On the spectra of direct sums and Kronecker products of side length 2 hypermatrices

Yuval Filmus [*], Edinah K. Gnang [†]

January 5, 2016

Abstract

Our main result is an elementary derivation of the spectral decomposition of hypermatrices generated by arbitrary combinations of Kronecker products and direct sums of cubic hypermatrices of side length 2. The method is based on a generalization of Parseval's identity. We use this general formulation of Parseval's identity to introduce hypermatrix Fourier transforms and discrete Fourier hypermatrices. We extend to hypermatrices a variant of the Gram–Schmidt orthogonalization process as well as Sylvester's classical Hadamard matrix construction. We conclude the paper with illustrations of spectral decompositions of adjacency hypermatrices of finite groups and a short proof the hypermatrix Rayleigh quotient inequality.

1 Introduction

Hypermatrices are multidimensional arrays of complex numbers which generalize matrices. Formally, we may define a hypermatrix to be a finite set of complex numbers whose members are indexed by distinct elements of some fixed integer Cartesian product set of the form

$$\{0, 1, 2, \cdots, n_1\} \times \{0, 1, 2, \cdots, n_2\} \times \cdots \times \{0, 1, 2, \cdots, n_m\}.$$

Such a hypermatrix is said to be of order m and of size $(n_1 + 1) \times (n_2 + 1) \times \cdots \times (n_m + 1)$. Such a hypermatrix is said to be cubic of side length n if $n_1 = n_2 = \cdots = n_m = n$. In particular, matrices are second order hypermatrices. Hypermatrix algebras arise from attempts to extend to hypermatrices classical matrix algebra concepts and algorithms [MB94, GKZ94, Ker08, GER11]. Hypermatrices are common occurrences in applications relating to computer science, statistics and physics often embedded into multilinear forms associated with objective functions to be minimized or maximized. While many hypermatrix algebras have been proposed in the tensor/hypermatrix

[*]Computer Science Department, Technion - Israel Institute of Technology
[†]Department of Mathematics, Purdue University, email: egnang@purdue.edu

literature [Lim13], our discussion here focuses on the Bhattacharya-Mesner (BM) algebra first developed in [MB90, MB94] and the general BM algebra first proposed in [GER11]. The general BM product is of interest because it encompasses as special cases many other hypermatrix products in the literature including the usual matrix product, the Segre outer product, the contraction product, the higher order SVD, and the multilinear matrix multiplication. In [GER11], the authors devised a non-constructive proof of existence for the spectral decompositions of some special hypermatrices as conjectured by P. Bhattacharya [Bha95]. Our main result here is a constructive method for obtaining the spectral decomposition of hypermatrices generated by arbitrary combinations of Kronecker products and direct sums of cubic hypermatrices of side length 2. The method is based on a generalization of Parseval's identity. We use this general formulation of Parseval's identity to introduce hypermatrix Fourier transforms and discrete Fourier hypermatrices. We extend to hypermatrices a variant of the Gram–Schmidt orthogonalization process as well as Sylvester's classical Hadamard matrix construction. We conclude the paper with illustrations of spectral decompositions of adjacency hypermatrices of finite groups and a short proof of the hypermatrix Rayleigh quotient inequality.

This article is accompanied by an extensive Sage [S$^+$15] symbolic hypermatrix algebra package which implements the various features of the general BM 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. We would like to thank Vladimir Retakh and Ahmed Elgammal for patiently introducing us to the theory of hypermatrices. We are grateful to Doron Zeilberger and Saugata Basu for helpful suggestions. We would like to thank Ori Parzanchevski for helping get the hypermatrix package started. The second author was supported by the National Science Foundation, and is grateful for the hospitality of the Institute for Advanced Study and the Department of Mathematics at Purdue University.

2 Overview of the BM algebra

The Bhattacharya-Mesner product, or BM product for short, was first developed in [MB90, MB94]. The BM product provides a natural generalization of the classical matrix product. The BM product of second order hypermatrices therefore corresponds the usual matrix product. For notational consistency, we will on occasion refer to the matrix product $\mathbf{A}^{(1)} \cdot \mathbf{A}^{(2)}$ as $\text{Prod}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}\right)$. The BM product is best introduced to the unfamiliar reader by first describing the BM product of third and fourth order hypermatrices. Note that the BM product of second order hypermatrices is a binary operation, the BM product of third order hypermatrices is a ternary operation, the BM product of fourth order hypermatrices takes four operands, and so on.

The BM product of three third order hypermatrices $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$, denoted $\text{Prod}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\right)$, is defined if

$$\mathbf{A}^{(1)} \text{ is } n_1 \times k \times n_3,\ \mathbf{A}^{(2)} \text{ is } n_1 \times n_2 \times k \text{ and } \mathbf{A}^{(3)} \text{ is } k \times n_2 \times n_3.$$

The result is a hypermatrix of size $n_1 \times n_2 \times n_3$, and specified entry-wise by

$$\left[\text{Prod}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\right)\right]_{i_1,i_2,i_3} = \sum_{0\le j<k} a^{(1)}_{i_1\, j\, i_3}\, a^{(2)}_{i_1\, i_2\, j}\, a^{(3)}_{j\, i_2\, i_3}.$$

