On the Bhattacharya-Mesner rank of third order hypermatrices

Edinah K. Gnang †, Yuval Filmus ‡

June 21, 2017

Abstract

We introduce the Bhattacharya-Mesner rank of third order hypermatrices as a relaxation to the tensor rank and devise some upper bounds for the rank. We extend to third order hypermatrices the relation between linear dependence and their rank. We also derive necessary and sufficient conditions for the existence of third order hypermatrix inverse pair. Finally we use inverse pair to extend to third order hypermatrices the fundamental theorem of linear algebra.

1 Introduction

Third order hypermatrices are three dimensional generalization of matrices. Formally, a hypermatrix is a finite set of complex numbers indexed by distinct elements of some fixed Cartesian product set of the form

\[ \{0, 1, 2, \cdots, n_1\} \times \{0, 1, 2, \cdots, n_2\} \times \{0, 1, 2, \cdots, n_3\}. \]

Such a hypermatrix is said to be of size \((n_1 + 1) \times (n_2 + 1) \times (n_3 + 1)\). The hypermatrix is cubic and of side length \(n\) if \(n_1 = n_2 = n_3 = n - 1\). Hypermatrix algebras arise from attempts to extend to hypermatrices classical matrix algebra concepts and algorithms [MB94, GKZ 94, Ker08, GER11, FG17]. Our discussion focuses on the Bhattacharya-Mesner algebra first developed in [MB90, MB94] and the general Bhattacharya-Mesner algebra first proposed in [GER11]. The general Bhattacharya-Mesner product encompasses as special cases many other hypermatrix products and decompositions discussed in the literature including the usual matrix product, the Segre outer product, the contraction product, the higher order singular value decomposition, and the multilinear matrix multiplication [Lim13]. Third order hypermatrices are also motivated by systems of linear equations with coefficients from a skew field.

This article is accompanied by an extensive and actively maintained Sage [S+15] symbolic hypermatrix algebra package which implements various features of the general Bhattacharya-Mesner algebra. The package is made available at the link https://github.com/gnang/HypermatrixAlgebraPackage.

Acknowledgement: We would like to thank Andrei Gabrielov for providing guidance while preparing this manuscript. The first author was supported by the National Science Foundation grant DMS–1161629, and is grateful for the hospitality of the Institute for Advanced Study and the Department of Mathematics at Purdue University.

*Department of Applied Mathematics and Statistics, Johns Hopkins University, email: egang1@jhu.edu
†Computer Science Department, Technion - Israel Institute of Technology
2 Overview of the Bhattacharya-Mesner algebra of third order hypermatrices

The Bhattacharya-Mesner product first developed in [MB90, MB94] provides a natural generalization to the matrix product. For notational consistency, we will sometimes use the notation Prod\((A^{(1)}, A^{(2)})\) to refer to the matrix product \(A^{(1)} \cdot A^{(2)}\). The Bhattacharya-Mesner product is best introduced to the unfamiliar reader by emphasizing similarities with the matrix product. Recall that given

\[
A^{(1)} \in \mathbb{C}^{n_1 \times k} \quad \text{and} \quad A^{(2)} \in \mathbb{C}^{k \times n_2},
\]

the product Prod\((A^{(1)}, A^{(2)})\)\(\in \mathbb{C}^{n_1 \times n_2}\) is specified entry-wise by

\[
\text{Prod}\left( A^{(1)}, A^{(2)} \right) [i_1, i_2] = \sum_{0 \leq j < k} A^{(1)} [i_1, j] \cdot A^{(2)} [j, i_2].
\]

While the matrix product is a binary operation, the product of third order hypermatrices is ternary. The product of third order hypermatrices \(A^{(1)}, A^{(2)}\) and \(A^{(3)}\), denoted Prod\((A^{(1)}, A^{(2)}, A^{(3)})\), is defined if

\[
A^{(1)} \in \mathbb{C}^{n_1 \times k \times n_3}, \quad A^{(2)} \in \mathbb{C}^{n_1 \times n_2 \times k} \quad \text{and} \quad A^{(3)} \in \mathbb{C}^{k \times n_2 \times n_3}.
\]

The result Prod\((A^{(1)}, A^{(2)}, A^{(3)})\)\(\in \mathbb{C}^{n_1 \times n_2 \times n_3}\) is specified entry-wise by

\[
\text{Prod}\left( A^{(1)}, A^{(2)}, A^{(3)} \right) [i_1, i_2, i_3] = \sum_{0 \leq j < k} A^{(1)} [i_1, j, i_3] \cdot A^{(2)} [i_2, j] \cdot A^{(3)} [j, i_2, i_3].
\]

The product of third order hypermatrices is also motivated by the study of general systems of linear equations over a skew field. Linear constraints of a general system are such that variables admit both left and a right coefficient from the skew field. Such systems are expressed as a Bhattacharya-Mesner product of the form

\[
\left\{ \sum_{0 \leq j < k} A [0, j, i_3] \cdot x [0, 0, j] \cdot B [j, 0, i_3] = c [0, 0, i_3] \right\}_{0 \leq i_3 < p},
\]

where \(A\) has size \(1 \times k \times n_3\), \(x\) has size \(1 \times 1 \times k\), \(B\) has size \(k \times 1 \times n_3\) and \(c\) has size \(1 \times 1 \times k\). The entries of \(A, B\) and \(c\) are elements of the skew field. Note that constraints in [1] are not expressible as a single left or right action of some coefficient matrix on \(x\).

\[
\text{Prod}\left( A, x, B \right) = c. \quad (2)
\]

In special cases such systems can be solved using quasi-determinants [GR91, GRW05]. A triplet of hypermatrices \((A^{(1)}, A^{(2)}, A^{(3)})\) for which the product is possible is called conformable.

We recall a variant of the Bhattacharya-Mesner product called the general Bhattacharya-Mesner product. The general product is of particular interest to our discussion because it enables considerable notational simplifications. The general product is defined for any conformable triple \((A^{(1)}, A^{(2)}, A^{(3)})\) and an additional cubic hypermatrix \(B\) called the \textit{background hypermatrix} of side length \(k\) (the contracted dimension). The general Bhattacharya-Mesner product of hypermatrices \(A^{(1)}, A^{(2)}\) and \(A^{(3)}\) with background \(B\), denoted \(\text{Prod}_B\left( A^{(1)}, A^{(2)}, A^{(3)} \right)\), is defined if

\[
A^{(1)} \text{ is } n_1 \times k \times n_3, \quad A^{(2)} \text{ is } n_1 \times n_2 \times k, \quad A^{(3)} \text{ is } k \times n_2 \times n_3 \quad \text{and} \quad B \text{ is } k \times k \times k.
\]
The result Prod\(_B\) (\(A^{(1)}\), \(A^{(2)}\), \(A^{(3)}\)) is of size \(n_1 \times n_2 \times n_3\) and specified entry-wise by

\[
\text{Prod}\_B \left( A^{(1)}, A^{(2)}, A^{(3)} \right) [i_1, i_2, i_3] = \sum_{0 \leq j_1, j_2, j_3 < k} A^{(1)} [i_1, j_1, i_3] \cdot A^{(2)} [i_1, i_2, j_2] \cdot A^{(3)} [j_3, i_2, i_3] \cdot B [j_1, j_2, j_3].
\]  

The original Bhattacharya-Mesner product is recovered from the general product by taking the background hyper-matrix \(B\) to be the Kronecker delta hypermatrix denoted \(\Delta\), whose entries are specified by

\[
\Delta [i_1, i_2, i_3] = \begin{cases} 
1 & \text{if } 0 \leq i_1 = i_2 = i_3 < n \\
0 & \text{otherwise}
\end{cases}
\]

Let us recall for the reader’s convenience the definition of the hypermatrix transpose operations. Let \(A \in \mathbb{C}^{n_1 \times n_2 \times n_3}\), the corresponding transpose, denoted \(A^\top \in \mathbb{C}^{n_2 \times n_3 \times n_1}\), is specified entry-wise by

\[
A^\top [i_1, i_2, i_3] = A [i_3, i_1, i_2].
\]

