathcal{S}(e_2) = (\alpha, \beta)$ then $Q(e) := Q(e_1) \cup Q(e_2)$. We assume here that e_1 and e_2 have the same free variables. This is without loss of generality. Indeed, as an example, suppose that we have $e_1(v_1, v_2)$ and $e_2(v_2, v_3)$. Then, we can replace e_1 by $e_1(v_1, v_2, v_3) = (v_3^T \cdot v_3) \times e_1(v_1, v_2)$ and similarly, e_2 by $e_2(v_1, v_2, v_3) = (v_1^T \cdot v_1) \times e_2(v_2, v_3)$, where in addition we replace scalar multiplication with its simulation using $f_\odot$ and the ones vector, as mentioned earlier.
- If $e = f_\odot(e_1, \ldots, e_k)$ with $\mathcal{S}(e_i) = \mathcal{S}(e_j)$ for all $i, j \in [1, k]$, then $Q(e) := Q(e_1) \bowtie \cdots \bowtie Q(e_k)$.
- If $e = e_1 \cdot e_2$ with $\mathcal{S}(e_1) = (\alpha, \gamma)$ and $\mathcal{S}(e_2) = (\gamma, \beta)$, we have two cases. If $\gamma = 1$ then $Q(e) := Q(e_1) \bowtie Q(e_2)$. If $\gamma \neq 1$ then
$$Q(e) := \pi_{\{\mathsf{row}_\alpha, \mathsf{col}_\beta, \gamma_1, \ldots, \gamma_k\}}\Big(\rho_{C \to \mathsf{col}_\gamma}(Q(e_1)) \bowtie \rho_{C \to \mathsf{row}_\gamma}(Q(e_2))\Big),$$
when $\mathrm{Rel}(\mathcal{S})(Q(e_1)) = \{\mathsf{row}_\alpha, \mathsf{col}_\gamma, \gamma_1', \ldots, \gamma_\ell'\}$, $\mathrm{Rel}(\mathcal{S})(Q(e_2)) = \{\mathsf{row}_\gamma, \mathsf{col}_\beta, \gamma_1'', \ldots, \gamma_\ell''\}$ and $\{\gamma_1, \ldots, \gamma_k\} = \{\gamma_1', \ldots, \gamma_k', \gamma_1'', \ldots, \gamma_m''\}$.
- If $e(v_1, \ldots, v_{p-1}, v_{p+1}, \ldots, v_k) = \Sigma v_p.e_1(v_1, \ldots, v_k)$ where $\mathcal{S}(e_1) = (\alpha, \beta)$ and $\mathcal{S}(V) = (\gamma, 1)$. Then we do
$$Q(e) := \pi_{\mathrm{Rel}(\mathcal{S})(Q(e_1)) \setminus \{\gamma_p\}} Q(e_1).$$
Indeed, by induction we know that
$$[\![Q(e_1)]\!]_{\mathrm{Rel}(\mathcal{I})}(t) = [\![e]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])_{i,j}$$
for tuple $t(\mathrm{row}_\alpha) = i$, $t(\mathrm{col}_\beta) = j$ and $t(\gamma_s) = i_s$ for $s = 1, \ldots, k$. Hence, for $t(\mathrm{row}_\alpha) = i$, $t(\mathrm{col}_\beta) = j$ and $t(\gamma_s) = i_s$ for $s = 1, \ldots, k$ and $s \neq p$,
$$[\![Q(e_1)]\!]_{\mathrm{Rel}(\mathcal{I})}(t) := \bigoplus_{i_p = 1, \ldots, \mathcal{D}(\gamma)} [\![e_1]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])_{i,j},$$
which precisely corresponds to
$$[\![\Sigma v_p.e_1(v_1, \ldots, v_k)]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_{p-1} \leftarrow b_{p-1}, v_{p+1} \leftarrow b_{p+1}, \ldots, v_k \leftarrow b_k])_{i,j}.$$

All other cases, when expressions have type $(\alpha, 1)$, $(1, \beta)$ or $(1, 1)$ can be dealt with in a similar way. □