Similarly, the BM product of fourth order hypermatrices $\mathbf{A}^{(1)}$,$\mathbf{A}^{(2)}$,$\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$, denoted $\text{Prod}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}, \mathbf{A}^{(4)}\right)$ is defined if

$$\mathbf{A}^{(1)} \text{ is } n_1 \times k \times n_3 \times n_4, \mathbf{A}^{(2)} \text{ is } n_1 \times n_2 \times k \times n_4,$$

$$\mathbf{A}^{(3)} \text{ is } n_1 \times n_2 \times n_3 \times k, \text{ and } \mathbf{A}^{(4)} \text{ is } k \times n_2 \times n_3 \times n_4.$$

The result $\text{Prod}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}, \mathbf{A}^{(4)}\right)$ will be of size $n_1 \times n_2 \times n_3 \times n_4$ and specified entry-wise by

$$\left[\text{Prod}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}, \mathbf{A}^{(4)}\right)\right]_{i_1,i_2,i_3,i_4} = \sum_{0\le j<k} a^{(1)}_{i_1\, j\, i_3 i_4}\, a^{(2)}_{i_1 i_2\, j\, i_4}\, a^{(3)}_{i_1 i_2 i_3\, j}\, a^{(4)}_{j\, i_2 i_3 i_4}.$$

The reader undoubtedly has already discerned the general pattern, but for the sake of completeness we express the entries of the BM product of order m hypermatrices

$$\left[\text{Prod}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(t)}, \cdots, \mathbf{A}^{(m)}\right)\right]_{i_1,\cdots,i_m} = \sum_{0\le j<k} a^{(1)}_{i_1\, j\, i_3\cdots i_m} \cdots a^{(t)}_{i_1\cdots i_t\, j\, i_{t+2}\cdots i_m} \cdots a^{(m)}_{j\, i_2\cdots i_m}. \tag{1}$$

An arbitrary m-tuple of order m hypermatrices $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$ for which the BM product is defined is called *BM conformable*.

We recall a recent variant of the BM product called the *general BM product*. The general BM product was first proposed in [GER11]. The general BM product encompasses as special cases many other hypermatrix products in the literature, including the usual matrix product, the Segre outer product, the contraction product, the higher order SVD, and the multilinear matrix multiplication [Lim13]. In addition, the general BM product is of particular interest to our discussion because it enables considerable notational simplification. The general BM product of order m hypermatrices is defined for any BM conformable m-tuple $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$ and an additional cubic hypermatrix $\mathbf{B}$ called the background hypermatrix with side length k (the contracted dimension). This product is denoted $\text{Prod}_{\mathbf{B}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$ and has entries given by

$$\left[\text{Prod}_{\mathbf{B}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)\right]_{i_1,\cdots,i_m} = \sum_{0\le j_1,\cdots,j_m<k} a^{(1)}_{i_1\, j_1\, i_3\cdots i_m} \cdots a^{(t)}_{i_1\cdots i_t\, j_t\, i_{t+2}\cdots i_m} \cdots a^{(m)}_{j_m\, i_2\cdots i_m} b_{j_1\cdots j_m}. \tag{2}$$

For example, the general BM product of third order hypermatrices $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$ with the background hypermatrix $\mathbf{B}$ denoted $\text{Prod}_{\mathbf{B}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\right)$ is defined if

$$\mathbf{A}^{(1)} \text{ is } n_1 \times k \times n_3,\ \mathbf{A}^{(2)} \text{ is } n_1 \times n_2 \times k,\ \mathbf{A}^{(3)} \text{ is } k \times n_2 \times n_3 \text{ and } \mathbf{B} \text{ is } k \times k \times k.$$

The result is a hypermatrix be of size $n_1 \times n_2 \times n_3$ with entries given by

$$\left[\text{Prod}_{\mathbf{B}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\right)\right]_{i_1,i_2,i_3} = \sum_{0\le j_1,j_2,j_3<k} a^{(1)}_{i_1\, j_1\, i_3}\, a^{(2)}_{i_1 i_2\, j_2}\, a^{(3)}_{j_3\, i_2 i_3}\, b_{j_1 j_2 j_3}.$$

Note that the original BM product of order m hypermatrices is recovered from the general BM product by taking the background hypermatrix $\mathbf{B}$ to be the m-th order Kronecker delta hypermatrix denoted $\mathbf{\Delta}$, whose entries are specified by

$$[\mathbf{\Delta}]_{i_1,\cdots,i_m} = \begin{cases} 1 & \text{if } 0 \le i_1 = \cdots = i_m < n \\ 0 & \text{otherwise} \end{cases}.$$

In particular, Kronecker delta matrices correspond to identity matrices.