The transpose operation performs a cyclic permutation of the indices. For notational convenience we adopt the convention

\[
A^\top^2 := \left( A^\top \right)^\top, \quad A^\top^3 := \left( A^\top^2 \right)^\top = A.
\]

By this convention,

\[
A^\top^i = A^\top^j \quad \text{if } i \equiv j \mod 3.
\]

It follows from the definition that

\[
\text{Prod} \left( A^{(1)}, A^{(2)}, A^{(3)} \right)^\top = \text{Prod} \left( \left( A^{(2)} \right)^\top, \left( A^{(3)} \right)^\top, \left( A^{(1)} \right)^\top \right). \tag{4}
\]

The identity (4) generalizes the matrix identity

\[
\left( A^{(1)} \cdot A^{(2)} \right)^\top = \left( A^{(2)} \right)^\top \cdot \left( A^{(1)} \right)^\top.
\]

We describe the use of the general BM product to express trilinear forms. Let \(A \in \mathbb{C}^{m \times n \times p}\), \(x \in \mathbb{C}^{m \times 1 \times 1}\), \(y \in \mathbb{C}^{n \times 1 \times 1}\) and \(z \in \mathbb{C}^{p \times 1 \times 1}\). The corresponding trilinear form is given by

\[
\text{Prod}_A \left( x^\top^2, y^\top^1, z^0 \right) = \sum_{0 \leq i < m} \sum_{0 \leq j < n} \sum_{0 \leq k < p} A [i, j, k] \cdot x^\top^2 [0, i, 0] \cdot y^\top [0, 0, j] \cdot z [k, 0, 0].
\]
3 The Bhattacharya-Mesner rank

The Bhattacharya-Mesner outer product extends to hypermatrices the matrix outer product operation. Similarly to the matrix case the outer product is a special instance of the usual product. Fortunately, the general Bhattacharya-Mesner product provides a convenient way to express outer products. For an arbitrary conformable triple \((A^{(1)}, A^{(2)}, A^{(3)})\), an outer product corresponds to a product of oriented matrix slices of the form

\[
\text{Prod} \left( A^{(1)} [:, t, :], A^{(2)} [:, :, t] , A^{(3)} [t, :, :] \right) .
\]

(5)

Recall that in the colon notation, \(A^{(1)} [:, t, :]\) refers to a hypermatrix slice of size \(n_1 \times 1 \times n_3\), where the second index is fixed to \(t\) while all other indices are allowed to vary within their prescribed ranges. Hypermatrix outer products are conveniently expressed as general BM products. The corresponding background hypermatrices are denoted \(\{\Delta(t)\}_{0 \leq t < n}\) and specified entry-wise by

\[
\Delta(t) [i_1, i_2, i_3] = \begin{cases} 
1 & \text{if } 0 \leq t = i_1 = i_2 = i_3 < n \\
0 & \text{otherwise}
\end{cases}.
\]

The outer product in (5) is more conveniently expressed as \(\text{Prod}_{\Delta(t)} \left( A^{(1)}, A^{(2)}, A^{(3)} \right)\). The Bhattacharya-Mesner outer product also suggest a natural generalization to the notion of hypermatrix rank. We show this by emphasizing the similarities with the matrix case. Recall from linear algebra that a matrix \(A\) is of rank \(r\) (over \(\mathbb{C}\)) if there exists a conformable matrix pair \(X^{(1)}, X^{(2)}\) such that

\[
A = \sum_{0 \leq t < r} \text{Prod}_{\Delta(t)} \left( X^{(1)}, X^{(2)} \right),
\]

and crucially, there exists no conformable matrix pair \(Y^{(1)}, Y^{(2)}\) for which

\[
A = \sum_{0 \leq t < r - 1} \text{Prod}_{\Delta(t)} \left( Y^{(1)}, Y^{(2)} \right),
\]

where

\[
\Delta(t) [i_1, i_2] = \begin{cases} 
1 & \text{if } 0 \leq t = i_1 = i_2 < n \\
0 & \text{otherwise}
\end{cases}.
\]

In other words, the matrix rank is the minimum number of outer product which add up to \(A\). This notion of rank extends verbatim to hypermatrices and is called the Bhattacharya-Mesner rank. Throughout our discussion the rank will refer to the Bhattacharya-Mesner rank where no confusion arises. A hypermatrix \(A\) has rank \(r\) (over \(\mathbb{C}\)) if there exists a conformable triple \((X^{(1)}, X^{(2)}, X^{(3)})\) such that

\[
A = \sum_{0 \leq t < r} \text{Prod}_{\Delta(t)} \left( X^{(1)}, X^{(2)}, X^{(3)} \right),
\]

(6)

and crucially there exists no BM conformable triple \((Y^{(1)}, Y^{(2)}, Y^{(3)})\) for which

\[
A = \sum_{0 \leq t < r - 1} \text{Prod}_{\Delta(t)} \left( Y^{(1)}, Y^{(2)}, Y^{(3)} \right),
\]
where
\[
\Delta^{(t)} [i_1, i_2, i_3] = \begin{cases} 
1 & \text{if } 0 \leq t = i_1 = i_2 = i_3 < n \\
0 & \text{otherwise}
\end{cases}
\]

In other words, the rank of \( A \) is the minimum number of outer products which add up to \( A \). Note that the usual notions of tensor/hypermatrix rank discussed in the literature [Lim13], including the canonical polyadic rank, correspond to special instances of the Bhattacharya-Mesner rank where additional constraints are imposed on the hypermatrices in the triple \((X^{(1)}, X^{(2)}, X^{(3)})\) in \[ \mathbb{C} \] as illustrated by the following proposition.

**Proposition 1**: The canonical polyadic decomposition rank of a third order hypermatrices \( A \in \mathbb{C}^{m \times n \times p} \) upper bounds the Bhattacharya-Mesner rank of \( A \).

**Proof**: Recall the definition of the vector outer product denoted \((x \otimes y \otimes z) \in \mathbb{C}^{m \times n \times p} \) of vectors \( x \in \mathbb{C}^{m \times 1 \times 1} \), \( y \in \mathbb{C}^{1 \times n \times 1} \), \( z \in \mathbb{C}^{1 \times 1 \times p} \) associated with the canonical polyadic decomposition specified entry wise by
\[
(x \otimes y \otimes z) [i, j, k] = x [i, 0, 0] \cdot y [0, j, 0] \cdot z [0, 0, k].
\]

We justify the upper bound claim by establishing that vector outer product above is a constrained Bhattacharya-Mesner outer product instance. Let \( X \in \mathbb{C}^{m \times 1 \times p}, Y \in \mathbb{C}^{m \times n \times 1} \) and \( Z \in \mathbb{C}^{1 \times n \times p} \) such that
\[
\forall 0 \leq i_1, i_2 < p, \quad X [i, 0, k_1] = x [i, 0, 0] = X [i, 0, k_2],
\]
\[
\forall 0 \leq i_1, i_2 < m, \quad Y [i_1, j, 0] = y [0, j, 0] = Y [i_2, j, 0],
\]
\[
\forall 0 \leq j_1, j_2 < n, \quad Z [0, j_1, k] = z [0, 0, k] = Z [0, j_2, k].
\]

We see that
\[
x \otimes y \otimes z = \text{Prod}_{\Delta^{(0)}} (X, Y, Z) = \text{Prod} (X, Y, Z).
\]

Alternatively for \( X \in \mathbb{C}^{m \times 1 \times p}, Y \in \mathbb{C}^{m \times n \times 1} \) and \( Z \in \mathbb{C}^{1 \times n \times p} \) such that
\[
\forall 0 \leq i_1, i_2 < m, \quad X [i_1, 0, k] = z [0, 0, k] = X [i_2, 0, k],
\]
\[
\forall 0 \leq j_1, j_2 < n, \quad Y [i, j_1, 0] = x [i, 0, 0] = Y [i, j_2, 0],
\]
\[
\forall 0 \leq k_1, k_2 < p, \quad Z [0, j, k_1] = y [0, j, 0] = Z [0, j, k_2].
\]

we also have
\[
x \otimes y \otimes z = \text{Prod}_{\Delta^{(0)}} (X, Y, Z) = \text{Prod} (X, Y, Z).
\]

The gap between the canonical polyadic rank and the Bhattacharya-Mesner rank is well illustrated by the fact that the hypermatrix
\[
\sum_{0 \leq t < r} \Delta^{(t)}
\]
has canonical polyadic rank \( r \) while its Bhattacharya-Mesner rank is 1 for all \( 0 < r < n \). So that the gap between the canonical polyadic rank and the Bhattacharya-Mesner rank can be as at least as large as \( n - 1 \) for some \( n \times n \times n \).
The following proposition emphasize the similarity of between the Bhattacharya-Mesner rank and the matrix rank.