E.2 From RA_K^+ to sum-MATLANG

Let $\mathcal{R}$ be binary relational schema. For each $R \in \mathcal{R}$ we associate a matrix variable V_R such that, if R is a binary relational signature, then V_R represents a (square) matrix, if R is unary, then V_R represents a vector and if $|R| = 0$ then V_R represents a constant. Formally, fix a symbol $\alpha \in \mathsf{Symb} \setminus \{1\}$. Let $\mathrm{Mat}(\mathcal{R})$ denote the MATLANG schema $(\mathcal{M}_\mathcal{R}, \mathsf{size}_\mathcal{R})$ such that $\mathcal{M}_\mathcal{R} = \{V_R \mid R \in \mathcal{R}\}$ and $\mathsf{size}_\mathcal{R}(V_R) = (\alpha, \alpha)$ whenever $|R| = 2$, $\mathsf{size}_\mathcal{R}(V_R) = (\alpha, 1)$ whenever $|R| = 1$ and $\mathsf{size}_\mathcal{R}(V_R) = (1, 1)$ whenever $|R| = 0$. Let $\mathcal{J}$ be the K-instance of $\mathcal{R}$ and suppose that $\mathrm{adom}(\mathcal{J}) = \{d_1, \ldots, d_n\}$ is the active domain (with arbitrary order) of $\mathcal{J}$. Define the matrix instance $\mathrm{Mat}(\mathcal{J}) = (\mathcal{D}_\mathcal{J}, \mathtt{mat}_\mathcal{J})$ such that $\mathcal{D}_\mathcal{J}(\alpha) = n$, $\mathtt{mat}_\mathcal{J}(V_R)_{i,j} = R^\mathcal{J}((d_i, d_j))$ whenever $|R| = 2$, $\mathtt{mat}_\mathcal{J}(V_R)_i = R^\mathcal{J}((d_i))$ whenever $|R| = 1$, and $\mathtt{mat}_\mathcal{J}(V_R)_{1,1} = R^\mathcal{J}$ whenever $|R| = 0$. Note that we consider the active domain of the whole K-instance.

We next translate RA_K^+ expressions in to sum-MATLANG expressions over an extended schema. More specifically, for each attribute $A \in \mathbb{A}$ we define a vector variable v_A of type $(\alpha, 1)$. Then for each RA_K^+ expression Q with attributes $A_1, \ldots, A_k$ we define a sum-MATLANG expression $e_Q(v_{A_1}, \ldots, v_{A_k})$ of type $(1, 1)$ such that the following inductive hypothesis holds:
$$[\![e_Q]\!](\mathrm{Mat}(\mathcal{J})[v_{A_1} \leftarrow b_{i_1}, \ldots, v_{A_k} \leftarrow b_{i_k}]) = [\![Q]\!]_\mathcal{J}(t) \qquad (*)$$
where $t(A_s) = i_s$ for $s = 1, \ldots, k$. The proof of this claim follows by induction on the structure of expressions:

- If $Q = R$, then $e_Q := v_{A_1}^T \cdot V_R \cdot v_{A_2}$ if $\mathcal{R}(R) = \{A_1, A_2\}$ with $A_1 < A_2$; $e_Q := V_R^T \cdot v_A$ if $\mathcal{R}(R) = \{A\}$; and $e_Q := V_R$ if $\mathcal{R}(R) = \{\}$.

- If $Q = Q_1 \cup Q_2$ then $e_Q := e_{Q_1} + e_{Q_2}$.
- If $Q = \pi_Y(Q_1)$ for $Y \subseteq \mathcal{R}(Q_1)$ and $\{B_1, \ldots, B_l\} = \mathcal{R}(Q_1) \setminus Y$ then
$$e_Q := \Sigma v_{B_1}.\, \Sigma v_{B_2}.\, \ldots \Sigma v_{B_l}.\, e_{Q_1}$$
- If $Q = \sigma_Y(Q_1)$ with $Y \subseteq \mathcal{R}(Q_1)$ then
$$e_Q := e_{Q_1} \cdot \prod_{A,B \in Y} (v_A^T \cdot v_B).$$
Here Π is the matrix multiplication of expressions of type $(1,1)$.
- If $Q = \rho_{X \mapsto Y}(Q_1)$ then
$$e_Q := e_{Q_1}[v_B \leftarrow v_A \mid A \in X, B \in Y, A \mapsto B].$$
In other words, we rename variable v_B with variable v_B in all the expression e_{Q_1}.
- If $Q = Q_1 \bowtie Q_2$ then $e_Q := e_{Q_1} \cdot e_{Q_1}$ where the product is over expression of type $(1,1)$.