We also recall for the reader's convenience the definition of the hypermatrix transpose operations. Let $\mathbf{A}$ be a hypermatrix of size $n_1 \times n_2 \times \cdots \times n_m$ whose entries are

$$[\mathbf{A}]_{i_1,i_2,\cdots,i_{m-1},i_m} = a_{i_1\, i_2 \,\cdots\, i_{m-1}\, i_m}.$$

The corresponding transpose, denoted $\mathbf{A}^T$, is a hypermatrix of size $n_2 \times n_3 \times \cdots \times n_m \times n_1$ whose entries are given by

$$\left[\mathbf{A}^T\right]_{i_1,i_2,\cdots,i_{m-1},i_m} = a_{i_m\, i_1 \,\cdots\, i_{m-2}\, i_{m-1}}.$$

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

$$\mathbf{A}^{T^2} := \left(\mathbf{A}^T\right)^T,\ \mathbf{A}^{T^3} := \left(\mathbf{A}^{T^2}\right)^T,\ \cdots,\ \mathbf{A}^{T^m} := \left(\mathbf{A}^{T^{(m-1)}}\right)^T = \mathbf{A}.$$

By this convention

$$\mathbf{A}^{T^i} = \mathbf{A}^{T^j} \ \text{ if } \ i \equiv j \mod m.$$

It follows from the definition of the transpose that

$$\text{Prod}\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right)^T = \text{Prod}\left(\left(\mathbf{A}^{(2)}\right)^T,\cdots,\left(\mathbf{A}^{(m)}\right)^T,\left(\mathbf{A}^{(1)}\right)^T\right), \tag{3}$$

The identity 3 generalizes the matrix transpose identity

$$\left(\mathbf{A}^{(1)}\cdot\mathbf{A}^{(2)}\right)^T = \left(\mathbf{A}^{(2)}\right)^T\cdot\left(\mathbf{A}^{(1)}\right)^T.$$

Finally, for further notational convenience, we briefly discuss the use of the general BM product to express multilinear forms and outer products.

Let $\mathbf{A} \in \mathbb{C}^{n_0\times\cdots\times n_{m-1}}$ denote an arbitrary order m hypermatrix and consider an arbitrary set of m vectors $\left\{\mathbf{x}_j \in \mathbb{C}^{n_j\times 1\times\cdots\times 1}\right\}_{0\le j<m}$. The general BM product

$$\text{Prod}_{\mathbf{A}}\left(\mathbf{x}_0^{T^{(m-1)}},\mathbf{x}_1^{T^{(m-2)}},\cdots,\mathbf{x}_{m-j-1}^{T^j},\cdots,\mathbf{x}_{m-2}^{T^1},\mathbf{x}_{m-1}^{T^0}\right),$$

expresses the mulitlinear form associated with the hypermatrix $\mathbf{A}$. As illustration, consider an arbitrary third order hypermatrix $\mathbf{A} \in \mathbb{C}^{m\times n\times p}$ and three vectors $\mathbf{x} \in \mathbb{C}^{m\times 1\times 1}$, $\mathbf{y} \in \mathbb{C}^{n\times 1\times 1}$ and $\mathbf{z} \in \mathbb{C}^{p\times 1\times 1}$. The corresponding multilinear form of degree 3 is expressed as

$$\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2},\mathbf{y}^{T^1},\mathbf{z}^{T^0}\right) = \sum_{0\le i<m}\sum_{0\le j<n}\sum_{0\le k<p} a_{ijk}\, x_i\, y_j\, z_k.$$

Similarly, for an arbitrary matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ and pair of vectors $\mathbf{x} \in \mathbb{C}^{m\times 1}$, $\mathbf{y} \in \mathbb{C}^{n\times 1}$, the corresponding quadratic form is expressed by

$$\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^1}, \mathbf{y}^{T^0}\right) = \sum_{0\le i<m} \sum_{0\le j<n} a_{ij}\, x_i\, y_j = \mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{y}.$$

The general BM product also provides a convenient way to express outer products. For an arbitrary BM conformable m-tuple $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$, a BM outer product corresponds to a product of the form

$$\text{Prod}\left(\mathbf{A}^{(1)}[:, t, :, \cdots, :], \mathbf{A}^{(2)}[:, :, t, \cdots, :], \cdots, \mathbf{A}^{(m-1)}[:, :, \cdots, t], \mathbf{A}^{(m)}[t, :, \cdots, :]\right). \qquad (4)$$

In the product above we used the colon notation. In the colon notation, $\mathbf{A}^{(1)}[:, t, :, \cdots, :]$ refers to a hypermatrix *slice* of size $n_1 \times 1 \times n_3 \times \cdots \times n_m$ where the second index is fixed to t while all other indices all allowed to vary within their prescribed ranges. Hypermatrix outer products will be a common occurrence throughout our discussion. Fortunately, hypermatrix outer products are more conveniently expressed in terms of the hypermatrices noted $\left\{\boldsymbol{\Delta}^{(t)}\right\}_{0\le t<n}$ with entries such that

$$\left[\boldsymbol{\Delta}^{(t)}\right]_{i_1,\cdots,i_m} = \begin{cases} 1 & \text{if } 0 \le t = i_1 = \cdots = i_m < k \\ 0 & \text{otherwise} \end{cases},$$

$$\Rightarrow \boldsymbol{\Delta} = \sum_{0\le t<k} \boldsymbol{\Delta}^{(t)}.$$

Where $\boldsymbol{\Delta}$ denotes here the order m Kronecker delta of side length k. Using the general BM product we may more conveniently express the BM outer product in 4 as $\text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$. This more convenient description of outer products also provides a natural extension to hypermatrices of the notion of rank. Using this notation, we recall from linear algebra that a matrix $\mathbf{B}$ is of rank r (over $\mathbb{C}$) if there exists a conformable matrix pair $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$ such that

$$\mathbf{B} = \sum_{0\le t<r} \text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right),$$

and crucially there exists no conformable matrix pair $\mathbf{Y}^{(1)}$, $\mathbf{Y}^{(2)}$ such that

$$\mathbf{B} = \sum_{0\le t<r-1} \text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}\right).$$