Proposition 2: The Bhattacharya-Mesner rank of \( A \in \mathbb{C}^{m \times n \times p} \) is upper bounded by \( \min\{m, n, p\} \).

Proof: We prove the assertion by considering three different cases:

The first case is concerned with \( A \in \mathbb{C}^{m \times \ell \times p} \) where \( \ell \leq \min\{m, p\} \). Let \( I_1 \in \mathbb{C}^{m \times \ell \times \ell} \) and \( I_2 \in \mathbb{C}^{\ell \times \ell \times p} \) with entries specified by

\[
\forall 0 \leq m_1, m_2 < m, \quad I_1[m_1, n, k] = I_1[m_2, n, k] = \begin{cases} 
1 & \text{if } 0 \leq n = k \leq \ell \\
0 & \text{otherwise}
\end{cases},
\]

\[
\forall 0 \leq p_1, p_2 < p, \quad I_2[k, n, p_1] = I_2[k, n, p_2] = \begin{cases} 
1 & \text{if } 0 \leq k = n \leq \ell \\
0 & \text{otherwise}
\end{cases},
\]

consequently we have

\[
A = \text{Prod}(A, I_1, I_2) = \sum_{0 \leq t < \ell} \text{Prod}_{\Delta(t)}(A, I_1, I_2).
\]

This establishes that the rank of \( A \) must be less or equal to \( \ell \).

The second case is concerned with \( B \in \mathbb{C}^{m \times n \times \ell} \) where \( \ell \leq \min\{m, n\} \). Let \( J_1 \in \mathbb{C}^{m \times \ell \times \ell} \) and \( J_2 \in \mathbb{C}^{\ell \times n \times \ell} \) with entries specified by

\[
\forall 0 \leq m_1, m_2 < m, \quad J_1[m_1, k, p] = J_1[m_2, k, p] = \begin{cases} 
1 & \text{if } 0 \leq k = p \leq \ell \\
0 & \text{otherwise}
\end{cases},
\]

\[
\forall 0 \leq n_1, n_2 < n, \quad J_2[k, n_1, p] = J_2[k, n_2, p] = \begin{cases} 
1 & \text{if } 0 \leq k = p \leq \ell \\
0 & \text{otherwise}
\end{cases},
\]

consequently we have

\[
B = \text{Prod}(J_1, B, J_2) = \sum_{0 \leq t < \ell} \text{Prod}_{\Delta(t)}(J_1, B, J_2)
\]

which establishes that the rank of \( B \) must be less or equal to \( \ell \).

The last case is concerned with \( C \in \mathbb{C}^{\ell \times n \times p} \) where \( \ell \leq \min\{n, p\} \). Let \( K_1 \in \mathbb{C}^{\ell \times \ell \times p} \) and \( K_2 \in \mathbb{C}^{\ell \times n \times \ell} \) with entries specified by

\[
\forall 0 \leq p_1, p_2 < p, \quad K_1[m, k, p_1] = K_1[m, k, p_2] = \begin{cases} 
1 & \text{if } 0 \leq m = k \leq \ell \\
0 & \text{otherwise}
\end{cases},
\]

\[
\forall 0 \leq n_1, n_2 < n, \quad K_2[m, n_1, k] = K_2[m, n_2, k] = \begin{cases} 
1 & \text{if } 0 \leq m = k \leq \ell \\
0 & \text{otherwise}
\end{cases},
\]
we have that
\[ C = \text{Prod} (K_1, K_2, C) = \sum_{0 \leq t < \ell} \text{Prod}_{\Delta(t)} (K_1, K_2, C) \]
which establishes that the rank of $C$ must be less or equal to $\ell$.

$A \in \mathbb{C}^{m \times n \times p}$ is full rank if $\text{rank}(A) = \min\{m, n, p\}$, otherwise $A$ is rank deficient.

### 3.1 Rank inequality

We introduce here the notion of matrix diagonal dependence. Let $A \in \mathbb{C}^{m \times \ell \times p}$, $B \in \mathbb{C}^{m \times n \times \ell}$, $C \in \mathbb{C}^{\ell \times n \times p}$ and $D \in \mathbb{C}^{m \times n \times p}$ be given such that $\ell < \min\{m, n, p\}$ and $D = \text{Prod}(A, B, C)$. The set of all $m \times n \times \ell$ hypermatrices $X$ subject to $D = \text{Prod}(A, X, C) \iff D[i, j, k] = \sum_{0 \leq t < \ell} A[i, t, k] X[i, j, t] C[t, j, k]$ is determined by solving for the entries of $X$ in the general linear system

\[
\left\{ \sum_{0 \leq t < \ell} \text{diag}(A[:, t, k]) \cdot \text{Mat}(X[:, :, t]) \cdot \text{diag}(C[t, :, k]) = \text{Mat}(D[:, :, k]) \right\}_{0 \leq k < p}.
\]

The set of matrices $\{M_t\}_{0 \leq t < p} \subset \mathbb{C}^{m \times n}$ are said to be left-right diagonally dependent if there exist diagonal matrices $\{\text{diag}(x_t)\}_{0 \leq t < p} \subset \mathbb{C}^{m \times m}$ and $\{\text{diag}(y_t)\}_{0 \leq t < p} \subset \mathbb{C}^{n \times n}$ such that

\[
\sum_{0 \leq t < p} \text{diag}(x_t) \cdot M_t \cdot \text{diag}(y_t) = 0_{m \times n},
\]

where at for at least two distinct index $t$ the associated left and right coefficients pair $(\text{diag}(x_t), \text{diag}(y_t))$ are both nonzero.

**Theorem 3** : Let $H \in \mathbb{C}^{m \times n \times p}$ be of rank $\ell < p = \min\{m, n, p\}$. Then the matrices

\[
\{\text{Mat}(H[:, :, k]) \in \mathbb{C}^{m \times n}\}_{0 \leq k < p}
\]

are left-right diagonally dependent.

**Proof** : Let $H$ have rank $\ell$ and $U \in \mathbb{C}^{m \times \ell \times p}$, $V \in \mathbb{C}^{m \times n \times \ell}$, $W \in \mathbb{C}^{\ell \times n \times p}$ such that $H = \text{Prod}(U, V, W)$.

Consequently, each one of the $p$ depth slices of $H$ corresponds to a diagonal linear combination of the $\ell$ depth slices of $V$. The corresponding constraints are given by

\[
\left\{ \sum_{0 \leq t < \ell} \text{diag}(U[:, t, k]) \cdot \text{Mat}(V[:, :, t]) \cdot \text{diag}(W[t, :, k]) = \text{Mat}(H[:, :, k]) \right\}_{0 \leq k < p}.
\]
We fix the hypermatrices $\mathbf{U}$ and $\mathbf{V}$ and solve for all hypermatrices $\mathbf{X} \in \mathbb{C}^{m \times n \times \ell}$ subject to

$$\mathbf{H} = \text{Prod} (\mathbf{U}, \mathbf{X}, \mathbf{W}).$$

In this form, the constraints express a general non-commutative system of linear equations with coefficients hypermatrices $\mathbf{A}$ of size $1 \times \ell \times p$, and $\mathbf{B}$ of size $\ell \times 1 \times p$, vector of variables $\mathbf{x}$ of size $1 \times 1 \times \ell$, and the right-hand side $\mathbf{c}$ of size $1 \times 1 \times p$. The corresponding system of linear equations is summarized as

$$\text{Prod} (\mathbf{A}, \mathbf{x}, \mathbf{B}) = \mathbf{c} \iff \left\{ \sum_{0 \leq t < \ell} \mathbf{A} [0, t, k] \cdot \mathbf{x} [0, 0, t] \cdot \mathbf{B} [t, 0, k] = \mathbf{c} [0, 0, k] \right\}_{0 \leq k < p}.$$  \hspace{1cm} (7)

Note that in the reformulation the entries of $\mathbf{A}$ and $\mathbf{B}$ are themselves diagonal matrices of size $m \times m$ and $n \times n$, respectively

$$\forall 0 \leq t < \ell, \quad 0 \leq k < p, \quad \mathbf{A} [0, t, k] = \text{diag} (\mathbf{U} [:, t, k]), \quad \mathbf{B} [t, 0, k] = \text{diag} (\mathbf{W} [t :, k]).$$

The entries of $\mathbf{x}$ are unknown $m \times n$ matrices and the entries of $\mathbf{c}$ are known $m \times n$ matrices specified by

$$\forall 0 \leq k < 3, \quad \mathbf{c} [0, 0, k] = \text{Mat} (\mathbf{H} [::, k]).$$

Since $\mathbf{H}$ has rank exactly $\ell$, for every $0 \leq t < \ell$ there exists $0 \leq k < p$ such that $\mathbf{A} [0, t, k]$ and $\mathbf{B} [t, 0, k]$ are non-zero.

