DISSERTATION

ALGEBRAIC INVARIANTS OF TENSORS: ALGORITHMS AND DECOMPOSITIONS

Submitted by

Chris Liu

Department of Mathematics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2026

Doctoral Committee:

Advisor: James B. Wilson

Alexander Hulpke
Christopher Peterson
Louis-Noël Pouchet

ABSTRACT

ALGEBRAIC INVARIANTS OF TENSORS: ALGORITHMS AND DECOMPOSITIONS

This dissertation studies algorithmic techniques and decomposition theories for the algebraic invariants of tensors. Tensors are multiway grids of numbers encoding multilinear data, with widespread use across the sciences. Algebraic invariants of tensors, such as its derivation algebra, give structural insights to the tensor. The first result is a novel algorithm for computing these algebraic invariants. For tensors satisfying a regularity condition, our algorithm computes their derivation algebra in $O(n^{4.5})$ arithmetic operations, improving on the currently known $O(n^7)$ algorithm and coming within a factor of $n^{1/2}$ of the deterministic verification cost. The second result is a decomposition theorem for tensors arising as a product of two tensors. This theorem characterizes the derivation algebra of such a tensor using the algebraic invariants of its factors.

ACKNOWLEDGEMENTS

I would first like to acknowledge my parents - Jiaqiong Joan Ran (冉家琼) and Yexin Liu (刘业新). Your love, dedication, and continued passion for learning has forged me to where I am today.

I am grateful to my friends from all the phases of my life in Calgary, New York, the Bay Area, and now Colorado. My life would not be in color without you all, either in person or remotely.

In this journey, I also have my eternal gratitude towards my advisor James Wilson. The opportunties to learn about the many fields of math, sample out different research questions, and collaboratively do math across the world - they are memories I will take with me forever.

Lastly, thank you, my love Adrianne, for being you, your creative mind, and loving spirit.

I acknowledge the use of generative AI in the writing of this dissertation. Specifically, GPT-5.4 was used for proofreading and for assistance in writing the code used in Section 3.5.

TABLE OF CONTENTS

Chapter 1

Introduction

A scalar records one measurement. A vector often records a magnitude and direction, so it has one axis, where an axis means an independent mode of access. A matrix has two axes, and many scientific and engineering settings record measurements with three or more axes. Tensors encapsulate all of these abstractions in one unifying concept. For instance, a matrix may be called a 2-tensor, a cube of numbers a 3-tensor, and so on. The number of axes is also called the *valence* of the tensor. This tensor abstraction is used in various disciplines to record information for some fixed reference frame. Some complementary perspectives studying tensors include [Bro97] [KB09] [Lan12] [RS18] [DLDMV00] [Tuc66]. Below we give two real world examples.

The first is the stress tensor in mechanics. For an object such as a steel beam supporting traffic, we can track how the internal force depends on the orientation of a surface element inside the material. At each point, with respect to a coordinate system, we have a 3×3 Cauchy stress tensor, where diagonal entries capture normal compression and tension, while off-diagonal entries capture shear [Irg08, p. 52]. Higher valence tensors are needed to model constitutive laws. For instance the generalized Hooke's Law for the stress tensor T and strain tensor E is a relationship given by $T_{ij} = S_{ijkl}E_{kl}$, where S is a 4-tensor called the elasticity tensor [Irg08, Eq 7.2.23].

The second comes from recommender systems, of which the Netflix tensor is the most famous. The starting point is to store a user $\times$ movie matrix of ratings. The factorization of this matrix was central to recommenders competing for the Netflix Prize [KBV09]. But recommendation systems may need more axes, for example, by considering preferences over time, and considering multiple devices. This necessitates higher valence tensors, for instance with axes user $\times$ movie $\times$ time. Tensor factorization methods explicitly model and work with these higher valence tensors [Kar+10].

Now we motivate studying tensors using algebraic techniques. For the rest of this introduction $\mathbb{F}$ will denote a field. Let A be a $\mathbb{F}$-algebra, meaning it is a $\mathbb{F}$-vector space with a bilinear multiplication $A \times A \rightarrowtail A$ ($\rightarrowtail$ for bilinear). After choosing a basis of the underlying vector space, the structure constants associated to the multiplication form a 3-tensor. The algebra A has invariants that are independent of the basis choice. For instance, does A have zero divisors or split as a direct product of smaller algebras? Looking at the multiplication tensor for A allows us to understand them. For example, let $\mathbb{F} = \mathbb{F}_{13}$ and let $t \in \mathbb{F}^{2\times2\times2}$ be the tensor with slices

$$t_{**1} = \begin{bmatrix} 10 & 5 \\ 5 & 6 \end{bmatrix} \qquad t_{**2} = \begin{bmatrix} 11 & 8 \\ 8 & 0 \end{bmatrix}.$$

This is the multiplication tensor of the algebra $\mathbb{F}[x]/(x^2 - 3x + 2)$ in the basis $\{3 + 4x, 7 + 10x\}$. If we instead use the basis $\tilde{b} = \{1, x\}$, then using the basis change matrix with columns $3 + 4x$ and $7 + 10x$ written in the $\tilde{b}$ coordinates, we get a tensor $\tilde{t}$ with slices

$$\tilde{t}_{**1} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \qquad \tilde{t}_{**2} = \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix}.$$

We display it more suggestively to showcase the multiplication rules.

	1	x
1	1	x
x	x	$3x - 2$

Finally, $x^2 - 3x + 2$ factors as $(x - 1)(x - 2)$ over $\mathbb{F}$. If $e = x - 1$ and $f = 2 - x$, then $e + f = 1$, $e^2 = e$, $f^2 = f$, and $ef = 0$. Hence e and f are complementary orthogonal idempotents, and by the Chinese Remainder Theorem, $\mathbb{F}[x]/(x^2 - 3x + 2) \cong \mathbb{F}[x]/(x - 1) \times \mathbb{F}[x]/(x - 2) \cong \mathbb{F} \times \mathbb{F}$. In the adapted basis $\{e, f\}$, the multiplication tensor is the table

$$\begin{array}{|c|cc|} \hline & e & f \\ \hline e & e & 0 \\ f & 0 & f \\ \hline \end{array}.$$

The above example is found through an algebraic invariant of the tensor called the centroid [Wil16, Proposition 6.4] [Mya90, Proposition 3.1]. More generally, the structure in the tensors representing not just multiplication tables is seen through these algebraic invariants. For example over $\mathbb{F} = \mathbb{F}_{997}$, let $t \in \mathbb{F}^{2\times 2\times 2}$ with

$$t_{**1} = \begin{bmatrix} 978 & 724 \\ 700 & 739 \end{bmatrix}, \qquad t_{**2} = \begin{bmatrix} 682 & 660 \\ 51 & 689 \end{bmatrix}$$

Using the adjoint algebra (middle nucleus), we find change of bases matrices M and N satisfying

$$M = \begin{bmatrix} 643 & 1 \\ 421 & 1 \end{bmatrix}, \; N = \begin{bmatrix} 111 & 836 \\ 886 & 836 \end{bmatrix}, \; M(t_{**1})N^{-1} = \begin{bmatrix} 516 & 0 \\ 0 & 443 \end{bmatrix}, \; M(t_{**2})N^{-1} = \begin{bmatrix} 35 & 0 \\ 0 & 332 \end{bmatrix}$$

so in the language of [BKW24], t splits into two independent $1\times 1\times 2$ face blocks. The Multilinear Algebra library in Magma [BCP97] and the Julia package OpenDleto [The26] provide algorithms to find these algebraic invariants and decompose tensors.

1.1 Main Results

We briefly describe results on the two main problems tackled by my dissertation.

The first is on computing of algebraic invariants of tensors, which have been key to discovering basis-independent cluster patterns in tensors [BKW24], decomposing p-groups [Wil09], computing endomorphisms of modules [BL08], and advancing isomorphism testing [IQ19], [BMW17] [BW12] [BMW22].

Suppose we are given a 3-tensor $T \in \mathbb{F}^{m\times n\times r}$ which represents trilinear measurements in a fixed basis. Let the (i, j, k)th component be denoted $T[i, j, k]$, and let $T_1, \ldots, T_r \in \mathbb{F}^{m\times n}$ be slices using the Matlab inspired slice notation, where $T_i := T[:,:,i]$. For $X \in \mathbb{F}^{m\times m}$, $Y \in \mathbb{F}^{n\times n}$, and $Z = (z_{ji}) \in \mathbb{F}^{r\times r}$, define tensors XT, TY, and T^Z by their slices

$$(XT)_i := XT_i, \qquad (TY)_i := T_iY, \qquad (T^Z)_i := \sum_{j=1}^{r} T_j z_{ji}.$$

Equations for the key algebraic invariants we study are as

$$\begin{aligned}
\text{(left nucleus)} \quad & \mathcal{L}(T) := && \{(X,Z) \in \mathbb{F}^{m\times m} \times \mathbb{F}^{r\times r} : XT = T^Z\},\\
\text{(middle nucleus)} \quad & \mathcal{M}(T) := && \{(X,Y) \in \mathbb{F}^{m\times m} \times \mathbb{F}^{n\times n} : XT = TY\},\\
\text{(right nucleus)} \quad & \mathcal{R}(T) := && \{(Y,Z) \in \mathbb{F}^{n\times n} \times \mathbb{F}^{r\times r} : TY = T^Z\},\\
\text{(centroid)} \quad & \mathcal{C}(T) := && \{(X,Y,Z) \in \mathbb{F}^{m\times m} \times \mathbb{F}^{n\times n} \times \mathbb{F}^{r\times r} : XT = TY = T^Z\},\\
\text{(derivation)} \quad & \mathrm{Der}(T) := && \{(X,Y,Z) \in \mathbb{F}^{m\times m} \times \mathbb{F}^{n\times n} \times \mathbb{F}^{r\times r} : XT + TY = T^Z\}.
\end{aligned} \tag{1.1}$$

Each corresponds to a linear system. Our first main result, Theorem B, shows that for tensors $T \in \mathbb{F}^{n\times n\times n}$ satisfying a regularity condition, the derivation algebra of T can be computed in $O(n^{4.5})$ time. This is an improvement over the currently known $O(n^7)$ algorithm.

The second result is about understanding structure when we combine two tensors into one via multiplication. Multiplication is a natural way to inductively build up data from smaller pieces. For instance, if A and B are $\mathbb{F}$-algebras, then $A \otimes B$ is again a $\mathbb{F}$-algebra. An example of this is the extensions of scalars construction $A \otimes L$ given a $\mathbb{F}$-algebra A and a field extension $L/\mathbb{F}$.

We define a product of tensors that serves the same role by combining interactions while preserving multilinearity. Our definition is most naturally stated for bilinear maps $b : U \times V \rightarrowtail W$. After choosing bases of U, V, and W, the map b is encoded by its multiplication table, and stored as a 3-tensor. We define the product bilinear map for $s : U_2 \times U_1 \rightarrowtail U_0$ and $t : V_2 \times V_1 \rightarrowtail V_0$ as

$$s \otimes t : (U_2 \otimes V_2) \times (U_1 \otimes V_1) \rightarrowtail U_0 \otimes V_0$$

on pure tensors by $(s \otimes t)(u \otimes x,\ v \otimes y) := s(u,v) \otimes t(x,y)$, and extended bilinearly.

Example 1.1. A classical result for Hamilton's quaternion algebra $\mathbb{H}$ is the $\mathbb{C}$-algebra isomorphism $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \cong M_2(\mathbb{C})$, given on basis elements by

$$1 \otimes 1 \mapsto \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad i \otimes 1 \mapsto \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad j \otimes 1 \mapsto \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad k \otimes 1 \mapsto \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Let $u = i_{\mathbb{H}} \otimes i_{\mathbb{C}}$. Then $u^2 = (i_{\mathbb{H}})^2 \otimes (i_{\mathbb{C}})^2 = (-1) \otimes (-1) = 1$, so u is a splitting element. Hence $e = \frac{1+u}{2}$ and $f = \frac{1-u}{2}$ are orthogonal idempotents with $e^2 = e$, $f^2 = f$, $ef = fe = 0$, and $e + f = 1$.

Concretely, u corresponds to the matrix $E_{22} - E_{11}$, so e corresponds to E_{22} and f to E_{11}, and we recover E_{12} and $-E_{21}$ by multiplying with the image of $j \otimes 1$. Thus extending the coefficients can turn a division algebra into a matrix algebra. This splitting phenomenon is studied by the Brauer group of central simple algebras (simple finite-dimensional associative algebras whose center is the base field) [Pie82].

Example 1.2. Below we give a bigger computational example using the multiplication structure. Consider the algebras $\mathbb{F}[x]/(x^2+1)$ and $\mathbb{F}[y]/(y^2+5y+10)$. Their tensor product is the algebra $\mathbb{F}[x,y]/(x^2+1, y^2+5y+10)$.

Similarly, their multiplication tables are each given as a $2 \times 2 \times 2$ table, and their combined multiplication is as shown below.

$$\begin{array}{|c|cc|} \hline & 1 & x \\ \hline 1 & 1 & x \\ x & x & -1 \\ \hline \end{array} \otimes \begin{array}{|c|cc|} \hline & 1 & y \\ \hline 1 & 1 & y \\ y & y & -5y-10 \\ \hline \end{array} = \begin{array}{|c|cccc|} \hline & 1 & x & y & xy \\ \hline 1 & 1 & x & y & xy \\ x & x & -1 & xy & -y \\ y & y & xy & -5y-10 & -5xy-10x \\ xy & xy & -y & -5xy-10x & 5y+10 \\ \hline \end{array}$$

The product bilinear map's multiplication tensor is on the right, displayed as a $4 \times 4 \times 4$ tensor, with four slices indexed by the basis $\{1, x, y, xy\}$.

Our second main result, Theorem D, states that given bilinear maps s and t, we can understand $\mathrm{Der}(s \otimes t)$ by individually understanding the algebraic invariants of s and t. This allows for characterizing $\mathrm{Der}(s \otimes t)$ using much less information than what one would otherwise expect. The bilinear map $s \otimes t$ preserves multilinearity at the cost of multiplicative dimension growth. For instance, given $s : \mathbb{F}^m \times \mathbb{F}^n \rightarrowtail \mathbb{F}^l$ and $t : \mathbb{F}^p \times \mathbb{F}^q \rightarrowtail \mathbb{F}^r$, the product $s \otimes t : (\mathbb{F}^m \otimes \mathbb{F}^p) \times (\mathbb{F}^n \otimes \mathbb{F}^q) \rightarrowtail \mathbb{F}^l \otimes \mathbb{F}^r$ requires $(mp)(nq)(lr)$ coordinates to describe. Yet our theorem allows for the individual understanding of the algebraic invariants of s , which is specified using only mnl coordinates, and t, which is specified using only pqr coordinates, to fully characterize $\mathrm{Der}(s \otimes t)$.

This result extends earlier results in matrix algebras by Benkart and Osborn [BO81] as well as in non-associative algebras by Brešar [Bre17]. Recall that for an algebra A, its derivation, $\mathrm{Der}(A)$, consists of linear maps $\delta \in \mathrm{Hom}(A, A)$ such that $\delta(xy) = (\delta(x))y + x(\delta(y))$. Benkart and Osborn show that $\mathrm{Der}(\mathbb{M}_n(A))$ is a sum of two pieces. First, derivations of A applied entrywise to matrices, and second, inner derivations coming from commutators over the nucleus of A - meaning in the algebra $\mathbb{M}_n(\mathrm{Nuc}(A))$ [BO81, Corollary 4.9].

As we show in Remark 5.5, the Benkart-Osborn result is a special case of Theorem D when s is the multiplication tensor of the algebra A and t is the multiplication tensor of $\mathbb{M}_n(\mathbb{F})$. Then $s \otimes t$ is the multiplication tensor of $\mathbb{M}_n(A) \cong A \otimes_{\mathbb{F}} \mathbb{M}_n(\mathbb{F})$, and the diagonal pieces of the tensor derivations of $s \otimes t$ correspond to the algebra derivation pieces of $\mathbb{M}_n(A)$.

Chapter 2

Preliminaries and Prior Work

Given $\mathbb{F}$-vector spaces U, V and W, a function $f : U \times V \to W$ is additive if $f(u + u', v) = f(u, v) + f(u', v)$, and $f(u, v + v') = f(u, v) + f(u, v')$ for all u, u' in U and v, v' in V. A function f is $\mathbb{F}$-bilinear if f is additive and satisfies $f(cu, v) = f(u, cv) = cf(u, v)$ for all $c \in \mathbb{F}$, $u \in U$, and $v \in V$. We write $f : U \times V \rightarrowtail W$ ($\rightarrowtail$ for bilinear). When context is clear, we elide the prefix $\mathbb{F}$.

Example 2.1. Examples of bilinear maps include:

1. If A is an $\mathbb{F}$-algebra, then multiplication map $\mu : A \times A \rightarrowtail A$ is bilinear.

2. For positive integers m, n, p, the rectangular matrix multiplication map $\mathbb{F}^{m\times n} \times \mathbb{F}^{n\times p} \rightarrowtail \mathbb{F}^{m\times p}$ given by $(X, Y) \mapsto XY$ is bilinear.

3. For $U = \mathbb{F}^2$ and $V = \mathbb{F}^3$, any bilinear form $b : U \times V \rightarrowtail \mathbb{F}$ can be represented by a matrix $B \in \mathbb{F}^{2\times 3}$ via $b(u, v) = u^T Bv$ for a fixed choice of bases of U and V.

The above extends to explain $\mathbb{F}$-trilinear and $\mathbb{F}$-multilinear in general. Given a bilinear map $b : U \times V \rightarrowtail W$, fix bases for U, V, and W. Then evaluating b on basis vectors of U and V, and evaluating the output against the dual basis in W^*, we may identify b with a trilinear form in $\mathbb{F}^m \times \mathbb{F}^n \times \mathbb{F}^l \rightarrowtail \mathbb{F}$ where $m = \dim U$, $n = \dim V$, $l = \dim W$. This trilinear form is identified as a 3-tensor (as an element of $\mathbb{F}^{m\times n\times l}$), and thus throughout this dissertation, $\mathbb{F}$-bilinear, trilinear, and multilinear maps are also called *tensors*.

Definition 2.2. For vector spaces U and V, the *tensor product* of U and V is the vector space $U \otimes V$ with the canonical map $\varphi : U \times V \rightarrowtail U \otimes V$, such that for every bilinear map f with domain $U \times V$, there is a unique induced linear map $\hat{f}$ satisfying $f = \hat{f} \circ \varphi$. [Gre12, p. 10]

For vectors $s \in U$ and $t \in V$, the notation $s \otimes t$ denotes the image of (s, t) under φ. The elements $s \otimes t$ are called *pure tensors*, and $U \otimes V$ is spanned by pure tensors.

Remark 2.3. Let $U = \mathbb{F}^m$ and $V = \mathbb{F}^n$ with their standard bases. Then $U \otimes V$ may be realized as $m \times n$ matrices by sending $e_i \otimes f_j$ to the elementary matrix E_{ij}.

Next we describe subspaces of a tensor product space. For subsets $A \subseteq U$ and $B \subseteq V$, define

$$[A \otimes B] \coloneqq \operatorname{span}\{a \otimes b : a \in A,\ b \in B\} \leq U \otimes V.$$

The notation $[A \otimes B]$ indicates this space is built internally as a subspace of $U \otimes V$. This distinction is needed as if $A \leq U$ and $B \leq V$ are subspaces, then there is also the *external* tensor product space $(A \otimes B, \varphi : A \times B \rightarrowtail A \otimes B)$.

Lemma 2.4. Let U, V be vector spaces, and $A \leq U$ and $X \leq V$ be subspaces. Then $[A \otimes X] = [U \otimes X] \cap [A \otimes V]$ as subspaces of $U \otimes V$.

Proof. Let $\iota_A : A \hookrightarrow U$ and $\iota_X : X \hookrightarrow V$ be the inclusion maps. The intersection $[U \otimes X] \cap [A \otimes V]$ is the pullback of the external tensor product spaces $U \otimes X$ and $A \otimes V$ along their inclusion maps into $U \otimes V$. Hence it suffices to show the following square is also a pullback:

$$\begin{tikzcd}
A \otimes X \arrow[r, "\iota_A \otimes \mathrm{id}_X"] \arrow[d, "\mathrm{id}_A \otimes \iota_X"'] & U \otimes X \arrow[d, "\mathrm{id}_U \otimes \iota_X"] \\
A \otimes V \arrow[r, "\iota_A \otimes \mathrm{id}_V"'] & U \otimes V
\end{tikzcd}$$

Let T be a vector space with maps $q_1 : T \to A \otimes V$ and $q_2 : T \to U \otimes X$ such that $(\iota_A \otimes \mathrm{id}_V) \circ q_1 = (\mathrm{id}_U \otimes \iota_X) \circ q_2$. We must construct the unique u making both triangles commute in the below diagram.

$$\begin{tikzcd}
T \arrow[dr, dashed, "u"] \arrow[drr, bend left, "q_2"] \arrow[ddr, bend right, "q_1"'] & & \\
& A \otimes X \arrow[r, "\iota_A \otimes \mathrm{id}_X"] \arrow[d, "\mathrm{id}_A \otimes \iota_X"'] & U \otimes X \arrow[d, "\mathrm{id}_U \otimes \iota_X"] \\
& A \otimes V \arrow[r, "\iota_A \otimes \mathrm{id}_V"'] & U \otimes V
\end{tikzcd}$$

This is done by showing $\mathrm{Im}(q_1) \leq \mathrm{Im}(\mathrm{id}_A \otimes \iota_X)$. Since $\mathrm{id}_A \otimes \iota_X$ is injective, any such factorization is unique. Since tensoring over a field preserves exact sequences, tensoring $0 \to X \xrightarrow{\iota_X} V \xrightarrow{q} V/X \to 0$ with A gives

$$0 \to A \otimes X \xrightarrow{\mathrm{id}_A \otimes \iota_X} A \otimes V \xrightarrow{\mathrm{id}_A \otimes q} A \otimes (V/X) \to 0.$$

So we must show $\mathrm{Im}(q_1) \leq \mathrm{Ker}(\mathrm{id}_A \otimes q)$, or $(\mathrm{id}_A \otimes q) \circ q_1 = 0$. This fact follows by the below commutative square.

$$\begin{tikzcd}
A \otimes V \arrow[r, "\iota_A \otimes \mathrm{id}_V"] \arrow[d, "\mathrm{id}_A \otimes q"'] & U \otimes V \arrow[d, "\mathrm{id}_U \otimes q"] \\
A \otimes (V/X) \arrow[r, "\iota_A \otimes \mathrm{id}_{V/X}"'] & U \otimes (V/X)
\end{tikzcd}$$

Indeed, as elements in $\mathrm{Hom}(A \otimes V, U \otimes (V/X))$,

$$(\mathrm{id}_U \otimes q) \circ (\iota_A \otimes \mathrm{id}_V) = (\iota_A \otimes \mathrm{id}_{V/X}) \circ (\mathrm{id}_A \otimes q), \tag{2.1}$$

because both sides send the pure tensor $a \otimes v$ to $\iota_A(a) \otimes q(v)$.

Applying $\mathrm{id}_U \otimes q$ to $(\iota_A \otimes \mathrm{id}_V) \circ q_1 = (\mathrm{id}_U \otimes \iota_X) \circ q_2$, the right-hand side vanishes because $(\mathrm{id}_U \otimes q) \circ (\mathrm{id}_U \otimes \iota_X) = \mathrm{id}_U \otimes (q \circ \iota_X) = 0$. Hence $(\mathrm{id}_U \otimes q) \circ (\iota_A \otimes \mathrm{id}_V) \circ q_1 = 0$, and by (2.1), $(\iota_A \otimes \mathrm{id}_{V/X}) \circ (\mathrm{id}_A \otimes q) \circ q_1 = 0$. Since $\iota_A \otimes \mathrm{id}_{V/X}$ is injective, $(\mathrm{id}_A \otimes q) \circ q_1 = 0$ as needed. Thus we find the unique u such that $q_1 = (\mathrm{id}_A \otimes \iota_X) \circ u$.

Finally, for u to make the upper triangule commute, we must show $q_2 = (\iota_A \otimes \mathrm{id}_X) \circ u$. We have

$$\begin{aligned}
(\mathrm{id}_U \otimes \iota_X) \circ (\iota_A \otimes \mathrm{id}_X) \circ u &= (\iota_A \otimes \mathrm{id}_V) \circ (\mathrm{id}_A \otimes \iota_X) \circ u \\
&= (\iota_A \otimes \mathrm{id}_V) \circ q_1 \\
&= (\mathrm{id}_U \otimes \iota_X) \circ q_2,
\end{aligned}$$

Therefore $(\mathrm{id}_U \otimes \iota_X) \circ \big((\iota_A \otimes \mathrm{id}_X) \circ u - q_2\big) = 0$. As $\mathrm{id}_U \otimes \iota_X$ is injective, we get $q_2 = (\iota_A \otimes \mathrm{id}_X) \circ u$ as needed. $\square$

Remark 2.5. For $\iota_A : A \hookrightarrow U$ and $\iota_B : B \hookrightarrow V$ inclusion maps, the induced map $\iota_A \otimes \iota_B : A \otimes B \to U \otimes V$ is injective, and we identify the external tensor product space $A \otimes B$ with its image, which is precisely the span of $\{a \otimes b : a \in A, b \in B\} = [A \otimes B]$. This clarifies the notation $A \otimes B$ when A and B are subspaces of ambient vector spaces.

2.1 Algebras and Actions on bilinear maps

Existing work such as [Jac37], [Mya90], [BW14], and [Wil16] highlights the role of the algebras in Equation (1.1). In this subsection, we expand on that and describe some operations that we need for the future sections.

Let U, V, and W be finite-dimensional $\mathbb{F}$-vector spaces for the remainder of this section. Let $\mathrm{End}(U) = \mathrm{Hom}(U, U)$ be the $\mathbb{F}$-linear endomorphisms of a vector space U, and let $b : U \times V \rightarrowtail W$ be a bilinear map. We will also use the notation $\mathrm{Bil}(U, V; W)$ for the $\mathbb{F}$-vector space of bilinear maps $U \times V \rightarrowtail W$. For $\alpha \in \mathrm{End}(U)$, $\beta \in \mathrm{End}(V)$, and $\gamma \in \mathrm{End}(W)$, define the bilinear maps resulting from operators acting on axes as

$$\begin{aligned}
b \bullet_2 \alpha &: U \times V \rightarrowtail W, & (u, v) &\mapsto b(\alpha u, v),\\
b \bullet_1 \beta &: U \times V \rightarrowtail W, & (u, v) &\mapsto b(u, \beta v),\\
b \bullet_0 \gamma &: U \times V \rightarrowtail W, & (u, v) &\mapsto \gamma(b(u, v)).
\end{aligned}$$

Definition 2.6. The left, middle, and right nuclei of b are respectively

$$\begin{aligned}
\mathcal{L}(b) &\coloneqq \{(\alpha, \gamma) \in \mathrm{End}(U) \times \mathrm{End}(W) : b \bullet_2 \alpha = b \bullet_0 \gamma\},\\
\mathcal{M}(b) &\coloneqq \{(\alpha, \beta) \in \mathrm{End}(U)^{\mathrm{op}} \times \mathrm{End}(V) : b \bullet_2 \alpha = b \bullet_1 \beta\},\\
\mathcal{R}(b) &\coloneqq \{(\beta, \gamma) \in \mathrm{End}(V) \times \mathrm{End}(W) : b \bullet_1 \beta = b \bullet_0 \gamma\}.
\end{aligned}$$

Here $\mathrm{End}(U)^{\mathrm{op}}$ is the opposite algebra of $\mathrm{End}(U)$, meaning composition order is reversed. The reason for using $\mathrm{End}(U)^{\mathrm{op}}$ in $\mathcal{M}(b)$ is that if $(\alpha,\beta),(\alpha',\beta') \in \mathcal{M}(b)$, then $b((\alpha'\alpha)u,v) = b(\alpha u,\beta' v) = b(u,(\beta\beta')v)$, so the product is $(\alpha,\beta)(\alpha',\beta') = (\alpha'\alpha,\beta\beta')$. Notice $\mathcal{M}(b)$ is the coordinate free description of the adjoint algebra in Equation (1.1).

The centroid of b is

$$\mathcal{C}(b) \coloneqq \{(\alpha,\beta,\gamma) \in \mathrm{End}(U) \times \mathrm{End}(V) \times \mathrm{End}(W) : b \bullet_2 \alpha = b \bullet_1 \beta = b \bullet_0 \gamma\},$$

and the derivation algebra of b is

$$\mathrm{Der}(b) \coloneqq \{(\alpha,\beta,\gamma) \in \mathrm{End}(U) \times \mathrm{End}(V) \times \mathrm{End}(W) : b \bullet_2 \alpha + b \bullet_1 \beta = b \bullet_0 \gamma\}.$$

Remark 2.7. Each of $\mathcal{L}(b)$, $\mathcal{M}(b)$, $\mathcal{R}(b)$, $\mathcal{C}(b)$, $\mathrm{Der}(b)$ is an $\mathbb{F}$-vector subspace of the corresponding product of endomorphism spaces. Furthermore, $\mathcal{L}(b)$, $\mathcal{M}(b)$, $\mathcal{R}(b)$, and $\mathcal{C}(b)$ are associative algebras, with $\mathcal{C}(b)$ also commutative when b is fully nondegenerate, while $\mathrm{Der}(b)$ is a Lie algebra. [BMW20, Section 1.2].

2.2 Radicals and non-degeneracy of bilinear maps

Let $b : U \times V \rightarrowtail W$ be a bilinear map. For $u \in U$ and $v \in V$, define

$$\begin{aligned} b(u,V) &\coloneqq \{b(u,v') : v' \in V\}, \qquad b(U,v) \coloneqq \{b(u',v) : u' \in U\},\\ b(U,V) &\coloneqq \mathrm{span}\{b(u',v') : u' \in U,\ v' \in V\}. \end{aligned}$$

The *left radical*, *right radical*, and *coradical* of a bilinear map b are respectively $\mathrm{Rad}_L(b) \coloneqq \{u \in U : b(u,V) = 0\}$, $\mathrm{Rad}_R(b) \coloneqq \{v \in V : b(U,v) = 0\}$, and $\mathrm{Corad}(b) \coloneqq W/b(U,V)$.

We call b *fully nondegenerate* if $\mathrm{Rad}_L(b) = \mathrm{Rad}_R(b) = 0$ and $\mathrm{Corad}(b) = 0$. Now define the *axis actions*

$$\begin{aligned}
\Lambda_2^b &: \mathrm{End}(U) \to \mathrm{Bil}(U, V; W), \quad \alpha \mapsto b \bullet_2 \alpha, \\
\Lambda_1^b &: \mathrm{End}(V) \to \mathrm{Bil}(U, V; W), \quad \beta \mapsto b \bullet_1 \beta, \\
\Lambda_0^b &: \mathrm{End}(W)^{\mathrm{op}} \to \mathrm{Bil}(U, V; W), \quad \gamma \mapsto b \bullet_0 \gamma.
\end{aligned}$$

The superscript is suppressed when context is clear.

Lemma 2.8. If $b : U \times V \rightarrowtail W$ is fully nondegenerate, then Λ_2, Λ_1, and Λ_0 are injective.

Proof. If $\Lambda_2(\alpha) = 0$, then $b(\alpha u, v) = 0$ for all $u \in U$ and $v \in V$, so $\alpha u \in \mathrm{Rad}_L(b)$ for every u. Since $\mathrm{Rad}_L(b) = 0$, we get $\alpha = 0$. The proof for Λ_1 is analogous. For Λ_0, if $\Lambda_0(\gamma) = 0$, then $\gamma(b(u, v)) = 0$ for all u, v, so γ vanishes on $b(U, V) = W$. Hence $\gamma = 0$. $\square$

Fact 2.9. [Wil15, Section 8.3] For any bilinear map $b : U \times V \rightarrowtail W$, there is a canonical induced fully-nondegenerate bilinear map

$$b/\sqrt{b} : \big(U/\mathrm{Rad}_L(b)\big) \times \big(V/\mathrm{Rad}_R(b)\big) \rightarrowtail b(U, V),$$

defined by $\big(b/\sqrt{b}\big)\big(u + \mathrm{Rad}_L(b), v + \mathrm{Rad}_R(b)\big) \coloneqq b(u, v)$.

The algebraic invariants of $b/\sqrt{b}$ gives us an understanding of the invariants of b. For instance, the left nuclei gives the following exact sequence of vector spaces

$$0 \to \mathrm{Hom}(U, \mathrm{Rad}_L(b)) \times \mathrm{Hom}(\mathrm{Corad}(b), W) \to \mathcal{L}(b) \xrightarrow{\rho_L} \mathcal{L}(b/\sqrt{b}) \to 0, \tag{2.2}$$

where $\rho_L(\alpha, \gamma) \coloneqq (\bar{\alpha}, \gamma|_{b(U,V)})$ and $\bar{\alpha}$ is induced by α on $U/\mathrm{Rad}_L(b)$. We justify this exact sequence. First, for well-definedness of ρ_L, we prove $(\alpha, \gamma) \in \mathcal{L}(b)$ satisfies $\alpha(\mathrm{Rad}_L(b)) \leq \mathrm{Rad}_L(b)$, and $\gamma\big(b(U, V)\big) \leq b(U, V)$. Let $u \in \mathrm{Rad}_L(b)$. Then $b(u, v) = 0$ for all v, so $\gamma(b(u, v)) =$

$\gamma(0) = 0$. But $\gamma(b(u,v)) = b(\alpha u, v)$ as well, so $\alpha u \in \mathrm{Rad}_L(b)$. The proof for $\gamma(b(U,V)) \leq b(U,V)$ is analogous.

To see the image of ρ_L lies in $\mathcal{L}(b/\sqrt{b})$, we check for all $u \in U$, $v \in V$,

$$\begin{aligned}\left(b/\sqrt{b}\right)\left(\bar{\alpha}(u+\mathrm{Rad}_L(b)), v+\mathrm{Rad}_R(b)\right) &= b(\alpha u, v) = \gamma(b(u,v)) \\ &= \gamma|_{b(U,V)}\Big(\left(b/\sqrt{b}\right)\left(u+\mathrm{Rad}_L(b), v+\mathrm{Rad}_R(b)\right)\Big),\end{aligned}$$

so $(\bar{\alpha}, \gamma|_{b(U,V)}) \in \mathcal{L}(b/\sqrt{b})$. We now show ρ_L is surjective. Given $(\bar{\alpha}, \gamma_0) \in \mathcal{L}(b/\sqrt{b})$, choose the lift $\alpha \in \mathrm{End}(U)$ of $\bar{\alpha}$ with $\alpha|_{\mathrm{Rad}_L(b)} = 0$, and choose $\gamma \in \mathrm{End}(W)$ with $\gamma|_{b(U,V)} = \gamma_0$ and $\gamma = 0$ on a chosen complement of $b(U,V)$ in W. Then for all $u \in U$ and $v \in V$,

$$\begin{aligned}b(\alpha u, v) &= \left(b/\sqrt{b}\right)\left(\bar{\alpha}(u+\mathrm{Rad}_L(b)), v+\mathrm{Rad}_R(b)\right) \\ &= \gamma_0\Big(\left(b/\sqrt{b}\right)\left(u+\mathrm{Rad}_L(b), v+\mathrm{Rad}_R(b)\right)\Big) = \gamma(b(u,v)).\end{aligned}$$

Hence $(\alpha, \gamma) \in \mathcal{L}(b)$ as needed. The kernel of ρ_L are all pairs of endomorphisms where the first component is the zero morphism modulo $\mathrm{Rad}_L(b)$ (hence $\alpha(U) \subseteq \mathrm{Rad}_L(b)$) and the second component is zero restricted to $b(U,V)$. Thus, $\mathrm{Ker}(\rho_L) = \{\alpha \in \mathrm{End}(U) : \alpha(U) \subseteq \mathrm{Rad}_L(b)\} \times \{\gamma \in \mathrm{End}(W) : \gamma|_{b(U,V)} = 0\} \cong \mathrm{Hom}(U, \mathrm{Rad}_L(b)) \times \mathrm{Hom}(\mathrm{Corad}(b), W)$, and Equation (2.2) is exact.