One can check, by induction over the construction, that the inductive hypothesis $(*)$ holds in each case. Now we can obtain proposition 6.4.

Proposition 6.4. *Let $\mathcal{R}$ be a binary relational schema. For each RA_K^+ expression Q over $\mathcal{R}$ such that $|\mathcal{R}(Q)| = 2$, there exists a* sum-MATLANG *expression $\Psi(Q)$ over* MATLANG *schema $Mat(\mathcal{R})$ such that for any K-instance $\mathcal{J}$ with $\mathrm{adom}(\mathcal{J}) = \{d_1, \ldots, d_n\}$ over $\mathcal{R}$,*
$$[\![Q]\!]_{\mathcal{J}}((d_i, d_j)) = [\![\Psi(Q)]\!](Mat(\mathcal{J}))_{i,j}.$$
Similarly for when $|\mathcal{R}(Q)| = 1$, or $|\mathcal{R}(Q)| = 0$ respectively.

Proof. As a consequence of the previous discussion above, when Q is a RA_K^+ expression such that $\mathcal{R}(Q) = \{A_1, A_2\}$ with $A_1 < A_2$ then we define
$$\Psi(Q) \;=\; \Sigma v_{A_1}.\, \Sigma v_{A_2}.\, e_Q \cdot (v_{A_1} \cdot v_{A_2}^T).$$
Instead, when $\mathcal{R}(Q) = \{A\}$ we have
$$\Psi(Q) \;=\; \Sigma v_A.\, (v_A \cdot e_Q).$$
And when $\mathcal{R}(Q) = \{\}$ we have
$$\Psi(Q) \;=\; e_Q.$$
By using the inductive hypothesis $(*)$ one can check that $\Psi(Q)$ works in each case as expected. □

E.3 Weighted logics and FO-MATLANG

We prove proposition 6.7:

Proposition 6.7. *Weighted logics over Γ and* FO-MATLANG *over $\mathcal{S}$ have the same expressive power. More specifically,*

- *for each* FO-MATLANG *expression e over $\mathcal{S}$ such that $\mathcal{S}(e) = (1,1)$, there exists a WL-formula $\Phi(e)$ over $WL(\mathcal{S})$ such that for every instance $\mathcal{I}$ of $\mathcal{S}$, $[\![e]\!](\mathcal{I}) = [\![\Phi(e)]\!]_{WL(\mathcal{I})}$.*
- *for each WL-formula φ over Γ without free variables, there exists a* FO-MATLANG *expression $\Psi(\varphi)$ such that for any structure $\mathcal{A}$ over $Mat(\Gamma)$, $[\![\varphi]\!]_{\mathcal{A}} = [\![\Psi(\varphi)]\!](Mat(\mathcal{A}))$.*

Proof. Both directions are proved by induction on the structure of expressions.

(FO-MATLANG to WL) First, let $\mathcal{S} = (\mathcal{M}, \mathtt{size})$ be a schema of square matrices, that is, there exists an α such that $\mathtt{size}(V) \in \{1, \alpha\} \times \{1, \alpha\}$ for every $V \in \mathcal{M}$. Define the relational vocabulary $\mathrm{WL}(\mathcal{S}) = \{R_V \mid V \in \mathcal{M}\}$ such that $\mathrm{arity}(R_V) = 2$ if $\mathtt{size}(V) = (\alpha, \alpha)$, $\mathrm{arity}(R_V) = 1$ if $\mathtt{size}(V) \in \{(\alpha, 1), (1, \alpha)\}$, and $\mathrm{arity}(R_V) = 0$ otherwise. Then given a matrix instance $\mathcal{I} = (\mathcal{D}, \mathtt{mat})$ over $\mathcal{S}$ with $\mathcal{D}(\alpha) = n$ define the structure $\mathrm{WL}(\mathcal{I}) = (\{1, \ldots, n\}, \{R_V^{\mathcal{I}}\})$ such that $R_V^{\mathcal{I}}(i,j) = \mathtt{mat}(V)_{i,j}$ if $\mathtt{size}(V) = (\alpha, \alpha)$, $R_V^{\mathcal{I}}(i) = \mathtt{mat}(V)_i$ if $\mathtt{size}(V) \in \{(\alpha, 1), (1, \alpha)\}$, and $R_V^{\mathcal{I}} = \mathtt{mat}(V)$ if $\mathtt{size}(V) = (1,1)$.