A variant of Gaussian elimination which avoids division allows for the determination of solutions to systems of the form

$$\left\{ \sum_{0 \leq t < \ell} \mathbf{A} [0, t, k] \cdot \mathbf{x} [0, 0, t] \cdot \mathbf{B} [t, 0, k] = \mathbf{c} [0, 0, k] \right\}_{0 \leq k < p}.$$  \hspace{1cm} (7)

The procedure is easily understood from the first round of elimination. The first sequence of row operation is given by

$$\forall 0 < k < p, \quad (-1) \mathbf{A} [0, 0, k] \mathbf{R}_k \mathbf{B} [0, 0, k] + \mathbf{A} [0, 0, 0] \mathbf{R}_k \mathbf{B} [0, 0, 0] \rightarrow \mathbf{R}_k.$$  \hspace{1cm} (8)

The very first constraint of the system is rewritten for consistency of the variables as

$$\sum_{0 \leq t < \ell} (\mathbf{I}_2 \otimes \mathbf{A} [0, 0, 0]) (\mathbf{I}_2 \otimes \mathbf{x} [0, 0, t]) (\mathbf{I}_2 \otimes \mathbf{B} [t, 0, 0]) = (\mathbf{I}_2 \otimes \mathbf{c} [0, 0, 0]),$$

the other constraints are updated according to (8) to obtain

$$\left\{ \sum_{0 \leq t < \ell} -\mathbf{A} [0, 0, k] \mathbf{A} [0, t, 0] \mathbf{x} [0, 0, t] \mathbf{B} [t, 0, 0] \mathbf{B} [0, 0, k] + \mathbf{A} [0, 0, 0] \mathbf{A} [0, t, k] \mathbf{x} [0, 0, t] \mathbf{B} [t, 0, k] \mathbf{B} [0, 0, 0] \right. \mathbf{B} [0, 0, k]$$

$$\mathbf{c} [0, 0, k] + \mathbf{B} [0, 0, 0] \mathbf{c} [0, 0, k] \mathbf{B} [0, 0, 0] \} \rightarrow \mathbf{R}_k.$$  \hspace{1cm} (8)
with $y$ and since some of them are eliminated (because $\ell < p$).

In the process of Gaussian elimination, the right-hand side $s$ are certain "linear" combinations of the coefficients, until the constraints are put in upper triangular form or Row Echelon Form.

By our hypothesis, the system $\mathbf{U}$ admits at least one solution which is associated with the assignments

$$
\forall 0 \leq t < \ell, \quad x [0,0,t] = \text{Mat} (\mathbf{V} [,;:,t]).
$$

Because there are more constraints then there are variables in $x$ (i.e. $p > \ell$), the procedure will produce at least $(p - \ell)$ trivial constraints of the form $0 = 0$ for some assignment of the free variables and thereby establishes the existence of a diagonal combination which annihilates the depth slices of $\mathbf{D}$.

**Theorem 4**: If $\mathbf{A} \in \mathbb{C}^{m \times n \times (r+1)}$, where $r + 1 \leq \min (m, n)$, has no zero entries then

$$
\text{rank} (\mathbf{A}) = (r + 1) = \min (m + n), \quad (r + 1) \leq \binom{m}{2} \cdot \binom{n}{2}.
$$

**Proof**: The argument here builds on Proposition 3 which establishes that $\text{rank} (\mathbf{A}) = r + 1$ if the depth slices of $\mathbf{A}$ are diagonally independent. We describe here a general method for investigating the diagonal dependence of depth slices. Let $\mathbf{A} \in \mathbb{C}^{m \times n \times (r+1)}$, where $r + 1 \leq \min (m, n)$. Diagonal dependence of the depth slices of $\mathbf{A}$ is expressed by the equality

$$
\text{Prod} (\mathbf{X}, \mathbf{A}, \mathbf{Y}) = \mathbf{0}_{m \times n \times 1},
$$

where $\mathbf{X} \in \mathbb{C}^{m \times (r+1) \times 1}$ and $\mathbf{Y} \in \mathbb{C}^{(r+1) \times n \times 1}$. Assuming with some loss of generality that for all $0 \leq i < m, 0 \leq j < n$ we have

$$
\mathbf{X} [i,r,0] \neq 0 \quad \text{and} \quad 0 \neq \mathbf{Y} [r,j,0].
$$
allows us to rewrite the constraints as

$$\forall \ 0 \leq i < m, 0 \leq j < n, \sum_{0 \leq t < r} \left( \frac{X[i, t, 0]}{X[i, r, 0]} A[i, j, t] \left( \frac{Y[t, j, 0]}{Y[r, j, 0]} \right) \right) = -A[i, j, r].$$

The constraints can be re-expressed in terms of $X \in \mathbb{C}^{m \times r \times 1}$ and $Y \in \mathbb{C}^{r \times n \times 1}$ as

$$\text{Prod}(X, A[:, :, r-1], Y) = -A[:, :, r] \iff \left\{ \sum_{0 \leq t < r} X[i, t, 0] A[i, j, t] Y[t, j, 0] = -A[i, j, r] \right\}_{0 \leq i < m, 0 \leq j < n}.$$

We further cast these constraints as a system of linear equations of the form

$$\forall \ 0 \leq i < m, 0 \leq j < n, \sum_{0 \leq u+i+v<m-n} M[n \cdot i + j, n \cdot u + v] (X[u, 0, 0] Y[0, v, 0]) =$$

$$- \left( A[i, j, r] + \sum_{0 \leq t < r} X[i, t, 0] A[i, j, t] Y[t, j, 0] \right).$$

Cramer’s rule yields polynomial expressions for the monomials of $X \cdot Y[0, v, 0]$ in terms of the remaining variables. Note that the monomials of $X \cdot Y[0, v, 0]$ correspond to entries of an $m \times n$ rank one matrix. The problem is thus reduced to a system of $\binom{m}{2} \cdot \binom{n}{2}$ constraints in the remaining $(m+n)\cdot(r-1)$ variable entries of $X$ and $Y$. These constraints correspond to expressions deduced from equalities of the form

$$\forall \ 0 \leq i_0 < i_1 < m, 0 \leq j_0 < j_1 < n, \det \begin{pmatrix} X[i_0, 0, 0] & Y[0, j_1, 0] \\ X[i_1, 0, 0] & Y[0, j_0, 0] \end{pmatrix} = 0. \quad (10)$$

If these constraints have some non-zero solution, then we can find an assignment for $X[i, 0, 0], Y[0, j, 0]$ such that $X[i, 0, 0] Y[0, j, 0] = X[i, 0, 0] Y[0, j, 0]$. Indeed, suppose that $X[i_0, 0, 0] Y[0, j_0, 0] \neq 0$, and define $X[i, 0, 0] = \frac{X[i, 0, 0] Y[0, j_0, 0]}{X[i_0, 0, 0] Y[0, j_0, 0]}$ and $Y[0, j, 0] = X[i_0, 0, 0] Y[0, j, 0]$. Then

$$X[i, 0, 0] Y[0, j, 0] = \frac{X[i_0, 0, 0] Y[0, j, 0] X[i, 0, 0] Y[0, j_0, 0]}{X[i_0, 0, 0] Y[0, j_0, 0]} = X[i, 0, 0] Y[0, j, 0].$$