This definition of rank extends verbatim to hypermatrices and is called the BM rank. An order m hypermatrix $\mathbf{B}$ has BM rank r (over $\mathbb{C}$) if there exists a BM conformable m-tuple $\left(\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\right)$ such that

$$\mathbf{B} = \sum_{0\le t<r} \text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\right),$$

and crucially there exists no BM conformable m-tuple $\left(\mathbf{Y}^{(1)}, \cdots, \mathbf{Y}^{(m)}\right)$ such that

$$\mathbf{B} = \sum_{0 \le t < r-1} \text{Prod}_{\mathbf{\Delta}^{(t)}} \left(\mathbf{Y}^{(1)}, \cdots, \mathbf{Y}^{(m)}\right).$$

Note that the usual notions of tensor/hypermatrix rank as well as the canonical polyadic rank discussed in the literature [Lim13] correspond to constrained versions of the BM rank where additional constraints are imposed on the hypermatrices in the m-tuple $\left(\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\right)$.

3 General Parseval identity and Fourier transforms

3.1 Hypermatrix Parseval identity

The classical matrix Parseval identity states that if $\mathbf{U}$ is an $n \times n$ unitary matrix then for every complex vectors $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)} \in \mathbb{C}^{n \times 1}$ of length n,

$$\left(\overline{\mathbf{x}^{(1)}}\right)^T \cdot \mathbf{x}^{(2)} = \left(\overline{\mathbf{U} \cdot \mathbf{x}^{(1)}}\right)^T \cdot \left(\mathbf{U} \cdot \mathbf{x}^{(2)}\right).$$

When generalizing this to hypermatrices we can't quite form the matrix-vector products $\mathbf{U} \cdot \mathbf{x}^{(1)}, \mathbf{U} \cdot \mathbf{x}^{(2)}$. Instead, notice that $\mathbf{y}^{(1)} = \mathbf{U} \cdot \mathbf{x}^{(1)}$ and $\mathbf{y}^{(2)} = \mathbf{U} \cdot \mathbf{x}^{(2)}$ satisfy

$$\overline{y_k^{(1)}}\, y_k^{(2)} = \left[\overline{\mathbf{U} \cdot \mathbf{x}^{(1)}}\right]_k^T \left[\mathbf{U} \cdot \mathbf{x}^{(2)}\right]_k = \text{Prod}_{\mathbf{\Delta}^{(k)}} \left(\left(\overline{\mathbf{U} \cdot \mathbf{x}^{(1)}}\right)^T, \mathbf{U} \cdot \mathbf{x}^{(2)}\right)$$

This formulation generalizes to the hypermatrix case.

An m-tuple $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$ of cubic order m hypermatrices each of side length n, forms an *uncorrelated tuple* if the corresponding BM product equals the Kronecker delta hypermatrix :

$$\text{Prod}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right) = \mathbf{\Delta}.$$

In some sense, uncorrelated tuples extend to hypermatrices the notion of matrix inverse pair. Furthermore, a cubic m-th order hypermatrix $\mathbf{Q}$ of side length n is *orthogonal* if the following holds :

$$\text{Prod}\left(\mathbf{Q}, \mathbf{Q}^{T^{(m-1)}}, \cdots, \mathbf{Q}^{T^k}, \cdots, \mathbf{Q}^{T^2}, \mathbf{Q}^T\right) = \mathbf{\Delta}.$$

Finally, a cubic hypermatrix $\mathbf{U}$ of even order say $2m$ and of side length n is *unitary* if the following holds

$$\text{Prod}\left(\mathbf{U}, \overline{\mathbf{U}}^{T^{(2m-1)}}, \cdots \overline{\mathbf{U}}^{T^{2k+1}}, \mathbf{U}^{T^{2k}}, \cdots, \mathbf{U}^{T^2}, \overline{\mathbf{U}}^T\right) = \mathbf{\Delta}.$$

Both orthogonal and unitary hypermatrices yield special uncorrelated tuples.

For an arbitrary uncorrelated m-tuple $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$, each of side length n, and some arbitrary m-tuple of vectors $(\mathbf{x}_0, \cdots, \mathbf{x}_{m-1})$, each of size $n \times 1 \times \cdots \times 1$, the associated *Parseval identity* is formulated in terms of BM outer products

$$\forall\, 0 \le k < n, \quad \mathbf{P}_k = \text{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right).$$

Proposition 1 : Let $\left\{\mathbf{x}^{(j)}\right\}_{0 \le j < m} \subset \mathbb{C}^{n \times 1 \times \cdots \times 1}$ and $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$ an arbitrary uncorrelated m-tuple, then for all vectors $\left\{\mathbf{y}^{(j)}\right\}_{0 \le j < m} \subset \mathbb{C}^{n \times 1 \times \cdots \times 1}$ such that

$$\forall\, 0 \le k < n, \quad \left(\prod_{0 \le j < m} y_k^{(j)}\right) = \text{Prod}_{\mathbf{P}_k}\left(\left(\mathbf{x}^{(m-1)}\right)^{T^{(m-1)}}, \cdots, \left(\mathbf{x}^{(j)}\right)^{T^j}, \cdots, \left(\mathbf{x}^{(0)}\right)^{T^0}\right),$$

where $\mathbf{P}_k = \text{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$ and $y_k^{(j)}$ denotes the k-th entry of the vector $\mathbf{y}^{(j)}$, we have

$$\text{Prod}\left(\left(\mathbf{y}^{(m-1)}\right)^{T^{(m-1)}}, \cdots, \left(\mathbf{y}^{(j)}\right)^{T^j}, \cdots, \left(\mathbf{y}^{(0)}\right)^{T^0}\right)$$
$$=$$
$$\text{Prod}\left(\left(\mathbf{x}^{(m-1)}\right)^{T^{(m-1)}}, \cdots, \left(\mathbf{x}^{(j)}\right)^{T^j}, \cdots, \left(\mathbf{x}^{(0)}\right)^{T^0}\right).$$