Similar to the proof of Proposition 6.3, for each expression $e(v_1, \ldots, v_k)$ of type (α, α) we must encode in WL the α and the vector variables $v_1, \ldots, v_k$. For this, we use variables x_α^{row}, x_α^{col}, and x_{v_i} for each variable $v_1, \ldots, v_k$. Then we use the following inductive hypothesis (similar to Proposition 6.3):

- If $e(v_1, \ldots, v_k)$ is of type (α, α) then there exists a WL formula $\varphi_e(x_\alpha^{\mathsf{row}}, x_\alpha^{\mathsf{col}}, x_{v_1}, \ldots, x_{v_k})$ such that
$$[\![\varphi_e]\!]_{\mathrm{WL}(\mathcal{I})}(\sigma) \;=\; [\![e]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])_{i,j}$$
for assignment σ with $\sigma(x_\alpha^{\mathsf{row}}) = i$, $\sigma(x_\alpha^{\mathsf{col}}) = j$ and $\sigma(x_{v_s}) = i_s$ for $s = 1, \ldots, k$.
- If $e(v_1, \ldots, v_k)$ is of type $(\alpha, 1)$ then there exists a WL formula $\varphi_e(x_\alpha^{\mathsf{row}}, x_{v_1}, \ldots, x_{v_k})$ such that
$$[\![\varphi_e]\!]_{\mathrm{WL}(\mathcal{I})}(\sigma) \;=\; [\![e]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])_i$$
for assignment σ with $\sigma(x_\alpha^{\mathsf{row}}) = i$ and $\sigma(x_{v_s}) = i_s$ for $s = 1, \ldots, k$. And similarly for when e is type $(1, \alpha)$.

- If $e(v_1, \ldots, v_k)$ is of type $(1,1)$ then there exists a WL formula $\varphi_e(x_{v_1}, \ldots, x_{v_k})$ such that

$$[\![\varphi_e]\!]_{\mathrm{WL}(\mathcal{I})}(\sigma) \;=\; [\![e]\!](\mathcal{I}[v_1 \leftarrow b_{i_1}, \ldots, v_k \leftarrow b_{i_k}])$$

  for assignment σ with $\sigma(x_{v_s}) = i_s$ for $s = 1, \ldots, k$.

If we prove the previous statement we are done, because the last bullet is what we want to show when e has no free vector variables. Then rest of the proof is to go by induction on the structure of FO-MATLANG expressions. For a WL-formula φ and FO-variables x, y, we will write $\varphi[x \mapsto y]$ the formula φ when x is replaced with y all over the formula (syntactically). Let e be a FO-MATLANG expression.

- If $e := V$ and $\mathcal{S}(e) = (\alpha, \alpha)$ then $\varphi_e := R_V(x_\alpha^{\mathsf{row}}, x_\alpha^{\mathsf{col}})$. Similarly, if $\mathcal{S}(e)$ is of type $(\alpha, 1)$, $(1, \alpha)$, or $(1,1)$, then $\varphi_e := R_V(x_\alpha^{\mathsf{row}})$, $\varphi_e := R_V(x_\alpha^{\mathsf{col}})$, and $\varphi_e := R_V$, respectively.

- If $e := v$, for $v \in \{v_1, \ldots, v_k\}$, and $\mathcal{S}(v) = (\alpha, 1)$ then $\varphi_e := x_\alpha^{\mathsf{row}} = x_v$. Similarly, if $\mathcal{S}(v) = (1, \alpha)$ then $\varphi_e := x_\alpha^{\mathsf{col}} = x_v$.