Consequently $A$ stands a chance to have rank $r+1$ if the number of constraints is at least the number of remaining variables. For if this was not the case elimination via the method of resultants would lead to the existence over $\mathbb{C}$ of a non-trivial solution to the diagonal dependence problem. The condition is thus expressed by the inequality

$$(m+n) \cdot (r-1) \leq \binom{m}{2} \cdot \binom{n}{2}.$$ 

In particular, it is easy to see that for a non-zero $A \in \mathbb{C}^{2 \times 2 \times 2}$, where $r = 1$, we see that the inequality is trivially satisfied. Furthermore $A$ has rank 1 if

$$A[0, 0, 0] A[0, 1, 0] A[1, 0, 1] A[1, 1, 0] - A[0, 0, 1] A[0, 1, 0] A[1, 0, 0] A[1, 1, 1] = 0,$$
otherwise \( A \) is full rank. This follows from solving for \( X \in \mathbb{C}^{m \times 1 \times p} \), \( Y \in \mathbb{C}^{m \times n \times 1} \) and \( Z \in \mathbb{C}^{1 \times n \times p} \) in the rank one factorization equation

\[
\text{Prod}(X, Y, Z) = A.
\]

Fortunately, such constraints are multiplicative linear constraints as discussed in [FG17], and the statement above follows from the fact that the rank of the adjacency matrix of the 3-dimensional cube is 7. By contrast, the situation is very different for \( 3 \times 3 \times 3 \) hypermatrices, as shown in Proposition 8.

### 3.2 Towards hypermatrix rank revealing factorization

In order to emphasize the analogy with the matrix case, we start by recalling the key idea underlying the Gaussian elimination procedure for obtaining rank revealing factorizations.

**Theorem 6**: Let \( X \in \mathbb{C}^{m \times \ell} \) and \( Y \in \mathbb{C}^{\ell \times n} \) be matrices such that for some scalars \( u_t \) there exist an index \( \tau \) such that,

\[
\text{Prod}_{\Delta(\tau)}(X, Y) = \text{Prod}\left(X[:, \tau], \sum_{0 \leq t \neq \tau < \ell} u_t Y[t, :)\right),
\]

then \( \text{Prod}(X, Y) \) has rank at most \( \ell - 1 \).

**Proof**: We describe here the proof in the case \( \tau = \ell - 1 \). The argument is exactly the same for any other choice of \( 0 \leq \tau < \ell \). The product of \( X \in \mathbb{C}^{m \times \ell} \) and \( Y \in \mathbb{C}^{\ell \times n} \) is a sum of \( \ell \) outer products. In particular, the outer product of the last column of \( X \) with the last row of \( Y \) is given by

\[
\text{Prod}_{\Delta(\ell-1)}(X, Y) = \text{Prod}(X[:, \ell - 1], Y[\ell - 1, :]).
\]

By our assumption the row space of \( Y \) has at most dimension \( \ell - 1 \), and more explicitly we have

\[
\text{Prod}_{\Delta(\ell-1)}(X, Y) = \text{Prod}\left(X[:, \ell - 1], \sum_{0 \leq t < \ell - 1} u_t Y[t, :)\right).
\]

Given the assumption, we express \( \text{Prod}(X, Y) \) as a sum of fewer outer products as follows:

\[
\text{Prod}_{\Delta(\ell-1)}(X, Y) = \sum_{0 \leq t < \ell - 1} \text{Prod}(u_t X[:, \ell - 1], Y[t, :]),
\]

\[
\Rightarrow \text{Prod}(X, Y) = \left( \sum_{0 \leq t < \ell - 1} \text{Prod}_{\Delta(t)}(X, Y) \right) + \text{Prod}\left(X[:, \ell - 1], \sum_{0 \leq t < \ell - 1} u_t Y[t, :)\right),
\]

and hence

\[
\text{Prod}(X, Y) = \sum_{0 \leq t < \ell - 1} \text{Prod}\left((X[:, t] + u_t X[:, \ell - 1]), Y[t, :)\right). \quad (11)
\]
expresses the elementary column operations

\[ X[:, t] + u_t X[:, \ell - 1] \rightarrow X[:, t], \quad \forall 0 \leq t < \ell - 1. \]

We now extend the argument above to third order hypermatrices.

**Theorem 7:** Let \( X \in \mathbb{C}^{m \times \ell \times p} \), \( Y \in \mathbb{C}^{m \times n \times \ell} \) and \( Z \in \mathbb{C}^{\ell \times n \times p} \) such that for some index \( \tau \) the following holds:

\[ \forall 0 \leq k < \ell, \quad \text{Prod}_{\Delta_{(\tau)}} (X, Y, Z)[:, :, k] = \sum_{0 \leq t < \tau < \ell} [\text{diag} (u_t) \cdot \text{diag} (X[:, \tau, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[\tau, :, k]) \cdot \text{diag} (v_t) + \text{diag} (u_t) \cdot \text{diag} (X[:, \tau, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[t, :, k]) + \text{diag} (X[:, t, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[\ell - 1, :, k]) \cdot \text{diag} (v_t)]. \]

then \( \text{Prod} (X, Y, Z) \) has rank at most \( \ell - 1 \).

**Proof:** We describe here the proof in the case \( \tau = \ell - 1 \). The argument is exactly the same for any choice \( 0 \leq \tau < \ell \). The product of \( X \in \mathbb{C}^{m \times \ell \times p} \), \( Y \in \mathbb{C}^{m \times n \times \ell} \) and \( Z \in \mathbb{C}^{\ell \times n \times p} \) is a sum of \( \ell \) outer products. In particular, consider the outer product of the last column slice of \( X \) with the last depth slice of \( Y \) and the last row slice of \( Z \) given by

\[ \text{Prod}_{\Delta_{(\ell - 1)}} (X, Y, Z) = \text{Prod} (X[:, \ell - 1, :], Y[:, :, \ell - 1], Z[\ell - 1, :, :]). \]

By analogy to the matrix case, an assumption is made on the last outer product in the sum

\[ \text{Prod} (X[:, \ell - 1, :], Y[:, :, \ell - 1], Z[\ell - 1, :, :]). \]

The explicit assumption is that

\[ \forall 0 \leq k < \ell, \quad \text{Prod}_{\Delta_{(\ell - 1)}} (X, Y, Z)[:, :, k] = \sum_{0 \leq t < \ell - 1} [\text{diag} (u_t) \cdot \text{diag} (X[:, \ell - 1, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[\ell - 1, :, k]) \cdot \text{diag} (v_t) + \text{diag} (u_t) \cdot \text{diag} (X[:, \ell - 1, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[t, :, k]) + \text{diag} (X[:, t, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[\ell - 1, :, k]) \cdot \text{diag} (v_t)]. \]

Given our assumption we express \( \text{Prod}(X, Y, Z) \) as a sum of fewer outer products as follows:

\[ \forall 0 \leq k < p, \quad \text{Prod} (X, Y, Z)[:, :, k] = \sum_{0 \leq t < \ell - 1} \text{diag} (X[:, t, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[t, :, k]) + \sum_{0 \leq t < \ell - 1} [\text{diag} (u_t) \cdot \text{diag} (X[:, \ell - 1, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[\ell - 1, :, k]) \cdot \text{diag} (v_t) + \text{diag} (u_t) \cdot \text{diag} (X[:, \ell - 1, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[t, :, k]) + \text{diag} (X[:, t, k]) \cdot Y[:, :, t] \cdot \text{diag} (Z[\ell - 1, :, k]) \cdot \text{diag} (v_t)]. \]
Let us start by addressing the rank 1 case. The Bhattacharya-Mesner rank of a generic hypermatrix $A \in \mathbb{C}^{n \times n \times n}$ is expressed as a sum of the smallest possible number of outer products. Hence

\[
\forall 0 \leq k < p, \quad \text{Prod} (X, Y, Z)[:, :, k] = \sum_{0 \leq t < \ell - 1} [(\text{diag}(u_t) \text{diag}(X[:, \ell - 1, k]) + \text{diag}(X[:, t, k])) Y[:, :, t] (\text{diag}(Z[t :, k]) + \text{diag}(Z[\ell - 1 :, k]) \text{diag}(v_t))].
\]

We obtain rank revealing factorizations of $\text{Prod}(X, Y, Z)$ by successively solving for diagonal matrices $\{\text{diag}(u_t), \text{diag}(v_t)\}_{0 \leq t < \ell - 1}$ which reduce each time by one the number of outer product summands. The corresponding the elementary slice operation is prescribed by

\[
\begin{cases}
(\text{diag}(u_t) X[:, \ell - 1, k]) + X[:, t, k] & \rightarrow X[:, t, k] \\
(Z[t :, k] + Z[\ell - 1 :, k]) \text{diag}(v_t) & \rightarrow Z[t :, k]
\end{cases} \quad \forall 0 \leq t < \ell - 1.
\]

We propose to solve such constraints using the method of resultants. The procedure is repeated until $\text{Prod}(X, Y, Z)$ is expressed as a sum of the smallest possible number of outer products.