2.3 Product of tuples of endomorphisms

For $k \in \{0,1,2\}$, let U_k and V_k be finite-dimensional $\mathbb{F}$-vector spaces. Let $\kappa_k : \mathrm{End}(U_k) \otimes \mathrm{End}(V_k) \to \mathrm{End}(U_k \otimes V_k)$, defined by $\kappa_k(\alpha \otimes f)(u \otimes v) := \alpha u \otimes f v$ for all $u \in U_k$ and $v \in V_k$ denote the Kronecker isomorphisms. We introduce the following for combining two tuples of endomorphisms.

Definition 2.10. Let K be finite. For each $k \in K$, let U_k, V_k be finite-dimensional $\mathbb{F}$-vector spaces. Define

$$\boxtimes : \Big(\prod_{k\in K} \mathrm{End}(U_k)\Big) \times \Big(\prod_{k\in K} \mathrm{End}(V_k)\Big) \to \prod_{k\in K} \mathrm{End}(U_k \otimes V_k),$$

by the mapping for $a = (\alpha_k)_{k\in K}$ and $b = (f_k)_{k\in K}$ as $a \boxtimes b := (\kappa_k(\alpha_k \otimes f_k))_{k\in K}$.

For example, let $K = \{2, 1, 0\}$, then for $a = (\alpha_2, \alpha_1, \alpha_0) \in \mathrm{End}(U_2) \times \mathrm{End}(U_1) \times \mathrm{End}(U_0)$ and $b = (f_2, f_1, f_0) \in \mathrm{End}(V_2) \times \mathrm{End}(V_1) \times \mathrm{End}(V_0)$, then $a \boxtimes b = \big(\kappa_2(\alpha_2 \otimes f_2), \kappa_1(\alpha_1 \otimes f_1), \kappa_0(\alpha_0 \otimes f_0)\big)$, and is an element of $\mathrm{End}(U_2 \otimes V_2) \times \mathrm{End}(U_1 \otimes V_1) \times \mathrm{End}(U_0 \otimes V_0)$.

For subspaces $A \le \prod_{k\in K} \mathrm{End}(U_k)$ and $B \le \prod_{k\in K} \mathrm{End}(V_k)$, define

$$A \boxtimes B := \mathrm{span}\{a \boxtimes b : a \in A,\ b \in B\} \le \prod_{k\in K} \mathrm{End}(U_k \otimes V_k).$$

2.4 Product of bilinear maps

As previously described, there is a need to combine two bilinear maps into one using a product construction.

Definition 2.11. Fix two bilinear maps $s : U_2 \times U_1 \rightarrowtail U_0$, and $t : V_2 \times V_1 \rightarrowtail V_0$. The *tensor product bilinear map*

$$s \otimes t : (U_2 \otimes V_2) \times (U_1 \otimes V_1) \rightarrowtail U_0 \otimes V_0$$

is the unique bilinear map defined in pure tensors as $(s \otimes t)((u_2 \otimes v_2), (u_1 \otimes v_1)) = s(u_2, u_1) \otimes t(v_2, v_1)$ and extended bilinearly. When context is clear or we want to reserve the word *tensor product*, we will also refer to $s \otimes t$ as the *product bilinear map*.

Lemma 2.12. If s and t are fully-nondegenerate bilinear maps, then $s \otimes t$ is fully nondegenerate.

Proof. Define the curried functions $\lambda_s : U_2 \to \mathrm{Hom}(U_1, U_0)$ by $\lambda_s(u_2)(u_1) := s(u_2, u_1)$ and $\lambda_t : V_2 \to \mathrm{Hom}(V_1, V_0)$ by $\lambda_t(v_2)(v_1) := t(v_2, v_1)$. Then $\mathrm{Rad}_L(s) = \mathrm{Ker}(\lambda_s)$ and $\mathrm{Rad}_L(t) = \mathrm{Ker}(\lambda_t)$, so full nondegeneracy of s and t implies that λ_s and λ_t are injective. Identify $\mathrm{Hom}(U_1, U_0) \otimes \mathrm{Hom}(V_1, V_0)$ with $\mathrm{Hom}(U_1 \otimes V_1,\ U_0 \otimes V_0)$ by the Kronecker isomorphism. Under this identification, the curried map $U_2 \otimes V_2 \to \mathrm{Hom}(U_1 \otimes V_1, U_0 \otimes V_0)$ associated to $s \otimes t$ is the map $\lambda_s \otimes \lambda_t$. Since tensoring over a field preserves injectivity, $\lambda_s \otimes \lambda_t$ is injective, so $\mathrm{Rad}_L(s \otimes t) = 0$. The argument for $U_1 \otimes V_1$ is analogous.

Since $\mathrm{Corad}(s) = \mathrm{Corad}(t) = 0$, we have $s(U_2, U_1) = U_0$ and $t(V_2, V_1) = V_0$. Thus for arbitrary $u_0 \in U_0$ and $v_0 \in V_0$, write $u_0 = \sum_i s(u_{2,i}, u_{1,i})$ and $v_0 = \sum_j t(v_{2,j}, v_{1,j})$. Then

$$u_0 \otimes v_0 = \left(\sum_i s(u_{2,i}, u_{1,i})\right) \otimes \left(\sum_j t(v_{2,j}, v_{1,j})\right) = \sum_{i,j} (s \otimes t)((u_{2,i} \otimes v_{2,j}), (u_{1,i} \otimes v_{1,j})),$$

so $(s \otimes t)(U_2 \otimes V_2, U_1 \otimes V_1) = U_0 \otimes V_0$. This concludes $\mathrm{Corad}(s \otimes t) = 0$. $\square$

Remark 2.13. If s and t are fully nondegenerate, then the axis action maps $\Lambda_2^{s\otimes t}$, $\Lambda_1^{s\otimes t}$, and $\Lambda_0^{s\otimes t}$ are injective by Lemmas 2.8 and 2.12.

Next we come to a key correspondence that allows us to analyze axis actions on product bilinear maps by *linear algebra* techniques. Recall for X, Y vector spaces, and the tensor product space of X and Y, denoted $(X \otimes Y, \varphi : X \times Y \rightarrowtail X \otimes Y)$, the element $\varphi(x, y) \in X \otimes Y$ is the pure tensor most often written as $x \otimes y$. As the notation $x \otimes y$ can be ambiguous when x and y are themselves bilinear maps, we use $\varphi(x, y)$ when referring to the pure tensor element in the tensor product vector space.

Let $\mathcal{U} := \mathrm{Bil}(U_2, U_1; U_0)$, $\mathcal{V} := \mathrm{Bil}(V_2, V_1; V_0)$, and define the bilinear map

$$\Theta : \mathcal{U} \times \mathcal{V} \to \mathrm{Bil}(U_2 \otimes V_2, U_1 \otimes V_1; U_0 \otimes V_0), \quad \Theta(s, t) := s \otimes t \text{ as the product bilinear map.}$$

The universal property of the tensor product space $\mathcal{U} \otimes \mathcal{V}$ gives a unique induced linear map

$$\widetilde{\Theta} : \mathcal{U} \otimes \mathcal{V} \to \mathrm{Bil}(U_2 \otimes V_2, U_1 \otimes V_1; U_0 \otimes V_0).$$

By currying, there are isomorphisms $\mathcal{U} \cong \mathrm{Hom}(U_2 \otimes U_1, U_0)$, $\mathcal{V} \cong \mathrm{Hom}(V_2 \otimes V_1, V_0)$, and $\mathrm{Bil}(U_2 \otimes V_2, U_1 \otimes V_1; U_0 \otimes V_0) \cong \mathrm{Hom}((U_2 \otimes V_2) \otimes (U_1 \otimes V_1), U_0 \otimes V_0)$. So $\widetilde{\Theta}$ is the Kronecker isomorphism composed with the associator and swap isomorphisms $(U_2 \otimes U_1) \otimes (V_2 \otimes V_1) \cong (U_2 \otimes V_2) \otimes (U_1 \otimes V_1)$, and thus is a bijection.

We fix the notation $\Xi := \widetilde{\Theta}^{-1}$. Also recall the Kronecker isomorphism $\kappa_k : \mathrm{End}(U_k) \otimes \mathrm{End}(V_k) \to \mathrm{End}(U_k \otimes V_k)$ given by $\kappa_k(\alpha \otimes f)(u \otimes v) = \alpha u \otimes f v$.

Lemma 2.14. Following the notation above, let $s \in \mathcal{U}$, $t \in \mathcal{V}$ be bilinear maps, and $(\mathcal{U} \otimes \mathcal{V}, \varphi : \mathcal{U} \times \mathcal{V} \rightarrowtail \mathcal{U} \otimes \mathcal{V})$ the tensor product space of $\mathcal{U}$ with $\mathcal{V}$. For all $k \in \{0, 1, 2\}$, $\alpha \in \mathrm{End}(U_k)$, and $f \in \mathrm{End}(V_k)$,

$$\Xi\big((s \otimes t) \bullet_k \kappa_k(\alpha \otimes f)\big) = \varphi(s \bullet_k \alpha, t \bullet_k f).$$

Proof. By definition of $\widetilde{\Theta}$, we have $\widetilde{\Theta}\big(\varphi(s \bullet_k \alpha, t \bullet_k f)\big) = \Theta(s \bullet_k \alpha, t \bullet_k f)$. Evaluating both the bilinear map $\Theta(s \bullet_k \alpha, t \bullet_k f)$ and $(s \otimes t) \bullet_k \kappa_k(\alpha \otimes f)$ on pure tensors shows they agree, so applying Ξ gives the claim. □

Corollary 2.15. For $k \in \{0, 1, 2\}$ and $x \in \mathrm{End}(U_k \otimes V_k)$ with decomposition $x = \sum_j \kappa_k(\alpha_j \otimes f_j)$, linearity gives

$$\Xi\big((s \otimes t) \bullet_k x\big) = \sum_j \Xi\big((s \otimes t) \bullet_k \kappa_k(\alpha_j \otimes f_j)\big) = \sum_j \varphi(s \bullet_k \alpha_j, t \bullet_k f_j). \tag{2.3}$$

2.5 Prior work

Existing research shows that the algebras in Equation (1.1) encode structural information about tensors. Myasnikov, Wilson, and others in [Mya90], [Wil16], [Wil12], [MM10], [Wil09] prove that, for a bilinear map b, the algebras $\mathrm{Adj}(b)$ and $\mathrm{Cen}(b)$ control direct sum decompositions and automorphisms of b, using them to prove properties for the originating algebraic structures. The most famous special case for the middle nucleus equation is the theory of Kronecker modules, which in our terminology corresponds to tensors in $\mathbb{F}^{m \times n \times 2}$ [Kro90]. Recent work [BKW24] shows that derivations also control the original tensor's null patterns, limiting the subspaces on which the tensor is non-zero.

There has also been prior work on understanding derivations of tensors coming from the product construction in Definition 2.11. Leger and Luks study generalized derivations, quasiderivations, and quasicentroids as extensions of algebra derivations [LL00], and Novotný and Hřivnák study (α, β, γ)-derivations together with associated invariant functions [NH08]. Closer to our results, Brešar studies derivations of tensor products of nonassociative algebras [Bre17]. Benkart and Osborn study derivations and automorphisms of nonassociative matrix algebras [BO81], and Azam studies derivations of tensor products of algebras [Aza08].

Chapter 3

Algorithms for nuclei and derivation equations

The descriptions of the algebras in Equation (1.1) are linear in the unknown matrix entries. So each can be solved as a large system of linear equations. In collaboration with Joshua Maglione and James B. Wilson, we obtain an asymptotically faster algorithm for computing the adjoint algebra in certain scenarios, inspired by analogous results for matrices known as the Bartels-Stewart algorithm [BS72]. In this chapter, I state and prove results from this joint work using a novel approach. I also state and prove my own original result, which concerns the derivation algebra, following a similar approach.

Theorem A. Let $\mathcal{F}$ be a Sylvester regular family of tensors as defined by Definition 3.34, meaning each problem's solution space is completely determined by a restricted subsystem. Furthermore, assume we have a black-box algorithm to find this subsystem. For each $(R, S, T) \in \mathcal{F}$ with $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, and $T \in \mathbb{F}^{a\times b\times c}$, set $n := a + b + c + r + s$. Then SimultaneousSylvesterSystem instances from $\mathcal{F}$ are solved in arithmetic cost $O(n^4)$. This gives an algorithm for the Sylvester regular family with the same order of arithmetic cost as deterministic verification, and an improvement to the $O(n^7)$ cost of solving the full linear system.

Theorem B. Let $\mathcal{F}$ be a derivation regular family of tensors as defined by Definition 3.41, meaning each problem's solution space is completely determined by a restricted subsystem. Furthermore, assume we have a black-box algorithm to find this subsystem. For each $(R, S, T) \in \mathcal{F}$ with $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, and $T \in \mathbb{F}^{a\times b\times t}$, set $n := a+b+c+r+s+t$. Then DerivationSystem instances from $\mathcal{F}$ are solved in arithmetic cost $O(n^{4.5})$. This gives an algorithm for the derivation regular family within a factor $n^{1/2}$ of deterministic verification cost $O(n^4)$, and an improvement to the $O(n^7)$ cost of solving the full linear system.

In both cases, the black-box algorithm for choosing the restricted subsystem can be implemented efficiently in practice with heuristics or randomized techniques. Furthermore, tensors in the regular families satisfy a condition that contains many natural classes of tensors. Examples include generic cubic $(n \times n \times n)$ tensors, tensors arising from algebras with bounded solution dimension, and tensors without solutions whose inconsistency can be detected on a smaller portion of the coordinates. In the rest of this section, we will reduce each problem to a set of smaller linear algebra problems. Their solutions are affine subspaces, which we define below.

3.1 Preliminaries

An *affine space* is a pair (A, U) where A is a non-empty set and U is a vector space acting on A by a fixed-point free and transitive action. Write the action as $+ : A \times U \to A$, $(a, u) \mapsto a + u$. We call U the *direction space* of A. When U is clear from context, we also say "A is an affine space".

Lemma 3.1. Let (A, U) be an affine space. For any $a, b \in A$, there is a unique $u \in U$ such that $a + u = b$. We denote this vector by $b - a$.

Proof. Existence follows from transitivity. If $a+u = b = a+v$, then $a+(u-v) = a$. Fixed-point freeness gives $u - v = 0$, so $u = v$. □

Let (A, U) be a d-dimensional affine space. An *affine frame* is a tuple $(a_0, \ldots, a_d)$ in A such that $\{a_1 - a_0, \ldots, a_d - a_0\}$ is a basis of U. Every $a \in A$ has unique coordinates $\lambda = (\lambda_1, \ldots, \lambda_d) \in \mathbb{F}^d$ with $a = a_0 + \sum_{i=1}^{d} \lambda_i (a_i - a_0)$.

An affine frame may also be described by a base point $a_0 \in A$ together with a basis of U.

Convention 3.2. For computation, we encode an affine space by an affine frame.

Let (A, U) and (B, V) be affine spaces. A map $f : A \to B$ is an *affine map* if there exists a linear map $L : U \to V$ such that $f(a + u) = f(a) + L(u)$ for all $a \in A$ and $u \in U$. We call L the underlying linear map of f. Below we list some properties of affine spaces and maps that we will use in following sections without further comment.

1. Every vector space V is also an affine space with direction space V.

2. If A, B are vector spaces and $f : A \to B$ is a linear map, then f is also an affine map.

3. If (A, U) and (B, V) are affine spaces, then $(A \times B, U \times V)$ is an affine space with action $(a, b) + (u, v) := (a + u, b + v)$.

4. An affine map $f : A \times B \to C$ has an underlying linear map $L : U \times V \to W$ with $f\big((a, b) + (u, v)\big) = f(a, b) + L(u, v)$.

5. Let U and V be vector spaces, and let $L : U \to V$ be linear. If (A, W) is an affine subspace of U, then $L|_A : A \to V$ is affine, with underlying linear map $L|_W$.

Example 3.3. Let $A = \mathbb{R}^3$, $a_0 = (1, 2, 3)$, and $W = \{(x, y, 0) : x, y \in \mathbb{R}\} \leq \mathbb{R}^3$. Then $a_0 + W$ is the plane $z = 3$, an affine subspace of $(\mathbb{R}^3, \mathbb{R}^3)$.

As an example of (5) above, if $L : \mathbb{R}^3 \to \mathbb{R}$ is the linear map $L(x, y, z) = x + y + z$, then $L|_{a_0+W}$ is the affine map $(x, y, 3) \mapsto x+y+3$, with underlying linear map $L|_W : (x, y, 0) \mapsto x+y$.

Finally, we describe how we analyze algorithms and our notation for indexing and accessing tensors. Throughout, for n a positive natural number, $[n]$ denotes the set $\{1, \ldots, n\}$.

Convention 3.4. For functions $f, g : \mathbb{N} \to \mathbb{N}$, we write $f \in O(g)$ if there exist constants $C, n_0 \in \mathbb{N}$ such that $f(n) \leq C\, g(n)$ for all $n \geq n_0$.

Convention 3.5. We always use the standard basis for coordinate vector spaces $\mathbb{F}^n$. For $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$, we use the bilinear pairing $\langle u, v \rangle := \sum_{i=1}^{n} u_i v_i$.

Convention 3.6. We analyze algorithms by their arithmetic complexity, meaning the number of field operations over $\mathbb{F}$ (addition, subtraction, multiplication, and division).

Definition 3.7. Let $T \in \mathbb{F}^{a \times b \times c}$ be a tensor, let $I \subseteq [a]$, $J \subseteq [b]$, and let $k \in [c]$. We use the slice notation

$$T[I, :, k] \in \mathbb{F}^{|I| \times b}, \qquad T[:, J, k] \in \mathbb{F}^{a \times |J|}, \qquad T[I, J, k] \in \mathbb{F}^{|I| \times |J|}.$$

For $M \in \mathbb{F}^{a \times b}$ a matrix, we write

$$M[I, :] \in \mathbb{F}^{|I| \times b}, \qquad M[:, J] \in \mathbb{F}^{a \times |J|}.$$

The indexing of sliced tensors and matrices is by the fixed ordering of the slicing sets. For instance, for $I = \{i_1 < \cdots < i_t\} \subset [a]$, the restricted matrix $X[I, :] \in \mathbb{F}^{t \times r}$ has $X[I, :][j, k] = X[i_j, k]$. We use this convention for all restricted tensor and matrix indexing.

Example 3.8. Let $T \in \mathbb{F}^{2 \times 2 \times 2}$ be defined by $T[:, :, 1] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $T[:, :, 2] = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$,. We have $T[\{1\}, :, 1] = \begin{bmatrix} 1 & 2 \end{bmatrix}$, and $T[:, \{2\}, 1] = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$. In the matrix case, let $M = \begin{bmatrix} 9 & 10 & 11 \\ 12 & 13 & 14 \end{bmatrix} \in \mathbb{F}^{2 \times 3}$. Then $M[:, \{2, 3\}] = \begin{bmatrix} 10 & 11 \\ 13 & 14 \end{bmatrix} \in \mathbb{F}^{2 \times 2}$.

Definition 3.9. We define vectorization for matrices and tensors. For $M = (m_{ij}) \in \mathbb{F}^{a\times b}$, $\mathrm{vec}(M) \in \mathbb{F}^{ab}$ for $(i,j) \in [a] \times [b]$ by

$$\mathrm{vec}(M)_{i+(j-1)a} = m_{ij}.$$

For $T = (t_{ijk}) \in \mathbb{F}^{a\times b\times c}$, define $\mathrm{vec}(T) \in \mathbb{F}^{abc}$ for $(i,j,k) \in [a] \times [b] \times [c]$ by

$$\mathrm{vec}(T)_{i+(j-1)a+(k-1)ab} = t_{ijk}.$$

Example 3.10. For

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{F}^{2\times 3}, \qquad T \in \mathbb{F}^{2\times 2\times 2}, \quad T[:,:,1] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad T[:,:,2] = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix},$$

we have $\mathrm{vec}(M) = (1,4,2,5,3,6)$, and $\mathrm{vec}(T) = (1,3,2,4,5,7,6,8)$.

3.2 Linear Algebra subroutines

Lemma 3.11. Let (A,U) and (B,V) be affine spaces, let $f : A \to B$ be affine with underlying linear map $L : U \to V$, and let (S,W) be an affine subspace of B. Then the preimage, $f^{-1}(S) \subseteq A$, is either empty or an affine subspace of A.

Proof. If $f^{-1}(S) = \varnothing$, there is nothing to prove. Assume $f^{-1}(S) \neq \varnothing$, choose $a_0 \in f^{-1}(S)$, and set $U_0 := L^{-1}(W) \leq U$. For $u \in U$, $f(a_0+u) = f(a_0) + L(u)$. Since $f(a_0) \in S$ and S has direction space W,

$$f(a_0+u) \in S \iff L(u) \in W \iff u \in U_0.$$

Hence $f^{-1}(S) = a_0 + U_0$, so $f^{-1}(S)$ is an affine subspace. $\square$

Problem LinSolve($\mathbb{F}$).

Given $\mathbb{F}$-vector spaces U, V, a linear map $L : U \to V$, and $v \in V$.

Return If a solution exists, a computational encoding of the affine solution space $\{u \in U : L(u) = v\}$. Otherwise $\varnothing$.

Proposition 3.12. LinSolve($\mathbb{F}$) returns either an empty set or an affine subspace of U.

Proof. The map L is affine and $\{v\}$ is an affine subspace of V with direction space $\{0\}$, so the solution set is $L^{-1}(\{v\})$. Applying Lemma 3.11 concludes the proposition. $\square$

Proposition 3.13. A coordinatized instance of LinSolve($\mathbb{F}$), with

$$U = \mathbb{F}^{n\times\ell}, \qquad V = \mathbb{F}^{m\times\ell}, \qquad L(X) = MX, \qquad M \in \mathbb{F}^{m\times n},$$

is solvable in cost $\mathtt{T}_{p1}(m, n, \ell) \in O(\min\{m, n\}^2 \max\{m, n\} + mn\ell)$.

This is the standard dense Gaussian-elimination bound for one shared coefficient matrix and ℓ right-hand sides. Next is the workhorse for our solution strategy, allowing us to lift subproblem solutions to solutions of the whole problem.

Problem LinearEqualsAffine($\mathbb{F}$).

Given $\mathbb{F}$-vector spaces U, W, a linear map $L : U \to W$, an affine space (Θ, V), and an affine map $N : \Theta \to W$.

Return If a solution exists, a computational encoding of the affine solution space $\{(u, \theta) \in U\times\Theta : L(u) = N(\theta)\}$. Otherwise $\varnothing$.

Proposition 3.14. The solution set of LinearEqualsAffine($\mathbb{F}$) is either empty or an affine subspace of $U \times \Theta$.

Proof. Define $F : U \times \Theta \to W$ by $F(u, \theta) := L(u) - N(\theta)$. Let $M : V \to W$ be the underlying linear map of N. Then for all $(u, \theta) \in U \times \Theta$ and $(u', v) \in U \times V$,

$$F\big((u, \theta) + (u', v)\big) = L(u + u') - N(\theta + v) = F(u, \theta) + \big(L(u') - M(v)\big),$$

so F is affine with underlying linear map $(u', v) \mapsto L(u') - M(v)$. The solution set is $F^{-1}(\{0\})$, which is empty or an affine subspace by Lemma 3.11. $\square$

Proposition 3.15. For an instance of LinearEqualsAffine($\mathbb{F}$), let $\rho : W \to W/\operatorname{im}(L)$ be the quotient map and define the *feasible parameters*

$$\Theta_{\text{feas}} := \{\theta \in \Theta : \rho(N(\theta)) = 0\}.$$

Then (u, θ) satisfies $L(u) = N(\theta)$ if and only if $\theta \in \Theta_{\text{feas}}$ and $u \in u_0(\theta) + \operatorname{Ker}(L)$ for some fixed $u_0(\theta)$ such that $L(u_0(\theta)) = N(\theta)$.

Proof. For fixed θ, solvability is $N(\theta) \in \operatorname{im}(L)$, which is exactly $\rho(N(\theta)) = 0$. If $u_0(\theta)$ is one solution, then for any other u, $L(u) = N(\theta)$ if and only if $L(u - u_0(\theta)) = 0$ if and only if $u - u_0(\theta) \in \operatorname{Ker}(L)$. $\square$

Proposition 3.16. Given a coordinatized instance of LinearEqualsAffine($\mathbb{F}$), with

$$U = \mathbb{F}^{n\times\ell}, \qquad W = \mathbb{F}^{m\times\ell}, \qquad L(X) = MX, \qquad M \in \mathbb{F}^{m\times n},$$

and with affine-frame coordinates $\theta \in \mathbb{F}^{\xi}$ for Θ and $N(\theta) = N_0 + \sum_{i=1}^{\xi} \theta_i N_i$, having $N_i \in W$, where

$$\kappa := \dim \operatorname{Ker}(M^T), \qquad \eta := \dim \Theta_{\text{feas}} \text{ if } \Theta_{\text{feas}} \neq \varnothing, \text{ and } 0 \text{ otherwise.}$$

The problem is solvable in cost

$$\mathtt{T}_{p2}(m, n, \ell, \xi, \eta, \kappa) \in O\Big(\underbrace{mn^2 + \xi\kappa m\ell + (\kappa\ell)\xi^2}_{\text{feasible parameters}} + \underbrace{mn^2 + n^2\ell}_{\text{kernel of } L} + \underbrace{(\eta+1)\xi\, m\ell + mn^2 + (\eta+1)mn\ell}_{\text{affine frame solve}}\Big).$$

Proof. The first term is the cost to solve for feasible parameters. Compute a basis $w_1, \dots, w_\kappa$ of $\operatorname{Ker}(M^\top)$ in $O(mn^2)$. Since $\operatorname{Ker}(M^\top) = \operatorname{im}(M)^\perp$ under the standard pairing,

$$x \in \operatorname{im}(M) \iff \langle w_i, x \rangle = 0 \text{ for all } i \in [\kappa].$$

Let C be the matrix with rows $w_1, \dots, w_\kappa$. Applying columnwise to $N(\theta)$ means $\rho(N(\theta)) = 0$ exactly when $CN(\theta) = 0$. Recall $N(\theta)$ is coordinatized as $N(\theta) = N_0 + \sum_{i=1}^{\xi} \theta_i N_i$. Thus we solve the inhomogeneous system $A_0 + \sum_{i=1}^{\xi} \theta_i A_i = 0$ with $A_i := CN_i$. Hence the three costs in the feasible-parameters term are $O(mn^2)$ to compute w_i, $O(\xi\kappa m\ell)$ to form A_i, and $O((\kappa\ell)\xi^2)$ to solve the $(\kappa\ell) \times \xi$ system.

The second term is to compute a basis for the kernel of L. Compute a basis $u_1, \dots, u_\nu$ of $\operatorname{Ker}(M)$ in $O(mn^2)$, where $\nu := \dim \operatorname{Ker}(M)$. Since $L(X) = MX$ acts columnwise, a basis for $\operatorname{Ker}(L)$ is given by the matrices $u_i e_j^\top$ for $i \in [\nu]$ and $j \in [\ell]$. This costs $O(\nu n\ell) \subseteq O(n^2\ell)$.

The third term is to solve on an affine frame of the feasible parameters. We evaluate N on this affine frame of $(\eta + 1)$ points in $O((\eta + 1)\xi\, m\ell)$, row-reduce M once in $O(mn^2)$, and solve the $(\eta + 1)$ right-hand sides with that reduction in $O((\eta + 1)mn\ell)$. $\square$

Example 3.17. In the following example, we have $m = 3, n = 2, \xi = 2$, and $\ell = 2$. And set

$$M = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{3\times 2}, \qquad N : \mathbb{F}^2 \to \mathbb{F}^{3\times 2}, \qquad N(\theta_1, \theta_2) = N_0 + \theta_1 N_1 + \theta_2 N_2,$$

$$N_0 = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ -1 & -2 \end{bmatrix}, \qquad N_1 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 1 & 2 \end{bmatrix}, \qquad N_2 = \begin{bmatrix} 0 & -1 \\ 0 & 2 \\ 1 & 2 \end{bmatrix}.$$

In the three-part process

1. **Feasible parameters.** Here $\mathrm{im}(M) = \{(u,v,0)^{\mathsf{T}} : u,v \in \mathbb{F}\}$, so feasibility is exactly vanishing of the third row of $N(\theta_1,\theta_2)$. This is accomplished by the condition $\theta_1+\theta_2=1$. Hence $\Theta_{\text{feas}} = \{(\theta_1,\theta_2)\in\mathbb{F}^2 : \theta_1+\theta_2=1\}$.

2. **Kernel.** $\mathrm{Ker}(M)=\{0\}$, so $\mathrm{Ker}(L)=\{0\}$.

3. **Solve on an affine frame of Θ_{feas}.** Take frame points $\theta^{(0)}=(1,0)$ and $\theta^{(1)}=(0,1)$. Solving $MX=N(\theta^{(0)})$ and $MX=N(\theta^{(1)})$ gives

$$X^{(0)} = \begin{bmatrix} 3 & -5 \\ 0 & 4 \end{bmatrix}, \qquad X^{(1)} = \begin{bmatrix} 0 & -10 \\ 1 & 6 \end{bmatrix}.$$

   Therefore parametrizing $\theta\in\Theta_{\text{feas}}$ as $\theta=(1-\lambda,\lambda)$, we conclude $X(\theta)=X^{(0)}+\lambda\big(X^{(1)}-X^{(0)}\big)$.

Remark 3.18. If L has a left inverse G, then for each $\theta\in\Theta_{\text{feas}}$, $L(u)=N(\theta)$ has exactly one solution, $u=G(N(\theta))$.

The last linear algebra subroutine solves the compatibility criterion that arises when lifting subproblem solutions.

Problem AffinePullback($\mathbb{F}$).

Given $\mathbb{F}$-vector spaces U,V,W and affine maps $f:U\to W$, $g:V\to W$.

Return If a solution exists, a computational encoding of the solution space $P:=\{(u,v)\in U\times V : f(u)=g(v)\}$, and otherwise $\varnothing$.

Remark 3.19. The solution space P is affine. Define the map $h(u,v):=f(u)-g(v)$. Then $P=h^{-1}(\{0\})$. So it suffices to prove h is an affine map. Let L_f and L_g be the underlying linear maps of f and g, and define $L_h: U\times V\to W$ by $L_h(u',v'):=L_f(u')-L_g(v')$. Then $h((u,v)+(u',v'))=f(u+u')-g(v+v')=f(u)+L_f(u')-(g(v)+L_g(v'))=h(u,v)+L_h(u',v')$. This concludes h is affine and that P is either an empty set or an affine subspace.

Proposition 3.20. Given a coordinate instance of AffinePullback($\mathbb{F}$) with

$$U = \mathbb{F}^d, \quad V = \mathbb{F}^e, \quad W = \mathbb{F}^m, \quad f(u) = Au + \alpha, \quad g(v) = Bv + \beta,$$

given by $A \in \mathbb{F}^{m\times d}$ and $B \in \mathbb{F}^{m\times e}$, it is solvable in cost $\mathtt{T}_{p1}(m, d+e, 1)$.

Proof. The condition $f(u) = g(v)$ is the block linear system $\begin{bmatrix} A & -B \end{bmatrix} \begin{bmatrix} u & v \end{bmatrix}^{\mathsf{T}} = \beta - \alpha$, which is one instance of LinSolve($\mathbb{F}$). $\square$

3.3 Simultaneous Sylvester Equations for nuclei of tensors

We work towards Theorem A in this section. Throughout, a, b, c, r, s will be natural numbers, and we use the notation

$$\mathcal{X} := \mathbb{F}^{a\times r}, \qquad \mathcal{Y} := \mathbb{F}^{s\times b},$$

to denote vector spaces. This notation may have subscripts to indicate restricted solution spaces. As an example, for $I \subseteq [a]$,

$$\mathcal{X}_I := \mathbb{F}^{|I|\times r}.$$

Problem SimultaneousSylvesterSystem.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, and $T \in \mathbb{F}^{a\times b\times c}$.

Return If a solution exists, the affine subspace

$$\{(X, Y) \in \mathcal{X} \times \mathcal{Y} : (\forall k \in [c])\ X\,R[:,:,k] + S[:,:,k]\,Y = T[:,:,k]\},$$

and otherwise $\varnothing$.

This problem is named as a generalization to the single matrix Sylvester equation $XA + BX = C$ [Syl84]. It can be solved using LinSolve($\mathbb{F}$) with $m = abc$, $n = ar + bs$, and $\ell = 1$. For a tensor $A \in \mathbb{F}^{a\times b\times c}$, taking $(R, S, T) = (A, -A, 0)$ with $r = a$ and $s = b$ computes the middle nucleus of A. To compute the left nucleus, apply the Knuth-Liebler shuffle [FMW20, Section 7.1.5] to A along its second and third axes to get $\widetilde{A} \in \mathbb{F}^{a\times c\times b}$. Then taking $(R, S, T) = (\widetilde{A}, -\widetilde{A}, 0)$ with $r = a$ and $s = c$ computes the left nucleus of A. The right nucleus is computed in the same fashion, by first swapping the first and third axes to get $\widetilde{A} \in \mathbb{F}^{c\times b\times a}$ and taking $(R, S, T) = (\widetilde{A}, -\widetilde{A}, 0)$ with $r = c$ and $s = b$.

Example 3.21. Below is an example of laying out the equations for SimultaneousSylvesterSystem to solve using LinSolve($\mathbb{F}$). Take $a = 2$, $b = 2$, $c = 2$, $r = 3$, and $s = 4$, with slices

$$R[:,:,1] = \begin{bmatrix} 1 & 5 \\ 0 & -1 \\ 2 & 11 \end{bmatrix}, \qquad R[:,:,2] = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix},$$

$$S[:,:,1] = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 1 & 1 & 10 & -1 \end{bmatrix}, \qquad S[:,:,2] = \begin{bmatrix} 0 & 1 & 7 & -1 \\ 1 & -1 & -4 & 1 \end{bmatrix},$$

$$T[:,:,1] = \begin{bmatrix} 1 & 19 \\ -11 & 4 \end{bmatrix}, \qquad T[:,:,2] = \begin{bmatrix} -10 & 4 \\ 8 & 2 \end{bmatrix}.$$

The corresponding augmented linear system is shown in Figure 3.1.

x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}	y_{11}	y_{21}	y_{31}	y_{41}	y_{12}	y_{22}	y_{32}	y_{42}	t
1	0	2				1	0	3	0					1
0	1	−1				0	1	7	−1					−10
5	−1	11								1	0	3	0	19
1	1	1								0	1	7	−1	4
			1	0	2	1	1	10	−1					−11
			0	1	−1	1	−1	−4	1					8
			5	−1	11					1	1	10	−1	4
			1	1	1					1	−1	−4	1	2

Figure 3.1: An augmented matrix for the Simultaneous Sylvester System in Example 3.21.