In particular, in the matrix case where $\mathbf{x}^{(0)}, \mathbf{x}^{(1)} \in \mathbb{C}^{n \times 1}$ and $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ are $n \times n$ matrix inverse pairs, Parseval's identity asserts that

$$\forall\, \mathbf{y}^{(0)}, \mathbf{y}^{(1)} \in \mathbb{C}^{n \times 1}$$

such that

$$\forall\, 0 \le k < n, \quad y_k^{(1)}\, y_k^{(0)} = \text{Prod}_{\mathbf{P}_k}\left(\left(\mathbf{x}^{(1)}\right)^{T}, \mathbf{x}^{(0)}\right) \text{ and } \mathbf{P}_k = \text{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}\right)$$

we have

$$\text{Prod}\left(\left(\mathbf{y}^{(1)}\right)^{T}, \mathbf{y}^{(0)}\right) = \text{Prod}\left(\left(\mathbf{x}^{(1)}\right)^{T}, \mathbf{x}^{(1)}\right).$$

Proof : The proof follows from the identity

$$\text{Prod}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right) = \sum_{0 \le k < n} \text{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right).$$

Consequently

$$\left(\sum_{0 \le k < n} \prod_{0 \le j < m} y_k^{(j)}\right) = \sum_{0 \le k < n} \text{Prod}_{\mathbf{P}_k}\left(\left(\mathbf{x}^{(m-1)}\right)^{T^{(m-1)}}, \cdots, \left(\mathbf{x}^{(j)}\right)^{T^j}, \cdots, \left(\mathbf{x}^{(0)}\right)^{T^0}\right) =$$

$$\mathrm{Prod}_{\left(\sum_{0\le k<n}\mathbf{P}_k\right)}\left(\left[\mathbf{x}^{(m-1)}\right]^{T^{(m-1)}},\cdots,\left[\mathbf{x}^{(j)}\right]^{T^j},\cdots,\left[\mathbf{x}^{(0)}\right]^{T^0}\right),$$

where $\mathbf{P}_k = \mathrm{Prod}_{\mathbf{\Delta}^{(k)}}\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right)$. This yields the desired result

$$\mathrm{Prod}\left(\left(\mathbf{y}^{(m-1)}\right)^{T^{(m-1)}},\cdots,\left(\mathbf{y}^{(j)}\right)^{T^j},\cdots,\left(\mathbf{y}^{(0)}\right)^{T^0}\right)$$
$$=$$
$$\mathrm{Prod}\left(\left(\mathbf{x}^{(m-1)}\right)^{T^{(m-1)}},\cdots,\left(\mathbf{x}^{(j)}\right)^{T^j},\cdots,\left(\mathbf{x}^{(0)}\right)^{T^0}\right).$$

3.2 Hypermatrix orthogonalization and constrained uncorrelated tuples

Applications of the proposed generalization of Parseval's identity are predicated on the existence of non-trivial[1] orthogonal, unitary and uncorrelated hypermatrix tuples. We present here an algorithmic proof of existence of non-trivial orthogonal and uncorrelated hypermatrices of all orders and side lengths. The main argument will be akin to proving the existence of non-trivial orthogonal matrices by showing that the Gram–Schmidt process derives non-trivial orthogonal matrices from generic input matrices.
More generally, we call *orthogonalization procedures* any algorithms which take as input some generic hypermatrices and output either orthogonal, unitary, or uncorrelated hypermatrix tuples.

The first variant of the Gram–Schmidt process which extends to hypermatrices was proposed in [Gna13]. We will show here that this variant of the Gram–Schmidt process yields an algorithmic proof of existence of non trivial orthogonal and non-trivial uncorrelated hypermatrix tuples.

Matrix orthogonalization problem:
Derive from a generic input matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$ a matrix $\mathbf{X}$ of the same size subject to

$$\left(\mathbf{1}_{n\times n}-\mathbf{I}_n\right)\circ\left(\mathbf{X}\cdot\mathbf{X}^T\right)=\mathbf{0}_{n\times n},$$

where $\circ$ denotes the entry-wise product also called the Hadamard product, and $\mathbf{1}_{n\times n}$ denotes the $n\times n$ all one matrix. (Equivalently, the rows $x_1,\ldots,x_n$ of X are pairwise orthogonal.)