**Proposition 8**: The Bhattacharya-Mesner rank of a generic $A \in \mathbb{C}^{n \times n \times n}$ is at most 2 if $n = 2$ and at most $n - 1$ if $n \in \{3, 4, 5\}$.

**Proof**: Let us start by addressing the rank 1 case. A hypermatrix $A \in \{\mathbb{C} \setminus \{0\}\}^{m \times n \times p}$ has rank of 1 if a solution exists to the system

\[
X[i, 0, k] Y[i, j, 0] Z[0, j, k] = A[i, j, k],
\]

where the unknowns $X$, $Y$, $Z$ are respectively of size $m \times 1 \times p$, $m \times n \times 1$ and $1 \times n \times p$. Fortunately such constraints are easily solvable via a slight variant of Gaussian elimination as discussed in [FG17]. For instance one derives that when $m = n = p = 2$, $A$ has rank 1 if

\[
(A[0, 0, 0] A[0, 1, 1] A[1, 0, 1] A[1, 1, 0] - A[0, 0, 1] A[0, 1, 0] A[1, 0, 0] A[1, 1, 1]) = 0.
\]

For $n > 2$, the hypermatrix $A \in \mathbb{C}^{n \times n \times n}$ has rank $n - 1$ if there exists a solution to the system

\[
\text{Prod} (U, V, W) = A
\]

in the unknowns $U \in \mathbb{C}^{n \times (n-1) \times n}$, $V \in \mathbb{C}^{n \times n \times (n-1)}$ and $W \in \mathbb{C}^{(n-1) \times n \times n}$. The constraints can be written in terms of the depth slices as follows

\[
\left\{ \sum_{0 \leq j < n-1} \text{diag}(U[:, j, i]) \cdot \text{Mat}(V[:, :, j]) \cdot \text{diag}(W[j :, i]) = \text{Mat}(A[:, :, i]) \right\}_{0 \leq i < n}
\]

\footnote{A generic hypermatrix is one whose entries do not satisfy any non-trivial algebraic relation. In particular, all of its entries are non-zero.}
which in turn we rewrite as
\[
S = \left\{ \sum_{0 \leq j < n-1} (\text{vec}_{n \times 1} (U[:,j,i]) \cdot \text{vec}_{1 \times n} (W[j,:,i]) \circ \text{Mat} (V[:,,:,j])) = \text{Mat} (A[:,,:,i]) \right\}_{0 \leq i < n}
\]

Note that the subsystem
\[
\left\{ \sum_{0 \leq j < n-1} (\text{vec}_{n \times 1} (U[:,j,i]) \cdot \text{vec}_{1 \times n} (W[j,:,i]) \circ \text{Mat} (V[:,,:,j])) = \text{Mat} (A[:,,:,i]) \right\}_{0 \leq i < n-1}
\]

expresses a system of \(n - 1\) equations in the \(n - 1\) unknowns \(\{\text{Mat} (V[:,,:,j])\}_{0 \leq k < n-1}\) in a commutative module over \(n \times n\) matrices, where the product operation in the module corresponds to the element-wise product, also called the Hadamard product, and symbolized by the \(\circ\) operator. Using Gaussian elimination without division, we eliminate from the system \(S\) all entries of \(V\). The existence of a solution to the resulting system of \(n^2\) polynomial equations in the variable entries of \(U \in \mathbb{C}^{n \times (n-1) \times n}\), and \(W \in \mathbb{C}^{(n-1) \times n \times n}\) is determined via computation of resultants \([Stu02]\) for \(n \in \{3, 4, 5\}\).

**Conjecture**: The Bhattacharya-Mesner rank of a generic
\(A \in \mathbb{C}^{n \times n \times n}\) is at most \(n - 1\) for \(n > 2\).

**Heuristic method**: Given a generic \(A \in \{\mathbb{C}\\backslash\{0\}\}^{n \times n \times n}\) \(n > 2\) determine the diagonal dependence between the depth slices which enables a reduction of the number of outer products in the canonical factorization from \(n\) to \(n-1\).

**Method**:
Considering the canonical factorization of \(A \in \mathbb{C}^{n \times n \times n}\) expressed in terms of the identity pair \(J_1, J_2 \in \mathbb{C}^{n \times n \times n}\) such that
\[
A = \text{Prod} (J_1, A, J_2),
\]
where
\[
\forall 0 \leq i < n, \quad J_1[i,t,p] = \begin{cases} 1 & \text{if } 0 \leq t = p \leq n \\ 0 & \text{otherwise} \end{cases},
\]
\[
\forall 0 \leq j < n, \quad J_2[t,j,p] = \begin{cases} 1 & \text{if } 0 \leq t = p \leq n \\ 0 & \text{otherwise} \end{cases}.
\]

By construction,
\[
\text{Prod}_{\Delta(k)} (J_1, A, J_2)[:,;:,j] = \begin{cases} A[:,;:k] & \text{if } j = k \\ 0_{n \times n} & \text{otherwise} \end{cases}.
\]

\(^2\)A generic hypermatrix is one whose entries do not satisfy any non-trivial algebraic relation. In particular, all of its entries are non-zero.
By Theorem 7 the number of outer products can be reduced by one in the canonical factorization if there exist \( X, Y \in \mathbb{C}^{n \times n} \) as well as an index \( \tau \) such that

\[
\forall 0 \leq k < n, \quad \text{Prod}_{\Delta^{(\tau)}} (J_1, A, J_2) [:, :, k] = \\
\sum_{0 \leq t \neq \tau < n} [\text{diag} (X [:, t]) \text{diag} (J_1 [\cdot, \cdot, \cdot, \tau, t]) A [\cdot, :, :, \cdot, \cdot] \text{diag} (Y [\cdot, \cdot, \cdot, \cdot, t])] + \\
\text{diag} (X [:, t]) \text{diag} (J_1 [\cdot, \cdot, \cdot, \tau, t]) A [\cdot, :, :, \cdot, \cdot] \text{diag} (J_2 [\cdot, \cdot, \cdot, \cdot, t]) A [\cdot, :, :, \cdot, \cdot] \text{diag} (Y [\cdot, \cdot, \cdot, \cdot, t])] \\
\Rightarrow \left( \sum_{0 \leq t \neq \tau < n} \text{diag} (X [:, t]) A [\cdot, :, :, \cdot, \cdot] \text{diag} (Y [\cdot, \cdot, \cdot, \cdot, t]) \right) = A [\cdot, :, :, \cdot, \cdot].
\]

It therefore suffices to establish the existence of a pair of non-zero matrices \( X, Y \in \mathbb{C}^{n \times n} \) subject to

\[
\sum_{0 \leq t \leq n} \text{diag} (X [:, t]) A [\cdot, :, :, \cdot, \cdot] \text{diag} (Y [\cdot, \cdot, \cdot, \cdot, t]) = 0_{n \times n},
\]

and there exist some index \( 0 \leq \tau < n \) for which \( \det \{ \text{diag} (X [\cdot, \cdot, \cdot, \cdot, \tau]) \text{diag} (Y [\cdot, \cdot, \cdot, \cdot, \tau]) \} \neq 0 \). These conditions express necessary and sufficient condition for the depth slice \( A [\cdot, :, :, \cdot, \cdot] \) to be expressible as a left-right diagonal linear combination of the remaining slices \( \{ A [\cdot, :, :, \cdot, \cdot] \}_{0 \leq t \neq \tau < n} \).