We describe an alternative approach, culminating in the following.

Theorem A. Let $\mathcal{F}$ be a Sylvester regular family of tensors as defined by Definition 3.34, meaning each problem's solution space is completely determined by a restricted subsystem. Furthermore, assume we have a black-box algorithm to find this subsystem. For each $(R, S, T) \in \mathcal{F}$ with $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, and $T \in \mathbb{F}^{a\times b\times c}$, set $n := a + b + c + r + s$. Then SimultaneousSylvesterSystem instances from $\mathcal{F}$ are solved in arithmetic cost $O(n^4)$. This gives an algorithm for the Sylvester regular family with the same order of arithmetic cost as deterministic verification, and an improvement to the $O(n^7)$ cost of solving the full linear system.

Example 3.23. We use the following running example throughout Section 3.3. Work over $\mathbb{F} = \mathbb{Q}$, with $a = b = r = s = 2$ and $c = 2$. Define

$$R[:, :, 1] = S[:, :, 1] = T[:, :, 1] = \mathrm{diag}(1, 0),$$
$$R[:, :, 2] = S[:, :, 2] = \mathrm{diag}(0, 1), \quad T[:, :, 2] = \mathrm{diag}(0, 2),$$

where $\mathrm{diag}(u, v)$ is a 2×2 matrix with the elements u and v on the diagonal. The full solution family is 2-dimensional, parametrized by $\alpha, \beta \in \mathbb{F}$ as $X = \mathrm{diag}(\alpha, \beta)$, and $Y = \mathrm{diag}(1-\alpha, 2-\beta)$.

Our strategy for SimultaneousSylvesterSystem is to solve smaller systems with a subset of the constraints, and then lift them. We start with the following.

Problem DoubleRestricted.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times c}$, and index sets $I \subseteq [a]$, $J \subseteq [b]$ with $|I| = a'$ and $|J| = b'$.

Return If a solution exists, the affine subspace

$$\mathcal{S}_{\mathrm{DR}} := \{(X_I, Y_J) \in \mathcal{X}_I \times \mathcal{Y}_J : (\forall k \in [c])\ X_I R[:, J, k] + S[I, :, k]Y_J = T[I, J, k]\},$$

and otherwise $\varnothing$.

To solve, vectorizing yields an instance of LinSolve($\mathbb{F}$) which costs $\mathtt{T}_{p1}(m = a'b'c,\ n = a'r + b's,\ \ell = 1)$. We refer to this cost as $\mathtt{T}_{\mathrm{DR}}(a, b, c, r, s, a', b')$, or when context is clear, just $\mathtt{T}_{\mathrm{DR}}$. By Proposition 3.12, the returned solution set is affine.

Example 3.24. In the running example, take $I = J = \{1\}$. Then $a' = b' = 1$, and

$$X_I = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \in \mathbb{F}^{1\times 2}, \qquad Y_J = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{F}^{2\times 1}.$$

Here DoubleRestricted gives the constraint $x_1 + y_1 = 1$, with no other constraints on (x_2, y_2). Hence

$$\mathcal{S}_{\mathrm{DR}} = \left\{ \left(\begin{bmatrix} \alpha & \beta \end{bmatrix}, \begin{bmatrix} 1-\alpha \\ \gamma \end{bmatrix} \right) : \alpha, \beta, \gamma \in \mathbb{F} \right\}.$$

Problem RowRestricted.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times c}$, and index sets $I \subseteq [a]$, $J \subseteq [b]$ with $|I| = a'$ and $|J| = b'$.

Return If a solution exists, the affine subspace

$$\mathcal{S}_{\mathrm{R}} \coloneqq \{(X_I, Y) \in \mathcal{X}_I \times \mathcal{Y} : (\forall k \in [c])\ X_I R[:,:,k] + S[I,:,k]Y = T[I,:,k]\},$$

and otherwise $\varnothing$.

Proposition 3.25. For fixed $I \subseteq [a]$ and $J \subseteq [b]$, RowRestricted is solved by solving DoubleRestricted and then LinearEqualsAffine($\mathbb{F}$), with cost

$$O\big(\mathtt{T}_{p1}(m = a'b'c,\ n = a'r + b's,\ \ell = 1) + d\,c\,a'r(b-b') \\ + \mathtt{T}_{p2}(m = ca',\ n = s,\ \ell = b-b',\ \xi = d,\ \eta = \eta_R,\ \kappa = \kappa_R)\big).$$

where $a' = |I|$, $b' = |J|$, d is the dimension of the solution set of DoubleRestricted ($d = \dim \mathcal{S}_{\mathrm{DR}}$), and (η_R, κ_R) are the feasible parameter dimension and the left nullspace dimension respectively in the LinearEqualsAffine($\mathbb{F}$) instance (d, η_R, and κ_R are all 0 if $\varnothing$ is returned for DoubleRestricted)

Proof. **Algorithm. Step 1.** Solve DoubleRestricted. If it returns $\varnothing$, return $\varnothing$ and stop.

Step 2. Set $\widehat{J} := [b] \setminus J$ and define an affine map $N_R : \mathcal{S}_{\mathrm{DR}} \to \mathbb{F}^{(ca') \times (b-b')}$ as well as an element $M_R \in \mathbb{F}^{(ca') \times s}$ by

$$N_R(X_I, Y_J) := \begin{bmatrix} T[I, \widehat{J}, 1] - X_I R[:, \widehat{J}, 1] \\ \vdots \\ T[I, \widehat{J}, c] - X_I R[:, \widehat{J}, c] \end{bmatrix} \in \mathbb{F}^{(ca') \times (b-b')}, \qquad M_R := \begin{bmatrix} S[I, :, 1] \\ \vdots \\ S[I, :, c] \end{bmatrix} \in \mathbb{F}^{(ca') \times s} \tag{3.1}$$

Apply LinearEqualsAffine($\mathbb{F}$) to

$$U = \mathcal{Y}_{\widehat{J}}, \quad W = \mathbb{F}^{(ca') \times (b-b')},$$
$$(L_R : U \to W, \quad \widehat{Y} \mapsto M_R \widehat{Y}), \quad \Theta = \mathcal{S}_{\mathrm{DR}}, \quad N = N_R.$$

This returns either $\varnothing$, in which case our algorithm stops, or an affine subspace $\mathcal{A}_R \subseteq \mathcal{S}_{\mathrm{DR}} \times \mathcal{Y}_{\widehat{J}}$ of pairs $((X_I, Y_J), \widehat{Y})$ satisfying $M_R \widehat{Y} = N_R(X_I, Y_J)$.

Step 3. For each $((X_I, Y_J), \widehat{Y}) \in \mathcal{A}_R$, form the unique $Y \in \mathcal{Y}$ such that $Y[:, J] = Y_J$ and $Y[:, \widehat{J}] = \widehat{Y}$. Return the affine subspace of $\mathcal{X}_I \times \mathcal{Y}$ obtained by sending each $((X_I, Y_J), \widehat{Y}) \in \mathcal{A}_R$ to (X_I, Y).

Timing. Step 1 is one call to DoubleRestricted, costing $\mathtt{T}_{\mathrm{DR}}(a, b, c, r, s, a', b')$. Step 2 has two contributions. Recall $\dim(\mathcal{S}_{\mathrm{DR}}) = d$, and we are given an affine frame $(X_{I,i}, Y_{J,i})_{i=0}^{d}$. Building the resulting coordinatized affine map N_R from Equation (3.1) costs $O\big(d\,c\,a' r(b-b')\big)$, because each of the d matrices has size $a' \times r$, each $R[:, \widehat{J}, k]$ has size $r \times (b-b')$, and we compute c multiplications between them.

Invoking Proposition 3.16 on (M_R, N_R) adds $\mathtt{T}_{p2}(m = ca',\ n = s,\ \ell = b - b',\ \xi = d,\ \eta = \eta_R,\ \kappa = \kappa_R)$. Adding the three steps gives the desired cost.

Correctness. Step 1 gives exactly $\mathcal{S}_{\mathrm{DR}}$, i.e. all (X_I, Y_J) that satisfy the (I, J)-restricted equations. Step 2 gives exactly the elements $\widehat{Y}$ for which the combined Y_J and $\widehat{Y}$ satisfy the row restricted equation. Step 3 combines $(Y_J, \widehat{Y})$ into Y, so $\mathcal{S}_{\mathrm{R}}$ is exactly the set of row-restricted solutions. $\square$

Example 3.26. Continue the running example with $I = J = \{1\}$, so $\widehat{J} = \{2\}$. From Example 3.24, write

$$X_I = \begin{bmatrix} \alpha & \beta \end{bmatrix}, \qquad Y_J = \begin{bmatrix} 1-\alpha \\ \gamma \end{bmatrix}.$$

Step 2 of Proposition 3.25 computes

$$M_R = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad N_R(X_I, Y_J) = \begin{bmatrix} 0 \\ -\beta \end{bmatrix}.$$

Solving $M_R \widehat{Y} = N_R(X_I, Y_J)$ forces $\widehat{Y} = \begin{bmatrix} 0 \\ \delta \end{bmatrix}$ and $\beta = 0$. Therefore

$$\mathcal{S}_{\mathrm{R}} = \left\{ \left(\begin{bmatrix} \alpha & 0 \end{bmatrix}, \begin{bmatrix} 1-\alpha & 0 \\ \gamma & \delta \end{bmatrix} \right) : \alpha, \gamma, \delta \in \mathbb{F} \right\}$$

.

Next we describe and solve an analogous variant of the row restricted problem, where instead, we restrict the columns of the original tensors.

Problem ColumnRestricted.

Given Tensors $R \in \mathbb{F}^{r \times b \times c}$, $S \in \mathbb{F}^{a \times s \times c}$, $T \in \mathbb{F}^{a \times b \times c}$, and index sets $I \subseteq [a]$, $J \subseteq [b]$ with $|I| = a'$ and $|J| = b'$.

Return If a solution exists, the affine subspace

$$\mathcal{S}_{\mathrm{C}} := \{(X, Y_J) \in \mathcal{X} \times \mathcal{Y}_J : (\forall k \in [c]) \; XR[:, J, k] + S[:, :, k]Y_J = T[:, J, k]\},$$

and otherwise $\varnothing$.

Proposition 3.27. For fixed $I \subseteq [a]$ and $J \subseteq [b]$, ColumnRestricted is solved by solving DoubleRestricted and then LinearEqualsAffine($\mathbb{F}$), in cost

$$O\big(\mathtt{T}_{p1}(m = a'b'c,\ n = a'r + b's,\ \ell = 1) + d\,c\,(a-a')sb' \\ + \mathtt{T}_{p2}(m = cb',\ n = r,\ \ell = a-a',\ \xi = d,\ \eta = \eta_C,\ \kappa = \kappa_C)\big).$$

where $a' = |I|$, $b' = |J|$, d is the dimension of the solution set of DoubleRestricted ($d = \dim \mathcal{S}_{\mathrm{DR}}$), and (η_C, κ_C) are the feasible parameter dimension and the left nullspace dimension respectively in the LinearEqualsAffine($\mathbb{F}$) instance (d, η_C, and κ_C are all 0 if $\varnothing$ is returned)

Proof. **Algorithm. Step 1.** Solve DoubleRestricted. If it returns $\varnothing$, return $\varnothing$ and stop.

Step 2. Set $\widehat{I} := [a] \setminus I$ and define an affine map $N_C : \mathcal{S}_{\mathrm{DR}} \to \mathbb{F}^{(cb') \times (a-a')}$ as well as an element $M_C \in \mathbb{F}^{(cb') \times r}$ by

$$N_C(X_I, Y_J) := \begin{bmatrix} (T[\widehat{I}, J, 1] - S[\widehat{I}, :, 1]Y_J)^{\mathsf{T}} \\ \vdots \\ (T[\widehat{I}, J, c] - S[\widehat{I}, :, c]Y_J)^{\mathsf{T}} \end{bmatrix} \in \mathbb{F}^{(cb') \times (a-a')}, \quad M_C := \begin{bmatrix} (R[:, J, 1])^{\mathsf{T}} \\ \vdots \\ (R[:, J, c])^{\mathsf{T}} \end{bmatrix} \in \mathbb{F}^{(cb') \times r}.$$

Apply LinearEqualsAffine($\mathbb{F}$) to

$$U = \mathbb{F}^{r \times (a-a')}, \quad W = \mathbb{F}^{(cb') \times (a-a')},$$
$$(L_C : U \to W, \quad \widehat{X} \mapsto M_C \widehat{X}), \quad \Theta = \mathcal{S}_{\mathrm{DR}}, \quad N = N_C.$$

This returns either $\varnothing$, in which case our algorithm stops, or an affine subspace $\mathcal{A}_C \subseteq \mathcal{S}_{\mathrm{DR}} \times \mathbb{F}^{r \times (a-a')}$ of pairs $((X_I, Y_J), \widehat{X})$ satisfying $M_C \widehat{X} = N_C(X_I, Y_J)$.

Step 3. For each $((X_I, Y_J), \widehat{X}) \in \mathcal{A}_C$, form the unique $X \in \mathcal{X}$ such that $X[I, :] = X_I$ and $X[\widehat{I}, :] = \widehat{X}^{\mathsf{T}}$. Return $\mathcal{S}_{\mathrm{C}}$, the affine subspace of $\mathcal{X} \times \mathcal{Y}_J$ obtained by sending each $((X_I, Y_J), \widehat{X}) \in \mathcal{A}_C$ to (X, Y_J).

Timing. The same cost accounting as Proposition 3.25. Step 1 costs $\mathtt{T}_{\mathrm{DR}}(a, b, c, r, s, a', b')$, Step 2 has cost in $O\big(d\, c\, (a - a')sb' + \mathtt{T}_{p2}(m = cb',\ n = r,\ \ell = a - a',\ \xi = d,\ \eta = \eta_C,\ \kappa = \kappa_C)\big)$. Adding the three steps gives the desired cost.

Correctness. Same as Proposition 3.25. □

Example 3.28. Continue the running example with $I = J = \{1\}$, so $\widehat{I} = \{2\}$. From Example 3.24, write

$$X_I = \begin{bmatrix} \alpha & \beta \end{bmatrix}, \qquad Y_J = \begin{bmatrix} 1-\alpha \\ \gamma \end{bmatrix}.$$

Step 2 of Proposition 3.27 gives

$$M_C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad M_C \widehat{X} = N_C(X_I, Y_J) = \begin{bmatrix} 0 \\ -\gamma \end{bmatrix},$$

with $\widehat{X}$ as the unknown. Solving gives $\widehat{X} = \begin{bmatrix} 0 \\ \epsilon \end{bmatrix}$ and $\gamma = 0$. Therefore

$$\mathcal{S}_{\mathrm{C}} = \left\{ \left(\begin{bmatrix} \alpha & \beta \\ 0 & \epsilon \end{bmatrix}, \begin{bmatrix} 1-\alpha \\ 0 \end{bmatrix} \right) : \alpha, \beta, \epsilon \in \mathbb{F} \right\}$$

.

The final problem is the following.

Problem LiftRestricted.

Given Solutions to RowRestricted and ColumnRestricted.

Return If a solution exists, the affine solution space of SimultaneousSylvesterSystem, and otherwise ∅.

Solving RowRestricted and ColumnRestricted gives $\mathcal{S}_{\mathrm{R}} \subseteq \mathcal{X}_I \times \mathcal{Y}$ and $\mathcal{S}_{\mathrm{C}} \subseteq \mathcal{X} \times \mathcal{Y}_J$. For $z_R = (X_I, Y) \in \mathcal{S}_{\mathrm{R}}$ and $z_C = (X, Y_J) \in \mathcal{S}_{\mathrm{C}}$, define projections

$$\pi_J : \mathcal{X}_I \times \mathcal{Y} \to \mathcal{X}_I \times \mathcal{Y}_J, \qquad \pi_J(z_R) := (X_I, Y[:, J]),$$
$$\pi_I : \mathcal{X} \times \mathcal{Y}_J \to \mathcal{X}_I \times \mathcal{Y}_J, \qquad \pi_I(z_C) := (X[I, :], Y_J),$$

for which we want restricted solutions to agree on.

Proposition 3.29. Let $\mathcal{S}_{\mathrm{R}} \subseteq \mathcal{X}_I \times \mathcal{Y}$, and $\mathcal{S}_{\mathrm{C}} \subseteq \mathcal{X} \times \mathcal{Y}_J$ be the solutions of RowRestricted and ColumnRestricted. Let the pullback of the projection maps be defined as

$$P := \{(z_R, z_C) \in \mathcal{S}_{\mathrm{R}} \times \mathcal{S}_{\mathrm{C}} : \pi_J(z_R) = \pi_I(z_C)\},$$

the output map be defined as

$$\Psi : P \to \mathcal{X} \times \mathcal{Y}, \qquad \Psi\big((X_I, Y), (X, Y_J)\big) := (X, Y),$$

and the error map be defined as

$$\mathrm{Err} : \mathcal{X} \times \mathcal{Y} \to \mathbb{F}^{a \times b \times c}, \qquad \mathrm{Err}(X, Y) := \big(XR[:, :, k] + S[:, :, k]Y - T[:, :, k]\big)_{k \in [c]}.$$

Then $\{\Psi(z) : z \in P \text{ and } \mathrm{Err}(\Psi(z)) = 0\}$ is the solution set to LiftRestricted.

Proof. The restricted pairs agreeing from $\mathcal{S}_{\mathrm{R}}$ and $\mathcal{S}_{\mathrm{C}}$ are exactly P. For each $z \in P$, $\Psi(z)$ is a candidate solution. This candidate is in the solution set of SimultaneousSylvesterSystem exactly when its residual is zero. □

Proposition 3.30. Let $\mathcal{S}_{\mathrm{R}}, \mathcal{S}_{\mathrm{C}}$ be the outputs of RowRestricted and ColumnRestricted. If $\mathcal{S}_{\mathrm{R}} \neq \varnothing$ and $\mathcal{S}_{\mathrm{C}} \neq \varnothing$, let

$$d := \max\{\dim \mathcal{S}_{\mathrm{R}}, \dim \mathcal{S}_{\mathrm{C}}\}, \qquad P := \{(z_R, z_C) \in \mathcal{S}_{\mathrm{R}} \times \mathcal{S}_{\mathrm{C}} : \pi_J(z_R) = \pi_I(z_C)\},$$

$$\tau := \dim P \text{ if } P \neq \varnothing, \text{ and } 0 \text{ otherwise.}$$

Otherwise set both d and τ equal to 0.

As a reminder, $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, and $T \in \mathbb{F}^{a\times b\times c}$ are the input tensors to the problem, the unknown matrices are $X \in \mathbb{F}^{a\times r}$ and $Y \in \mathbb{F}^{s\times b}$, and the chosen index sets $I \subseteq [a]$ and $J \subseteq [b]$ have sizes a' and b'. Then LiftRestricted is solvable in cost

$$\begin{aligned}\mathtt{T}_{\text{lift}}(a, b, c, r, s, a', b', d, \tau) \in O\big((a'r + b's)d^2 + (\tau + 1)\, c\, a\, b\, (r + s)\\ + \mathtt{T}_{p1}(m = abc,\ n = \tau,\ \ell = 1)\big).\end{aligned}$$

Proof. **Algorithm.**

If either RowRestricted or ColumnRestricted returns $\varnothing$, then we return $\varnothing$ and stop.

Step 1 (compute the candidates). From the outputs of RowRestricted and ColumnRestricted, we get the affine frames $(z_{R,0}, \dots, z_{R,d_R}) \subseteq \mathcal{S}_{\mathrm{R}}$, and $(z_{C,0}, \dots, z_{C,d_C}) \subseteq \mathcal{S}_{\mathrm{C}}$.

Solve AffinePullback($\mathbb{F}$) with

$$\begin{aligned}&U = \mathbb{F}^{d_R}, \quad V = \mathbb{F}^{d_C}, \quad W = \mathcal{X}_I \times \mathcal{Y}_J,\\ &f : U \to W,\ \lambda \mapsto \pi_J(\chi_R(\lambda)), \quad g : V \to W,\ \mu \mapsto \pi_I(\chi_C(\mu)),\end{aligned}$$

where $\chi_R : \mathbb{F}^{d_R} \to \mathcal{S}_{\mathrm{R}}$ and $\chi_C : \mathbb{F}^{d_C} \to \mathcal{S}_{\mathrm{C}}$ are the coordinate maps. If $P = \varnothing$, return $\varnothing$. Otherwise output the frame $(p_0, \dots, p_\tau) \subseteq \mathcal{S}_{\mathrm{R}} \times \mathcal{S}_{\mathrm{C}}$ from the coordinate pullback solutions.

Step 2 (compute the error map on a frame). For each $t \in \{0, \dots, \tau\}$, compute

$$(X_t, Y_t) := \Psi(p_t), \qquad E_t := \mathrm{Err}(X_t, Y_t) \in \mathbb{F}^{a\times b\times c}, \qquad e_t := \mathrm{vec}(E_t) \in \mathbb{F}^{abc}.$$

If $\tau = 0$, this means at most one solution exists. We check if $e_0 = 0$. In the case $e_0 \neq 0$, return $\varnothing$. When $e_0 = 0$, the solution is the unique point $\Psi(p_0)$.

Else, $\tau > 0$, so we solve the inhomogeneous linear system $Mx = -e_0$, where the columns of $M \in \mathbb{F}^{abc\times\tau}$ are given by vectors $e_i - e_0$ for $i \in [\tau]$ via LinSolve($\mathbb{F}$). If it returns $\varnothing$, return $\varnothing$. Otherwise output the frame of solutions in $\mathcal{X} \times \mathcal{Y}$ by applying Ψ to the coordinate solutions.

Timing. Step 1 is one instance of AffinePullback($\mathbb{F}$), with $d = d_R$, $e = d_C$, and $m = a'r + sb'$. Under the assumptions that d_R, d_C are $O(d)$, Proposition 3.20 takes $O(d^2(a'r + sb'))$ operations. Evaluating Err at one point costs $O(c\,a\,b\,(r+s))$, so all $\tau+1$ evaluations cost $O((\tau+1)\,c\,a\,b\,(r+s))$. Solving $Mx = -e_0$ is one call to LinSolve($\mathbb{F}$) with $m = abc$, $n = \tau$, and $\ell = 1$, so it costs $\mathtt{T}_{p1}(m = abc,\ n = \tau,\ \ell = 1)$. □

Example 3.31. From Examples 3.26 and 3.28, a compatible pair in the pullback has the form $X = \begin{bmatrix} \alpha & 0 \\ 0 & \epsilon \end{bmatrix}, Y = \begin{bmatrix} 1-\alpha & 0 \\ 0 & \delta \end{bmatrix}$, for $\alpha, \epsilon, \delta \in \mathbb{F}$. Step 2 checks

$$\mathrm{Err}(X,Y) = \big(XR[:,:,k] + S[:,:,k]Y - T[:,:,k]\big)_{k\in[2]},$$

and the remaining condition is $\epsilon + \delta = 2$. Therefore LiftRestricted returns the same solution as solving the full linear system after renaming ϵ to β, namely $X = \mathrm{diag}(\alpha, \beta)$, and $Y = \mathrm{diag}(1 - \alpha, 2 - \beta)$ for $\alpha, \beta \in \mathbb{F}$.

Now we combine the above to solve SimultaneousSylvesterSystem.

Theorem 3.32. Let $I \subseteq [a]$ and $J \subseteq [b]$. Then SimultaneousSylvesterSystem is solvable in cost

$$\mathtt{T}_{\mathrm{DR}} + \mathtt{T}_{\mathrm{RR}} + \mathtt{T}_{\mathrm{CR}} + \mathtt{T}_{\mathrm{lift}},$$

where $\mathtt{T}_{\mathrm{DR}}$ is the doubly restricted solve cost defined above, $\mathtt{T}_{\mathrm{RR}}$ and $\mathtt{T}_{\mathrm{CR}}$ are the remaining row and column restricted costs from Propositions 3.25 and 3.27, not including the doubly restricted solve, and $\mathtt{T}_{\mathrm{lift}}$ is the lift cost from Proposition 3.30.

Proof. By correctness of DoubleRestricted, we obtain the doubly-restricted solution family for (I, J). By Propositions 3.25 and 3.27, the row- and column-restricted families computed from that data are exactly the solutions of RowRestricted and ColumnRestricted. Then LiftRestricted is applied to these two families. Its correctness means we solve SimultaneousSylvesterSystem. □

An end-to-end example can be found in Example A.1.

While Theorem 3.32 gives the correct solution for any choice of I and J, we describe a class of problems on which this solve-and-lift algorithm performs best.

Definition 3.33. Let $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times c}$, and C be a positive integer. We say that (R, S, T) is *Sylvester-C-regular* if there exist index sets $I \subseteq [a]$ and $J \subseteq [b]$, with $|I| = a'$ and $|J| = b'$, such that $a' \leq C$, $b' \leq C$, $a'b'c \geq a'r + b's$, and either $\mathcal{S}_{\text{DR}} = \varnothing$, or else

$$\dim \mathcal{S}_{\text{DR}} \leq C, \quad N_R(\mathcal{S}_{\text{DR}}) \subseteq \operatorname{im}(M_R), \quad N_C(\mathcal{S}_{\text{DR}}) \subseteq \operatorname{im}(M_C), \quad \operatorname{Ker}(M_R) = \operatorname{Ker}(M_C) = \{0\}.$$

These conditions ensure the doubly-restricted system is either infeasible, or the solution space has bounded dimension. They also ensure each doubly-restricted solution admits a unique row lift and a unique column lift.

The existential dependence on the index sets I and J keeps the regularity condition weak, since we only require some subsystem to have the desired lifting properties. To obtain the claimed arithmetic complexity speedup for SimultaneousSylvesterSystem requires computing those index sets, which is the following problem.

Problem ChooseSylvesterRestriction.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times c}$, and a positive integer C.

Return If (R, S, T) is Sylvester-C-regular, index sets $I \subseteq [a]$ and $J \subseteq [b]$ satisfying the requirements of (R, S, T) being Sylvester-C-regular, otherwise $\varnothing$.

For the complexity statements, we assume a black-box algorithm for the above problem. The reason is that we do not presently have a deterministic strategy beyond brute-force search for finding precisely which subsets $I \subseteq [a]$ and $J \subseteq [b]$ satisfy the regularity condition. In practice, these index sets may be efficiently chosen by heuristic or randomized methods given some knowledge of the source of the tensor data.

Definition 3.34. A collection $\mathcal{F}$ of tensor triples is called a *Sylvester regular family* if there exists a positive integer C such that every $(R, S, T) \in \mathcal{F}$ is Sylvester-C-regular.

Now we are ready to prove Theorem A, which states that Sylvester regular family instances can be solved in lower arithmetic complexity as compared to solving the full linear system, assuming we have a black-box algorithm for finding the restricted index sets.

Proof of Theorem A. Let $\mathcal{F}$ be a Sylvester regular family, and let $(R, S, T) \in \mathcal{F}$ with $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, and $T \in \mathbb{F}^{a\times b\times c}$. Let C be as in Definition 3.34, and apply the black-box algorithm for ChooseSylvesterRestriction to choose $I \subseteq [a]$ and $J \subseteq [b]$ such that (R, S, T) is Sylvester-C-regular. Recall that $|I| = a'$ and $|J| = b'$. For the doubly-restricted step, Sylvester-C-regularity gives $a'b'c \geq a'r + b's$, and hence

$$\mathtt{T}_{\mathrm{DR}} \in O\big((a'b'c)(a'r + b's)^2\big).$$

Let $n := a+b+c+r+s$. The above cost is $O(n^3)$ since $a', b' \leq C$. If the doubly-restricted system has no solution, then the full system has no solution as well, and the algorithm stops here within the claimed $O(n^4)$ bound. Otherwise, let $\mathcal{S}_{\mathrm{DR}}$ denote its affine solution space. Since (R, S, T) is Sylvester-C-regular, both feasible parameter spaces equal $\mathcal{S}_{\mathrm{DR}}$, meaning $\eta_R = \eta_C = \dim \mathcal{S}_{\mathrm{DR}} \leq C$. For the row and column restricted steps, after excluding the already counted doubly-restricted solve, the remaining costs are

stage	$\mathtt{T}_{p2}$ **parameters**	**additional term**	**cost**
RowRestricted	$(m, n, \ell, \xi, \eta, \kappa) = (ca',\ s,\ b - b',\ d_{\mathrm{DR}},\ \eta_R,\ \kappa_R)$	$d_{\mathrm{DR}}\, c\, a'r\, (b - b')$	$O(n^3)$
ColRestricted	$(m, n, \ell, \xi, \eta, \kappa) = (cb',\ r,\ a - a',\ d_{\mathrm{DR}},\ \eta_C,\ \kappa_C)$	$d_{\mathrm{DR}}\, c\, (a - a')sb'$	$O(n^3)$

Here $a, b, c, r, s \leq n$, so in each case above $m, n, \ell \in O(n)$, $\xi, \eta \leq C$, and $\kappa \in O(n)$. Recall from Proposition 3.16 that $\mathtt{T}_{p2}$ has cost bounded by sum of the terms mn^2, $\xi\kappa m\ell$, $(\kappa\ell)\xi^2$, $n^2\ell$, $(\eta+1)\xi m\ell$, and $(\eta+1)mn\ell$. In both the row and column restricted case, mn^2, $\xi\kappa m\ell$, $n^2\ell$, $(\eta +$

$1)mn\ell \in O(n^3)$, and $(\kappa\ell)\xi^2,\ (\eta+1)\xi m\ell \in O(n^2)$, so $\mathtt{T}_{p2} \in O(n^3)$. Each displayed additional term is also $O(n^3)$, so the cost of each stage is $O(n^3)$, as reflected in the cost column. Therefore RowRestricted ($\mathtt{T}_{\mathrm{RR}}$) and ColRestricted ($\mathtt{T}_{\mathrm{CR}}$) each cost $O(n^3)$.

Being Sylvester-C-regular means each point of $\mathcal{S}_{\mathrm{DR}}$ has a unique row lift and a unique column lift, so $\mathcal{S}_{\mathrm{R}}$ and $\mathcal{S}_{\mathrm{C}}$ agree on the overlap, meaning the pullback P has the same dimension. Hence $\dim \mathcal{S}_{\mathrm{R}} = \dim \mathcal{S}_{\mathrm{C}} = \dim P = \dim \mathcal{S}_{\mathrm{DR}} \leq C$, and therefore $\tau \leq C$. Recall $d = \max\{\dim \mathcal{S}_{\mathrm{R}}, \dim \mathcal{S}_{\mathrm{C}}\}$ in Proposition 3.30, we conclude

$$\begin{aligned}\mathtt{T}_{\mathrm{lift}}(a,b,c,r,s,a',b',d,\tau) \in O\big((a'r+b's)d^2 + (\tau+1)\,c\,a\,b\,(r+s) \\ + \mathtt{T}_{p1}(m=abc,\ n=\tau,\ \ell=1)\big).\end{aligned}$$

Here $(a'r+b's)d^2 \in O(n)$, $(\tau+1)\,c\,a\,b\,(r+s) \in O(n^4)$, and $\mathtt{T}_{p1}(m=abc,\ n=\tau,\ \ell=1) \in O(n^3)$, so $\mathtt{T}_{\mathrm{lift}} \in O(n^4)$. This gives the total solve-and-lift cost as $O(n^4)$.

Full deterministic verification checks all abc constraints and has cost $O(abc(r+s)) = O(n^4)$. Therefore, instances of SimultaneousSylvesterSystem from a Sylvester regular family $\mathcal{F}$ are solved in arithmetic complexity $O(n^4)$. □

3.4 Derivation Equations for derivations of tensors

In this section, we work towards Theorem B with a similar strategy as the Sylvester case. Throughout, a, b, c, r, s, t will be natural numbers, and we denote the spaces

$$\mathcal{X} := \mathbb{F}^{a\times r}, \qquad \mathcal{Y} := \mathbb{F}^{s\times b}, \qquad \mathcal{Z} := \mathbb{F}^{t\times c}.$$

This notation may have subscripts to indicate restricted solution spaces. As an example, for $I \subseteq [a]$, $\mathcal{X}_I$ is the space $\mathbb{F}^{|I|\times r}$. For a matrix $Z \in \mathbb{F}^{t\times c}$, define T^Z, the outer action of Z on T, by specifying its slices for $k \in [c]$ as

$$(T^Z)[:,:,k] := \sum_{j \in [t]} T[:,:,j]\, Z[j,k].$$

For $K \subseteq [c]$ with $|K| = c'$ and $Z_K := Z[:,K] \in \mathbb{F}^{t \times c'}$, we likewise write $T^{Z_K} \in \mathbb{F}^{a \times b \times c'}$, with slices indexed by $[c']$ in the induced order on K.

Problem DerivationSystem.

Given Tensors $R \in \mathbb{F}^{r \times b \times c}$, $S \in \mathbb{F}^{a \times s \times c}$, $T \in \mathbb{F}^{a \times b \times t}$.

Return If a solution exists, the affine subspace

$$\left\{(X,Y,Z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} : (\forall k \in [c])\ X\, R[:,:,k] + S[:,:,k]\, Y = (T^Z)[:,:,k]\right\},$$

and otherwise $\varnothing$.

This problem can be solved using LinSolve($\mathbb{F}$) with $m = abc$, $n = ar + bs + ct$, and $\ell = 1$. For a tensor $A \in \mathbb{F}^{a \times b \times c}$, taking all three input tensors equal to A and setting $r = a$, $s = b$, and $t = c$ recovers the derivation algebra of A. We describe an alternative approach, culminating in the following.

Theorem B. Let $\mathcal{F}$ be a derivation regular family of tensors as defined by Definition 3.41, meaning each problem's solution space is completely determined by a restricted subsystem. Furthermore, assume we have a black-box algorithm to find this subsystem. For each $(R,S,T) \in \mathcal{F}$ with $R \in \mathbb{F}^{r \times b \times c}$, $S \in \mathbb{F}^{a \times s \times c}$, and $T \in \mathbb{F}^{a \times b \times t}$, set $n := a + b + c + r + s + t$. Then DerivationSystem instances from $\mathcal{F}$ are solved in arithmetic cost $O(n^{4.5})$. This gives an algorithm for the derivation regular family within a factor $n^{1/2}$ of deterministic verification cost $O(n^4)$, and an improvement to the $O(n^7)$ cost of solving the full linear system.

As in the above section, we solve the derivation system by solving on a smaller subsystem, and then lifting those solutions.

Problem TripleRestrictedDer.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times t}$ and index sets $I \subseteq [a]$, $J \subseteq [b]$, and $K = \{k_1 < \cdots < k_{c'}\} \subseteq [c]$, with $|I| = a'$, $|J| = b'$, and $|K| = c'$.