- if $e := e_1^T$ and $\mathcal{S}(e) = (\alpha, \alpha)$ then

$$\varphi_e := \varphi_{e_1}[x_\alpha^{\mathsf{row}} \mapsto x_\alpha^{\mathsf{col}}, x_\alpha^{\mathsf{col}} \mapsto x_\alpha^{\mathsf{row}}].$$

  Similarly, if $\mathcal{S}(e)$ is equal to $(\alpha, 1)$ or $(1, \alpha)$ then $\varphi_e := \varphi_{e_1}[x_\alpha^{\mathsf{row}} \mapsto x_\alpha^{\mathsf{col}}]$ and $\varphi_e := \varphi_{e_1}[x_\alpha^{\mathsf{col}} \mapsto x_\alpha^{\mathsf{row}}]$, respectively.

- If $e = e_1 + e_2$ with $\mathcal{S}(e_1) = \mathcal{S}(e_2)$, then $\varphi_e := \varphi_{e_1} \oplus \varphi_{e_2}$.

- If $e = f_\odot(e_1, \ldots, e_k)$ with $\mathcal{S}(e_i) = \mathcal{S}(e_j)$ for all $i, j \in [1, k]$, then $\varphi_e := \varphi_{e_1} \odot \varphi_{e_2} \cdots \odot \varphi_{e_k}$.

- If $e = e_1 \cdot e_2$ with $\mathcal{S}(e_1) = \mathcal{S}(e_2) = (\alpha, \alpha)$, then $\varphi_e := \Sigma y.\, \varphi_{e_1}[x_\alpha^{\mathsf{col}} \mapsto y] \odot \varphi_{e_2}[x_\alpha^{\mathsf{row}} \mapsto y]$ where y is a fresh variable not mentioned in φ_{e_1} or φ_{e_2}. Instead, if $\mathcal{S}(e_1) = (\alpha', 1)$ and $\mathcal{S}(e_2) = (1, \alpha'')$ with $\alpha', \alpha'' \in \{\alpha, 1\}$, then $\varphi_e := \varphi_{e_1} \odot \varphi_{e_2}$.

- If $e = \Sigma v.e_1(v)$, then we define $\varphi_e := \Sigma x_v.\, \varphi_{e_1}(x_v)$.

- If $e = \Pi^\circ v.e_1(v)$, then $\varphi_e := \Pi x_v.\, \varphi_{e_1}(x_v)$.

From the construction it is now straightforward to check that the inductive hypothesis holds for all cases. To conclude this direction, we have to define $\Phi(e) := \varphi_e$ for every expression e and we are done.

(WL to FO-MATLANG) We now encode weighted structures into matrices and vectors. Let Γ be a relational vocabulary where $\mathrm{arity}(R) \leq 2$. Define $\mathrm{Mat}(\Gamma) = (\mathcal{M}_\Gamma, \mathtt{size}_\Gamma)$ such that $\mathcal{M}_\Gamma = \{V_R \mid R \in \Gamma\}$ and $\mathtt{size}_\Gamma(V_R)$ is equal to (α, α), $(\alpha, 1)$, or $(1,1)$ if $\mathrm{arity}(R) = 2$, $\mathrm{arity}(R) = 1$, or $\mathrm{arity}(R) = 0$, respectively, for some $\alpha \in \mathtt{Symb}$. Similarly, let $\mathcal{A} = (A, \{R^{\mathcal{A}}\}_{R \in \Gamma})$ be a structure with $A = \{a_1, \ldots, a_n\}$, ordered arbitrarily. Then we define the matrix instance $\mathrm{Mat}(\mathcal{A}) = (\mathcal{D}, \mathtt{mat})$ such that $\mathcal{D}(\alpha) = n$, $\mathtt{mat}(V_R)_{i,j} = R^{\mathcal{A}}(a_i, a_j)$ if $\mathrm{arity}(R) = 2$, $\mathtt{mat}(V_R)_{i,1} = R^{\mathcal{A}}(a_i)$ if $\mathrm{arity}(R) = 1$, and $\mathtt{mat}(V_R)_{1,1} = R^{\mathcal{A}}$ otherwise.