More explicitly the diagonal dependence is expressed as follows

\[
S := \left\{ 0 = \sum_{0 \leq j < n} (X [i, j] A [i, k, j]) Y [j, k] \right\}_{0 \leq i, j < n}.
\]

\( S \) can be partitioned into \( n \) separate systems such that

\[
S = \bigcup_{0 \leq k < n} S_k,
\]

where

\[
\forall 0 \leq k < n, \quad S_k = \left\{ 0 = \sum_{0 \leq j < n} (X [i, j] A [i, k, j]) Y [j, k] \right\}_{0 \leq i < n}.
\]

The \( n \times n \) coefficient matrix \( M_k \) associated with each system \( S_k \) is given by

\[
M_k := X [\cdot, :] \circ A [\cdot, k, :] \iff M_k [i, j] = X [i, j] A [i, k, j].
\]

The \( n^2 \times n^2 \) coefficient matrix \( M \) associated with \( S \) is given by

\[
M = \bigoplus_{0 \leq k < n} M_k.
\]

\( S \) admits a solution matrix \( Y \) having no zero columns if

\[
\forall 0 \leq k < n, \quad 0 = \det (M_k).
\]

15
The system \( \{ \det (M_k) = 0 \}_{0 \leq k < n} \) is linear in the entries of any arbitrary column \( \tau \) of \( X \). More explicitly, using the cofactor expansion we write

\[
0 \leq k < n
\]

\[
0 = \sum_{0 \leq i < n} A [i, k, \tau] \text{Cofactor}_{i, \tau} (M_k) X [i, \tau].
\]

Note that in the case \( n = 2 \) we have that for generic \( A \),

\[
(A [0, 0, 0] A [0, 1, 1] A [1, 0, 1] A [1, 1, 0] - A [0, 0, 1] A [0, 1, 0] A [1, 0, 0] A [1, 1, 1]) \neq 0
\]

However the condition \( \{ 0 = \det (M_k) \}_{0 \leq k < 2} \) is explicitly given by

\[
\begin{pmatrix}
 a_{001} a_{100} & -a_{000} a_{101} \\
 a_{011} a_{110} & -a_{010} a_{111}
\end{pmatrix}
\begin{pmatrix}
 x_{01} x_{10} \\
 x_{00} x_{11}
\end{pmatrix}
= \begin{pmatrix}
 0 \\
 0
\end{pmatrix}
\]

for which the existence of non trivial solution leads to the assertion that

\[
\Rightarrow (A [0, 0, 0] A [011] A [1, 1, 0] A [1, 1, 0] - A [0, 0, 1] A [0, 1, 0] A [1, 0, 0] A [1, 1, 1]).
\]

Therefore ruling out the fact that \( 2 \times 2 \times 2 \) hypermatrices have rank 1 in the case \( n > 2 \). The system \( \{ \det (M_k) = 0 \}_{0 \leq k < n} \) is linear in the entries of any particular column \( \tau \) of \( X \). More explicitly, using the cofactor expansion of the determinant we write

\[
0 = \sum_{0 \leq i < n} A [i, k, \tau] \text{Cofactor}_{i, \tau} (M_k) X [i, \tau].
\]

Let \( F_\tau \) denote the associated \( n \times n \) coefficient matrix given by

\[
\forall 0 \leq k, i < n, \quad F_\tau [k, i] = A [i, k, \tau] \text{Cofactor}_{i, \tau} (M_k).
\]

The multivariate polynomial \( \det (F_\tau) \) in the remaining \( n (n - 1) \) entries of \( \{ X [:, t] \}_{0 \leq t \neq \tau < n} \) has degree \( n - 1 \) in each of the variables. The vanishing of \( \det (F_\tau) \) is necessary and sufficient to ensure that \( S \) admits a solution matrix \( X \) having a non-trivial column \( X [:, \tau] \). When \( n < 6 \) the roots of \( \det (F_\tau) = 0 \) in one of the remaining variable can be expressed symbolically in terms of radical expressions of the remaining \( n (n - 1) - 1 \) entries of \( \{ X [:, t] \}_{0 \leq t \neq \tau < n} \).

When \( n \notin \{ 3, 4, 5 \} \), symbolic expressions for the roots may be found via A-hypergeometric series [Stu00].

By substituting into the entries of \( F_\tau \) the symbolic solutions to \( \det (F_\tau) = 0 \) the matrix \( F_\tau \) will have rank \( r \) with \( 0 \leq r < n \). For generic \( A \), \( r > 0 \). For notational convenience assume that the \( r \times r \) sub-matrix of full rank is \( F_\tau [:, r \tau] \). Consequently the column \( X [:, \tau] \) can be expressed via Cramer’s rule as follows

\[
X [:, r \tau] = \frac{\text{adj} (F_\tau [:, r \tau])}{\det (F_\tau [:, r \tau])} \sum_{r < j < n} (-X [j, \tau]) F_\tau [:, r \tau].
\]

If it turns out that if \( F_\tau [:, r \tau] \) is not the full rank sub-matrix of \( F_\tau \) then a similar expression can be derived in terms of the full rank \( r \times r \) sub-matrix of \( F_\tau \). Note that in the expression above \( \{ X [j, \tau] \}_{r < j < n} \) are free variables with the exception of the variable which expresses symbolically the root. Having obtained the desired symbolic parametrization of \( X \) we substitute back into \( M \) to obtain the parametrization for \( Y \). By Theorem 7 the diagonal dependence allows for the reduction from \( n \) to \( n - 1 \) the number of outer-product in the canonical factorization.
3.3 Necessary and sufficient conditions for inverse pairs

Inverse pairs, introduced in [MB90], [MB94], extend to hypermatrices the notion of matrix inverse. Recall from linear algebra that given \( A, B \in \mathbb{C}^{n \times n} \) are inverse to one another if

\[
\forall X \in \mathbb{C}^{n \times m}, \quad \text{Prod}(B, \text{Prod}(A, X)) = X.
\]

It follows by associativity that

\[
\forall X \in \mathbb{C}^{n \times m}, \quad \text{Prod}(B, \text{Prod}(A, X)) = X \Rightarrow \text{Prod}(A, \text{Prod}(B, X)) = X.
\]

Similarly, given a hypermatrix pair \( A \in \mathbb{C}^{m \times p \times p}, B \in \mathbb{C}^{p \times n \times p} \), the corresponding outer-inverse pair \( C \in \mathbb{C}^{m \times p \times p}, D \in \mathbb{C}^{p \times n \times p} \) are defined by

\[
\forall X \in \mathbb{C}^{m \times n \times p}, \quad \text{Prod}(C, \text{Prod}(A, X, B), D) = X.
\]

Note that the hypermatrix pair \( A, B \) are the inner-inverse pair of the hypermatrix pair \( C, D \). Entry-wise the inverse pair relation yields

\[
X[i, j, k] = \sum_{0 \leq s, t < p} C[i, t, k] \text{Prod}(A, X, B)[i, j, t] D[t, j, k]. 
\]  \hspace{1cm} (12)

By substituting

\[
\text{Prod}(A, X, B)[i, j, t] = \sum_{0 \leq s < p} A[i, s, t] X[i, j, s] B[s, j, t]
\]

into (12) we get

\[
X[i, j, k] = \sum_{0 \leq s, t < p} C[i, t, k] A[i, s, t] X[i, j, s] B[s, j, t] D[t, j, k].
\]

We regroup the factors in the summands as follows

\[
X[i, j, k] = \sum_{0 \leq s, t < p} (C[i, t, k] D[t, j, k])(A[i, s, t] B[s, j, t]) X[i, j, s].
\]