Return If a solution exists, the affine subspace

$$\mathcal{D}_{I,J,K} := \{(X_I, Y_J, Z_K) \in \mathcal{X}_I \times \mathcal{Y}_J \times \mathcal{Z}_K : \\ (\forall j \in [c'])\; X_I\, R[:, J, k_j] + S[I, :, k_j]\, Y_J = (T^{Z_K})[I, J, j]\},$$

and otherwise $\varnothing$.

This is a dense linear system with $ra' + sb' + tc'$ unknowns and $a'b'c'$ equations. Hence it is solvable in cost

$$\mathtt{T}_{p1}(m = a'b'c',\ n = ra' + sb' + tc',\ \ell = 1). \tag{3.2}$$

We now describe three restricted systems which give solution families for X, Y, and Z, respectively.

Problem ColDepthRestrictedDer.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times t}$, and index sets $I \subseteq [a]$, $J \subseteq [b]$, and $K = \{k_1 < \cdots < k_{c'}\} \subseteq [c]$ with $|I| = a'$, $|J| = b'$, and $|K| = c'$.

Return If a solution exists, the affine subspace

$$\mathcal{D}_{[a],J,K} := \{(X, Y_J, Z_K) \in \mathcal{X} \times \mathcal{Y}_J \times \mathcal{Z}_K : \\ (\forall j \in [c'])\; X\, R[:, J, k_j] + S[:, :, k_j]\, Y_J = (T^{Z_K})[:, J, j]\},$$

and otherwise $\varnothing$.

Problem RowDepthRestrictedDer.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times t}$, and index sets $I \subseteq [a]$, $J \subseteq [b]$, and $K = \{k_1 < \cdots < k_{c'}\} \subseteq [c]$ with $|I| = a'$, $|J| = b'$, and $|K| = c'$.

Return If a solution exists, the affine subspace

$$\mathcal{D}_{I,[b],K} := \{(X_I, Y, Z_K) \in \mathcal{X}_I \times \mathcal{Y} \times \mathcal{Z}_K : \\ (\forall j \in [c'])\ X_I\, R[:,:,k_j] + S[I,:,k_j]\, Y = (T^{Z_K})[I,:,j]\big\},$$

and otherwise $\varnothing$.

Problem RowColRestrictedDer.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times t}$, and index sets $I \subseteq [a]$, $J \subseteq [b]$, $K \subseteq [c]$ with $|I| = a'$, $|J| = b'$, and $|K| = c'$.

Return If a solution exists, the affine subspace

$$\mathcal{D}_{I,J,[c]} := \{(X_I, Y_J, Z) \in \mathcal{X}_I \times \mathcal{Y}_J \times \mathcal{Z} : \\ (\forall k \in [c])\ X_I\, R[:,J,k] + S[I,:,k]\, Y_J = (T^{Z})[I,J,k]\big\},$$

and otherwise $\varnothing$.

Proposition 3.36. Each of the problems ColDepthRestrictedDer, RowDepthRestrictedDer, and RowColRestrictedDer are solved by TripleRestrictedDer followed by LinearEqualsAffine($\mathbb{F}$).

Proof. All three restricted problems are solved by first computing $\mathcal{D}_{I,J,K}$ using TripleRestrictedDer (this is only done once and the solution is shared), and then LinearEqualsAffine($\mathbb{F}$).

We detail RowDepthRestrictedDer. Let $\widehat{J} := [b] \setminus J$, and write $K = \{k_1 < \cdots < k_{c'}\}$. Define the affine map $N_R : \mathcal{D}_{I,J,K} \to \mathbb{F}^{(a'c')\times(b-b')}$ and $M_R \in \mathbb{F}^{a'c'\times s}$ by

$$N_R(X_I, Y_J, Z_K) := \begin{bmatrix} (T^{Z_K})[I, \widehat{J}, 1] - X_I\, R[:, \widehat{J}, k_1] \\ \vdots \\ (T^{Z_K})[I, \widehat{J}, c'] - X_I\, R[:, \widehat{J}, k_{c'}] \end{bmatrix} \in \mathbb{F}^{(a'c')\times(b-b')}, \quad M_R := \begin{bmatrix} S[I, :, k_1] \\ \vdots \\ S[I, :, k_{c'}] \end{bmatrix} \in \mathbb{F}^{(a'c')\times s}. \tag{3.3}$$

Then, for each $(X_I, Y_J, Z_K) \in \mathcal{D}_{I,J,K}$, we solve the corresponding LinearEqualsAffine($\mathbb{F}$) instance with coefficient matrix M_R and right-hand side $N_R(X_I, Y_J, Z_K)$. The solution triple (X_I, Y, Z_K) is formed by combining Y_J and $\widehat{Y}$, in a fashion identical to the Sylvester case.

The problem ColDepthRestrictedDer is solved similarly with matrices N_C and M_C, while Row-ColRestrictedDer requires some new notation to set up the system. Write $\widehat{K} = \{\ell_1 < \cdots < \ell_{c-c'}\} = [c] \setminus K$. Define a matrix $M_D \in \mathbb{F}^{a'b'\times t}$ by specifying its jth column $M_D[:, j] (j \in [t])$ as $\mathrm{vec}(T[I, J, j]) \in \mathbb{F}^{a'b'}$.

Then we define the affine map $N_D : \mathcal{D}_{I,J,K} \to \mathbb{F}^{a'b'\times(c-c')}$ by specifying the jth column of its output:

$$N_D(X_I, Y_J, Z_K)[:, j] := \mathrm{vec}\big(X_I\, R[:, J, \ell_j] + S[I, :, \ell_j]\, Y_J\big) \qquad (j \in [c - c'])$$

Now solving for $\widehat{Z} \in \mathbb{F}^{t\times(c-c')}$ in $M_D\widehat{Z} = N_D(X_I, Y_J, Z_K)$ using LinearEqualsAffine($\mathbb{F}$) concludes the solution. □

Proposition 3.37. Let $d := \dim \mathcal{D}_{I,J,K}$ and let (η_R, κ_R), (η_C, κ_C), and (η_Z, κ_Z) be the corresponding feasible parameter dimension / left nullspace dimension parameter pairs in the three LinearEqualsAffine($\mathbb{F}$) instances solving Proposition 3.36. Write $\mathtt{T}_{\mathrm{TR}} := \mathtt{T}_{p1}(m = a'b'c',\ n = a'r + sb' + tc',\ \ell = 1)$. Then the cost of each stage is as follows.

stage	step 1	step 2	step 3
RowDepth	$\mathtt{T}_{\mathrm{TR}}$	$d\,c'\,a'(b-b')\,(r+t)$	$\mathtt{T}_{p2}(m=a'c', n=s, \ell=b-b', \xi=d, \eta=\eta_R, \kappa=\kappa_R)$
ColDepth	$\mathtt{T}_{\mathrm{TR}}$	$d\,c'\,(a-a')\,b'\,(s+t)$	$\mathtt{T}_{p2}(m=b'c', n=r, \ell=a-a', \xi=d, \eta=\eta_C, \kappa=\kappa_C)$
RowCol	$\mathtt{T}_{\mathrm{TR}}$	$d\,(c-c')\,a'b'\,(r+s)$	$\mathtt{T}_{p2}(m=a'b', n=t, \ell=c-c', \xi=d, \eta=\eta_Z, \kappa=\kappa_Z)$

where summing each stage gives the cost of the corresponding restricted problem.

Proof. Each stage's cost is the sum of the contributions from the three steps. The first step is the shared triple restricted solve, the second is cost to build the affine map, and third the cost of the corresponding LinearEqualsAffine($\mathbb{F}$) instance.

We explain the step 2 cost for RowDepthRestrictedDer. Using Equation (3.3), we get for each $j \in [c']$, $X_I \in \mathbb{F}^{a' \times r}$ multiplies $R[:, \widehat{J}, k_j] \in \mathbb{F}^{r \times (b-b')}$, and $T^{Z_K}[I, \widehat{J}, j]$ is an $a' \times (b-b')$ block formed from $Z_K \in \mathbb{F}^{t \times c'}$ and the t slices of T. Hence each block costs $a'(b-b')(r+t)$, with c' blocks per frame point and d frame points, giving $d\,c'\,a'(b-b')(r+t)$. The ColDepth RestrictedDer and RowColRestrictedDer bounds are obtained in the same way from their stacked block dimensions, and summing across gives the stated total cost. □

Problem LiftRestrictedDer.

Given Solutions to ColDepthRestrictedDer, RowDepthRestrictedDer, and RowColRestrictedDer.

Return If a solution exists, the affine solution space of DerivationSystem, and otherwise $\varnothing$.

The overall strategy is the same as for LiftRestricted. Each lifted solution space has a restriction map to $\mathcal{D}_{I,J,K}$, for which we want the overall solution to agree on. Namely, define restriction maps

$$\rho_X : \mathcal{D}_{[a],J,K} \to \mathcal{D}_{I,J,K}, \quad (X, Y_J, Z_K) \mapsto (X[I,:], Y_J, Z_K),$$
$$\rho_Y : \mathcal{D}_{I,[b],K} \to \mathcal{D}_{I,J,K}, \quad (X_I, Y, Z_K) \mapsto (X_I, Y[:,J], Z_K),$$
$$\rho_Z : \mathcal{D}_{I,J,[c]} \to \mathcal{D}_{I,J,K}, \quad (X_I, Y_J, Z) \mapsto (X_I, Y_J, Z[:,K]).$$

We first form the 2-way pullback as an instance of AffinePullback($\mathbb{F}$) by solving $Q_{XY} := \{(u,v) \in \mathcal{D}_{[a],J,K} \times \mathcal{D}_{I,[b],K} : \rho_X(u) = \rho_Y(v)\}$. This gives compatible partial solution triples (X, Y, Z_K).

Then as a second pullback calculation, we solve for $Q := \{((u,v),w) \in Q_{XY} \times \mathcal{D}_{I,J,[c]} : \rho_X(u) = \rho_Y(v) = \rho_Z(w)\}$.

Define $\Psi_{\mathrm{Der}} : Q \to \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ by $\Psi_{\mathrm{Der}}\big(((X, Y_J, Z_K), (X_I, Y, Z_K)), (X_I, Y_J, Z)\big) = (X, Y, Z)$, which outputs candidate solution triples. What remains is to find candidate triples whose solutions are correct on all the constraints. This is done by defining

$$\mathrm{Err} : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to \mathbb{F}^{a \times b \times c}, \qquad \mathrm{Err}(X,Y,Z)[:,:,k] := X\,R[:,:,k] + S[:,:,k]\,Y - (T^Z)[:,:,k],$$

and solving for the candidate solutions satisfying $\mathrm{Err} = 0$ in a linear system.

An end-to-end example can be found in Example A.2.

Proposition 3.38. Let $\mathcal{D}_{I,J,K}$, $\mathcal{D}_{[a],J,K}$, $\mathcal{D}_{I,[b],K}$, and $\mathcal{D}_{I,J,[c]}$ be the outputs of TripleRestrictedDer, ColDepthRestrictedDer, and RowDepthRestrictedDer, and RowColRestrictedDer respectively. If none of these outputs is empty, let Q_{XY} and Q be the two pullbacks built in LiftRestrictedDer, and

$$\begin{array}{llll} d := \dim \mathcal{D}_{I,J,K}, & d_X := \dim \mathcal{D}_{[a],J,K}, & d_Y := \dim \mathcal{D}_{I,[b],K}, & d_Z := \dim \mathcal{D}_{I,J,[c]}, \\ m_{\mathrm{TR}} := a'r + sb' + tc', & d_{XY} := \dim Q_{XY}, & \tau := \dim Q \text{ if } Q \neq \varnothing, & \text{and } 0 \text{ otherwise.} \end{array}$$

Otherwise, let each of d, d_X, d_Y, d_Z, m_{TR}, d_{XY}, and τ equal to 0.

As a reminder, $R \in \mathbb{F}^{r \times b \times c}$, $S \in \mathbb{F}^{a \times s \times c}$, and $T \in \mathbb{F}^{a \times b \times t}$ are the input tensors, the unknowns are $X \in \mathbb{F}^{a \times r}$, $Y \in \mathbb{F}^{s \times b}$, and $Z \in \mathbb{F}^{t \times c}$, the chosen index sets $I \subseteq [a]$, $J \subseteq [b]$, and $K \subseteq [c]$ have sizes a', b', and c' respectively.

Then LiftRestrictedDer is solvable in cost

$$\begin{aligned} \mathtt{T}_{\text{lift}} \in O\big(&\mathtt{T}_{p1}(m = m_{\mathrm{TR}},\ n = d_X + d_Y,\ \ell = 1) + \mathtt{T}_{p1}(m = m_{\mathrm{TR}},\ n = d_{XY} + d_Z,\ \ell = 1) \\ &+ (\tau + 1)\,abc(r + s + t) + \mathtt{T}_{p1}(m = abc,\ n = \tau,\ \ell = 1)\big). \end{aligned}$$

Proof. If any of the four input spaces is $\varnothing$, return $\varnothing$ and stop. The first two $\mathtt{T}_{p1}$ summands are the costs of the two applications of AffinePullback($\mathbb{F}$). In both pullbacks, the common codomain is the space $\mathcal{X}_I \times \mathcal{Y}_J \times \mathcal{Z}_K$ of dimension m_{TR}. After the second pullback, we evaluate the Err map on an affine frame of size $\tau + 1$ via Ψ_{Der}. Each evaluation costs $O(abc(r + s + t))$, since for each $k \in [c]$ we must compute $X\,R[:,:,k]$, $S[:,:,k]\,Y$, and $(T^Z)[:,:,k]$. Finally, solving for the coordinates of Q that satisfy $\mathrm{Err}(\Psi_{\mathrm{Der}}(q)) = 0$ requires an application of LinSolve($\mathbb{F}$) with $m = abc$, $n = \tau$, and $\ell = 1$. □

Theorem 3.39. Fix $I \subseteq [a]$, $J \subseteq [b]$, and $K \subseteq [c]$. Then DerivationSystem is solvable in cost

$$\mathtt{T}_{\mathrm{TR}} + \mathtt{T}_{\mathrm{CD}} + \mathtt{T}_{\mathrm{RD}} + \mathtt{T}_{\mathrm{RC}} + \mathtt{T}_{\mathrm{lift}}$$

where $\mathtt{T}_{\mathrm{TR}}$ is the TripleRestrictedDer cost from Equation (3.2), $\mathtt{T}_{\mathrm{CD}}$, $\mathtt{T}_{\mathrm{RD}}$, and $\mathtt{T}_{\mathrm{RC}}$ are the three remaining restricted lift costs from Proposition 3.37, not counting solving the triple restricted system, and $\mathtt{T}_{\mathrm{lift}}$ is the cost of computing compatible solutions from Proposition 3.38.

Proof. Same reasoning as Theorem 3.32. □

While Theorem 3.39 gives the correct solution for any choice of I, J and K, we describe a class of problems on which this solve-and-lift algorithm performs best.

Definition 3.40. Let $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times t}$, and let C be a positive integer. We say that (R, S, T) is *derivation-C-regular* if there exist index sets $I \subseteq [a]$, $J \subseteq [b]$, and $K \subseteq [c]$, with $|I| = a'$, $|J| = b'$, and $|K| = c'$, with $a' \le Ca^{0.5}$, $b' \le Cb^{0.5}$, $c' \le Cc^{0.5}$, and $a'b'c' \ge a'r + sb' + tc'$. We also require for the choice of I, J, K, that $\mathcal{D}_{I,J,K} = \varnothing$, or else

$$\begin{aligned}
&\dim \mathcal{D}_{I,J,K} \le C,\\
&N_R(\mathcal{D}_{I,J,K}) \subseteq \mathrm{im}(M_R),\ N_C(\mathcal{D}_{I,J,K}) \subseteq \mathrm{im}(M_C),\ N_D(\mathcal{D}_{I,J,K}) \subseteq \mathrm{im}(M_D),\\
&\mathrm{Ker}(M_R) = \mathrm{Ker}(M_C) = \mathrm{Ker}(M_D) = \{0\}.
\end{aligned}$$

These conditions ensure the triple-restricted system is either infeasible, or the solution space has bounded dimension. They also ensure each triple-restricted solution admits unique lifts to the three restricted families.

The existential dependence on the index sets I, J, and K keeps the regularity condition weak, since we only require some subsystem to have the desired lifting properties. To obtain the claimed arithmetic complexity speedup for DerivationSystem requires computing those index sets, which is the following problem.

Problem ChooseDerivationRestriction.

Given Tensors $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, $T \in \mathbb{F}^{a\times b\times t}$, and a positive integer C.

Return If (R, S, T) is derivation-C-regular, index sets $I \subseteq [a]$, $J \subseteq [b]$, and $K \subseteq [c]$ satisfying the requirements of (R, S, T) being derivation-C-regular, otherwise $\varnothing$.

For the complexity statements, we assume a black-box algorithm for the above problem. The reason is that we do not presently have a deterministic strategy beyond brute-force search for finding precisely which subsets $I \subseteq [a]$, $J \subseteq [b]$, and $K \subseteq [c]$ satisfy the regularity condition. In practice, these index sets may be efficiently chosen by heuristic or randomized methods given some knowledge of the source of the tensor data.

Definition 3.41. A collection $\mathcal{F}$ of tensor triples is called a *derivation regular family* if there exists a positive integer C such that every $(R, S, T) \in \mathcal{F}$ is derivation-C-regular.

Now we are ready to prove Theorem B, which states that derivation regular family instances can be solved in lower arithmetic complexity as compared to solving the full linear system, assuming we have a black-box algorithm for finding the restricted index sets.

Proof of Theorem B. Let $\mathcal{F}$ be a derivation regular family, and let $(R, S, T) \in \mathcal{F}$ with $R \in \mathbb{F}^{r\times b\times c}$, $S \in \mathbb{F}^{a\times s\times c}$, and $T \in \mathbb{F}^{a\times b\times t}$. Let C be as in Definition 3.41, and apply the black-box algorithm for ChooseDerivationRestriction to choose $I \subseteq [a]$, $J \subseteq [b]$, and $K \subseteq [c]$ such that (R, S, T) is derivation-C-regular. Recall that $|I| = a'$, $|J| = b'$, and $|K| = c'$, and let $n := a+b+c+r+s+t$. For the triple-restricted step, derivation-C-regularity gives $a'b'c' \geq a'r + sb' + tc'$, and hence

$$\mathtt{T}_{\mathrm{TR}} := \mathtt{T}_{p1}(m = a'b'c',\ n = a'r + sb' + tc',\ \ell = 1) \in O\big((a'b'c')(a'r + sb' + tc')^2\big).$$

Since $a', b', c' \in O(n^{1/2})$ with constant depending only on C, and $r, s, t \leq n$, the above cost is $O(n^{4.5})$. If the triple-restricted system has no solution, then the full system has no solution as well, and the algorithm stops here. Otherwise, let $\mathcal{D}_{I,J,K}$ denote its affine solution space. Since (R, S, T) is derivation-C-regular, each point of $\mathcal{D}_{I,J,K}$ has a unique lift to each restricted family. Hence the three feasible parameter spaces all have dimension equal to $\dim \mathcal{D}_{I,J,K}$, so $\eta_R = \eta_C = \eta_Z = \dim \mathcal{D}_{I,J,K} \leq C$. Also $d = \dim \mathcal{D}_{I,J,K} \leq C$. For the three restricted lift stages, after excluding the already counted triple-restricted solve, the remaining costs are

stage	$\mathtt{T}_{p2}$ **parameters**	**additional term**	**cost**
RowDepth	$(m, n, \ell, \xi, \eta, \kappa) = (a'c',\ s,\ b - b',\ d,\ \eta_R,\ \kappa_R)$	$d\,c'a'(b - b')(r + t)$	$O(n^3)$
ColDepth	$(m, n, \ell, \xi, \eta, \kappa) = (b'c',\ r,\ a - a',\ d,\ \eta_C,\ \kappa_C)$	$d\,c'(a - a')b'(s + t)$	$O(n^3)$
RowCol	$(m, n, \ell, \xi, \eta, \kappa) = (a'b',\ t,\ c - c',\ d,\ \eta_Z,\ \kappa_Z)$	$d(c - c')a'b'(r + s)$	$O(n^3)$

Here $a, b, c, r, s, t \leq n$, so in each row of the table $m, n, \ell \in O(n)$, $\xi, \eta \leq C$, and $\kappa \leq m \in O(n)$. Recall from Proposition 3.16 that $\mathtt{T}_{p2}$ has cost bounded by the sum of the terms mn^2, $\xi\kappa m\ell$, $(\kappa\ell)\xi^2$, $n^2\ell$, $(\eta + 1)\xi m\ell$, and $(\eta + 1)mn\ell$. In all three stages, $mn^2,\ \xi\kappa m\ell,\ n^2\ell,\ (\eta + 1)mn\ell \in O(n^3)$, and $(\kappa\ell)\xi^2,\ (\eta + 1)\xi m\ell \in O(n^2)$, so $\mathtt{T}_{p2} \in O(n^3)$. Each additional term is also $O(n^3)$, so the cost of each stage is $O(n^3)$, as reflected in the final column. Therefore RowDepth ($\mathtt{T}_{\mathrm{RD}}$), ColDepth ($\mathtt{T}_{\mathrm{CD}}$), and RowCol ($\mathtt{T}_{\mathrm{RC}}$) each cost $O(n^3)$.

For the final lifting stage, recall $m_{\mathrm{TR}} = a'r + sb' + tc'$, so $m_{\mathrm{TR}} \in O(n^{3/2})$. Because the three lifts extend the same triple-restricted solution, they agree on their common restricted data, so each point of $\mathcal{D}_{I,J,K}$ determines unique compatible points in Q_{XY} and Q. Hence $d_X = d_Y = d_Z = d_{XY} = \tau = \dim \mathcal{D}_{I,J,K} \leq C$. Using Proposition 3.38, we get

$$\mathtt{T}_{\mathrm{lift}} \in O\big(\mathtt{T}_{p1}(m = m_{\mathrm{TR}},\ n = d_X + d_Y,\ \ell = 1) + \mathtt{T}_{p1}(m = m_{\mathrm{TR}},\ n = d_{XY} + d_Z,\ \ell = 1)$$

$$+ (\tau + 1)\, abc(r + s + t) + \mathtt{T}_{p1}(m = abc,\ n = \tau,\ \ell = 1)\big).$$

Here the first two $\mathtt{T}_{p1}$ terms are $O(n^{3/2})$, $(\tau + 1)\, abc(r + s + t) \in O(n^4)$, and $\mathtt{T}_{p1}(m = abc,\ n = \tau,\ \ell = 1) \in O(n^3)$, so $\mathtt{T}_{\text{lift}} \in O(n^4)$. This gives the total solve-and-lift cost as $O(n^{4.5})$. Therefore, instances of DerivationSystem from a derivation regular family $\mathcal{F}$ are solved in arithmetic complexity $O(n^{4.5})$.

Full deterministic verification checks all abc constraints and has cost $O(abc(r+s+t)) = O(n^4)$, so the solve-and-lift algorithm is within an $n^{0.5}$ factor of deterministic verification. □

3.5 Performance Benchmarking

We conclude this section with implementation performance comparisons for the solve-and-lift algorithms against the baseline algorithm of solving the full linear system. The first two experiments draw tensors from the regular families described in Theorems A and B. To show the algorithm's applicability outside of the regular family regime, we include a third case where we use the Sylvester system to compute the mid-nucleus of a single tensor. This last comparison also allows us to compare versus OpenDleto [The26], which contains a currently unpublished algorithm that is faster than the baseline linear system solver.

The restricted index sets are taken to be $[a'] \subset [a]$, $[b'] \subset [b]$, and $[c'] \subset [c]$. The code to reproduce these experiments is available at `https://github.com/cliu587/fast-der-solver`, and the timings were collected on an Apple M1 CPU with 16 gigabytes of RAM.

3.5.1 Simultaneous Sylvester Equation Solving Results

We create instances from a Sylvester regular family (Definition 3.34). The performance using the solution strategy described in Theorem A is compared to solving the large linear system as laid out in Example 3.21. See Figure 3.2 for the performance numbers.

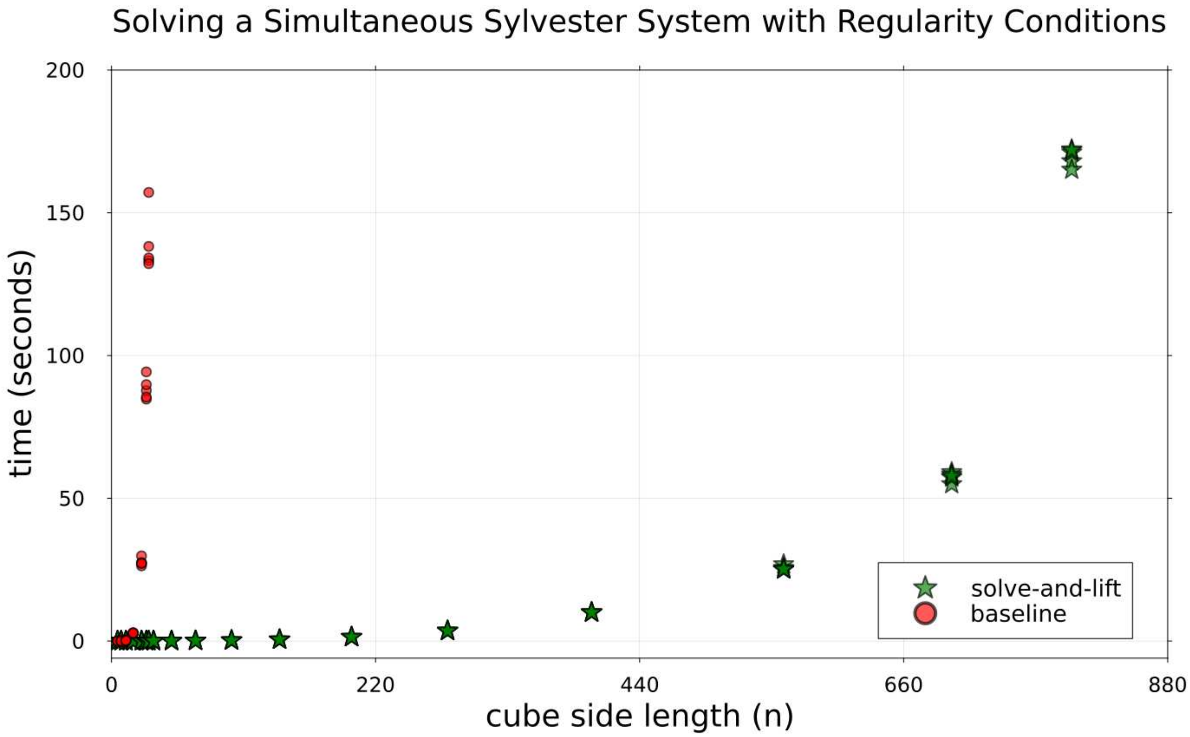


Figure 3.2: We see the performance improvements of the fast solve-and-lift algorithm as compared to solving the full linear system. Each algorithm was run 5 times for some chosen side lengths, with the results given as a scatterplot.

3.5.2 Derivation Equation Solving Results

We create instances from a derivation regular family (Definition 3.41). The performance using the solution strategy described in Theorem B is compared to solving the large linear system as laid out in an analogous fashion to Example 3.21. See Figure 3.3 for the performance numbers.

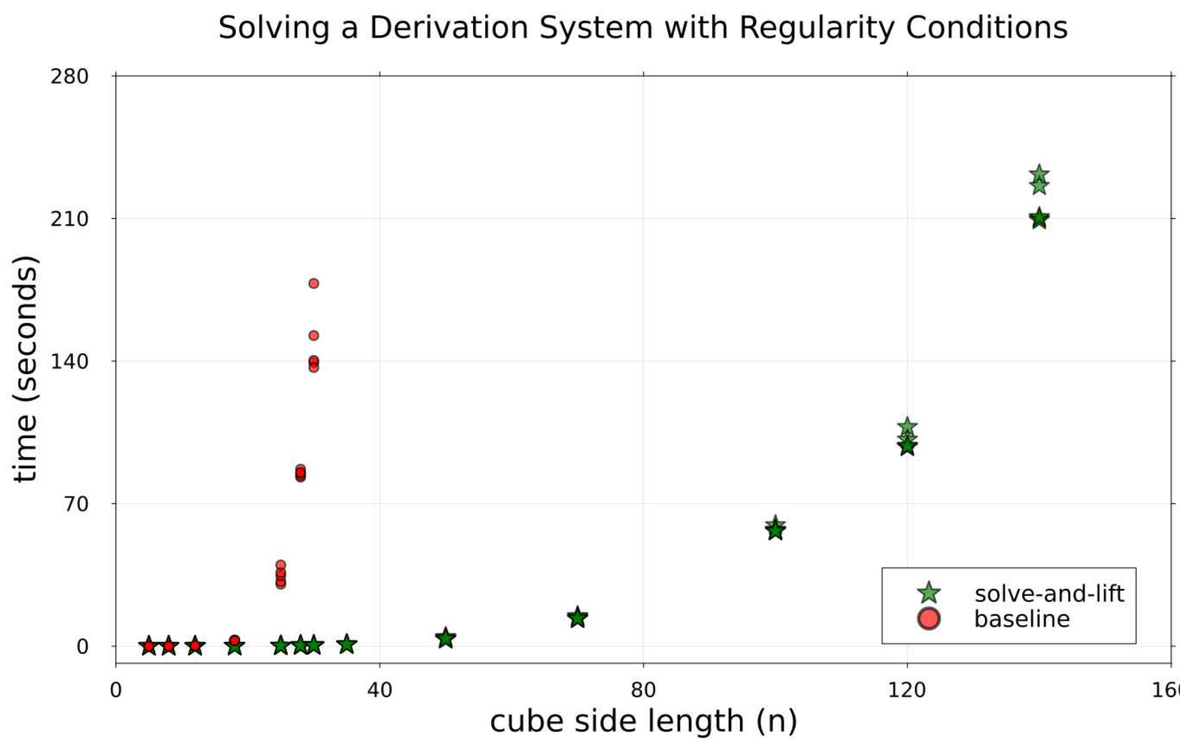


Figure 3.3: We see the performance improvements of the fast solve-and-lift algorithm as compared to solving the full linear system. Each algorithm was run 5 times for some chosen side lengths, with the results given as a scatterplot.

3.5.3 Non-Regular Family Mid Nucleus Solving Results

We also benchmark a non-regular family of instances where the solution dimension is unbounded. For each n, we choose an $n \times n$ matrix M whose minimal polynomial has degree n, which means there exists v such that $\{v, Mv, \ldots, M^{n-1}v\}$ is a basis of $\mathbb{F}^n$. Since this is a basis, for any $n \times n$ matrix X, there is a polynomial f satisfying $Xv = f(M)v$. If X commutes with M, then $XM^iv = M^iXv = f(M)M^iv$ for all i. Since X and $f(M)$ evaluate to the same outputs on a basis, $X = f(M)$. Because X is any matrix commuting with M, the centralizer of M is $\mathbb{F}[M]$.

We form an $n \times n \times n$ tensor A whose slices are sampled from $\mathbb{F}[M]$ so that they generate $\mathbb{F}[M]$, and then apply random row and column changes of basis to the slices. Then we let $R = A$, $S = -A$, and $T = 0$, so the solution space of the Sylvester system gives the mid nucleus of the tensor A, which corresponds to the centralizer of $\mathbb{F}[M]$, which is dimension n. Three methods are compared: the full linear system solving baseline, OpenDleto's unpublished solution strategy that reduces the number of constraints, and the solve-and-lift implementation.

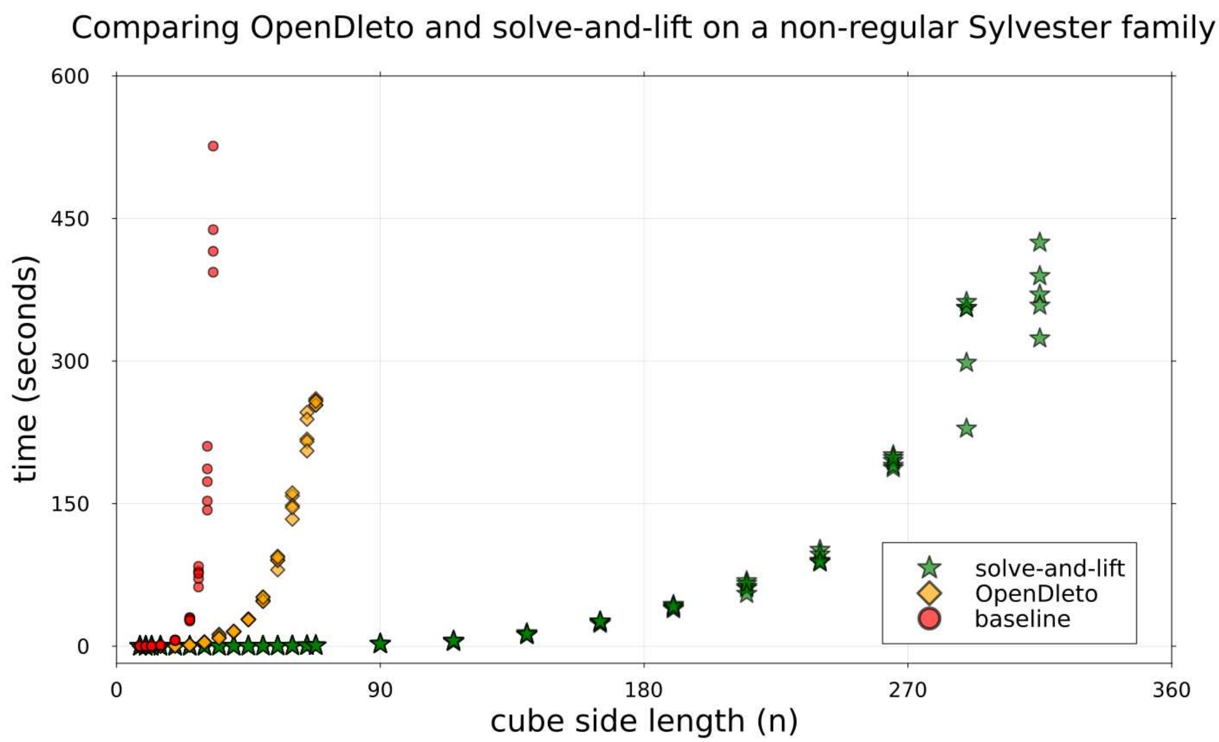


Figure 3.4: Implementation performance for a non-regular family of the solve-and-lift strategy compared to solving the full linear system as well as using OpenDleto to compute the mid nucleus of a tensor. The red show the full linear system baseline, the orange show OpenDleto, and the green show solve-and-lift. The plot shows that OpenDleto improves substantially over the baseline, while solve-and-lift remains faster on these instances even though the mid-nucleus dimension grows with n. Each algorithm was run 5 times for the chosen side lengths, with the results given as a scatterplot.

Chapter 4

Nuclei of Product of Bilinear Maps

As a reminder, for a bilinear map $b : U \times V \rightarrowtail W$, its left nucleus is $\mathcal{L}(b) = \{(\alpha, \gamma) \in \mathrm{End}(U) \times \mathrm{End}(W) : b \bullet_2 \alpha = b \bullet_0 \gamma\}$. The middle and right nucleus, $\mathcal{M}(b), \mathcal{R}(b)$, are defined analogously.