Hypermatrix orthogonalization problem:
Derive from a generic order m input hypermatrix $\mathbf{A}\in\mathbb{C}^{n\times\cdots\times n}$ a hypermatrix $\mathbf{X}$ of the same size subject to

$$\left(\mathbf{1}_{n\times\cdots\times n}-\mathbf{\Delta}\right)\circ\mathrm{Prod}\left(\mathbf{X},\mathbf{X}^{T^{(m-1)}},\cdots,\mathbf{X}^{T^2},\mathbf{X}^T\right)=\mathbf{0}_{n\times\cdots\times n},$$

where $\circ$ denotes the Hadamard product.

It is well-known that the Gram–Schmidt process yields a solution to the matrix orthogonalization

[1] Note that trivial orthogonal, unitary and uncorrelated hypermatrix tuples are obtained by considering diagonal hypermatrices whose nonzero entries are roots of unity located at entries such that all their indices are equal.

problem. We describe here a variant of the Gram–Schmidt process which extends to hypermatrices of all orders. Our proposed solution to the matrix orthogonalization problem is obtained by solving for the entries of $\mathbf{X}$ in the following system of $n\binom{n}{2}$ equations:

$$\left\{ x_{ut}\, x_{vt} = a_{ut}\, a_{vt} - n^{-1} \sum_{0\le s<n} a_{us}\, a_{vs} \right\}_{\substack{0\le t<n \\ 0\le u<v<n}} .$$

(It is not hard to check that any such solution is indeed orthogonal.)

For notational convenience, we rewrite the constraints above in terms of the general BM product. The system of $n\binom{n}{2}$ equations can be more simply expressed as

$$\forall\, 0\le t<n,\ (\mathbf{1}_{n\times n} - \mathbf{I}_n)\circ \text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{X}, \mathbf{X}^T\right) = \text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{A}, \mathbf{A}^T\right) - n^{-1}\text{Prod}\left(\mathbf{A}, \mathbf{A}^T\right), \quad (5)$$

where $\mathbf{1}_{n\times n}$ denotes the $n\times n$ all one matrix.

Similarly, a solution to the hypermatrix orthogonalization problem is obtained by solving for the entries of $\mathbf{X}$ in the hypermatrix formultation of the constraints in 5 given by

$$\forall\, 0\le t<n,\quad (\mathbf{1}_{n\times\cdots\times n} - \boldsymbol{\Delta})\circ \text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{X}, \mathbf{X}^{T^{(m-1)}}, \cdots, \mathbf{X}^{T^2}, \mathbf{X}^T\right) =$$

$$(\mathbf{1}_{n\times\cdots\times n} - \boldsymbol{\Delta})\circ\left(\text{Prod}_{\boldsymbol{\Delta}^{(t)}}\left(\mathbf{A}, \mathbf{A}^{T^{(m-1)}}, \cdots, \mathbf{A}^{T^2}, \mathbf{A}^T\right) - n^{-1}\text{Prod}\left(\mathbf{A}, \mathbf{A}^{T^{(m-1)}}, \cdots, \mathbf{A}^{T^2}, \mathbf{A}^T\right)\right). \quad (6)$$

(Again, it is not hard to check that any solution to this sytem is orthogonal.)

Both matrix and hypermatrix orthogonalization constraints in 5 and 6 turn out to be *monomial constraints*. General monomial constraints correspond to a system of equations which can be expressed in terms of a coefficient matrix $\mathbf{A}\in\mathbb{C}^{m\times n}$, a right-hand side vector $\mathbf{b}\in\mathbb{C}^{m\times 1}$, and an unknown vector $\mathbf{x}$ of size $n\times 1$. These constraints are of the form

$$\left\{ \left(\prod_{0\le t<n} x_t^{a_{it}} \right) = b_i \right\}_{1\le i\le m} . \quad (7)$$

Such constraints are in fact linear constraints as seen by taking the logarithm on both sides of the equation for each constraint. However, we solve such systems without using logarithms to avoid any difficulty related to branching of the logarithm. Instead, we solve such systems using a slight variation of the Gauss-Jordan elimination algorithm, prescribed by the following elementary row operations:

- Row exchange : $\text{R}_i \leftrightarrow \text{R}_j$
- Row scaling : $(\text{R}_i)^k \rightarrow \text{R}_i$
- Row linear combination : $(\text{R}_i)^k \cdot (\text{R}_j) \rightarrow \text{R}_j$

where $k \in \mathbb{C}$ and R_i denotes the particular constraint $\left(\prod_{0\le t<n} x_t^{a_{it}}\right) = b_i$.

Proposition 2a : The solution $\mathbf{X}$ to the orthogonalization constraints for a generic input hypermatrix $\mathbf{A}$ [2] yields a non-trivial orthogonal hypermatrix after normalization of rows of the solution matrix $\mathbf{X}$.

Proof : The proof follows directly from Gauss-Jordan elimination procedure. We deduce from the reduced row echelon form of the orthogonalization constraints 5,6 a criterion for existence of solutions in terms of a single polynomial in the entries of $\mathbf{A}$ which should be different from zero for some input. This condition will be generically satisfied, thereby establishing the desired result.