Consequently, the inverse pair relation asserts that

\[
\sum_{0 \leq t < p} (C[i, t, k] D[t, j, k])(A[i, s, t] B[s, j, t]) = \begin{cases} 
1 & \text{if } 0 \leq k = s < p \\
0 & \text{otherwise}
\end{cases},
\]

which expresses the product of a matrix with its inverse. Let \( F \) and \( F^{-1} \) denote these matrices of size \( m \cdot n \cdot p \times m \cdot n \cdot p \). The non-zero entries of \( F \) are expressed in terms of the entries of \( A \) and \( B \) as follows:

\[
\forall 0 \leq i < m, 0 \leq j < n, 0 \leq s, t < p, \quad F[i \cdot n \cdot p + j \cdot p + t, i \cdot n \cdot p + j \cdot p + s] = A[i, s, t] \cdot B[s, j, t].
\]
Assuming the pair $A \in \mathbb{C}^{m \times p \times p}$, $B \in \mathbb{C}^{p \times n \times p}$ is given, we seek the corresponding inverse pair $C \in \mathbb{C}^{m \times p \times p}$, $D \in \mathbb{C}^{p \times n \times p}$. Note that the associated matrices $F, F^{-1} \in \mathbb{C}^{m \cdot p \times m \cdot n \cdot p}$ both are made of a collection of $m \cdot n$ diagonal block matrices where each block is of size $p \times p$. Given that we know the entries of $F$ in terms of the entries of $A$ and $B$, we determine the entries of $F^{-1}$ as follows:

$$F^{-1} = \frac{\text{Adj}(F)}{\det(F)}.$$

Consequently the entries of $C$ and $D$ are determined by the multiplicative linear constraints

$$\{ C[i, t, k] \cdot D[t, j, k] = F^{-1}[i \cdot n \cdot p + j \cdot p + k, i \cdot n \cdot p + j \cdot p + l] \} \quad 0 \leq i < m, 0 \leq j < n \cdot$$

The constraints in (13) are multiplicative linear constraints which are solved using a variant of Gauss-Jordan elimination as described in [FG17]. Therefore the two necessary and sufficient conditions which guarantee the existence of an inverse pair $C \in \mathbb{C}^{m \times p \times p}$, $D \in \mathbb{C}^{p \times n \times p}$ for the pair $A \in \mathbb{C}^{m \times p \times p}$, $B \in \mathbb{C}^{p \times n \times p}$ are $\det(F) \neq 0$ and the fact that the multiplicative linear system (13) admits at least one solution.

### 3.4 Fundamental theorem of third order hypermatrix algebra

The fundamental theorem of third order hypermatrix algebra generalizes the well-known fundamental theorem of linear algebra [Str93]. For the sake of completeness, we recall the statement and a proof of the fundamental theorem of linear algebra. We subsequently extend the argument to third order hypermatrices. [pre-pending criteria for matrices and hypermatrices. The idea is to add free variables and determine the existence of inverse pairs.]

**Theorem 5**: If $A \in \mathbb{C}^{n \times n}$ has rank $r \leq n$ then the nullspace of $A$ has dimension $n - r$.

**Proof**: The proof exhibits a basis for the nullspace of $A$ having dimension $n - r$ given that $A$ has rank $r$.

$$\text{Rank}(A) = r \Rightarrow A = U \cdot V,$$

where $U \in \mathbb{C}^{n \times r}$, $V \in \mathbb{C}^{r \times n}$ and $\{ U[:, t] \}_{0 \leq t < r}$, $\{ V[t, :] \}_{0 \leq t < r}$ both form linearly independent sets of vectors, for if this was not the case the rank of $A$ could be further reduced. The matrix $A$ can therefore be expressed as the following outer product sum

$$A = \sum_{0 \leq t < r} U[:, t] \cdot V[t, :] + \sum_{0 \leq t < n} 0_{n \times 1} \cdot X[t, :],$$

where $X$ is an $(n - r) \times n$ matrix chosen so as to ensure that the matrix

$$\sum_{0 \leq t < r} I_n[:, t] \cdot V[t, :] + \sum_{0 \leq t < n} I_n[:, t] \cdot X[t, :]$$

has an inverse. The existence of $X$ follows from a polynomial method argument. It therefore follows that

$$A \cdot \left( \sum_{0 \leq t < r} I_n[:, t] \cdot V[t, :] + \sum_{0 \leq t < n} I_n[:, t] \cdot X[t, :] \right)^{-1} = \sum_{0 \leq t < r} U[:, t] \cdot I_n[t, :] + \sum_{0 \leq t < n} 0_{n \times 1} \cdot I_n[t, :].$$

18
which establishes that the dimension of the nullspace is at least \( n - r \). The fact that the dimension of the nullspace is precisely \( n - r \) follows from the linear independence of the vectors \( \{X[t, :]\}_{r \leq t < n} \), and completes our proof.

**Theorem 6**: Let \( A \in \mathbb{C}^{m \times n \times p} \), \( X_1 \in \mathbb{C}^{m \times p \times r} \) and \( X_2 \in \mathbb{C}^{p \times n \times r} \) such that

\[
\text{Prod}(X_1, A, X_2) = 0_{m \times n \times r},
\]

where \( p = \min \{m, n, p\} \) and \( X_1, X_2 \) are such that \( p - r \) non-zero depth slices can be prepended to both hypermatrices to form an invertible pairs of hypermatrices of size \( m \times p \times p \) and \( p \times n \times p \), respectively, denoted \( Y_1^{-1}, Y_2^{-1} \). Then \( A \) has rank at most \( p - r \).

**Proof**: The proof exhibits a factorization of \( A \). The hypothesis asserts that \( \text{Prod}(Y_1^{-1}, A, Y_2^{-1}) = B \in \mathbb{C}^{n \times n \times n} \) where the first \( r \) depth slices are all zero. Let \( Y_1, Y_2 \) denote inverse outer-pairs of \( Y_1^{-1}, Y_2^{-1} \). Consequently we have

\[
\text{Prod}(Y_1, \text{Prod}(Y_1^{-1}, A, Y_2^{-1}), Y_2) = A = \text{Prod}(Y_1, B, Y_2) = \sum_{r \leq t < p} \text{Prod}(Y_1[::], A[t, :], Y_2[::]),
\]

which expresses \( A \) as a sum of \( p - r \) outer products.

**References**

[FG17] Yuval Filmus and Edinah K. Gnang, *On the spectra of hypermatrix direct sum and kronecker products constructions*, Linear Algebra and its Applications (2017), –.

[GER11] E. K. Gnang, A. Elgammal, and V. Retakh, *A spectral theory for tensors*, Annales de la faculte des sciences de Toulouse Mathematiques 20 (2011), no. 4, 801–841.

[GKZ94] I. Gelfand, M. Kapranov, and A. Zelevinsky, *Discriminants, resultants and multidimensional determinant*, Birkhauser, Boston, 1994.

[GR91] I. M. Gelfand and V. S. Retakh, *Determinants of matrices over noncommutative rings*, Functional Analysis and Its Applications 25 (1991), no. 2, 91–102.

[GRW05] I. M. Gelfand, V. S. Retakh, and R. Wilson, *Quasideterminants*, Advances in Mathematics 193 (2005), no. 1, 56–141.

[Ker08] Richard Kerner, *Ternary and non-associative structures*, International Journal of Geometric Methods in Modern Physics 5 (2008), 1265–1294.

[Lim13] Lek-Heng Lim, *Tensors and hypermatrices*, Handbook of Linear Algebra (Leslie Hogben, ed.), CRC Press, 2013.
[MB90] D. M. Mesner and P. Bhattacharya, *Association schemes on triples and a ternary algebra*, Journal of combinatorial theory A55 (1990), 204–234.

[MB94] D. M. Mesner and P. Bhattacharya, *A ternary algebra arising from association schemes on triples*, Journal of algebra 164 (1994), 595–613.

[S+15] W. A. Stein et al., *Sage Mathematics Software (Version 6.9)*, The Sage Development Team, 2015, http://www.sagemath.org.

[Str93] Gilbert Strang, *The fundamental theorem of linear algebra*, The American Mathematical Monthly 100 (1993), no. 9, 848–855.

[Stu00] Bernd Sturmfels, *Solving algebraic equations in terms of a-hypergeometric series*, Discrete Mathematics 210 (2000), no. 1, 171 – 181.

[Stu02] B. Sturmfels, *Solving systems of polynomial equations*, Conference Board of the Mathematical Sciences regional conference series in mathematics, no. no. 97, American Mathematical Soc., 2002.