In this section, we work towards the below theorem.

Theorem C. Let $s : U_2 \times U_1 \rightarrowtail U_0$ and $t : V_2 \times V_1 \rightarrowtail V_0$ be fully nondegenerate bilinear maps. Then

$$\mathcal{L}(s\otimes t) = \mathcal{L}(s)\boxtimes\mathcal{L}(t), \quad \mathcal{M}(s\otimes t) = \mathcal{M}(s)\boxtimes\mathcal{M}(t), \quad \mathcal{R}(s\otimes t) = \mathcal{R}(s)\boxtimes\mathcal{R}(t).$$

As a result of this theorem, we also have the following.

Corollary 4.2. Let $s : U_2 \times U_1 \rightarrowtail U_0$ and $t : V_2 \times V_1 \rightarrowtail V_0$ be fully nondegenerate bilinear maps. Then $\mathcal{C}(s \otimes t) = \mathcal{C}(s) \boxtimes \mathcal{C}(t)$.

4.1 Previous results

Proposition 4.3 ([Wil09, Propositions 7.8 and 7.9])**.** Let $d : U \times U \rightarrowtail C$ be a nondegenerate Hermitian C-form and $b : V \times V \rightarrowtail W$ a k-bilinear map. Then $d \otimes b$ is Hermitian and $\mathrm{Adj}(d \otimes b) = \mathrm{Adj}(d) \otimes \mathrm{Adj}(b)$.

In our notation, $\mathrm{Adj}(b) = \mathcal{M}(b)$. In the Hermitian setting, the adjoint ring is faithfully represented on one axis, so the notation $\mathrm{Adj}(d) \otimes \mathrm{Adj}(b)$ is unambiguous. Theorem C recovers this case via $\mathcal{M}(d \otimes b) = \mathcal{M}(d) \boxtimes \mathcal{M}(b)$.

4.2 Algebra preliminaries

Definition 4.4. Let U, V be finite-dimensional $\mathbb{F}$-vector spaces, and $t \in U \otimes V$. Define

$$\Phi_L : U \otimes V \to \operatorname{Hom}(V^*, U), \qquad \Phi_L(u \otimes v)(\varphi) = \varphi(v)u,$$
$$\Phi_R : U \otimes V \to \operatorname{Hom}(U^*, V), \qquad \Phi_R(u \otimes v)(\psi) = \psi(u)v$$
$$\operatorname{Im}_L(t) \coloneqq \operatorname{Im}(\Phi_L(t)) \leq U, \qquad \operatorname{Im}_R(t) \coloneqq \operatorname{Im}(\Phi_R(t)) \leq V.$$

Proposition 4.5. Let U, V be finite-dimensional $\mathbb{F}$-vector spaces and $t \in U \otimes V$. Then $t \in \operatorname{Im}_L(t) \otimes \operatorname{Im}_R(t) \leq U \otimes V$.

Proof. Let $A \coloneqq \operatorname{Im}_L(t) \leq U$, and $X \coloneqq \operatorname{Im}_R(t) \leq V$. Fix a basis $v_1, \dots, v_n$ of V with dual basis $\varphi_1, \dots, \varphi_n$. Then $t = \sum_{j=1}^n \Phi_L(t)(\varphi_j) \otimes v_j \in A \otimes V$, since each $\Phi_L(t)(\varphi_j)$ lies in A.

Fix a basis $u_1, \dots, u_m$ of U with dual basis $\psi_1, \dots, \psi_m$. Then $t = \sum_{i=1}^m u_i \otimes \Phi_R(t)(\psi_i) \in U \otimes X$, since each $\Phi_R(t)(\psi_i)$ lies in X. Therefore by Lemma 2.4, $t \in (A \otimes V) \cap (U \otimes X) = A \otimes X$. $\square$

Remark 4.6. Let U, V be finite-dimensional $\mathbb{F}$-vector spaces. Let $\{u_i\}$ be an ordered basis of U and $\{v_j\}$ an ordered basis of V. Then t can be uniquely written as $t = \sum_{i,j} T_{ij}\, u_i \otimes v_j \in U \otimes V$. This identifies t with the matrix $T = (T_{ij})$.

Example 4.7. Let $U = \mathbb{F}^2$ with standard bases e_1, e_2 and $V = \mathbb{F}^3$ with standard basis f_1, f_2, f_3. Let $t = e_1 \otimes f_1 + e_2 \otimes f_1 + e_2 \otimes f_2$. Then t corresponds to the matrix $T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$.

Definition 4.8. Let U, V be finite-dimensional $\mathbb{F}$-vector spaces and $t \in U \otimes V$. The *rank* of t is the minimum number r such that t can be written as a sum $t = \sum_{i=1}^r u_i \otimes v_i$, for $u_i \in U$, and $v_i \in V$. The sum $\sum_{i=1}^r u_i \otimes v_i$ is called an *minimal rank decomposition* of t. [KB09, Section 3.1]

We'll use the observation that for $t \in U \otimes V$, the rank of t is less than or equal to the minimum of $\dim U$, and $\dim V$, as we can always expand t either in a basis of U or a basis of V.

Lemma 4.9. Let $t = \sum_{i=1}^r u_i \otimes v_i$ in $U \otimes V$, for $r = \operatorname{rank}(t)$. Then $\{u_1, \dots, u_r\}$ and $\{v_1, \dots, v_r\}$ are linearly independent, and $\operatorname{Im}_L(t) = \operatorname{span}\{u_1, \dots, u_r\}$, $\operatorname{Im}_R(t) = \operatorname{span}\{v_1, \dots, v_r\}$.

Proof. Set $W := \operatorname{span}\{u_1, \dots, u_r\}$ and $X := \operatorname{span}\{v_1, \dots, v_r\}$. Since $t \in W \otimes V$, the above observation gives $\operatorname{rank}(t) \le \dim W$. Since W is spanned by r vectors, $\dim W \le r$. Together with $\operatorname{rank}(t) = r$, this gives $\dim W = r$, so $\{u_1, \dots, u_r\}$ is linearly independent. An analogous argument gives $\dim X = r$, so $\{v_1, \dots, v_r\}$ is linearly independent.

Extend $v_1, \dots, v_r$ to a basis of V, and let $\varphi_1, \dots, \varphi_r \in V^*$ be the corresponding dual functionals. Then $\Phi_L(t)(\varphi_i) = u_i$, so $W \le \operatorname{Im}_L(t)$.

For the reverse inclusion, we see for all $\varphi \in V^*$, $\Phi_L(t)(\varphi) = \sum_{i=1}^r \varphi(v_i)\, u_i \in W$, so $\operatorname{Im}_L(t) \le W$. Hence $\operatorname{Im}_L(t) = W$. An analogous argument concludes $\operatorname{Im}_R(t) = X$. □

Remark 4.10. Suppose $\sum_{i=1}^r a_i \otimes x_i = \sum_{i=1}^r b_i \otimes y_i$ are two minimal rank decompositions of the same element $t \in U \otimes V$. Fix ordered bases of U and V, and let A, B be matrices whose columns are coordinates of a_i, b_i, and X, Y be matrices whose rows coordinates of x_i, y_i. In these fixed coordinates, we get the equality $AX = BY$. By Lemma 4.9, the columns of A and B are two ordered bases of the same space $\operatorname{Im}_L(t)$, so there is a unique $M \in \mathrm{GL}_r(\mathbb{F})$ such that $B = AM$. Then $AX = BY = AMY$, and since the columns of A are linearly independent, $X = MY$, so $Y = M^{-1}X$.

Example 4.11. Take $\mathbb{F} = \mathbb{F}_5$, $U = V = \mathbb{F}^3$, and $r = 2$. Define in the standard basis

$$A = \left[\begin{array}{c|c} & \\ a_1 & a_2 \\ & \end{array}\right] = \left[\begin{array}{c|c} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{array}\right], \quad X = \left[\begin{array}{c} x_1 \\ \hline x_2 \end{array}\right] = \left[\begin{array}{ccc} 1 & 2 & 0 \\ \hline 0 & 1 & 1 \end{array}\right]$$

$$B = \left[\begin{array}{c|c} & \\ b_1 & b_2 \\ & \end{array}\right] = \left[\begin{array}{c|c} 1 & 2 \\ 0 & 3 \\ 3 & 4 \end{array}\right], \quad Y = \left[\begin{array}{c} y_1 \\ \hline y_2 \end{array}\right] = \left[\begin{array}{ccc} 3 & 2 & 1 \\ \hline 4 & 0 & 2 \end{array}\right].$$

A direct multiplication verifies

$$AX = BY = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

meaning $a_1 \otimes x_1 + a_2 \otimes x_2 = b_1 \otimes y_1 + b_2 \otimes y_2$. As per Remark 4.10, we find

$$M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad M^{-1} = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix},$$

satisfying

$$AM = \begin{bmatrix} 1 & 2 \\ 0 & 3 \\ 3 & 4 \end{bmatrix} = B, \qquad M^{-1}X = \begin{bmatrix} 3 & 2 & 1 \\ 4 & 0 & 2 \end{bmatrix} = Y.$$

4.3 Key reduction

We'll use Lemma 2.14 to reduce an equality about the image of axis actions to a matrix factorization statement. As a reminder for U, V finite dimensional vector spaces, the map $\kappa : \mathrm{End}(U) \otimes \mathrm{End}(V) \to \mathrm{End}(U \otimes V)$ is the Kronecker isomorphism.

Lemma 4.12. Let $s : U_2 \times U_1 \rightarrowtail U_0$ and $t : V_2 \times V_1 \rightarrowtail V_0$ be fully nondegenerate bilinear maps. Let $K \subseteq \{2, 1, 0\}$ be nonempty, and $w \in \mathrm{Bil}(U_2, U_1; U_0) \otimes \mathrm{Bil}(V_2, V_1; V_0)$ be an element such that for each $k \in K$, there is some $x_k \in \mathrm{End}(U_k \otimes V_k)$ satisfying

$$w = \Xi\big((s \otimes t) \bullet_k x_k\big) \in \mathrm{Bil}(U_2, U_1; U_0) \otimes \mathrm{Bil}(V_2, V_1; V_0).$$

Fix a minimal rank decomposition $w = \sum_{i=1}^{r} u_i \otimes v_i$. Then for each $k \in K$, there exists $\{\alpha_{k,i} \in \mathrm{End}(U_k)\}_{i=1}^{r}$ and $\{f_{k,i} \in \mathrm{End}(V_k)\}_{i=1}^{r}$ such that for all $i \in \{1, \dots, r\}$, we have $u_i = s \bullet_k \alpha_{k,i}$, and $v_i = t \bullet_k f_{k,i}$, with $x_k = \sum_{i=1}^{r} \kappa_k(\alpha_{k,i} \otimes f_{k,i})$.

Proof. Fix $k \in K$ and any decomposition $x_k = \sum_j \kappa_k(\alpha_j \otimes f_j)$. Corollary 2.15 gives $w = \sum_j \varphi(s \bullet_k \alpha_j, t \bullet_k f_j)$, so $w \in \mathrm{Im}(\Lambda_k^s) \otimes \mathrm{Im}(\Lambda_k^t)$, with $\mathrm{Im}_L(w) \leq \mathrm{Im}(\Lambda_k^s)$, and $\mathrm{Im}_R(w) \leq \mathrm{Im}(\Lambda_k^t)$.

Since $w = \sum_i u_i \otimes v_i$ is a minimal rank decomposition, apply Lemma 4.9 to conclude for each i that the elements $u_i \in \mathrm{Im}_L(w) \leq \mathrm{Im}(\Lambda_k^s)$ and $v_i \in \mathrm{Im}_R(w) \leq \mathrm{Im}(\Lambda_k^t)$ are images of axis actions on s and t respectively. Specifically, there exist $\alpha_{k,i} \in \mathrm{End}(U_k)$ and $f_{k,i} \in \mathrm{End}(V_k)$ such that $u_i = s \bullet_k \alpha_{k,i}$, and $v_i = t \bullet_k f_{k,i}$.

Finally, for each k, set $x'_k := \sum_{i=1}^r \kappa_k(\alpha_{k,i} \otimes f_{k,i})$. Since $\Xi((s \otimes t) \bullet_k x_k) = w = \Xi((s \otimes t) \bullet_k x'_k)$ and Ξ is an isomorphism, the images of axis actions are equal. The injectivity of $\Lambda_k^{s \otimes t}$ then gives $x_k = x'_k$. □

4.4 Proof of Theorem C

Proof. We first prove the middle nucleus case, $\mathcal{M}(s \otimes t) = \mathcal{M}(s) \boxtimes \mathcal{M}(t)$.

($\subseteq$) Let $(x, y) \in \mathcal{M}(s \otimes t)$, meaning $(s \otimes t) \bullet_2 x = (s \otimes t) \bullet_1 y$. Set $w := \Xi\big((s \otimes t) \bullet_2 x\big) = \Xi\big((s \otimes t) \bullet_1 y\big) \in \mathrm{Bil}(U_2, U_1; U_0) \otimes \mathrm{Bil}(V_2, V_1; V_0)$. Fix a minimal rank decomposition $w = \sum_{i=1}^r u_i \otimes v_i$. Apply Lemma 4.12 with $K = \{2, 1\}$. We obtain $\alpha_{2,i} =: \alpha_i \in \mathrm{End}(U_2)$, $\alpha_{1,i} =: \beta_i \in \mathrm{End}(U_1)$, $f_{2,i} =: f_i \in \mathrm{End}(V_2)$, $f_{1,i} =: g_i \in \mathrm{End}(V_1)$ such that $s \bullet_2 \alpha_i = u_i = s \bullet_1 \beta_i$, $t \bullet_2 f_i = v_i = t \bullet_1 g_i$, and $x = \sum_{i=1}^r \kappa_2(\alpha_i \otimes f_i)$, $y = \sum_{i=1}^r \kappa_1(\beta_i \otimes g_i)$.

Hence $(\alpha_i, \beta_i) \in \mathcal{M}(s)$ and $(f_i, g_i) \in \mathcal{M}(t)$ for each i, so we have the desired conclusion

$$(x, y) = \sum_{i=1}^{r} \big((\alpha_i, \beta_i) \boxtimes (f_i, g_i)\big) \in \mathcal{M}(s) \boxtimes \mathcal{M}(t).$$

($\supseteq$) Let $(\alpha, \beta) \in \mathcal{M}(s)$ and $(f, g) \in \mathcal{M}(t)$. Then

$$(s \otimes t) \bullet_2 \kappa_2(\alpha \otimes f) = (s \bullet_2 \alpha) \otimes (t \bullet_2 f) = (s \bullet_1 \beta) \otimes (t \bullet_1 g) = (s \otimes t) \bullet_1 \kappa_1(\beta \otimes g).$$

So $(\alpha, \beta) \boxtimes (f, g) = (\kappa_2(\alpha \otimes f), \kappa_1(\beta \otimes g)) \in \mathcal{M}(s \otimes t)$. Since $\mathcal{M}(s) \boxtimes \mathcal{M}(t)$ is the span of such $\boxtimes$ terms, we get $\mathcal{M}(s) \boxtimes \mathcal{M}(t) \subseteq \mathcal{M}(s \otimes t)$. The left and right nuclei case are analogous.

□

A worked-out nucleus decomposition example is in the appendix (Example C.1).

Now we wrap up this chapter with the proof that $\mathcal{C}(s \otimes t) = \mathcal{C}(s) \boxtimes \mathcal{C}(t)$.

Proof of Corollary 4.2. We prove the direction $\mathcal{C}(s \otimes t) \subseteq \mathcal{C}(s) \boxtimes \mathcal{C}(t)$. Let $(x_2, x_1, x_0) \in \mathcal{C}(s \otimes t)$. Then $(s \otimes t) \bullet_2 x_2 = (s \otimes t) \bullet_1 x_1 = (s \otimes t) \bullet_0 x_0$. Let $w \coloneqq \Xi\big((s \otimes t) \bullet_2 x_2\big) \in \mathrm{Bil}(U_2, U_1; U_0) \otimes \mathrm{Bil}(V_2, V_1; V_0)$ be the element in common. Fix a minimal rank decomposition $w = \sum_{i=1}^{r} u_i \otimes v_i$, and apply Lemma 4.12 with $K = \{2, 1, 0\}$, then follow the same argument as the proof of Theorem C.

The direction $\mathcal{C}(s \otimes t) \supseteq \mathcal{C}(s) \boxtimes \mathcal{C}(t)$ is analogous to the proof of $\mathcal{M}(s \otimes t) \supseteq \mathcal{M}(s) \boxtimes \mathcal{M}(t)$ in Theorem C. $\square$

Chapter 5

Derivation of Product of Bilinear Maps

The main result of this section is the following.

Theorem D. Let s and t be fully nondegenerate bilinear maps. Then

$$\begin{aligned}\mathrm{Der}(s \otimes t) = &\mathrm{Der}(s) \boxtimes \mathcal{C}(t) + \mathcal{C}(s) \boxtimes \mathrm{Der}(t)\\ &+ \iota_{20}(\mathcal{L}(s) \boxtimes \mathcal{L}(t)) + \iota_{10}(\mathcal{R}(s) \boxtimes \mathcal{R}(t)) + \iota_{21}(\mathcal{M}(s) \boxtimes \mathcal{M}(t)),\end{aligned} \tag{5.1}$$

where $\iota_{20}(\alpha, \gamma) = (\alpha, 0, \gamma)$, $\iota_{10}(\beta, \gamma) = (0, \beta, \gamma)$, and $\iota_{21}(\alpha, \beta) = (-\alpha, \beta, 0)$.

The strategy is similar to the nuclei case, where we relate equations of axis actions on tensor product bilinear maps to equalities in the tensor product of vector spaces.

5.1 Previous results

In this subsection, algebras are finite-dimensional over $\mathbb{F}$, and not necessarily associative. For an algebra A, its derivation algebra refers to

$$\mathcal{D}_{\text{alg}}(A) \coloneqq \{d \in \mathrm{End}(A) : \forall x, y \in A,\ d(xy) = d(x)y + xd(y)\}.$$

Brešar proved the following.

Theorem 5.2 (Brešar, Theorem 3.1 [Bre17])**.** Let R and S be nonassociative unital algebras. Suppose that either at least one of R and S is finite-dimensional or they are both finitely generated. Then every derivation d of $R \otimes S$ can be written as

$$d = \mathrm{ad}(u) + \sum_{j=1}^{p} \lambda_{z_j} \otimes f_j + \sum_{i=1}^{q} g_i \otimes \lambda_{w_i}, \tag{5.2}$$

where $u \in \mathrm{Nuc}(R) \otimes \mathrm{Nuc}(S)$, $z_j \in Z(R)$, $w_i \in Z(S)$, $f_j \in \mathcal{D}_{\text{alg}}(S)$, and $g_i \in \mathcal{D}_{\text{alg}}(R)$.

Proposition 5.3. Let R and S be unital algebras over $\mathbb{F}$, with multiplication maps μ_R, μ_S. Then Theorem D implies Theorem 5.2 in the case both R and S are finite dimensional.

Proof. See Appendix B. □

Specialization to Matrix Algebras. Let A be a unital finite-dimensional $\mathbb{F}$-algebra. As a vector space, $M_n(A) \cong \mathbb{F}^{n\times n} \otimes A$, and the multiplication on $M_n(A)$ is defined by $(X \otimes a)(Y \otimes b) = XY \otimes ab$.

Theorem 5.4 (Benkart–Osborn, Corollary 4.9 [BO81]). Assume $\operatorname{char}\mathbb{F} \neq 2, 3$ and $n > 3$. Then for any unital algebra A, $\mathcal{D}_{\text{alg}}(M_n(A)) = \mathcal{D}_{\text{alg}}(L_n'(A)) = \mathcal{D}_{\text{alg}}(K_n(A)) = (\mathcal{D}_{\text{alg}}(A))^{\#} + \operatorname{ad}_{M_n(N)}$.

Remark 5.5. Notation wise, in Theorem 5.4, N denotes the nucleus of A, $L_n(A)$ is $M_n(A)$ under the Lie product, $L_n'(A) = [L_n(A), L_n(A)]$, and $K_n(A) = L_n'(A)/Z(L_n'(A))$. Also, $(\mathcal{D}_{\text{alg}}(A))^{\#}$ denotes entrywise application of derivations of A. Thus in the notation of this chapter,

$$N = \operatorname{Nuc}(A), \qquad (\mathcal{D}_{\text{alg}}(A))^{\#} = I_n \otimes \mathcal{D}_{\text{alg}}(A), \qquad \operatorname{ad}_{M_n(N)} = \operatorname{ad}(M_n(\operatorname{Nuc}(A))).$$

So Theorem 5.4 proves $\mathcal{D}_{\text{alg}}(M_n(A)) = \big(I_n \otimes \mathcal{D}_{\text{alg}}(A)\big) + \operatorname{ad}(M_n(\operatorname{Nuc}(A)))$.

Benkart–Osborn notes "It should be commented that in showing $\operatorname{Der} M_n(A) = \operatorname{Der}(A)^{\#} + \operatorname{ad}_{M_n(N)}$ no restrictions on $\operatorname{char}\Phi$ or on n are necessary." [BO81, Page 425]. As a result, we have the following.

Proposition 5.6. Theorem D implies Theorem 5.4.

Proof. Theorem D implies Theorem 5.2 in the finite dimensional case by Proposition 5.3, and Theorem 5.2 implies Theorem 5.4 by [Bre17, Cor. 4.2]. □

5.2 Preliminaries

The following decomposition definition features prominently. Throughout this subsection, U and V are always finite dimensional $\mathbb{F}$-vector spaces.

Definition 5.7. Let $p, q, r \in U \otimes V$ be nonzero tensors satisfying $p + q = r$. Let

$$A \coloneqq \mathrm{Im}_L(p), \quad B \coloneqq \mathrm{Im}_L(q), \quad C \coloneqq \mathrm{Im}_L(r),$$
$$X \coloneqq \mathrm{Im}_R(p), \quad Y \coloneqq \mathrm{Im}_R(q), \quad Z \coloneqq \mathrm{Im}_R(r),$$

where $p \in A \otimes X$, $q \in B \otimes Y$, $r \in C \otimes Z$ by Proposition 4.5. A (p, q, r) *locally independent unified decomposition* (LIU decomposition) consists of a natural number n, and vectors

$$a_i \in A,\ b_i \in B,\ c_i \in C, \qquad x_i \in X,\ y_i \in Y,\ z_i \in Z,$$

for $i \in [n]$, satisfying the following:

- **Decomposition equality.**

$$p = \sum_{i=1}^{n} a_i \otimes x_i, \qquad q = \sum_{i=1}^{n} b_i \otimes y_i, \qquad r = \sum_{i=1}^{n} c_i \otimes z_i.$$

- **Local equality.** For all $i \in [n]$, $a_i \otimes x_i + b_i \otimes y_i = c_i \otimes z_i$.

Example 5.8. Let $U = V = \mathbb{F}^2$ with bases u_1, u_2 and v_1, v_2. Define $p \coloneqq u_1 \otimes v_1 + u_1 \otimes v_2$, $q \coloneqq u_2 \otimes v_1 - u_1 \otimes v_2$, and $r \coloneqq u_1 \otimes v_1 + u_2 \otimes v_1$. Then $p + q = r$, and $A = \mathrm{span}\{u_1\}$, $B = U$, $C = U$ as subspaces of U, and $X = V$, $Y = V$, $Z = \mathrm{span}\{v_1\}$ as subspaces of V. Now let

$$a_1 = u_1,\ b_1 = u_2,\ c_1 = u_1 + u_2, \qquad x_1 = y_1 = z_1 = v_1,$$
$$a_2 = u_1,\ b_2 = -u_1,\ c_2 = 0, \qquad x_2 = y_2 = v_2,\ z_2 = v_1.$$

Then for each $i \in \{1, 2\}$, $a_i \otimes x_i + b_i \otimes y_i = c_i \otimes z_i$, and

$$p = a_1 \otimes x_1 + a_2 \otimes x_2, \quad q = b_1 \otimes y_1 + b_2 \otimes y_2, \quad r = c_1 \otimes z_1 + c_2 \otimes z_2.$$

Thus this is an LIU decomposition of $p + q = r$.

5.3 Existence of a LIU decomposition of the equation $p + q = r$

Convention 5.9. For the remainder of this subsection, whenever given nonzero tensors $p, q, r \in U \otimes V$ satisfying $p + q = r$, let

$$A \coloneqq \mathrm{Im}_L(p), \quad B \coloneqq \mathrm{Im}_L(q), \quad C \coloneqq \mathrm{Im}_L(r),$$
$$X \coloneqq \mathrm{Im}_R(p), \quad Y \coloneqq \mathrm{Im}_R(q), \quad Z \coloneqq \mathrm{Im}_R(r).$$

It is always the case that $A + B \leq U$ and $X + Y \leq V$. We also assume without loss of generality that $U \coloneqq A + B$, $V \coloneqq X + Y$, because it is always the case that

$$p \in A \otimes X \leq (A + B) \otimes (X + Y) \leq U \otimes V$$
$$q \in B \otimes Y \leq (A + B) \otimes (X + Y) \leq U \otimes V$$
$$r \in \mathrm{Im}_L(r) \otimes \mathrm{Im}_R(r).$$

Furthermore, $\mathrm{Im}_L(r), \mathrm{Im}_R(r)$ satisfies

$$\mathrm{Im}_L(r) = \mathrm{Im}_L(p + q) \leq \mathrm{Im}_L(p) + \mathrm{Im}_L(q) = A + B,$$
$$\mathrm{Im}_R(r) = \mathrm{Im}_R(p + q) \leq \mathrm{Im}_R(p) + \mathrm{Im}_R(q) = X + Y.$$

So we may redefine U to the space $A + B$, and redefine V to the space $X + Y$.

Lemma 5.10. Let U be a $\mathbb{F}$-vector space, and let $A, B, C \leq U$ with $U = A + B$. Define $U^{(1)} \leq U$ by

$$U^{(1)} \coloneqq A \cap B \cap C.$$

Define $U^{(2)} \leq A \cap B$, $U^{(3)} \leq A \cap C$, and $U^{(4)} \leq B \cap C$ by choosing complements

$$A \cap B = U^{(1)} \oplus U^{(2)}, \qquad A \cap C = U^{(1)} \oplus U^{(3)}, \qquad B \cap C = U^{(1)} \oplus U^{(4)}.$$

Define $U^{(5)} \le A$ and $U^{(6)} \le B$ by choosing complements

$$A = (U^{(1)} + U^{(2)} + U^{(3)}) \oplus U^{(5)}, \qquad B = (U^{(1)} + U^{(2)} + U^{(4)}) \oplus U^{(6)}.$$

Then

$$U = \bigoplus_{i=1}^{6} U^{(i)}, \qquad A = \bigoplus_{i \in \{1,2,3,5\}} U^{(i)}, \qquad B = \bigoplus_{i \in \{1,2,4,6\}} U^{(i)}.$$

Proof. Let $d_i := \dim U^{(i)}$ for $1 \le i \le 6$. By construction, $\dim(A \cap B) = d_1 + d_2$, $\dim(A \cap C) = d_1 + d_3$. Using the fact $\dim(X + Y) = \dim X + \dim Y - \dim(X \cap Y)$, with $X = A \cap B$, $Y = A \cap C$, we get

$$\dim\big((A \cap B) + (A \cap C)\big) = (d_1 + d_2) + (d_1 + d_3) - d_1 = d_1 + d_2 + d_3.$$

Since $A = ((A \cap B) + (A \cap C)) \oplus U^{(5)}$, this gives $\dim A = d_1 + d_2 + d_3 + d_5$. Similarly, from $B = ((A \cap B) + (B \cap C)) \oplus U^{(6)}$, $\dim B = d_1 + d_2 + d_4 + d_6$. Finally, as $U = A + B$, $\dim U = \dim(A + B) = \dim A + \dim B - \dim(A \cap B) = \sum_{i=1}^{6} d_i$.

Recall

$$U = A + B = \sum_{i=1}^{6} U^{(i)}, \qquad A = \sum_{i \in \{1,2,3,5\}} U^{(i)}, \qquad B = \sum_{i \in \{1,2,4,6\}} U^{(i)},$$

and the fact that for any finite family of subspaces $\{W_i\}$, $\dim(\sum_i W_i) \le \sum_i \dim W_i$, with equality if and only if the sum is direct. Applying this to the spaces U, A, B concludes

$$U = \bigoplus_{i=1}^{6} U^{(i)}, \qquad A = \bigoplus_{i \in \{1,2,3,5\}} U^{(i)}, \qquad B = \bigoplus_{i \in \{1,2,4,6\}} U^{(i)}. \qquad \square$$

The above lemma is to prepare for below, which is the key one in proving the existence of a LIU decomposition for the equation $p + q = r$.

Lemma 5.11. Let $p, q, r \in U \otimes V$ be nonzero elements satisfying $p + q = r$, and

$$A := \mathrm{Im}_L(p), \quad B := \mathrm{Im}_L(q), \quad C := \mathrm{Im}_L(r),$$
$$X := \mathrm{Im}_R(p), \quad Y := \mathrm{Im}_R(q), \quad Z := \mathrm{Im}_R(r).$$

Recall via Convention 5.9 that $U = A + B$, $V = X + Y$. With respect to the decomposition in Lemma 5.10, p, q, r satisfy:

$$\begin{aligned} p &\in (U^{(1)} \otimes V^{(1)}) \oplus (U^{(1)} \otimes V^{(\bullet)}) \oplus (U^{(\bullet)} \otimes V^{(1)}) \oplus (U^{(2)} \otimes V^{(2)}) \\ &\quad \oplus (U^{(3)} \otimes V^{(3)}), \\ q &\in (U^{(1)} \otimes V^{(1)}) \oplus (U^{(1)} \otimes V^{(\bullet)}) \oplus (U^{(\bullet)} \otimes V^{(1)}) \oplus (U^{(2)} \otimes V^{(2)}) \\ &\quad \oplus (U^{(4)} \otimes V^{(4)}), \\ r &\in (U^{(1)} \otimes V^{(1)}) \oplus (U^{(1)} \otimes V^{(\bullet)}) \oplus (U^{(\bullet)} \otimes V^{(1)}) \oplus (U^{(3)} \otimes V^{(3)}) \\ &\quad \oplus (U^{(4)} \otimes V^{(4)}), \end{aligned}$$

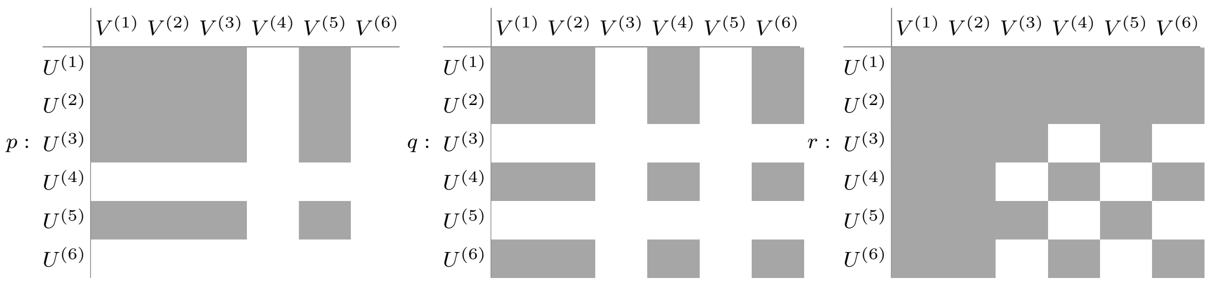
where $U^{(\bullet)} := U^{(2)} \oplus \cdots \oplus U^{(6)}$ and $V^{(\bullet)} := V^{(2)} \oplus \cdots \oplus V^{(6)}$.

Proof. Apply Lemma 5.10 to the spaces $A, B, C \leq U$ and to the spaces $X, Y, Z \leq V$. From the definitions, $A = U^{(1)} \oplus U^{(2)} \oplus U^{(3)} \oplus U^{(5)}$ and $B = U^{(1)} \oplus U^{(2)} \oplus U^{(4)} \oplus U^{(6)}$, and similarly $X = V^{(1)} \oplus V^{(2)} \oplus V^{(3)} \oplus V^{(5)}$ and $Y = V^{(1)} \oplus V^{(2)} \oplus V^{(4)} \oplus V^{(6)}$. Let $\pi_{ij} : U \otimes V \to U^{(i)} \otimes V^{(j)}$ be the block projections and write

$$\pi_{i\bullet} := \sum_{j=1}^{6} \pi_{ij}, \qquad \pi_{\bullet j} := \sum_{i=1}^{6} \pi_{ij}.$$

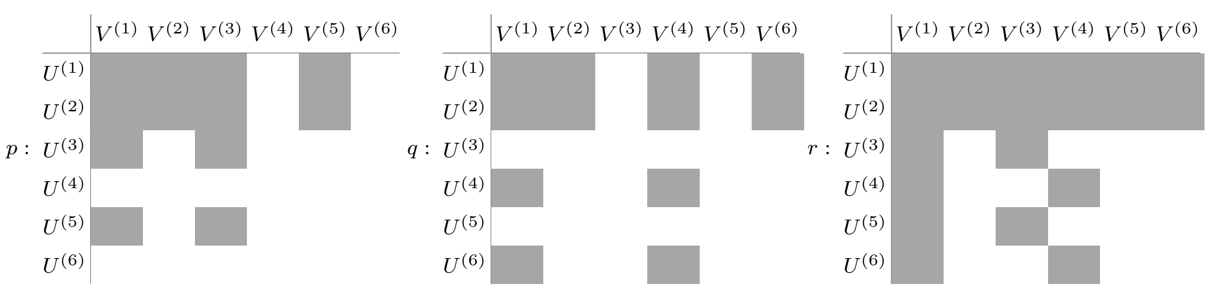
As $p \in A \otimes X$ and A is a direct sum of spaces not including $U^{(4)}$ and $U^{(6)}$, we have $\pi_{4\bullet}(p) = \pi_{6\bullet}(p) = 0$. Similarly, $\pi_{3\bullet}(q) = \pi_{5\bullet}(q) = 0$, $\pi_{\bullet 4}(p) = \pi_{\bullet 6}(p) = 0$ and $\pi_{\bullet 3}(q) = \pi_{\bullet 5}(q) = 0$.

For $1 \leq i, j \leq 6$, set $p_{ij} := \pi_{ij}(p)$, $q_{ij} := \pi_{ij}(q)$, and $r_{ij} := \pi_{ij}(r)$. Applying π_{ij} to $p + q = r$ gives for $(1 \leq i, j \leq 6)$, $p_{ij} + q_{ij} = r_{ij}$. This is a blockwise $p + q = r$ equality: if two entries in a block are zero, then the third is zero. Taking stock we visualize these relations below, where grey means *potentially nonzero*, white means *guaranteed to be zero*. For instance, the above equations says for p, the rows for $U^{(4)}$, and $U^{(6)}$, and the columns for $V^{(4)}$ and $V^{(6)}$ are 0, which are illustrated below.