For example, in the case of a 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix},$$

Gauss-Jordan elimination yields the solution

$$\mathbf{X} = \begin{pmatrix} \frac{a_{00}a_{10}-a_{01}a_{11}}{2\,x_{10}} & -\frac{a_{00}a_{10}-a_{01}a_{11}}{2\,x_{11}} \\ x_{10} & x_{11} \end{pmatrix}.$$

The rows of $\mathbf{X}$ can be normalized to form an orthogonal matrix if no division by zero occurs and

$$\left[\mathbf{X}\cdot\mathbf{X}^T\right]_{0,0} \neq 0\,,\left[\mathbf{X}\cdot\mathbf{X}^T\right]_{1,1} \neq 0 \Leftrightarrow \left(a_{00}a_{10}-a_{01}a_{11}\right)\left(x_{10}^2+x_{11}^2\right)x_{10}x_{11} \neq 0.$$

Similarly for a $2 \times 2 \times 2$ hypermatrix

$$\mathbf{A}\left[:,:,0\right] = \begin{pmatrix} a_{000} & a_{010} \\ a_{100} & a_{110} \end{pmatrix}, \quad \mathbf{A}\left[:,:,1\right] = \begin{pmatrix} a_{001} & a_{011} \\ a_{101} & a_{111} \end{pmatrix},$$

Gauss-Jordan elimination yields the solution

$$\mathbf{X}\left[:,:,0\right] = \begin{pmatrix} \frac{(a_{000}a_{001}a_{100}-a_{010}a_{011}a_{110})x_{101}}{a_{001}a_{100}a_{101}-a_{011}a_{110}a_{111}} & \frac{(a_{000}a_{001}a_{100}-a_{010}a_{011}a_{110})x_{111}}{a_{001}a_{100}a_{101}-a_{011}a_{110}a_{111}} \\ \frac{a_{001}a_{100}a_{101}-a_{011}a_{110}a_{111}}{2\,x_{001}x_{101}} & -\frac{a_{001}a_{100}a_{101}-a_{011}a_{110}a_{111}}{2\,x_{011}x_{111}} \end{pmatrix}, \mathbf{X}\left[:,:,1\right] = \begin{pmatrix} x_{001} & x_{011} \\ x_{101} & x_{111} \end{pmatrix}.$$

The rows of $\mathbf{X}$ can be normalized to form an orthogonal matrix if no division by zero occurs and

$$\left[\mathrm{Prod}\left(\mathbf{X},\mathbf{X}^{T^2},\mathbf{X}^T\right)\right]_{0,0,0} \neq 0,\ \left[\mathrm{Prod}\left(\mathbf{X},\mathbf{X}^{T^2},\mathbf{X}^T\right)\right]_{1,1,1} \neq 0$$

$$\Leftrightarrow$$

$$\left(a_{001}a_{100}a_{101}-a_{011}a_{110}a_{111}\right)\left(a_{000}a_{001}a_{100}-a_{010}a_{011}a_{110}\right)\left(x_{101}^3+x_{111}^3\right)\left(x_{001}x_{101}x_{011}x_{111}\right) \neq 0$$

[2] A generic hypermatrix, is such that its entries do not satisfy any particular algebraic relation

Note that the proposed orthogonalization procedure in the matrix case is somewhat more restrictive in comparison to the Gram-Schmidt procedure. This is seen by observing that $0 \neq \det(\mathbf{A})$ is not a sufficient condition to ensure the existence of solutions to the orthogonalization procedure. However, the proposed orthogonalization procedure has the benefit of obtaining a set of orthogonal row vectors which depend on the ordering of the rows of the input $\mathbf{A}$ only up to sign.

More interestingly, the proposed orthogonalization constraints are special instances of a more general problem called the constrained uncorrelated tuple problem. A solution to the constrained uncorrelated tuple problem provides a proof of existence of non-trivial uncorrelated tuples. The constrained uncorrelated tuple problem is specified as follows.

Constrained inverse pair problem:
Derive from a generic input matrix pair $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)} \in \mathbb{C}^{n\times n}$ matrices $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$ of the same size such that

$$(\mathbf{1}_{n\times n} - \mathbf{I}_n) \circ \text{Prod}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right) = \mathbf{0}_{n\times n},$$

which minimizes

$$\sum_{0\le t<n} \left\|(\mathbf{1}_{n\times n} - \mathbf{I}_n) \circ \left(\text{Prod}_{\mathbf{\Delta}^{(t)}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}\right) - \text{Prod}_{\mathbf{\Delta}^{(t)}}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right)\right)\right\|_{\ell_2}^2$$

where $\circ$ denotes the Hadamard product.

Constrained uncorrelated tuple problem:
Derive from a generic m-tuple of order m input hypermatrices $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right) \subset \mathbb{C}^{n\times\cdots\times n}$ an m-tulple of hypermatrices $\left(\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\right)$ of the same such that

$$(\mathbf{1}_{n\times\cdots\times n} - \mathbf{\Delta}) \circ \text{Prod}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(m)}\right) = \mathbf{0}_{n\times\cdots\times n},$$

which minimizes

$$\sum_{0\le t<n} \left\|(\mathbf{1}_{n\times\cdots\times n} - \mathbf{\Delta}) \circ \left(\text{Prod}_{\mathbf{\Delta}^{(t)}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \cdots, \mathbf{A}^{(m)}\right) - \text{Prod}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(m)}\right)\right)\right\|_{\ell_2}^2$$

where $\circ$ denotes the Hadamard product and the hypermatrix $\mathbf{\Delta}$ denotes the Kronecker delta.