Similar to the above direction, we have to encode the FO variables of a formula φ with vector variables in the equivalent FO-MATLANG expression e_φ. For this, for each FO variable x we define a vector variable v_x of type $(\alpha, 1)$. Then for each formula $\varphi(x_1, \ldots, x_k)$ we define an expression $e_\varphi(v_{x_1}, \ldots, v_{x_k})$ of type $(1,1)$ such that for every assignment σ of $x_1, \ldots, x_k$ we have:

$$[\![e_\varphi]\!](\mathrm{Mat}(\mathcal{A})[v_{x_1} \leftarrow b_{i_1}, \ldots, v_{x_1} \leftarrow b_{i_k}]) \;=\; [\![\varphi]\!]_{\mathcal{A}}(\sigma)$$

such that $\sigma(x_s) = i_s$ for every $s \leq k$. Note that when the formula has no free variables, the proof of the proposition is shown. Finally, we proceed by induction over the formula φ over Γ.

- If $\varphi := x = y$, then $e_\varphi := v_x^T \cdot v_y$.

- If $\varphi := R(x, y)$, then $e_\varphi := v_x^T \cdot V_R \cdot v_y$. Similarly, if $\varphi := R(x)$ or $\varphi := R$, then $e_\varphi := V_R^T \cdot v_x$ and $e_\varphi := V_R$, respectively.

- If $\varphi = \varphi_1 \oplus \varphi_2$, then $e_\varphi := e_{\varphi_1} + e_{\varphi_2}$.

- If $\varphi = \varphi_1 \odot \varphi_2$, then $e_\varphi := f_\odot(e_{\varphi_1}, e_{\varphi_2})$.

- If $\varphi = \Sigma x.\, \varphi_1$, then $e_\varphi := \Sigma v_x.\, e_{\varphi_1}$.

- If $\varphi = \Pi^\circ x.\varphi_1$, then $e_\varphi := \Pi v_x.\, e_{\varphi_1}$.

The inductive hypothesis can be proved following the above construction. To finish the proof, we define $\Psi(\varphi) := e_\varphi$ and the proposition is shown.

□

E.4 Matrix inversion in prod-MATLANG extended with order

We conclude by verifying that the fragment defined in Section 6.3, i.e, prod-MATLANG extended with order and $f_{>0}$, can perform matrix inversion and compute the determinant. To this aim, we verify that all order predicates in Section B.1 can be derived using Σ, Π, $f_{>0}$ and $e_{S_<}$. Given this, it suffices to observe that Csanky's algorithm, as shown in Section C.3, only relies on expressions using Σ and Π and order information on canonical vectors and $f_/$. As consequence, our fragment can perform matrix inversion and compute the determinant.

It remains to show that if we have $e_{S_<}$, using Σ and Π and $f_{>0}$ we can can define all order predicates from Section B.1. We note that due to the restricted for-loops in Σ and Π, we do not have access to the intermediate result in the iterations and as such, it is unclear whether order information can be computed. This is why we assume access to $e_{S_<}$.

We first remark that if we have $e_{S_<}$, we can also obtain $e_{S_\leq}$ by adding e_{Id}. Hence, we can compute succ and succ^+ as well. Furthermore,

$$e_{\mathsf{min}} := \Sigma v.\ [\Pi w.\mathsf{succ}(w, v)] \times v.$$

$$e_{\mathsf{max}} := \Sigma v.\ [\Pi w.\ (1 - \mathsf{succ}(w, v))] \times v.$$

Both expressions are only using Σ and Π and succ, so are in our fragment. Furthermore, if we have $f_{>0}$ then we can define

$$e_{\mathsf{Pred}} := e_{S_<} - f_{>0}(e_{S_<}^2)$$

Also, recall that $e_{\mathsf{Next}} := e_{\mathsf{Pred}}^T$. As a consequence, we can now define $\mathsf{prev}(v)$ and $\mathsf{next}(v)$ as in B.1. Similarly, it is readily verified that also $e_{\mathsf{getPrevMatrix}}(V)$, $e_{\mathsf{getNextMatrix}}(V)$, $e_{\mathsf{min}+i}$ and $e_{\mathsf{max}+i}$ can be expressed in our fragment.