Row inferences. Since $\pi_{3\bullet}(q) = 0$, define $s_3 := \pi_{3\bullet}(p) = \pi_{3\bullet}(r)$. Then $s_3 \in U^{(3)} \otimes X$ and $s_3 \in U^{(3)} \otimes Z$, hence $s_3 \in U^{(3)} \otimes (X \cap Z) = U^{(3)} \otimes \big(V^{(1)} \oplus V^{(3)}\big)$. So row $U^{(3)}$ of p, r has support only in columns $V^{(1)}, V^{(3)}$. The same argument with $U^{(5)}$ gives row $U^{(5)}$ of p, r also in $V^{(1)} \oplus V^{(3)}$.

Since $\pi_{4\bullet}(p) = 0$, define $t_4 := \pi_{4\bullet}(q) = \pi_{4\bullet}(r)$. Then $t_4 \in U^{(4)} \otimes Y$ and $t_4 \in U^{(4)} \otimes Z$, hence $t_4 \in U^{(4)} \otimes (Y \cap Z) = U^{(4)} \otimes \big(V^{(1)} \oplus V^{(4)}\big)$. So row $U^{(4)}$ of q, r has support only in columns $V^{(1)}, V^{(4)}$. The same argument with $U^{(6)}$ gives row $U^{(6)}$ of q, r also in $V^{(1)} \oplus V^{(4)}$. After these row inferences, the potentially nonzero blocks are:

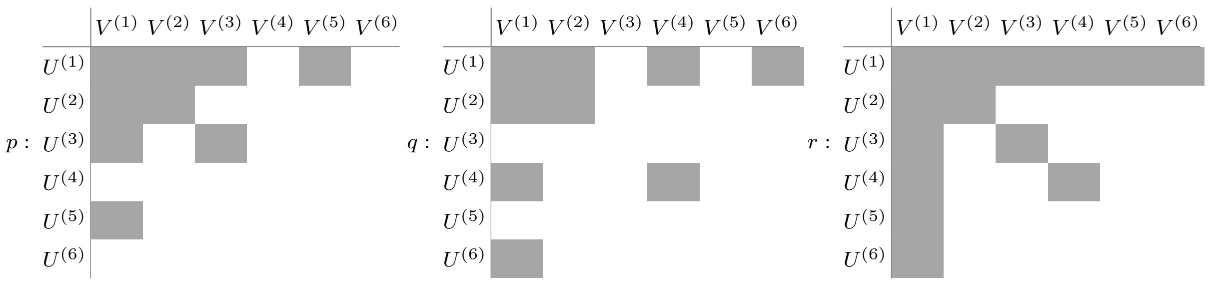


Column inferences. Since $\pi_{\bullet 3}(q) = 0$, define $u_3 := \pi_{\bullet 3}(p) = \pi_{\bullet 3}(r)$. Then $u_3 \in A \otimes V^{(3)}$ and $u_3 \in C \otimes V^{(3)}$, hence $u_3 \in (A \cap C) \otimes V^{(3)} = \big(U^{(1)} \oplus U^{(3)}\big) \otimes V^{(3)}$. So column $V^{(3)}$ of p, r has support only in rows $U^{(1)}, U^{(3)}$.

Since $\pi_{\bullet 4}(p) = 0$, define $u_4 := \pi_{\bullet 4}(q) = \pi_{\bullet 4}(r)$. Then $u_4 \in B \otimes V^{(4)}$ and $u_4 \in C \otimes V^{(4)}$, hence $u_4 \in (B \cap C) \otimes V^{(4)} = \big(U^{(1)} \oplus U^{(4)}\big) \otimes V^{(4)}$. So column $V^{(4)}$ of q, r has support only in rows $U^{(1)}, U^{(4)}$.

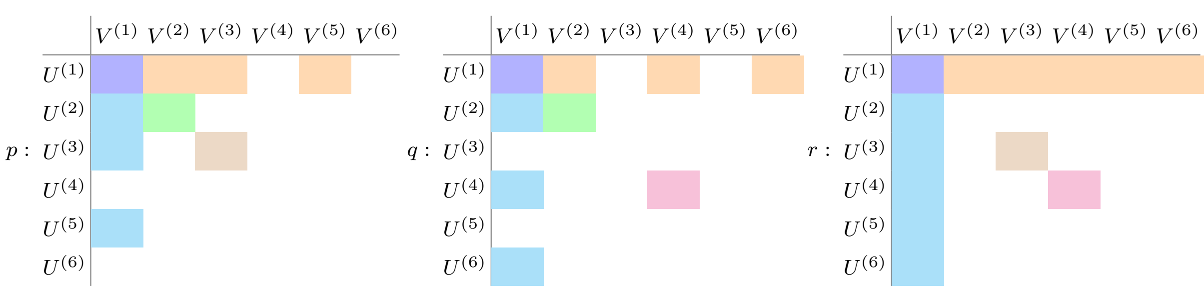
Since $\pi_{\bullet 5}(q) = 0$, define $u_5 := \pi_{\bullet 5}(p) = \pi_{\bullet 5}(r)$. Then $u_5 \in (A \cap C) \otimes V^{(5)} = \big(U^{(1)} \oplus U^{(3)}\big) \otimes V^{(5)}$. Combining this with the row-$U^{(3)}$ and row-$U^{(5)}$ restrictions above forces support in column $V^{(5)}$ to lie only in row $U^{(1)}$.

Since $\pi_{\bullet 6}(p) = 0$, define $u_6 := \pi_{\bullet 6}(q) = \pi_{\bullet 6}(r)$. Then $u_6 \in (B \cap C) \otimes V^{(6)} = \big(U^{(1)} \oplus U^{(4)}\big) \otimes V^{(6)}$. Combining this with the row-$U^{(4)}$ and row-$U^{(6)}$ restrictions above forces support in column $V^{(6)}$ to lie only in row $U^{(1)}$. After these column inferences, the potentially nonzero blocks are:



The $U^{(2)} \otimes V^{(2)}$ block relation. Set $v_2 := \pi_{2\bullet}(r)$. By the previous inferences, row $U^{(2)}$ of both p and q is supported in $V^{(1)} \oplus V^{(2)} = X \cap Y$, hence $v_2 \in U^{(2)} \otimes (X \cap Y)$. And also by definition, $r \in C \otimes Z$, so projecting, $\pi_{2\bullet}(r) = v_2 \in U^{(2)} \otimes Z$. Thus $v_2 \in U^{(2)} \otimes ((X \cap Y) \cap Z) = U^{(2)} \otimes V^{(1)}$. This means $r_{22} = 0$, which is the last zero to be inferred in the block equations.

What remains are the following support patterns, which we give as colored cells below. Here blue is $U^{(1)} \otimes V^{(1)}$, orange is $U^{(1)} \otimes V^{(\bullet)}$, cyan is $U^{(\bullet)} \otimes V^{(1)}$, green is $U^{(2)} \otimes V^{(2)}$, brown is $U^{(3)} \otimes V^{(3)}$, and pink is $U^{(4)} \otimes V^{(4)}$, and blank entries are zero.



This gives the claimed support pattern

$$p \in (U^{(1)} \otimes V^{(1)}) \oplus (U^{(1)} \otimes V^{(\bullet)}) \oplus (U^{(\bullet)} \otimes V^{(1)}) \oplus (U^{(2)} \otimes V^{(2)}) \oplus (U^{(3)} \otimes V^{(3)}),$$

$$q \in (U^{(1)} \otimes V^{(1)}) \oplus (U^{(1)} \otimes V^{(\bullet)}) \oplus (U^{(\bullet)} \otimes V^{(1)}) \oplus (U^{(2)} \otimes V^{(2)}) \oplus (U^{(4)} \otimes V^{(4)}),$$

$$r \in (U^{(1)} \otimes V^{(1)}) \oplus (U^{(1)} \otimes V^{(\bullet)}) \oplus (U^{(\bullet)} \otimes V^{(1)}) \oplus (U^{(3)} \otimes V^{(3)}) \oplus (U^{(4)} \otimes V^{(4)}). \qquad \square$$

Example 5.12. (This calculation was done by Magma, with the script available at https://gist.github.com/cliu587/0db323a55bec8f2d014e870844a18d71.)

Let $\mathbb{F} = \mathbb{F}_{997}$, let $U = V = \mathbb{F}^6$ with standard bases, and identify $U \otimes V$ with $\mathbb{F}^{6\times 6}$ via $u \otimes v \mapsto [u][v]^T$ where $[u]$ is the column vector of coordinates.

Let $p, q, r \in U \otimes V$ correspond to the 6×6 matrices below.

$$p = \begin{bmatrix} 373 & 831 & 66 & 724 & 159 & 68 \\ 772 & 812 & 372 & 663 & 579 & 209 \\ 155 & 549 & 499 & 757 & 500 & 64 \\ 218 & 733 & 547 & 132 & 342 & 122 \\ 859 & 708 & 25 & 767 & 78 & 648 \\ 992 & 398 & 894 & 611 & 169 & 953 \end{bmatrix} \quad q = \begin{bmatrix} 176 & 355 & 631 & 671 & 759 & 955 \\ 252 & 645 & 740 & 978 & 500 & 638 \\ 308 & 485 & 963 & 867 & 212 & 39 \\ 146 & 830 & 155 & 854 & 264 & 14 \\ 30 & 375 & 193 & 887 & 38 & 692 \\ 107 & 281 & 218 & 640 & 676 & 450 \end{bmatrix} \quad r = \begin{bmatrix} 549 & 189 & 697 & 398 & 918 & 26 \\ 27 & 460 & 115 & 644 & 82 & 847 \\ 463 & 37 & 465 & 627 & 712 & 103 \\ 364 & 566 & 702 & 986 & 606 & 136 \\ 889 & 86 & 218 & 657 & 116 & 343 \\ 102 & 679 & 115 & 254 & 845 & 406 \end{bmatrix}$$

so that $p + q = r$ in $\mathbb{F}^{6\times 6}$. Using Magma, we find change-of-basis matrices M and N so that the new coordinates are adapted to the decomposition $U = \bigoplus_{i=1}^{6} U^{(i)}$ and $V = \bigoplus_{i=1}^{6} V^{(i)}$.

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 310 & 34 & 931 & 898 & 971 & 773 \\ 104 & 130 & 918 & 780 & 84 & 647 \\ 795 & 834 & 145 & 316 & 939 & 574 \\ 386 & 64 & 11 & 618 & 494 & 754 \\ 732 & 237 & 949 & 667 & 810 & 127 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 655 & 49 & 786 & 153 & 994 \\ 0 & 757 & 318 & 920 & 875 & 310 \\ 0 & 304 & 386 & 307 & 423 & 363 \\ 0 & 988 & 354 & 652 & 88 & 712 \\ 0 & 780 & 227 & 987 & 917 & 291 \\ 0 & 59 & 338 & 600 & 869 & 429 \end{bmatrix},$$

the transformed matrices $p' = MpN$, $q' = MqN$, $r' = MrN$ have the support patterns as described by Lemma 5.11.

$$p' = \begin{array}{c|cccccc} & V^{(1)} & V^{(2)} & V^{(3)} & V^{(4)} & V^{(5)} & V^{(6)} \\ \hline U^{(1)} & 373 & 16 & 259 & 0 & 968 & 0 \\ U^{(2)} & 119 & 905 & 0 & 0 & 0 & 0 \\ U^{(3)} & 966 & 0 & 123 & 0 & 0 & 0 \\ U^{(4)} & 0 & 0 & 0 & 0 & 0 & 0 \\ U^{(5)} & 647 & 0 & 0 & 0 & 0 & 0 \\ U^{(6)} & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \quad q' = \begin{array}{c|cccccc} & V^{(1)} & V^{(2)} & V^{(3)} & V^{(4)} & V^{(5)} & V^{(6)} \\ \hline U^{(1)} & 176 & 828 & 0 & 553 & 0 & 244 \\ U^{(2)} & 607 & 92 & 0 & 0 & 0 & 0 \\ U^{(3)} & 0 & 0 & 0 & 0 & 0 & 0 \\ U^{(4)} & 68 & 0 & 0 & 815 & 0 & 0 \\ U^{(5)} & 0 & 0 & 0 & 0 & 0 & 0 \\ U^{(6)} & 970 & 0 & 0 & 0 & 0 & 0 \end{array} \quad r' = \begin{array}{c|cccccc} & V^{(1)} & V^{(2)} & V^{(3)} & V^{(4)} & V^{(5)} & V^{(6)} \\ \hline U^{(1)} & 549 & 844 & 259 & 553 & 968 & 244 \\ U^{(2)} & 726 & 0 & 0 & 0 & 0 & 0 \\ U^{(3)} & 966 & 0 & 123 & 0 & 0 & 0 \\ U^{(4)} & 68 & 0 & 0 & 815 & 0 & 0 \\ U^{(5)} & 647 & 0 & 0 & 0 & 0 & 0 \\ U^{(6)} & 970 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Lemma 5.13. Let $p, q, r \in U \otimes V$ be nonzero tensors satisfying $p + q = r$. Then (p, q, r) admits an LIU decomposition.

Proof. Recall from Convention 5.9 the spaces $A, B, C \leq U$, and $X, Y, Z \leq V$ where $U = A + B$ and $V = X + Y$. Decompose U and V by Lemma 5.10, getting

$$U \otimes V = \bigoplus_{i,j=1}^{6} U^{(i)} \otimes V^{(j)}.$$

Let $\pi_{ij} : U \otimes V \to U^{(i)} \otimes V^{(j)}$ be the corresponding projections. And use the same notation as Lemma 5.11, meaning we define

$$p_{11} \coloneqq \pi_{11}(p), \quad p_{1\bullet} \coloneqq \sum_{j=2}^{6} \pi_{1j}(p), \quad p_{\bullet 1} \coloneqq \sum_{i=2}^{6} \pi_{i1}(p), \quad p_{22} \coloneqq \pi_{22}(p), \quad p_{33} \coloneqq \pi_{33}(p),$$

and analogously define $q_{11},\ q_{1\bullet},\ q_{\bullet 1},\ q_{22},\ q_{44}$, and $r_{11},\ r_{1\bullet},\ r_{\bullet 1},\ r_{33},\ r_{44}$. By Lemma 5.11, all other blocks vanish, so $p = p_{11} + p_{1\bullet} + p_{\bullet 1} + p_{22} + p_{33}$, $q = q_{11} + q_{1\bullet} + q_{\bullet 1} + q_{22} + q_{44}$, and $r = r_{11} + r_{1\bullet} + r_{\bullet 1} + r_{33} + r_{44}$. Applying the relevant projections to $p + q = r$ gives equations

$$\begin{array}{lll} p_{11} + q_{11} = r_{11} & p_{1\bullet} + q_{1\bullet} = r_{1\bullet} & p_{\bullet 1} + q_{\bullet 1} = r_{\bullet 1}, \\ p_{22} + q_{22} = 0 & p_{33} = r_{33} & q_{44} = r_{44}. \end{array}$$

We build an LIU decomposition for each block to prove the lemma. In $U^{(1)} \otimes V^{(1)}$ (blue block), choose bases of $U^{(1)}$ and $V^{(1)}$. Writing p_{11}, q_{11}, r_{11} in that basis, gives coefficientwise local equalities. The $U^{(2)} \otimes V^{(2)}$ (green block), $U^{(3)} \otimes V^{(3)}$ (brown block), and $U^{(4)} \otimes V^{(4)}$ (pink block) cases are specializations of the $U^{(1)} \otimes V^{(1)}$ (blue block) argument.

In $U^{(1)} \otimes V^{(\bullet)}$ (orange block), choose a basis $\{u_m\}_{m=1}^{d}$ of $U^{(1)}$ and write

$$p_{1\bullet} = \sum_{m=1}^{d} u_m \otimes f_m, \qquad q_{1\bullet} = \sum_{m=1}^{d} u_m \otimes g_m, \qquad r_{1\bullet} = \sum_{m=1}^{d} u_m \otimes h_m.$$

Since $p_{1\bullet} + q_{1\bullet} = r_{1\bullet}$, we get that for all m in $[d]$, $u_m \otimes f_m + u_m \otimes g_m = u_m \otimes h_m$. This gives local equalities with a common left factor in $A \cap B \cap C$. The $U^{(\bullet)} \otimes V^{(1)}$ (cyan block) case is analogous. Summing these six blockwise families gives the required LIU decomposition. □

Example 5.14. We demonstrate this splitting for Example 5.12. Recall all coefficients are in $\mathbb{F}_{997}$, and let u_i and v_j be basis vectors with $U^{(i)} = \langle u_i \rangle$ and $V^{(j)} = \langle v_j \rangle$.

For $U^{(1)} \otimes V^{(1)}$ (blue block), $p_{11} = 373\, u_1 \otimes v_1$, $q_{11} = 176\, u_1 \otimes v_1$, and $r_{11} = 549\, u_1 \otimes v_1$, so $p_{11} + q_{11} = (373 + 176)\, u_1 \otimes v_1 = 549\, u_1 \otimes v_1 = r_{11}$.

For $U^{(1)} \otimes V^{(\bullet)}$ (orange block), $p_{1\bullet} = u_1 \otimes (16v_2 + 259v_3 + 968v_5)$, and $q_{1\bullet} = u_1 \otimes (828v_2 + 553v_4 + 244v_6)$, so $p_{1\bullet} + q_{1\bullet} = u_1 \otimes (844v_2 + 259v_3 + 553v_4 + 968v_5 + 244v_6) = r_{1\bullet}$.

Likewise, for $U^{(\bullet)} \otimes V^{(1)}$ (cyan block), $p_{\bullet 1} = (119u_2 + 966u_3 + 647u_5) \otimes v_1$, and $q_{\bullet 1} = (607u_2 + 68u_4 + 970u_6) \otimes v_1$, hence $p_{\bullet 1} + q_{\bullet 1} = (726u_2 + 966u_3 + 68u_4 + 647u_5 + 970u_6) \otimes v_1 = r_{\bullet 1}$.

5.4 Proof of Theorem D

We apply the LIU decomposition to the images of axis actions and match each block to a summand in Equation (5.1). First, we need to describe the partitioning of the index sets of a LIU decomposition.

Remark 5.15. Given a LIU decomposition of $p + q = r$ produced by Lemma 5.13, with

$$p = \sum_{i=1}^{n} a_i \otimes x_i, \qquad q = \sum_{i=1}^{n} b_i \otimes y_i, \qquad r = \sum_{i=1}^{n} c_i \otimes z_i,$$

and local equalities for all i in $[n]$, $a_i \otimes x_i + b_i \otimes y_i = c_i \otimes z_i$. Let $I_{11},\ I_{1\bullet},\ I_{\bullet 1},\ I_{22},\ I_{33},\ I_{44}$ be the index sets coming from the six blockwise decompositions in the proof of Lemma 5.13. Then these sets give a partition $[n] = I_{11} \sqcup I_{1\bullet} \sqcup I_{\bullet 1} \sqcup I_{22} \sqcup I_{33} \sqcup I_{44}$.

Proof. ($\subseteq$) We prove the inclusion

$$\begin{aligned} \mathrm{Der}(s \otimes t) \subseteq{}& \mathrm{Der}(s) \boxtimes \mathcal{C}(t) + \mathcal{C}(s) \boxtimes \mathrm{Der}(t) \\ &+ \iota_{20}(\mathcal{L}(s) \boxtimes \mathcal{L}(t)) + \iota_{10}(\mathcal{R}(s) \boxtimes \mathcal{R}(t)) + \iota_{21}(\mathcal{M}(s) \boxtimes \mathcal{M}(t)). \end{aligned}$$

Let $(\delta^{(2)}, \delta^{(1)}, \delta^{(0)}) \in \mathrm{Der}(s \otimes t)$, and define

$$p \coloneqq \Xi((s \otimes t) \bullet_2 \delta^{(2)}), \qquad q \coloneqq \Xi((s \otimes t) \bullet_1 \delta^{(1)}), \qquad r \coloneqq \Xi((s \otimes t) \bullet_0 \delta^{(0)}).$$

Then $p + q = r$ in $\mathrm{Bil}(U_2, U_1; U_0) \otimes \mathrm{Bil}(V_2, V_1; V_0)$.

If $r = 0$, then $(s \otimes t) \bullet_2 \delta^{(2)} = -(s \otimes t) \bullet_1 \delta^{(1)}$, so $(-\delta^{(2)}, \delta^{(1)}) \in \mathcal{M}(s \otimes t)$, and $\Lambda_0^{s\otimes t}$ is injective so $\delta^{(0)} = 0$. Thus $(\delta^{(2)}, \delta^{(1)}, \delta^{(0)}) = (\delta^{(2)}, \delta^{(1)}, 0) \in \iota_{21}(\mathcal{M}(s \otimes t))$. The cases $p = 0$ and $q = 0$ are analogous, obtaining elements of $\iota_{10}(\mathcal{R}(s \otimes t))$ and $\iota_{20}(\mathcal{L}(s \otimes t))$ respectively. Applying Theorem C proves the desired inclusion in all three cases. Hence for the remainder of the proof, assume p, q, r are all nonzero.

The strategy is to apply Lemma 5.13 to $p + q = r$, obtain a LIU decomposition, then lift the local equalities to the axis operators. As a reminder, we've defined

$$\begin{aligned} A &\coloneqq \mathrm{Im}_L(p), \quad B \coloneqq \mathrm{Im}_L(q), \quad C \coloneqq \mathrm{Im}_L(r), \\ X &\coloneqq \mathrm{Im}_R(p), \quad Y \coloneqq \mathrm{Im}_R(q), \quad Z \coloneqq \mathrm{Im}_R(r), \end{aligned}$$

where $p \in A \otimes X$, $q \in B \otimes Y$, and $r \in C \otimes Z$ by Proposition 4.5. Let $(a_i, b_i, c_i, x_i, y_i, z_i)_{i=1}^n$ be the LIU decomposition by applying Lemma 5.13. By Remark 5.15, the LIU decomposition index set $[n]$ is partitioned into $[n] = I_{11} \sqcup I_{1\bullet} \sqcup I_{\bullet 1} \sqcup I_{22} \sqcup I_{33} \sqcup I_{44}$. Below we give a summary of how we handle each index block as a summand in the decomposition of the derivation algebra of $s \otimes t$.

index block	local equality	support	corresponding summand
I_{11}	$a_i = b_i = c_i,\ x_i + y_i = z_i$	$U^{(1)} \otimes V^{(1)}$	$\mathcal{C}(s) \boxtimes \mathrm{Der}(t)$
$I_{1\bullet}$	$a_i = b_i = c_i,\ x_i + y_i = z_i$	$U^{(1)} \otimes V^{(\bullet)}$	$\mathcal{C}(s) \boxtimes \mathrm{Der}(t)$
$I_{\bullet 1}$	$x_i = y_i = z_i,\ a_i + b_i = c_i$	$U^{(\bullet)} \otimes V^{(1)}$	$\mathrm{Der}(s) \boxtimes \mathcal{C}(t)$
I_{22}	$a_i = b_i,\ x_i + y_i = 0$	$U^{(2)} \otimes V^{(2)}$	$\iota_{21}(\mathcal{M}(s) \boxtimes \mathcal{M}(t))$
I_{33}	$a_i = c_i,\ x_i = z_i$	$U^{(3)} \otimes V^{(3)}$	$\iota_{20}(\mathcal{L}(s) \boxtimes \mathcal{L}(t))$
I_{44}	$b_i = c_i,\ y_i = z_i$	$U^{(4)} \otimes V^{(4)}$	$\iota_{10}(\mathcal{R}(s) \boxtimes \mathcal{R}(t))$

- $i \in I_{1\bullet}$ ($U^{(1)} \otimes V^{(\bullet)}$) or $i \in I_{11}$ ($U^{(1)} \otimes V^{(1)}$): For each i, the local equality is

$$a_i = b_i = c_i, \qquad x_i + y_i = z_i,$$

with $a_i, b_i, c_i \in A \cap B \cap C$, $x_i \in X$, $y_i \in Y$, $z_i \in Z$. The I_{11} case deserves explaination. The coefficientwise decomposition gives a local equality $\lambda u \otimes v + \mu u \otimes v = \nu u \otimes v$, which we rewrite as $u \otimes (\lambda v) + u \otimes (\mu v) = u \otimes (\nu v)$.

By Lemma 2.14 and Proposition 4.5, we have

$$A \leq \mathrm{Im}(\Lambda_2^s), \quad B \leq \mathrm{Im}(\Lambda_1^s), \quad C \leq \mathrm{Im}(\Lambda_0^s),$$
$$X \leq \mathrm{Im}(\Lambda_2^t), \quad Y \leq \mathrm{Im}(\Lambda_1^t), \quad Z \leq \mathrm{Im}(\Lambda_0^t).$$

So each of $a_i, b_i, c_i, x_i, y_i, z_i$ has a preimage under the corresponding axis action maps. Choose

$$\alpha_i \in \mathrm{End}(U_2), \qquad \beta_i \in \mathrm{End}(U_1), \qquad \gamma_i \in \mathrm{End}(U_0)$$
$$f_i \in \mathrm{End}(V_2), \qquad g_i \in \mathrm{End}(V_1), \qquad h_i \in \mathrm{End}(V_0)$$

such that $a_i = s \bullet_2 \alpha_i,\ b_i = s \bullet_1 \beta_i,\ c_i = s \bullet_0 \gamma_i$, and $x_i = t \bullet_2 f_i,\ y_i = t \bullet_1 g_i,\ z_i = t \bullet_0 h_i$.

Then $(\alpha_i, \beta_i, \gamma_i) \in \mathcal{C}(s)$ and $(f_i, g_i, h_i) \in \mathrm{Der}(t)$, so $(\kappa_2(\alpha_i \otimes f_i), \kappa_1(\beta_i \otimes g_i), \kappa_0(\gamma_i \otimes h_i)) \in \mathcal{C}(s) \boxtimes \mathrm{Der}(t)$.

- $i \in I_{\bullet 1}$ $(U^{(\bullet)} \otimes V^{(1)})$: Here each local equality satisfies $x_i = y_i = z_i$ and $a_i + b_i = c_i$. Choose

$$\alpha_i \in \mathrm{End}(U_2), \qquad \beta_i \in \mathrm{End}(U_1), \qquad \gamma_i \in \mathrm{End}(U_0)$$
$$f_i \in \mathrm{End}(V_2), \qquad g_i \in \mathrm{End}(V_1), \qquad h_i \in \mathrm{End}(V_0)$$

  such that $a_i = s \bullet_2 \alpha_i,\ b_i = s \bullet_1 \beta_i,\ c_i = s \bullet_0 \gamma_i$ and $x_i = t \bullet_2 f_i,\ y_i = t \bullet_1 g_i,\ z_i = t \bullet_0 h_i$. Then $(\alpha_i, \beta_i, \gamma_i) \in \mathrm{Der}(s)$ and $(f_i, g_i, h_i) \in \mathcal{C}(t)$, so $(\kappa_2(\alpha_i \otimes f_i), \kappa_1(\beta_i \otimes g_i), \kappa_0(\gamma_i \otimes h_i)) \in \mathrm{Der}(s) \boxtimes \mathcal{C}(t)$.

- $i \in I_{22}, I_{33}, I_{44}$: these are the nucleus cases with one side equal to 0. We show the I_{22} case, which has $c_i \otimes z_i = 0$. The local equality is $a_i \otimes x_i + b_i \otimes y_i = 0$, with $a_i, b_i \in A \cap B$ and $x_i, y_i \in X \cap Y$. Choose operators

$$\alpha_i \in \mathrm{End}(U_2),\ \beta_i \in \mathrm{End}(U_1), \qquad f_i \in \mathrm{End}(V_2),\ g_i \in \mathrm{End}(V_1)$$

  satisfying $a_i = s \bullet_2 \alpha_i,\ b_i = s \bullet_1 \beta_i$ and $x_i = t \bullet_2 f_i,\ y_i = t \bullet_1 g_i$. Then for all i,

$$(s \otimes t) \bullet_2 \kappa_2(\alpha_i \otimes f_i) = a_i \otimes x_i = -\, b_i \otimes y_i = -\,(s \otimes t) \bullet_1 \kappa_1(\beta_i \otimes g_i).$$

  Hence $\big(-\kappa_2(\alpha_i \otimes f_i),\ \kappa_1(\beta_i \otimes g_i)\big) \in \mathcal{M}(s \otimes t)$, and therefore $\iota_{21}\big(-\kappa_2(\alpha_i \otimes f_i), \kappa_1(\beta_i \otimes g_i)\big) \in \iota_{21}(\mathcal{M}(s \otimes t))$. Applying Theorem C, $\iota_{21}(\mathcal{M}(s \otimes t)) = \iota_{21}(\mathcal{M}(s) \boxtimes \mathcal{M}(t))$, so the operator triple lies in one of the summands on the right-hand side of Equation (5.1). The I_{33} and I_{44} cases are analogous.

For each i, let $(\delta_i^{(2)}, \delta_i^{(1)}, \delta_i^{(0)}) \in \mathrm{End}(U_2 \otimes V_2) \times \mathrm{End}(U_1 \otimes V_1) \times \mathrm{End}(U_0 \otimes V_0)$ be the resulting operator triple constructed corresponding to each block. Summing over i,

$$\Xi\left((s \otimes t) \bullet_2 \sum_{i=1}^{n} \delta_i^{(2)}\right) = p = \Xi((s \otimes t) \bullet_2 \delta^{(2)}),$$
$$\Xi\left((s \otimes t) \bullet_1 \sum_{i=1}^{n} \delta_i^{(1)}\right) = q = \Xi((s \otimes t) \bullet_1 \delta^{(1)}),$$

$$\Xi\left((s\otimes t)\bullet_0 \sum_{i=1}^{n}\delta_i^{(0)}\right) = r = \Xi((s\otimes t)\bullet_0 \delta^{(0)}).$$

Since Ξ is an isomorphism and $\Lambda_2^{s\otimes t}, \Lambda_1^{s\otimes t}, \Lambda_0^{s\otimes t}$ are injective, we get

$$\sum_{i=1}^{n}\delta_i^{(2)} = \delta^{(2)}, \qquad \sum_{i=1}^{n}\delta_i^{(1)} = \delta^{(1)}, \qquad \sum_{i=1}^{n}\delta_i^{(0)} = \delta^{(0)}.$$

As each is summing over the same indices, $(\delta^{(2)}, \delta^{(1)}, \delta^{(0)}) = \sum_{i=1}^{n}(\delta_i^{(2)}, \delta_i^{(1)}, \delta_i^{(0)})$. This proves the desired inclusion.

($\supseteq$) We prove $\mathrm{Der}(s\otimes t) \supseteq \mathrm{Der}(s)\boxtimes\mathcal{C}(t) + \mathcal{C}(s)\boxtimes\mathrm{Der}(t) + \iota_{20}(\mathcal{L}(s)\boxtimes\mathcal{L}(t)) + \iota_{10}(\mathcal{R}(s)\boxtimes\mathcal{R}(t)) + \iota_{21}(\mathcal{M}(s)\boxtimes\mathcal{M}(t))$ by proving each summand on the right belongs to $\mathrm{Der}(s\otimes t)$.

For $\mathrm{Der}(s)\boxtimes\mathcal{C}(t)$, take $(\alpha,\beta,\gamma)\in\mathrm{Der}(s)$, and $(f,g,h)\in\mathcal{C}(t)$. This means $s\bullet_2\alpha + s\bullet_1\beta = s\bullet_0\gamma$ and $t\bullet_2 f = t\bullet_1 g = t\bullet_0 h$.

By definition, $(\alpha,\beta,\gamma)\boxtimes(f,g,h) = \big(\kappa_2(\alpha\otimes f), \kappa_1(\beta\otimes g), \kappa_0(\gamma\otimes h)\big)$, hence

$$\begin{aligned}
(s\otimes t)\bullet_2\kappa_2(\alpha\otimes f) + (s\otimes t)\bullet_1\kappa_1(\beta\otimes g) &= \big((s\bullet_2\alpha)\otimes(t\bullet_2 f)\big) + \big((s\bullet_1\beta)\otimes(t\bullet_1 g)\big)\\
&= \big((s\bullet_2\alpha)+(s\bullet_1\beta)\big)\otimes t\bullet_0 h\\
&= (s\bullet_0\gamma)\otimes(t\bullet_0 h)\\
&= (s\otimes t)\bullet_0\kappa_0(\gamma\otimes h).
\end{aligned}$$

Taking span, $\mathrm{Der}(s)\boxtimes\mathcal{C}(t)$ lies in $\mathrm{Der}(s\otimes t)$. Hence $\mathrm{Der}(s)\boxtimes\mathcal{C}(t)\subseteq\mathrm{Der}(s\otimes t)$ by linearity.

The argument for $\mathcal{C}(s)\boxtimes\mathrm{Der}(t)\subseteq\mathrm{Der}(s\otimes t)$ is identical, and the argument for nucleus terms is also analogous after setting one of the three terms to 0. $\square$

A full derivation decomposition example is the Appendix (Example C.2).

5.5 Non-existence for higher valence tensors

In this subsection, we demonstrate an obstruction to being able to decompose the product of trilinear maps' derivations using the techniques from this section. Let $\mathcal{U} \coloneqq \mathrm{Tri}(U_3, U_2, U_1; U_0)$, and $\mathcal{V} \coloneqq \mathrm{Tri}(V_3, V_2, V_1; V_0)$, where Tri indicates trilinear. Fix trilinear maps $t_1 \in \mathcal{U}$ and $t_2 \in \mathcal{V}$, and let $(\delta^{(3)}, \delta^{(2)}, \delta^{(1)}, \delta^{(0)}) \in \mathrm{Der}(t_1 \otimes t_2)$, where $t_1 \otimes t_2$ is the product of trilinear maps defined analogously to Definition 2.11. This gives tensors $p, q, r, s \in \mathcal{U} \otimes \mathcal{V}$ with $p + q + r = s$. We begin with the four-term definition of the LIU decomposition.

Definition 5.16. Let U and V be finite dimensional $\mathbb{F}$-vector spaces. Let $p, q, r, s \in U \otimes V$ be nonzero elements satisfying $p + q + r = s$. Let

$$A \coloneqq \mathrm{Im}_L(p), \quad B \coloneqq \mathrm{Im}_L(q), \quad C \coloneqq \mathrm{Im}_L(r), \quad D \coloneqq \mathrm{Im}_L(s),$$
$$X \coloneqq \mathrm{Im}_R(p), \quad Y \coloneqq \mathrm{Im}_R(q), \quad Z \coloneqq \mathrm{Im}_R(r), \quad W \coloneqq \mathrm{Im}_R(s),$$

where $p \in A \otimes X$, $q \in B \otimes Y$, $r \in C \otimes Z$, and $s \in D \otimes W$ by Proposition 4.5.

A (p, q, r, s) *LIU decomposition* consists of a natural number n, and vectors

$$a_i \in A,\ b_i \in B,\ c_i \in C,\ d_i \in D, \quad x_i \in X,\ y_i \in Y,\ z_i \in Z,\ w_i \in W,$$

for $i \in [n]$, satisfying the following:

- **Decomposition equality.**

$$p = \sum_{i=1}^{n} a_i \otimes x_i, \quad q = \sum_{i=1}^{n} b_i \otimes y_i, \quad r = \sum_{i=1}^{n} c_i \otimes z_i, \quad s = \sum_{i=1}^{n} d_i \otimes w_i.$$

- **Local equality.**

$$\forall i \in [n], \quad a_i \otimes x_i + b_i \otimes y_i + c_i \otimes z_i = d_i \otimes w_i.$$

We shall show when $p + q + r = s$, there may not exist a LIU decomposition of the quadruple. As a LIU decomposition corresponds to decomposing a derivation of a product of trilinear maps by the algebraic invariants of the individual pieces, its non-existence rules out an analogous decomposition theorem using this strategy.

A natural question is whether one can reduce the $p+q+r=s$ case to the three-term case by setting $\tilde{p}=p+q$, and applying Lemma 5.13 to both $\tilde{p}+r=s$ and $p+q=\tilde{p}$. However, separate LIU decompositions of $(\tilde{p},r,s)$ and $(p,q,\tilde{p})$ do not guarantee a LIU decomposition of (p,q,r,s). Applying Lemma 5.13 to $(\tilde{p},r,s)$ gives

$$\tilde{p}=\sum_i \tilde{a}_i\otimes\tilde{x}_i,\quad \tilde{a}_i\otimes\tilde{x}_i\in \mathrm{Im}_L(\tilde{p})\otimes\mathrm{Im}_R(\tilde{p}).$$

Choose the corresponding local terms for r and s, so for each i we have $\tilde{a}_i\otimes\tilde{x}_i+c_i\otimes z_i=d_i\otimes w_i$. The equation $p+q=\tilde{p}$ means $\mathrm{Im}_L(\tilde{p})\leq\mathrm{Im}_L(p)+\mathrm{Im}_L(q)$ and $\mathrm{Im}_R(\tilde{p})\leq\mathrm{Im}_R(p)+\mathrm{Im}_R(q)$.

To recover a (p,q,r,s) LIU decomposition, for each i we would have to refine this local equality, with $c_i\otimes z_i\in\mathrm{Im}_L(r)\otimes\mathrm{Im}_R(r)$ and $d_i\otimes w_i\in\mathrm{Im}_L(s)\otimes\mathrm{Im}_R(s)$, to a local equality $a_i\otimes x_i+b_i\otimes y_i+c_i\otimes z_i=d_i\otimes w_i$. Therefore the following equation must be satisfied for each i:

$$\tilde{a}_i\otimes\tilde{x}_i=a_i\otimes x_i+b_i\otimes y_i\quad (a_i\in\mathrm{Im}_L(p),\ x_i\in\mathrm{Im}_R(p),\ b_i\in\mathrm{Im}_L(q),\ y_i\in\mathrm{Im}_R(q)),$$

but this is not guaranteed. Rank-one terms in $(A+B)\otimes(X+Y)$ can contain cross blocks terms in $A\otimes Y$ and $B\otimes X$. For example, with $A=\langle e_1\rangle$, $B=\langle e_2\rangle$, $X=\langle f_1\rangle$, $Y=\langle f_2\rangle$,

$$(e_1+e_2)\otimes(f_1+f_2)=e_1\otimes f_1+e_1\otimes f_2+e_2\otimes f_1+e_2\otimes f_2,$$

so this rank-one tensor has nonzero $A\otimes Y$ and $B\otimes X$ parts and cannot be of the form $a\otimes x+b\otimes y$ with $a\in A$, $x\in X$, $b\in B$, $y\in Y$.

Lemma 5.17. Assume $s \neq 0$. Let p, q, r, s be as the above definitions. If a LIU decomposition exists, then the intersection space

$$H \coloneqq (A \otimes X + B \otimes Y + C \otimes Z) \cap (D \otimes W)$$

contains a nonzero rank-one tensor.

Proof. Each local equality has right-hand side $d_i \otimes w_i \in D \otimes W$ and left-hand side in $A \otimes X + B \otimes Y + C \otimes Z$, hence $d_i \otimes w_i \in H$ for every i. Since $s = \sum_{i=1}^{n} d_i \otimes w_i \neq 0$, at least one $d_i \otimes w_i$ is nonzero, giving a nonzero rank-one tensor in H. □

We now show that LIU decompositions may not exist.

Theorem 5.18. Let $\mathbb{F}$ be a field such that $\operatorname{char}(\mathbb{F}) \neq 2$, and let $J \in \mathbb{F}^{2\times 2}$ be a matrix with an irreducible quadratic minimal polynomial over $\mathbb{F}$. Let $U = V = \mathbb{F}^4$ be vector spaces. Then there exist nonzero elements $p, q, r, s \in U \otimes V$ satisfying $p+q+r = s$ such that the quadruple (p, q, r, s) admits no LIU decomposition.

Proof. Take $U = V = \mathbb{F}^4$ with standard bases $\{e_i\}$ and $\{f_i\}$, and subspaces $D = \operatorname{span}\{e_1, e_2\} \leq U$, $D' = \operatorname{span}\{e_3, e_4\} \leq U$, $W = \operatorname{span}\{f_1, f_2\} \leq V$, $W' = \operatorname{span}\{f_3, f_4\} \leq V$. Define isomorphisms

$$\iota_D : D \to D',\ e_1 \mapsto e_3,\ e_2 \mapsto e_4, \qquad \iota_W : W \to W',\ f_1 \mapsto f_3,\ f_2 \mapsto f_4.$$

For $J \in \mathbb{F}^{2\times 2}$, use the fixed bases (e_1, e_2) of D, (e_3, e_4) of D', (f_1, f_2) of W, and (f_3, f_4) of W' to define linear maps $J_D : D \to D'$ and $J_W : W \to W'$, where J_D acts on columns vectors of coordinates, and J_W on rows vectors of coordinates using the entries of J. Write $Jd \coloneqq J_D(d)$ and $wJ \coloneqq J_W(w)$. With this notation, define the subspaces

$$A = \{(d, \iota_D d) : d \in D\} \leq U, \qquad X = \{(w, \iota_W w) : w \in W\} \leq V,$$

$$B = \{(d, -\iota_D d) : d \in D\} \leq U, \qquad Y = \{(w, -\iota_W w) : w \in W\} \leq V,$$

$$C = \{(d, Jd) : d \in D\} \leq U, \qquad\qquad Z = \{(w, wJ) : w \in W\} \leq V.$$

As a reminder, we define the intersection space $H := (A \otimes X + B \otimes Y + C \otimes Z) \cap (D \otimes W)$. We prove H is not the trivial subspace, and nonzero elements in H are rank 2.

Let $P, Q, R, S \in \mathbb{F}^{2\times 2}$. Relative to $U = D \oplus D'$ and $V = W \oplus W'$, for $p \in A \otimes X$, $q \in B \otimes Y$, $r \in C \otimes Z$, and $s \in D \otimes W$, we have

$$p = \begin{bmatrix} P & P \\ P & P \end{bmatrix}, \qquad q = \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix}, \qquad r = \begin{bmatrix} R & RJ \\ JR & JRJ \end{bmatrix}, \qquad s = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}. \tag{5.3}$$

Hence $p + q + r = s$ is equivalent to the matrix equation

$$\begin{bmatrix} P+Q+R & P-Q+RJ \\ P-Q+JR & P+Q+JRJ \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix},$$

meaning

$$P+Q+R = S, \qquad P-Q = -RJ, \qquad P-Q = -JR, \qquad P+Q = -JRJ.$$

Using the second and third equations, $RJ - JR = 0$, so R and J commute. Using this fact on the last equation, we get $P + Q = -RJ^2$. Substitution $-RJ^2$ for $P + Q$ in the first equation, we conclude $S = R - RJ^2$. This defines P, Q, and R in terms of the fixed J and S as

$$P = -\frac{1}{2}RJ(I_2 + J), \qquad Q = \frac{1}{2}RJ(I_2 - J), \qquad R = S(I_2 - J^2)^{-1}. \tag{5.4}$$

Since $\mathrm{char}(\mathbb{F}) \neq 2$ and the minimal polynomial of J is an irreducible quadratic, it is separable, so every matrix in $\mathbb{F}^{2\times 2}$ commuting with J is a polynomial in J by [EG04, Thm. 3.1]. Hence $R \in \mathbb{F}[J]$ and since $\mathbb{F}[J]$ is a field, R is invertible whenever nonzero.

As well, the matrices J, $I_2 + J$, $I_2 - J$, and $I_2 - J^2$ are invertible as the minimal polynomial of J is an irreducible quadratic. Thus if $S \neq 0$, then also $R \neq 0$, so R, S, P, and Q are all invertible by Equation (5.4). So every nonzero element of H has rank 2.

To show H is not the trivial subspace, it is sufficient to find a solution with $s \neq 0$. Fix $S = I_2$. Then Equation (5.4) gives

$$P = -\frac{1}{2}(I_2 - J^2)^{-1}J(I_2 + J), \qquad Q = \frac{1}{2}(I_2 - J^2)^{-1}J(I_2 - J), \qquad R = (I_2 - J^2)^{-1}.$$

We now find the tensors p, q, r, s whose coefficients are the matrices P, Q, R, S in standard coordinates. For the matrix P, Equation (5.3) gives

$$\begin{aligned} p &= \begin{bmatrix} P & P \\ P & P \end{bmatrix} \\ &= P_{11}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + P_{12}\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} + P_{21}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} + P_{22}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \\ &= P_{11} a_1 \otimes x_1 + P_{12} a_1 \otimes x_2 + P_{21} a_2 \otimes x_1 + P_{22} a_2 \otimes x_2 \end{aligned}$$

where the final equality used $a_1 = (e_1, e_3)$, $a_2 = (e_2, e_4)$, $x_1 = (f_1, f_3)$, and $x_2 = (f_2, f_4)$. Similar equalities translate the coefficient matrices Q, R, S to the tensors q, r, s. Because P, Q, R, S are invertible, $\mathrm{Im}_L(p) = \mathrm{span}\{a_1, a_2\} = A$ and $\mathrm{Im}_R(p) = \mathrm{span}\{x_1, x_2\} = X$. Analogous arguments give $\mathrm{Im}_L(q) = B$, $\mathrm{Im}_R(q) = Y$, $\mathrm{Im}_L(r) = C$, $\mathrm{Im}_R(r) = Z$, $\mathrm{Im}_L(s) = D$, and $\mathrm{Im}_R(s) = W$. This gives a quadruple (p, q, r, s) whose intersection space has all nonzero elements of rank greater than 1. Applying Lemma 5.17, we conclude (p, q, r, s) admits no LIU decomposition. $\square$

Example 5.19. We apply Theorem 5.18 to get a specific example. Fix $\mathbb{F} = \mathbb{F}_5$, $J = \begin{bmatrix} 0 & 4 \\ 1 & 4 \end{bmatrix}$, and $S = I_2$. We compute P, Q, R from Equation (5.4):

$$P = \begin{bmatrix} 1 & 1 \\ 4 & 2 \end{bmatrix}, \qquad Q = \begin{bmatrix} 3 & 2 \\ 3 & 0 \end{bmatrix}, \qquad R = (I_2 - J^2)^{-1} = \begin{bmatrix} 2 & 2 \\ 3 & 4 \end{bmatrix}.$$

In the standard coordinates, we write down a quadruple (p, q, r, s) from Equation (5.3),

$$p = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 4 & 2 & 4 & 2 \\ \hline 1 & 1 & 1 & 1 \\ 4 & 2 & 4 & 2 \end{array}\right], \quad q = \left[\begin{array}{cc|cc} 3 & 2 & 2 & 3 \\ 3 & 0 & 2 & 0 \\ \hline 2 & 3 & 3 & 2 \\ 2 & 0 & 3 & 0 \end{array}\right], \quad r = \left[\begin{array}{cc|cc} 2 & 2 & 2 & 1 \\ 3 & 4 & 4 & 3 \\ \hline 2 & 1 & 1 & 2 \\ 4 & 3 & 3 & 3 \end{array}\right], \quad s = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

where for instance, $p = \sum_{i,j=1}^{2} P_{ij} a_i \otimes x_j$ where $a_i = (e_i, \iota_D e_i)$, $x_i = (f_i, \iota_W f_i)$ and so on. This example was checked by hand, as well as through a Magma computation, which is available at this gist.

Chapter 6

Conclusion

This dissertation studies algebraic invariants of tensors from two perspectives. First, for tensors from Sylvester and derivation regular families (Definitions 3.34 and 3.41), we develop new algorithms for solving simultaneous Sylvester and derivation systems at a lower arithmetic complexity than the status quo. Theorems A and B These include, as special cases, computing nuclei and derivation algebras of tensors. The main idea contained in those theorems is that solutions to the linear equations defining the algebraic invariants can be found by solving a smaller subsystem and then lifting those solutions to full solutions, instead of solving the full linear system.

Second, we study algebraic invariants of products of bilinear maps. Specifically, Theorem C proves that the nuclei of a product of bilinear maps are the $\boxtimes$-products (Definition 2.10) of the nuclei of the factors. Theorem D proves that the derivation algebra of the product of bilinear maps is specified precisely by the algebraic invariants of the factors as laid out in Equation (5.1). This result extends classical results in algebra [BO81, Corollary 4.9] when viewing algebras through their multiplication tensors. The main idea contained in those theorems is that equations on product bilinear maps can be analyzed by answering a easier decomposition problem in the tensor product of vector spaces. We also prove an impossibility theorem on extending our approach to trilinear maps in Theorem 5.18.

6.1 Future Directions

Algorithms for finding restricted index sets.

In Theorems A and B, we assume the existence of black-box algorithms for ChooseSylvester-Restriction and ChooseDerivationRestriction, which return the restricted index sets used by the solve and lift approach. A future direction is to find and analyze algorithms for computing these restricted index sets. The core challenge is being able to bound the arithmetic complexity deterministically without enumeration.

Randomized algorithms for Sylvester and Derivation equation solving.

The purpose of the Sylvester and derivation regular families in Definitions 3.34 and 3.41 is to be able to prove deterministic arithmetic complexity improvements using the solve and lift approach. Another paradigm is to analyze these algorithms allowing for randomization. For instance, Freivalds' algorithm for matrix product verification can be used in place of the deterministic verification procedure. It tests whether $AB = C$ in Monte Carlo $O(n^2)$ time [Fre79]. Analyzing the arithmetic complexity of randomized variants of our solve and lift algorithms, and of algorithms assuming an average case model of coefficients, is a related future direction.

Faster algorithms for algebraic invariants of higher valence tensors.

The solve and lift algorithms for Sylvester and derivation equations have been developed for 3-tensors. A natural question asks whether there exist analogous solve and lift algorithms for higher valence tensors. Notation and theory have been established for higher valence tensors in [BMW20].

Analyzing the equation $p + q = r$ by Kronecker modules

Kronecker modules classify the indecomposables of a pair of matrices M, N up to change of bases [Kro90]. The equation $p + q = r$ in the tensor product space $U \otimes V$ has data specified by the pair of matrices p and q. Thus the spaces given by the decomposition of Lemma 5.13 are constrained by the Kronecker module classification. More work is needed to understand what constraints are placed on the dimension of the spaces $U^{(i)}$ and $V^{(j)}$. Furthermore, classifying triples of matrices up to change of bases is considered wild [Rin10], meaning it is as hard as classifying representations over arbitrary finite-dimensional algebras. Does the contrast between classifying pairs of matrices as tame and classifying triples of matrices as wild under this equivalence relation relate to our finding that a LIU decomposition always exists for the equation $p + q = r$, but not for $p + q + r = s$?

Decomposing the product of bilinear maps.

Our work in Chapters 4 and 5 deals with decomposing the algebraic invariants of the product of bilinear maps. A next question is detecting whether a given bilinear map admits a decomposition

$r \cong s \otimes t$. There are examples in algebra that suggest uniqueness and classification challenges, for instance the algebra $\mathbb{M}_2(\mathbb{C})$ decomposes in two different ways as tensor products over $\mathbb{R}$: both $\mathbb{H} \otimes \mathbb{C}$ and $\mathbb{M}_2(\mathbb{R}) \otimes \mathbb{C}$ are isomorphic to $\mathbb{M}_2(\mathbb{C})$.

Generalized products of multilinear maps.

In [FMW20], a (P, Ω)-tensor product is defined that generalizes the usual binary tensor product. Given a sequence of vector spaces $U_1, \ldots, U_n$ and relations specified by polynomial recurrences using the ideal $P \subset \mathbb{F}[x_1, \ldots x_n]$ and operators from the set $\Omega \subset \prod_{i=1}^{n} \mathrm{End}(U_i)$, we form the vector subspace $\Xi(P, \Omega) \leq U_1 \otimes \cdots U_n$ described as the (P, Ω)-tensor product space. This method has been used to compress the search space for tensor isomorphism problems [BMW22].

Throughout this dissertation, we use the tensor product vector space in the special case $P = (x_2 - x_1)$ and $\Omega \cong \mathbb{F}$. Future work can dig into what kinds of tensors arise from (P, Ω)-tensor products, and whether there are decomposition theorems for algebraic invariants of those tensors.

On a related front, we only analyze the derivation of products of bilinear maps due to the obstruction found in the trilinear map case. Defining a generalized (P, Ω)-product of trilinear and higher valence tensors may allow their derivation and other algebraic invariants to be analyzed.

Bibliography

[Aza08] Saeid Azam. "Derivations of tensor product of algebras". In: *Communications in Algebra* 36 (2008), pp. 905–927.

[BCP97] Wieb Bosma, John Cannon, and Catherine Playoust. "The Magma algebra system. I. The user language". In: *Journal of Symbolic Computation* 24.3-4 (1997), pp. 235–265. DOI: 10.1006/jsco.1996.0125.

[BKW24] Peter A. Brooksbank, Martin D. Kassabov, and James B. Wilson. *Detecting null patterns in tensor data*. 2024. arXiv: 2408.17425 [math.NA].

[BL08] Peter A. Brooksbank and Eugene M. Luks. "Testing isomorphism of modules". In: *Journal of Algebra* 320.11 (2008), pp. 4020–4029.

[BMW17] Peter A. Brooksbank, Joshua Maglione, and James B. Wilson. "A fast isomorphism test for groups whose Lie algebra has genus 2". In: *J. Algebra* 473 (2017), pp. 545–590. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2016.12.007.

[BMW20] Peter A. Brooksbank, Joshua Maglione, and James B. Wilson. "Exact sequences of inner automorphisms of tensors". In: *Journal of Algebra* 545 (Mar. 2020), pp. 43–63. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2019.07.006.

[BMW22] Peter A. Brooksbank, Joshua Maglione, and James B. Wilson. "Tensor isomorphism by conjugacy of Lie algebras". In: *Journal of Algebra* 604 (Aug. 2022), pp. 790–807. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2022.04.014 .

[BO81] G. M. Benkart and J. M. Osborn. "Derivations and Automorphisms of Nonassociative Matrix Algebras". In: *Transactions of the American Mathematical Society* 263.2 (1981), pp. 411–430. ISSN: 00029947.

[Bre17] Matej Brešar. "Derivations of tensor products of nonassociative algebras". In: *Linear Algebra and Applications* 530 (2017), pp. 244–252.

[Bro97] Rasmus Bro. “PARAFAC. Tutorial and applications”. In: *Chemometrics and Intelligent Laboratory Systems* 38.2 (1997), pp. 149–171.

[BS72] R. H. Bartels and G. W. Stewart. “Algorithm 432 [C2]: Solution of the matrix equation AX + XB = C [F4]”. In: *Commun. ACM* 15.9 (Sept. 1972), pp. 820–826. ISSN: 0001-0782. DOI: 10.1145/361573.361582.

[BW12] Peter A. Brooksbank and James B. Wilson. “Intersecting two classical groups”. In: *J. Algebra* 353 (2012), pp. 286–297. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2011.12.004.

[BW14] Peter A. Brooksbank and James B. Wilson. “Groups acting on tensor products”. In: *Journal of Pure and Applied Algebra* 218.3 (2014), pp. 405–416. ISSN: 0022-4049. DOI: 10.1016/j.jpaa.2013.06.011.

[DLDMV00] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. “A Multilinear Singular Value Decomposition”. In: *SIAM Journal on Matrix Analysis and Applications* 21.4 (2000), pp. 1253–1278.

[EG04] Wayne Eberly and Mark Giesbrecht. “Efficient decomposition of separable algebras”. In: *Journal of Symbolic Computation* 37.1 (2004), pp. 35–81. DOI: 10.1016/S0747-7171(03)00082-6.

[FMW20] Uriya First, Joshua Maglione, and James B. Wilson. *A spectral theory for transverse tensor operators*. 2020. arXiv: 1911.02518 [math.SP].

[Fre79] Rūsiņš Freivalds. “Fast Probabilistic Algorithms”. In: *Mathematical Foundations of Computer Science 1979*. Ed. by Jiří Bečvář. Vol. 74. Lecture Notes in Computer Science. Springer Berlin, Heidelberg, 1979, pp. 57–69. DOI: 10.1007/3-540-09526-8_5.

[Gre12] W.H. Greub. *Multilinear Algebra*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2012. ISBN: 9783662007952.

[IQ19] Gábor Ivanyos and Youming Qiao. “Algorithms Based on *-Algebras, and Their Applications to Isomorphism of Polynomials with One Secret, Group Isomorphism, and Polynomial Identity Testing”. In: *SIAM Journal on Computing* 48.3 (2019), pp. 926–963.

[Irg08] Fridtjov Irgens. *Continuum Mechanics*. Springer, 2008. ISBN: 978-3-540-74297-5.

[Jac37] Nathan Jacobson. “Abstract derivation and Lie algebras”. In: *Transactions of the American Mathematical Society* 42.2 (1937), pp. 206–224. DOI: 10.1090/s0002-9947-1937-1501922-7.

[Kar+10] Alexandros Karatzoglou et al. “Multiverse Recommendation: N-dimensional Tensor Factorization for Context-Aware Collaborative Filtering”. In: *Proceedings of the Fourth ACM Conference on Recommender Systems*. RecSys ’10. 2010, pp. 79–86. DOI: 10.1145/1864708.1864727.

[KB09] Tamara G. Kolda and Brett W. Bader. “Tensor Decompositions and Applications”. In: *SIAM Review* 51.3 (2009), pp. 455–500.

[KBV09] Yehuda Koren, Robert Bell, and Chris Volinsky. “Matrix Factorization Techniques for Recommender Systems”. In: *Computer* 42.8 (2009), pp. 30–37. DOI: 10.1109/MC.2009.263.

[Kro90] Leopold Kronecker. “Algebraische Reduction der Schaaren bilinearer Formen”. In: *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* (1890), pp. 1225–1237.

[Lan12] J.M. Landsberg. *Tensors: Geometry and Applications*. Graduate studies in mathematics. American Mathematical Society, 2012. ISBN: 9780821884812.

[LL00] George F. Leger and Eugene M. Luks. “Generalized Derivations of Lie Algebras”. In: *Journal of Algebra* 228.1 (2000), pp. 165–203. ISSN: 0021-8693. DOI: 10.1006/jabr.1999.8250.

[MM10] Takanori Maehara and Kazuo Murota. "A numerical algorithm for block-diagonal decomposition of matrix *-algebras with general irreducible components". In: *Japan Journal of Industrial and Applied Mathematics* 27 (2010), pp. 263–293.

[Mya90] A. G. Myasnikov. "Definable invariants of bilinear mappings". In: *Siberian Mathematical Journal* 31.1 (1990), pp. 89–99.

[NH08] M. Novotný and J. Hřivnák. "On (α, β, γ)-derivations of Lie algebras and corresponding invariant functions". In: *Journal of Geometry and Physics* 58.2 (2008), pp. 208–217. DOI: 10.1016/j.geomphys.2007.09.005.

[Pie82] Richard S. Pierce. *Associative Algebras*. Vol. 88. Graduate Texts in Mathematics. Springer, 1982.

[Rin10] Claus Michael Ringel. *Indecomposable Representations of the Kronecker Quivers*. 2010. arXiv: 1009.5635 [math.RT].

[RS18] Elina Robeva and Anna Seigal. "Duality of graphical models and tensor networks". In: *Information and Inference: A Journal of the IMA* 8.2 (June 2018), pp. 273–288. ISSN: 2049-8772. DOI: 10.1093/imaiai/iay009.

[Sch95] Richard D. Schafer. *An Introduction to Nonassociative Algebras*. New York: Dover Publications, 1995.

[Syl84] J. J. Sylvester. "Sur l'équation en matrices $px = xq$". In: *Comptes Rendus de l'Académie des Sciences de Paris* 99 (1884), pp. 67–71, 115–116.

[The26] The Tensor Space. *OpenDleto.jl*. https://github.com/thetensor-space/OpenDleto. Accessed: 2026-04-22. 2026.

[Tuc66] L. R. Tucker. "Some mathematical notes on three-mode factor analysis". In: *Psychometrika* 31 (1966c), pp. 279–311.

[Wil09] James B. Wilson. "Decomposing p-groups via Jordan algebras". In: *Journal of Algebra* 322.8 (2009), pp. 2642–2679.

[Wil12] James B. Wilson. “Existence, algorithms, and asymptotics of direct product decompositions, I”. In: *Groups - Complexity - Cryptology* 4.1 (Jan. 2012). ISSN: 1869-6104. DOI: 10.1515/gcc-2012-0007.

[Wil15] James B. Wilson. *Skolem-Noether for nilpotent products*. 2015. arXiv: 1507.04406 [math.RT].

[Wil16] James B. Wilson. “On automorphisms of groups, rings, and algebras”. In: *Communications in Algebra* 45.4 (Oct. 2016), pp. 1452–1478. ISSN: 1532-4125. DOI: 10.1080/00927872.2016.1175617.

Appendix A

Examples for Simultaneous Sylvester and Derivation Equations

A.1 Simultaneous Sylvester Equation Example

↑

Example A.1. Let $\mathbb{F} = \mathbb{Q}$, with $a = b = r = s = 2$ and $c = 4$. Define

$$R[:,:,1] = \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}, \quad R[:,:,2] = \begin{bmatrix} 3 & 1 \\ 4 & 3 \end{bmatrix}, \quad R[:,:,3] = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}, \quad R[:,:,4] = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix},$$
$$S[:,:,1] = \begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix}, \quad S[:,:,2] = \begin{bmatrix} 3 & 2 \\ 5 & 2 \end{bmatrix}, \quad S[:,:,3] = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}, \quad S[:,:,4] = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix},$$
$$T[:,:,1] = \begin{bmatrix} 11 & 7 \\ 24 & 25 \end{bmatrix}, \quad T[:,:,2] = \begin{bmatrix} 19 & 10 \\ 44 & 23 \end{bmatrix}, \quad T[:,:,3] = \begin{bmatrix} 3 & -1 \\ 20 & 15 \end{bmatrix}, \quad T[:,:,4] = \begin{bmatrix} 20 & 7 \\ 20 & 11 \end{bmatrix}.$$

The full system has the unique solution

$$X = \begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}, \qquad Y = \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix}.$$

Now run the solve-and-lift algorithm with $I = J = \{1\}$.

Double restricted. Set $X_I = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$, and $Y_J = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^\mathsf{T}$. The restricted system is

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 3 & 4 & 3 & 2 \\ 3 & 2 & 1 & 0 \\ 0 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 11 \\ 19 \\ 3 \\ 20 \end{bmatrix}.$$

It has the unique solution $X_I = \begin{bmatrix} -1 & 1 \end{bmatrix}$, and $Y_J = \begin{bmatrix} 4 & 3 \end{bmatrix}^\mathsf{T}$. Hence $\mathcal{S}_{\text{DR}} = \{(X_I, Y_J)\}$.

Row restricted. With $\widehat{J} = \{2\}$,

$$M_R = \begin{bmatrix} 2 & 1 \\ 3 & 2 \\ 1 & 0 \\ 4 & 1 \end{bmatrix}, \qquad N_R(X_I, Y_J) = \begin{bmatrix} 5 \\ 8 \\ 2 \\ 9 \end{bmatrix}.$$

Solving $M_R\widehat{Y} = N_R(X_I, Y_J)$ gives $\widehat{Y} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $Y = \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix}$. So

$$\mathcal{S}_{\text{R}} = \left\{ \left(\begin{bmatrix} -1 & 1 \end{bmatrix}, \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix} \right) \right\}.$$

Column restricted. With $\widehat{I} = \{2\}$,

$$M_C = \begin{bmatrix} 1 & 1 \\ 3 & 4 \\ 3 & 2 \\ 0 & 1 \end{bmatrix}, \qquad N_C(X_I, Y_J) = \begin{bmatrix} 5 \\ 18 \\ 12 \\ 3 \end{bmatrix}.$$

Solving $M_C \widehat{X} = N_C(X_I, Y_J)$ gives $\widehat{X} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and $X = \begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}$. So

$$\mathcal{S}_{\mathrm{C}} = \left\{ \left(\begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix} \right) \right\}.$$

Lift to full. The pullback is $P = \{(z_R, z_C) \in \mathcal{S}_{\mathrm{R}} \times \mathcal{S}_{\mathrm{C}} : \pi_J(z_R) = \pi_I(z_C)\} = \{p_0\}$, where

$$p_0 := \big((X_I, Y), (X, Y_J)\big), \qquad X_I = X[I, :], \;\; Y_J = Y[:, J].$$

Then $\mathrm{Err}(\Psi(p_0)) = \big(XR[:, :, k] + S[:, :, k]Y - T[:, :, k]\big)_{k \in [4]} = 0_{2\times 2\times 4}$. Therefore LiftRestricted returns the unique solution (X, Y) above.

A.2 Derivation Equation

↑

Example A.2. We illustrate the solution strategy on a $2 \times 2 \times 2$ instance. Let $a = b = c = r = s = t = 2$, and $I = J = K = \{1\}$. Define the slices

$$R[:, :, 1] = S[:, :, 1] = T[:, :, 1] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad R[:, :, 2] = S[:, :, 2] = T[:, :, 2] = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

and let the unknowns be

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \qquad Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}, \qquad Z = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}.$$

We solve the derivation equations by solving the large linear system. The constraints are, for $k \in \{1, 2\}$, $X\, R[:, :, k] + S[:, :, k]\, Y = (T^Z)[:, :, k]$.

For $k = 1$, this is $X \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} Y = z_{11}T[:,:,1] + z_{21}T[:,:,2]$. This gives the relations

$$\begin{bmatrix} x_{11} + y_{11} & y_{12} \\ x_{21} & 0 \end{bmatrix} = \begin{bmatrix} z_{11} & 0 \\ 0 & z_{21} \end{bmatrix},$$

so $y_{12} = x_{21} = z_{21} = 0$, and $x_{11} + y_{11} = z_{11}$. And for $k = 2$,

$$X \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} Y = z_{12}T[:,:,1] + z_{22}T[:,:,2],$$

gives the relations

$$\begin{bmatrix} 0 & x_{12} \\ y_{21} & x_{22} + y_{22} \end{bmatrix} = \begin{bmatrix} z_{12} & 0 \\ 0 & z_{22} \end{bmatrix},$$

so $z_{12} = x_{12} = y_{21} = 0$, and $x_{22} + y_{22} = z_{22}$. Combined, the solution space is

$$X = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}, \qquad Y = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}, \qquad Z = \begin{bmatrix} \alpha + \gamma & 0 \\ 0 & \beta + \delta \end{bmatrix}, \qquad \alpha, \beta, \gamma, \delta \in \mathbb{F},$$

We now construct the solution space using the steps described in this section.

Step 1: Triple-restricted solve. With $I = J = K = \{1\}$, the unknown blocks are $X_I = \begin{bmatrix} x_{11} & x_{12} \end{bmatrix}$, $Y_J = \begin{bmatrix} y_{11} & y_{21} \end{bmatrix}^{\mathsf{T}}$, and $Z_K = \begin{bmatrix} z_{11} & z_{21} \end{bmatrix}^{\mathsf{T}}$, and the triple-restricted equation $X_I R[:,J,1] + S[I,:,1]Y_J = (T^{Z_K})[I,J,1]$ is just the $(1,1)$-entry of the $k = 1$ equation computed above. This is the equation $x_{11} + y_{11} = z_{11}$.

Step 2: Column-depth lift. Recall $\widehat{I} = [a] \setminus I = \{2\}$ and $\widehat{X} = X[\widehat{I},:]^{\mathsf{T}} = \begin{bmatrix} x_{21} & x_{22} \end{bmatrix}^{\mathsf{T}}$. Using the notation from above:

$$M_C = \begin{bmatrix} (R[:,J,1])^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad N_C(X_I, Y_J, Z_K) = \begin{bmatrix} (T^{Z_K})[\widehat{I},J,1] - S[\widehat{I},:,1]Y_J \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 0 \end{bmatrix}.$$

So the equation $M_C\widehat{X} = N_C(X_I, Y_J, Z_K)$ corresponds to the system $\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}x_{21} & x_{22}\end{bmatrix}^{\mathsf{T}} = \begin{bmatrix}0\end{bmatrix}$, concluding $x_{21} = 0$.

Step 3: Row-depth lift. Recall $\widehat{J} = [b] \setminus J = \{2\}$ and $\widehat{Y} = Y[:, \widehat{J}] = \begin{bmatrix}y_{12} & y_{22}\end{bmatrix}^{\mathsf{T}}$. The matrices to solve in the lift are

$$M_R = \begin{bmatrix}S[I, :, 1]\end{bmatrix} = \begin{bmatrix}1 & 0\end{bmatrix}, \qquad N_R(X_I, Y_J, Z_K) = \begin{bmatrix}(T^{Z_K})[I, \widehat{J}, 1] - X_I R[:, \widehat{J}, 1]\end{bmatrix} = \begin{bmatrix}0\end{bmatrix}.$$

So the equation $M_R\widehat{Y} = N_R(X_I, Y_J, Z_K)$ corresponds to the system $\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}y_{12} & y_{22}\end{bmatrix}^{\mathsf{T}} = \begin{bmatrix}0\end{bmatrix}$, concluding $y_{12} = 0$.

Step 4: Row-col lift. Recall $\widehat{K} = [c] \setminus K = \{2\}$ and $\widehat{Z} = Z[:, \widehat{K}] = \begin{bmatrix}z_{12} & z_{22}\end{bmatrix}^{\mathsf{T}}$. The row-col lift uses

$$\begin{aligned}
M_D &= \begin{bmatrix}\mathrm{vec}(T[I, J, 1]) & \mathrm{vec}(T[I, J, 2])\end{bmatrix} = \begin{bmatrix}1 & 0\end{bmatrix}, \\
N_D(X_I, Y_J, Z_K) &= \mathrm{vec}\big(X_I R[:, J, 2] + S[I, :, 2] Y_J\big) = \begin{bmatrix}0\end{bmatrix}.
\end{aligned}$$

So the equation $M_D\widehat{Z} = N_D(X_I, Y_J, Z_K)$ corresponds to the system $\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}z_{12} & z_{22}\end{bmatrix}^{\mathsf{T}} = \begin{bmatrix}0\end{bmatrix}$, concluding $z_{12} = 0$.

Step 5: Pullback and final residual solve. Using the three lifts and the pullback conditions, the compatible candidates are

$$X = \begin{bmatrix}x_{11} & x_{12} \\ 0 & x_{22}\end{bmatrix}, \qquad Y = \begin{bmatrix}y_{11} & 0 \\ y_{21} & y_{22}\end{bmatrix}, \qquad Z = \begin{bmatrix}x_{11} + y_{11} & 0 \\ z_{21} & z_{22}\end{bmatrix}.$$

As a recap, Step 1 gives $z_{11} = x_{11} + y_{11}$, Step 2 gives $x_{21} = 0$, Step 3 gives $y_{12} = 0$, and Step 4 gives $z_{12} = 0$. Now apply the final lift condition $\mathrm{Err}(X, Y, Z) = 0$, where $\mathrm{Err}(X, Y, Z)[:, :, k] =$

$X\,R[:,:,k] + S[:,:,k]Y - (T^Z)[:,:,k]$. For the candidate solution above,

$$X\,R[:,:,1] = \begin{bmatrix} x_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad S[:,:,1]Y = \begin{bmatrix} y_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad (T^Z)[:,:,1] = \begin{bmatrix} x_{11}+y_{11} & 0 \\ 0 & z_{21} \end{bmatrix},$$
$$X\,R[:,:,2] = \begin{bmatrix} 0 & x_{12} \\ 0 & x_{22} \end{bmatrix}, \quad S[:,:,2]Y = \begin{bmatrix} 0 & 0 \\ y_{21} & y_{22} \end{bmatrix}, \quad (T^Z)[:,:,2] = \begin{bmatrix} 0 & 0 \\ 0 & z_{22} \end{bmatrix},$$

Hence the $k = 1$ equation $X\,R[:,:,1] + S[:,:,1]Y = (T^Z)[:,:,1]$ forces $z_{21} = 0$, while the $k = 2$ equation $X\,R[:,:,2] + S[:,:,2]Y = (T^Z)[:,:,2]$ forces $x_{12} = 0$, $y_{21} = 0$, and $z_{22} = x_{22} + y_{22}$. Using this information, the full derivation space matches what we computed initially as

$$X = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}, \qquad Y = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}, \qquad Z = \begin{bmatrix} \alpha+\gamma & 0 \\ 0 & \beta+\delta \end{bmatrix}, \qquad \alpha,\beta,\gamma,\delta \in \mathbb{F},$$

Appendix B

Proof of Proposition 5.3

As a reminder, throughout, R and S are finite-dimensional unital non-associative algebras with multiplication maps μ_R and μ_S. We will use the following notation.

$$X \coloneqq \mathrm{Der}(\mu_R) \boxtimes \mathcal{C}(\mu_S), \qquad Y \coloneqq \mathcal{C}(\mu_R) \boxtimes \mathrm{Der}(\mu_S),$$
$$N \coloneqq \iota_{20}(\mathcal{L}(\mu_R) \boxtimes \mathcal{L}(\mu_S)) + \iota_{10}(\mathcal{R}(\mu_R) \boxtimes \mathcal{R}(\mu_S)) + \iota_{21}(\mathcal{M}(\mu_R) \boxtimes \mathcal{M}(\mu_S)).$$

The argument starts from the decomposition $\mathrm{Der}(\mu_R \otimes \mu_S) = X + Y + N$ given by Theorem D. Since algebra derivations are diagonal triples, it is enough to analyze the diagonal of $X + Y + N$. We first show that the diagonal of $X + Y + N$ is the sum of the diagonal triples of the three summands, and then identify each piece with the corresponding terms in Brešar's decomposition.

Lemma B.1. If A is a unital algebra and $\mu : A \times A \rightarrowtail A$ is its multiplication tensor, then μ is fully nondegenerate.

Proof. The equalities $\mu(a, 1) = a$ and $\mu(1, a) = a$ force both the left and right radical to be 0, and $\mathrm{span}(\mu(A, A)) = A$ since every $a \in A$ equals $\mu(a, 1)$. □

Applying Lemma B.1 to R and S, we conclude the multiplication tensors μ_R and μ_S are fully nondegenerate. The map $d \mapsto (d, d, d)$ identifies the derivation algebra of A, $\mathcal{D}_{\mathrm{alg}}(A)$, with the diagonal triples of $\mathrm{Der}(\mu_A)$. For $x \in A$, write $\lambda_x(a) \coloneqq xa$ and $\rho_x(a) \coloneqq ax$ for the left and right multiplication operators. Recall the definitions for a unital algebra A [Sch95, Ch. I, §2, p. 10]:

(left nucleus)	$\mathrm{Nuc}_\ell(A) \coloneqq \{x \in A : \forall a, b \in A,\ (xa)b = x(ab)\}$,
(middle nucleus)	$\mathrm{Nuc}_m(A) \coloneqq \{y \in A : \forall a, b \in A,\ (ay)b = a(yb)\}$,
(right nucleus)	$\mathrm{Nuc}_r(A) \coloneqq \{z \in A : \forall a, b \in A,\ a(bz) = (ab)z\}$,
(nucleus)	$\mathrm{Nuc}(A) \coloneqq \mathrm{Nuc}_\ell(A) \cap \mathrm{Nuc}_m(A) \cap \mathrm{Nuc}_r(A)$,

$$\text{(center)} \qquad Z(A) \coloneqq \{c \in \mathrm{Nuc}(A) : \forall a \in A,\ ca = ac\}.$$

We do not separately define the centroid of an algebra A, since in the unital case it identifies canonically with the center. The corresponding tensor invariants are as follows.

Proposition B.2. Let A be a unital algebra with multiplication tensor μ. Then

$$\mathcal{M}(\mu) = \{(\rho_y, \lambda_y) : y \in \mathrm{Nuc}_m(A)\}, \qquad \mathcal{L}(\mu) = \{(\lambda_x, \lambda_x) : x \in \mathrm{Nuc}_\ell(A)\},$$
$$\mathcal{R}(\mu) = \{(\rho_z, \rho_z) : z \in \mathrm{Nuc}_r(A)\}, \qquad \mathcal{C}(\mu) = \{(\lambda_c, \lambda_c, \lambda_c) : c \in Z(A)\}.$$

Proof. We prove $\mathcal{M}(\mu) = \{(\rho_y, \lambda_y) : y \in \mathrm{Nuc}_m(A)\}$.

For $\subseteq$, let $(\alpha, \beta) \in \mathcal{M}(\mu)$. Then for all a, b in A, $(\alpha(a))b = a(\beta(b))$. Setting $b = 1$ gives $\alpha(a) = a\,\beta(1)$, so $\alpha = \rho_y$ with $y = \beta(1)$. Setting $a = 1$ gives $\beta(b) = \alpha(1)\,b$, so $\beta = \lambda_y$ with $y = \alpha(1)$. Hence $y = \alpha(1) = \beta(1)$ and $(ay)b = a(yb)$, so $y \in \mathrm{Nuc}_m(A)$. Therefore $(\alpha, \beta) = (\rho_y, \lambda_y)$ is in the right-hand side.

For $\supseteq$, let $y \in \mathrm{Nuc}_m(A)$. Then for all $a, b \in A$, $(\rho_y(a))b = (ay)b = a(yb) = a(\lambda_y(b))$, so $(\rho_y, \lambda_y) \in \mathcal{M}(\mu)$.

The equations for $\mathcal{L}(\mu)$ and $\mathcal{R}(\mu)$ are proved identically. Finally,

$$\begin{aligned}\mathcal{C}(\mu) &= \{(\alpha, \beta, \gamma) : \mu \bullet_2 \alpha = \mu \bullet_1 \beta = \mu \bullet_0 \gamma\} \\ &= \{(\alpha, \beta, \gamma) : (\alpha, \beta) \in \mathcal{M}(\mu),\ (\alpha, \gamma) \in \mathcal{L}(\mu),\ (\beta, \gamma) \in \mathcal{R}(\mu)\}.\end{aligned}$$

Using the previous three parts, let $(\alpha, \beta, \gamma) \in \mathcal{C}(\mu)$. Then $(\alpha, \gamma) = (\lambda_x, \lambda_x)$, $(\beta, \gamma) = (\rho_z, \rho_z)$, and $(\alpha, \beta) = (\rho_y, \lambda_y)$ for some $x \in \mathrm{Nuc}_\ell(A)$, $z \in \mathrm{Nuc}_r(A)$, $y \in \mathrm{Nuc}_m(A)$. Hence $\lambda_x = \rho_z$, $\lambda_x = \rho_y$, and $\rho_z = \lambda_y$, so $x = y = z =: c$ and $\lambda_c = \rho_c$. Thus $c \in \mathrm{Nuc}_\ell(A) \cap \mathrm{Nuc}_m(A) \cap \mathrm{Nuc}_r(A)$ and $ca = ac$ for all a. This means $c \in Z(A)$.

Conversely, if $c \in Z(A)$ then $(\lambda_c, \lambda_c, \lambda_c) \in \mathcal{C}(\mu)$. $\square$

Definition B.3. Let V be a vector space. For a subspace $W \leq \mathrm{End}(V)^3$, define $W^{\mathrm{diag}} := \{d \in \mathrm{End}(V) : (d,d,d) \in W\} \leq \mathrm{End}(V)$.

For $u \in \mathrm{Nuc}(A)$, let $\mathrm{ad}(u) := \lambda_u - \rho_u$. A standard fact from algebra gives $\mathrm{ad}(u) \in \mathcal{D}_{\text{alg}}(A)$. The following proposition is key to identifying the $\mathrm{ad}(u)$ piece of Brešar's decomposition.

Proposition B.4. Let A be a unital algebra with multiplication tensor μ, and let $N_A := \iota_{20}(\mathcal{L}(\mu)) + \iota_{10}(\mathcal{R}(\mu)) + \iota_{21}(\mathcal{M}(\mu))$. Then $N_A^{\mathrm{diag}} = \mathrm{ad}(\mathrm{Nuc}(A))$.

Proof. For $\supseteq$, let $u \in \mathrm{Nuc}(A)$. By Proposition B.2, $(\lambda_u, \lambda_u) \in \mathcal{L}(\mu)$, $(\rho_u, \rho_u) \in \mathcal{R}(\mu)$, and $(\rho_u, \lambda_u) \in \mathcal{M}(\mu)$. Hence $\iota_{20}(\lambda_u, \lambda_u) - \iota_{10}(\rho_u, \rho_u) + \iota_{21}(\rho_u, \lambda_u) = (\mathrm{ad}(u), \mathrm{ad}(u), \mathrm{ad}(u)) \in N_A$, so $\mathrm{ad}(u) \in N_A^{\mathrm{diag}}$.

For $\subseteq$, let $d \in N_A^{\mathrm{diag}}$. Then $(d,d,d) \in N_A$, so $(d,d,d) = \iota_{20}(\alpha, \gamma) + \iota_{10}(\beta, \gamma') + \iota_{21}(\alpha', \beta')$. By Proposition B.2, there exist $a \in \mathrm{Nuc}_\ell(A)$, $b \in \mathrm{Nuc}_r(A)$, and $c \in \mathrm{Nuc}_m(A)$ with

$$(\alpha, \gamma) = (\lambda_a, \lambda_a), \quad (\beta, \gamma') = (\rho_b, \rho_b), \quad (\alpha', \beta') = (\rho_c, \lambda_c).$$

Therefore $(d,d,d) = (\lambda_a - \rho_c,\ \rho_b + \lambda_c,\ \lambda_a + \rho_b)$. Comparing components gives $\rho_b = -\rho_c$, and $\lambda_a = \lambda_c$. Since A is unital, left and right multiplication are injective, so $a = c$, and $b = -c$. Hence $c \in \mathrm{Nuc}_\ell(A) \cap \mathrm{Nuc}_m(A) \cap \mathrm{Nuc}_r(A) = \mathrm{Nuc}(A)$ and $(d,d,d) = (\mathrm{ad}(c), \mathrm{ad}(c), \mathrm{ad}(c))$, so $d \in \mathrm{ad}(\mathrm{Nuc}(A))$. $\square$

For the rest of this section, also let $N_R := \iota_{20}(\mathcal{L}(\mu_R)) + \iota_{10}(\mathcal{R}(\mu_R)) + \iota_{21}(\mathcal{M}(\mu_R))$, and $N_S := \iota_{20}(\mathcal{L}(\mu_S)) + \iota_{10}(\mathcal{R}(\mu_S)) + \iota_{21}(\mathcal{M}(\mu_S))$.

For a unital algebra A, define the *defect map*

$$\epsilon_A : \mathrm{End}(A)^3 \to A^2, \qquad \epsilon_A(\alpha, \beta, \gamma) := (\alpha(1_A), \beta(1_A)).$$

This definition extends to subspaces $U \leq \mathrm{End}(A)^3$ as $\epsilon_A(U) := \{\epsilon_A(u) : u \in U\}$. The following shows the triples in $\mathrm{Der}(\mu_A)$ with zero defect come from algebra derivations.

Lemma B.5. $(\mathrm{Der}(\mu_A) \cap \mathrm{Ker}(\epsilon_A))^{\mathrm{diag}} = \mathcal{D}_{\mathrm{alg}}(A)$.

Proof. Let $d \in (\mathrm{Der}(\mu_A) \cap \mathrm{Ker}(\epsilon_A))^{\mathrm{diag}}$. Then $(d,d,d) \in \mathrm{Der}(\mu_A)$ and $\epsilon_A(d,d,d) = 0$, so $d(1_A) = 0$. Since $(d,d,d) \in \mathrm{Der}(\mu_A)$, we have $d(x)y + xd(y) = d(xy)$ for all $x, y \in A$, hence $d \in \mathcal{D}_{\mathrm{alg}}(A)$.

Conversely, if $d \in \mathcal{D}_{\mathrm{alg}}(A)$, then $(d,d,d) \in \mathrm{Der}(\mu_A)$. Setting $y = 1_A$ in $d(xy) = d(x)y + xd(y)$ gives $d(x) = d(x) + xd(1_A)$ for all $x \in A$, so $xd(1_A) = 0$ for all x. Taking $x = 1_A$ gives $d(1_A) = 0$. Thus $(d,d,d) \in \mathrm{Ker}(\epsilon_A)$, so $d \in (\mathrm{Der}(\mu_A) \cap \mathrm{Ker}(\epsilon_A))^{\mathrm{diag}}$. $\square$

Remark B.6. For any subspace $W \leq \mathrm{Der}(\mu_A)$, $W \cap \mathrm{Ker}(\epsilon_A) = \{(d,d,d) : d \in W^{\mathrm{diag}}\}$. Indeed, if $(\alpha, \beta, \gamma) \in W \cap \mathrm{Ker}(\epsilon_A)$, then substituting 1_A in each slot of the derivation identity gives $\alpha = \gamma = \beta$.

For $A = R \otimes S$, if $(\alpha, \beta, \gamma) \in \mathrm{End}(R)^3$ and $(f, g, h) \in \mathrm{End}(S)^3$, then evaluating ϵ_A gives

$$\epsilon_A\big((\alpha, \beta, \gamma) \boxtimes (f, g, h)\big) = (\alpha(1_R) \otimes f(1_S),\ \beta(1_R) \otimes g(1_S)). \tag{B.1}$$

We shall also need the following for the proof.

Remark B.7. If $z \in Z(S)$, then $u \boxtimes (\lambda_z, \lambda_z, \lambda_z) \in N$ for every $u \in N_R$. Analogously, if $c \in Z(R)$, then $(\lambda_c, \lambda_c, \lambda_c) \boxtimes v \in N$ for every $v \in N_S$.

Proof. By Proposition B.2, $z \in Z(S)$ gives $(\lambda_z, \lambda_z) \in \mathcal{L}(\mu_S)$, $(\rho_z, \rho_z) \in \mathcal{R}(\mu_S)$, and $(\rho_z, \lambda_z) \in \mathcal{M}(\mu_S)$. Writing $u = u_{20} + u_{10} + u_{21}$ with $u_{20} \in \iota_{20}(\mathcal{L}(\mu_R))$, $u_{10} \in \iota_{10}(\mathcal{R}(\mu_R))$, and $u_{21} \in \iota_{21}(\mathcal{M}(\mu_R))$, each $u_{ij} \boxtimes (\lambda_z, \lambda_z, \lambda_z)$ lies in the matching summand of N, hence so does the sum. The second statement is analogous. $\square$

The following key lemma lets us understand the output of Theorem 3.39 in the context of algebra derivations.

Lemma B.8. With the notation above,

$$(X + Y + N)^{\mathrm{diag}} = X^{\mathrm{diag}} + Y^{\mathrm{diag}} + N^{\mathrm{diag}}.$$

Proof. The inclusion $X^{\text{diag}} + Y^{\text{diag}} + N^{\text{diag}} \subseteq (X + Y + N)^{\text{diag}}$ is immediate.

We now prove the converse. Let $d \in (X + Y + N)^{\text{diag}}$ and write $(d, d, d) = x + y + n$ with $x \in X$, $y \in Y$, and $n \in N$. Let $E_R := \epsilon_R(\mathrm{Der}(\mu_R))$, $F_R := \epsilon_R(N_R)$, $E_S := \epsilon_S(\mathrm{Der}(\mu_S))$, and $F_S := \epsilon_S(N_S)$. Thus E_R and E_S are the defects of the full derivation algebra, while F_R and F_S are the defects of the nucleus summands N_R and N_S. This gives the relations $F_R \leq E_R$ and $F_S \leq E_S$. Now view $E_R \otimes Z(S)$ as a subspace of A^2 by $(a, b) \otimes z \mapsto (a \otimes z, b \otimes z)$, and similarly for $Z(R) \otimes E_S$, $F_R \otimes S$, and $R \otimes F_S$. We prove a sequence of intermediate claims before putting it all together.

Claim 1. We prove

$$\epsilon_A(X) = E_R \otimes Z(S), \qquad \epsilon_A(Y) = Z(R) \otimes E_S. \tag{B.2}$$

We prove the direction $\epsilon_A(X) \leq E_R \otimes Z(S)$. Every element of $X = \mathrm{Der}(\mu_R) \boxtimes \mathcal{C}(\mu_S)$ is a sum of terms $\delta \boxtimes (\lambda_z, \lambda_z, \lambda_z)$ with $\delta = (\alpha, \beta, \gamma) \in \mathrm{Der}(\mu_R)$ and $z \in Z(S)$. By (B.1), $\epsilon_A\big(\delta \boxtimes (\lambda_z, \lambda_z, \lambda_z)\big) = (\alpha(1_R) \otimes z, \beta(1_R) \otimes z) \in E_R \otimes Z(S)$, so $\epsilon_A(X) \leq E_R \otimes Z(S)$.

Conversely, if $\epsilon_R(\delta) = (a, b) \in E_R$ and $z \in Z(S)$, then $(a \otimes z, b \otimes z) = \epsilon_A\big(\delta \boxtimes (\lambda_z, \lambda_z, \lambda_z)\big)$, and such pure tensors span $E_R \otimes Z(S)$. Hence $\epsilon_A(X) = E_R \otimes Z(S)$, and the formula for $\epsilon_A(Y)$ is analogous.

Claim 2. We prove

$$\epsilon_A(Y) + \epsilon_A(N) \leq F_R \otimes S, \qquad \epsilon_A(X) + \epsilon_A(N) \leq R \otimes F_S. \tag{B.3}$$

Let $Z(R)^2$ denote the subspace of R^2 whose two components lie in $Z(R)$. We show $Z(R)^2 \leq F_R$. Let $c, d \in Z(R) \subseteq \mathrm{Nuc}(R)$, then $(\lambda_c, \lambda_c) \in \mathcal{L}(\mu_R)$ and $(\rho_d, \rho_d) \in \mathcal{R}(\mu_R)$, so $(c, 0) = \epsilon_R(\iota_{20}(\lambda_c, \lambda_c))$ and $(0, d) = \epsilon_R(\iota_{10}(\rho_d, \rho_d))$, concluding $(c, d) \in F_R$. Every element of $Y = \mathcal{C}(\mu_R) \boxtimes \mathrm{Der}(\mu_S)$ is a sum of terms $(\lambda_c, \lambda_c, \lambda_c) \boxtimes \delta$ with $c \in Z(R)$ and $\delta = (f, g, h) \in \mathrm{Der}(\mu_S)$.

By (B.1),

$$\epsilon_A\big((\lambda_c,\lambda_c,\lambda_c)\boxtimes\delta\big)=(c\otimes f(1_S),c\otimes g(1_S))\in Z(R)^2\otimes S\subseteq F_R\otimes S,$$

so $\epsilon_A(Y)\leq F_R\otimes S$.

If $t=\iota_{20}(u\boxtimes v)$ with $u=(\alpha,\gamma)\in\mathcal{L}(\mu_R)$ and $v=(f,h)\in\mathcal{L}(\mu_S)$, then $\epsilon_A(t)=(\alpha(1_R)\otimes f(1_S),0)$. Since $\epsilon_R(\iota_{20}(u))=(\alpha(1_R),0)\in F_R$, this gives $\epsilon_A(t)\in F_R\otimes S$. The ι_{10} and ι_{21} cases are the same, and the $R\otimes F_S$ case is analogous.

Claim 3. We prove

$$\epsilon_A(X\cap N)=F_R\otimes Z(S),\qquad \epsilon_A(Y\cap N)=Z(R)\otimes F_S. \tag{B.4}$$

For the inclusion $\epsilon_A(X\cap N)\subseteq F_R\otimes Z(S)$, use $X\cap N\leq X$ and $X\cap N\leq N$ together with (B.2), (B.3), and Lemma 2.4:

$$\epsilon_A(X\cap N)\subseteq(E_R\otimes Z(S))\cap(F_R\otimes S)=F_R\otimes Z(S).$$

Let $u\in N_R$ and $z\in Z(S)$. Since $z\in Z(S)$, we have $\lambda_z=\rho_z$, so $(\lambda_z,\lambda_z,\lambda_z)\in\mathcal{C}(\mu_S)$. By Remark B.7, $u\boxtimes(\lambda_z,\lambda_z,\lambda_z)\in X\cap N$. Writing $\epsilon_R(u)=(a,b)$, (B.1) gives $\epsilon_A\big(u\boxtimes(\lambda_z,\lambda_z,\lambda_z)\big)=(a\otimes z,b\otimes z)\in\epsilon_A(X\cap N)$. Since such simple tensors span $F_R\otimes Z(S)$, this gives the reverse inclusion. The second equality is analogous.

These claims imply

$$\epsilon_A(X)\cap\big(\epsilon_A(Y)+\epsilon_A(N)\big)=\epsilon_A(X\cap N),\qquad \epsilon_A(Y)\cap\big(\epsilon_A(X)+\epsilon_A(N)\big)=\epsilon_A(Y\cap N). \tag{B.5}$$

Indeed, $\epsilon_A(X)\cap\big(\epsilon_A(Y)+\epsilon_A(N)\big)\subseteq(E_R\otimes Z(S))\cap(F_R\otimes S)=F_R\otimes Z(S)=\epsilon_A(X\cap N)$, and the reverse inclusion is immediate since $X\cap N\leq X$ and $X\cap N\leq N$. The second equality is analogous.

Putting it all together. Since $(d,d,d) \in \mathrm{Der}(\mu_A)$ is diagonal, the proof of Lemma B.5 gives $(d,d,d) \in \mathrm{Ker}(\epsilon_A)$, so $\epsilon_A(x) = -\epsilon_A(y) - \epsilon_A(n) \in \epsilon_A(Y) + \epsilon_A(N)$ and $\epsilon_A(y) = -\epsilon_A(x) - \epsilon_A(n) \in \epsilon_A(X) + \epsilon_A(N)$. Thus by (B.5), $\epsilon_A(x) \in \epsilon_A(X \cap N)$ and $\epsilon_A(y) \in \epsilon_A(Y \cap N)$. Choose $x_N \in X \cap N$ and $y_N \in Y \cap N$ with $\epsilon_A(x_N) = \epsilon_A(x)$ and $\epsilon_A(y_N) = \epsilon_A(y)$.

Let $x_0 := x - x_N$, $y_0 := y - y_N$, and $n_0 := n + x_N + y_N$. Then $x_0 \in X \cap \mathrm{Ker}(\epsilon_A)$, $y_0 \in Y \cap \mathrm{Ker}(\epsilon_A)$, and $n_0 \in N$ with

$$\epsilon_A(n_0) = \epsilon_A(n) + \epsilon_A(x_N) + \epsilon_A(y_N) = \epsilon_A(n) + \epsilon_A(x) + \epsilon_A(y) = \epsilon_A(x+y+n) = 0,$$

so $n_0 \in N \cap \mathrm{Ker}(\epsilon_A)$. By Remark B.6, $x_0 = (d_X, d_X, d_X)$, $y_0 = (d_Y, d_Y, d_Y)$, and $n_0 = (d_N, d_N, d_N)$ for $d_X \in X^{\mathrm{diag}}, d_Y \in Y^{\mathrm{diag}}, d_N \in N^{\mathrm{diag}}$. Hence $d = d_X + d_Y + d_N \in X^{\mathrm{diag}} + Y^{\mathrm{diag}} + N^{\mathrm{diag}}$ as needed. □

Lemma B.9. With the notation above, $X^{\mathrm{diag}} = \mathcal{D}_{\mathrm{alg}}(R) \otimes Z(S)$ and $Y^{\mathrm{diag}} = Z(R) \otimes \mathcal{D}_{\mathrm{alg}}(S)$.

Proof. We prove the formula for X^{diag}. The inclusion $\mathcal{D}_{\mathrm{alg}}(R) \otimes Z(S) \subseteq X^{\mathrm{diag}}$ is immediate, since in the unital case $c \mapsto (\lambda_c, \lambda_c, \lambda_c)$ identifies $Z(S)$ with $\mathcal{C}(\mu_S)$.

For the reverse inclusion, let $u \in X^{\mathrm{diag}}$, so $(u,u,u) \in X$. Choose a basis $z_1, \ldots, z_m$ of $Z(S)$. By Proposition B.2, the triples $(\lambda_{z_i}, \lambda_{z_i}, \lambda_{z_i})$ form a basis of $\mathcal{C}(\mu_S)$. Hence there exist $d_i = (d_{2,i}, d_{1,i}, d_{0,i}) \in \mathrm{Der}(\mu_R)$ such that

$$(u,u,u) = \Big(\sum_{i=1}^m d_{2,i} \otimes \lambda_{z_i}, \sum_{i=1}^m d_{1,i} \otimes \lambda_{z_i}, \sum_{i=1}^m d_{0,i} \otimes \lambda_{z_i} \Big).$$

Evaluating the three components on $r \otimes 1_S$ gives

$$\sum_{i=1}^m d_{2,i}(r) \otimes z_i = \sum_{i=1}^m d_{1,i}(r) \otimes z_i = \sum_{i=1}^m d_{0,i}(r) \otimes z_i$$

for all $r \in R$, so $d_{2,i} = d_{1,i} = d_{0,i}$ for each i. Thus each d_i is of the form $d_i = (\delta_i, \delta_i, \delta_i)$ with $\delta_i \in \mathcal{D}_{\mathrm{alg}}(R)$, and $u = \sum_{i=1}^m \delta_i \otimes z_i \in \mathcal{D}_{\mathrm{alg}}(R) \otimes Z(S)$. The formula for Y^{diag} is analogous. □

Proof of Proposition 5.3. By Theorem D, $\mathrm{Der}(\mu_R \otimes \mu_S) = X + Y + N$. By the proof of Lemma B.5, diagonal triples in $\mathrm{Der}(\mu_R \otimes \mu_S)$ are algebra derivations of $R \otimes S$. Using Lemmas B.8 and B.9,

$$\mathcal{D}_{\text{alg}}(R \otimes S) = (X + Y + N)^{\text{diag}} = \mathcal{D}_{\text{alg}}(R) \otimes Z(S) + Z(R) \otimes \mathcal{D}_{\text{alg}}(S) + N^{\text{diag}}.$$

By Theorem C, $N = \iota_{20}(\mathcal{L}(\mu_R \otimes \mu_S)) + \iota_{10}(\mathcal{R}(\mu_R \otimes \mu_S)) + \iota_{21}(\mathcal{M}(\mu_R \otimes \mu_S))$, so Proposition B.4 gives $N^{\text{diag}} = \mathrm{ad}(\mathrm{Nuc}(R \otimes S))$. Finally, the nucleus identity $\mathrm{Nuc}(R \otimes S) = \mathrm{Nuc}(R) \otimes \mathrm{Nuc}(S)$ [Bre17, p. 2] gives

$$N^{\text{diag}} = \mathrm{ad}(\mathrm{Nuc}(R \otimes S)) = \mathrm{ad}(\mathrm{Nuc}(R) \otimes \mathrm{Nuc}(S)).$$

Therefore $\mathcal{D}_{\text{alg}}(R \otimes S) = \mathcal{D}_{\text{alg}}(R) \otimes Z(S) + Z(R) \otimes \mathcal{D}_{\text{alg}}(S) + \mathrm{ad}(\mathrm{Nuc}(R) \otimes \mathrm{Nuc}(S))$, which is the subspace form of Equation (5.2). $\square$

Appendix C

Tensor Product Bilinear Map Decomposition Examples

C.1 Nucleus decomposition

↑

Example C.1. We decompose $(x, y) \in \mathcal{M}(s \otimes t)$ as sum of terms from $\mathcal{M}(s) \boxtimes \mathcal{M}(t)$. Let $\mathbb{F} = \mathbb{F}_5$. Let $s, t : \mathbb{F}^2 \times \mathbb{F}^2 \rightarrowtail \mathbb{F}$ be given by 2×2 matrices $S = E_{11} + E_{12} + E_{22}$ and $T = I_2$, where E_{ij} denote matrix units. Consider

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad Y = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

for which we verify $X(S \otimes T) = (S \otimes T)Y$. Let x, y be the operators corresponding to X and Y. Set $p \coloneqq (s \otimes t) \bullet_2 x$ and $q \coloneqq (s \otimes t) \bullet_1 y$, where $p = q$. Thus $(x, y) \in \mathcal{M}(s \otimes t)$. A direct calculation gives

$$p = q = (E_{11} + E_{12}) \otimes E_{11} + E_{22} \otimes E_{22}.$$

We read off local equalities:

$$p = q = u_1 \otimes v_1 + u_2 \otimes v_2, \quad u_1 = E_{11} + E_{12},\ v_1 = E_{11},\ u_2 = E_{22},\ v_2 = E_{22}.$$

Choose

$$(\alpha_1, \beta_1) = (E_{11}, E_{11} + E_{12}), \qquad (f_1, g_1) = (E_{11}, E_{11}),$$

$$(\alpha_2, \beta_2) = (E_{22}, 4E_{12} + E_{22}), \qquad (f_2, g_2) = (E_{22}, E_{22}).$$

It suffices to check that $u_i = s \bullet_2 \alpha_i = s \bullet_1 \beta_i$ and $v_i = t \bullet_2 f_i = t \bullet_1 g_i$ for each $i \in \{1, 2\}$. Therefore $x = \kappa_2(\alpha_1 \otimes f_1) + \kappa_2(\alpha_2 \otimes f_2)$, and $y = \kappa_1(\beta_1 \otimes g_1) + \kappa_1(\beta_2 \otimes g_2)$.

To summarize, we've decomposed (x, y) as the sum $(\alpha_1, \beta_1) \boxtimes (f_1, g_1) + (\alpha_2, \beta_2) \boxtimes (f_2, g_2)$ in $\mathcal{M}(s) \boxtimes \mathcal{M}(t)$.

C.2 Derivation decomposition

↑

Example C.2. We decompose $(\delta^{(2)}, \delta^{(1)}, \delta^{(0)}) \in \mathrm{Der}(s \otimes t)$ as sum of terms from $\mathrm{Der}(s) \boxtimes \mathcal{C}(t)$ and $\mathcal{M}(s) \boxtimes \mathcal{M}(t)$. Let $\mathbb{F} = \mathbb{F}_5$. Let $s, t : \mathbb{F}^2 \times \mathbb{F}^2 \rightarrowtail \mathbb{F}$ be given by 2×2 matrices $S = E_{11} + E_{21} + E_{22}$ and $T = I_2$, where E_{ij} denote matrix units. Consider $\Delta = (\Delta^{(2)}, \Delta^{(1)}, \Delta^{(0)})$.

$$\Delta^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad \Delta^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \Delta^{(0)} = \begin{bmatrix} 1 \end{bmatrix},$$

for which we verify $\Delta^{(2)}(S \otimes T) + (S \otimes T)\Delta^{(1)} = (S \otimes T)^{\Delta^{(0)}}$. Let $\delta = (\delta^{(2)}, \delta^{(1)}, \delta^{(0)}) \in \mathrm{Der}(s \otimes t)$ be the operators corresponding to Δ. Set $p := (s \otimes t) \bullet_2 \delta^{(2)}$, $q := (s \otimes t) \bullet_1 \delta^{(1)}$, and $r := (s \otimes t) \bullet_0 \delta^{(0)}$, where $p + q = r$. A direct calculation gives

$$p = (E_{11} + E_{21}) \otimes I_2 + E_{22} \otimes E_{11},$$

$$q = E_{22} \otimes I_2 - E_{22} \otimes E_{11},$$

$$r = (E_{11} + E_{21} + E_{22}) \otimes I_2 + 0 \otimes E_{11},$$

so $p+q=r$. We read off local equalities:

$$\begin{aligned}
p &= a_1 \otimes x_1 + a_2 \otimes x_2, && a_1 = E_{11}+E_{21},\ x_1 = I_2,\ a_2 = E_{22},\ x_2 = E_{11},\\
q &= b_1 \otimes y_1 + b_2 \otimes y_2, && b_1 = E_{22},\ y_1 = I_2,\ b_2 = -E_{22},\ y_2 = E_{11},\\
r &= c_1 \otimes z_1 + c_2 \otimes z_2, && c_1 = S,\ z_1 = I_2,\ c_2 = 0,\ z_2 = E_{11}.
\end{aligned}$$

They are chosen so $a_i \otimes x_i + b_i \otimes y_i = c_i \otimes z_i$ for all $i \in \{1,2\}$. Let $\alpha_D = E_{11}+E_{21},\ f_D = I_2,\ \beta_D = E_{22},\ g_D = I_2,\ \gamma_D = 1,\ h_D = 1$. Then

$$a_1 = s \bullet_2 \alpha_D,\ x_1 = t \bullet_2 f_D, \qquad b_1 = s \bullet_1 \beta_D,\ y_1 = t \bullet_1 g_D, \qquad c_1 = s \bullet_0 \gamma_D,\ z_1 = t \bullet_0 h_D,$$

with $(\alpha_D, \beta_D, \gamma_D) \in \mathrm{Der}(s)$ and $(f_D, g_D, h_D) \in \mathcal{C}(t)$.

Similarly, let $\alpha_N = 4E_{21}+E_{22},\ f_N = E_{11},\ \beta_N = E_{22},\ g_N = E_{11}$. Then $a_2 = s \bullet_2 \alpha_N,\ x_2 = t \bullet_2 f_N$, and $b_2 = -s \bullet_1 \beta_N,\ y_2 = t \bullet_1 g_N$, meaning $(\alpha_N, \beta_N) \in \mathcal{M}(s)$, $(f_N, g_N) \in \mathcal{M}(t)$. Therefore

$$(\delta^{(2)}, \delta^{(1)}, \delta^{(0)}) = (\alpha_D, \beta_D, \gamma_D) \boxtimes (f_D, g_D, h_D) - \iota_{21}\big((\alpha_N, \beta_N) \boxtimes (f_N, g_N)\big).$$

This writes $(\delta^{(2)}, \delta^{(1)}, \delta^{(0)})$ as the sum of a term from $\mathrm{Der}(s) \boxtimes \mathcal{C}(t)$ and a term from $\iota_{21}\big(\mathcal{M}(s) \boxtimes \mathcal{M}(t)\big)$.