Proposition 2b : A solution to the constrained uncorrelated tuple problem is obtained by solving for the entries of the m-tuple of hypermatrices $\left(\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\right)$ in the constraints

$$\forall\, 0 \le t < n, \quad (\mathbf{1}_{n\times\cdots\times n} - \mathbf{\Delta}) \circ \text{Prod}_{\mathbf{\Delta}^{(t)}}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(m)}\right) =$$

$$(\mathbf{1}_{n\times\cdots\times n} - \mathbf{\Delta}) \circ \left[\text{Prod}_{\mathbf{\Delta}^{(t)}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \cdots, \mathbf{A}^{(m)}\right) - n^{-1}\text{Prod}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \cdots, \mathbf{A}^{(m)}\right)\right]. \quad (8)$$

For generic input hypermatrices $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right) \subset \mathbb{C}^{n\times\cdots\times n}$ the rows of $\left(\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\right)$ can be normalized to obtain a non-trivial uncorrelated tuple.

Proof : The proof again follows directly from the Gauss-Jordan elimination procedure. The constraints in 8 correspond to a system of n^m *monomial constraints* in $m \cdot n^m$ variables. We solve such a system via Gauss-Jordan elimination. By the argument used in the Proposition 2a we know the hypermatrices $\left(\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\right)$ can be normalized to form non-trivial uncorrelated tuples. Finally the fact that the obtained solution minimizes the sum

$$\sum_{0 \le t < n} \left\| (\mathbf{1}_{n\times\cdots\times n} - \mathbf{\Delta}) \circ \left(\text{Prod}_{\mathbf{\Delta}^{(t)}} \left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \cdots, \mathbf{A}^{(m)} \right) - \text{Prod}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(m)} \right) \right) \right\|_{\ell_2}^2$$

follows from the fact that the right-hand side of equalities in 8 expresses an orthogonal projection.

Our proposed solution to the uncorrelated tuple problem therefore yields an algorithmic proof of existence of non trivial uncorrelated tuples. The following corollary follows from Proposition 2a

Corollary 2c : For every order $m \ge 2$ and every side length $n \ge 2$ there exists an orthogonal hypermatrix having no zero entries.

Proof : By Proposition 2a, The reduced echelon form of the constraints 5,6 yields a criterion for the existence of non trivial solution. The criterion is expressed as a non-zero single polynomial of degree d in the entries of both $\mathbf{A}$ and possibly some free variables, which must not evaluate to zero for our choice of input. A generic choice of $\mathbf{A}$ and the free variables will satisfy this.

Corollary 2d : For every positive integer n and every positive integer m, there exists an order m orthogonal hypermatrix of side length 2^n whose entries are roots of unity all of which are scaled by the same normalizing constant.

Proof : By Lemma 3, it suffices to prove there exists for every positive integer m a hypermatrix of side length 2 such that

$$(\mathbf{1}_{n\times\cdots\times n} - \mathbf{\Delta}) \circ \text{Prod}\left(\mathbf{X}, \mathbf{X}^{T^{(m-1)}}, \cdots, \mathbf{X}^{T^2}, \mathbf{X}^T \right) = \mathbf{0}_{n\times\cdots\times n}$$

and

$$\left[\text{Prod}\left(\mathbf{X}, \mathbf{X}^{T^{(m-1)}}, \cdots, \mathbf{X}^{T^2}, \mathbf{X}^T \right) \right]_{0,0,\cdots,0} = \left[\text{Prod}\left(\mathbf{X}, \mathbf{X}^{T^{(m-1)}}, \cdots, \mathbf{X}^{T^2}, \mathbf{X}^T \right) \right]_{1,1,\cdots,1}$$

The proof exploits the very special structure of the hypermatrix orthogonalization constraints 6. The first observation is that for all indices $(i_1, \cdots, i_m)$ in the constraints 6, the right-hand side expressions associated with constraints whose left-hand side expressions are given by

$$x_{i_1\, 0\, i_3 \cdots i_m}\, x_{i_2\, 0\, i_4 \cdots i_m i_1} \cdots x_{i_m\, 0\, i_2 \cdots i_{m-1}}$$

and

$$x_{i_1\,1\,i_3\cdots i_m}\,x_{i_2\,1\,i_4\cdots i_m i_1}\cdots x_{i_m\,1\,i_2\cdots i_{m-1}}$$

will be identical, up to sign. The second observation is that by symmetry, for all indices $(i_1, i_2 \cdots, i_{m-1}, i_m) \notin \{(0,0,\cdots,0), (1,1,\cdots,1)\}$ it suffices to consider only two equations associated with each cyclic orbit $(i_1, i_2 \cdots, i_{m-1}, i_m), (i_2, \cdots, i_{m-1}, i_m, i_1), \cdots, (i_m, i_1, i_2 \cdots, i_{m-1})$. The two equations associated with each orbit are given by

$$x_{i_1\,0\,i_3\cdots i_m}\,x_{i_2\,0\,i_4\cdots i_m i_1}\cdots x_{i_m\,0\,i_2\cdots i_{m-1}} =$$

$$\left[\left(\mathbf{1}_{n\times\cdots\times n} - \boldsymbol{\Delta}\right) \circ \left(\frac{\mathrm{Prod}_{\boldsymbol{\Delta}^{(0)}}\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right) - \mathrm{Prod}_{\boldsymbol{\Delta}^{(1)}}\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right)}{2}\right)\right]_{i_1,i_2\cdots,i_{m-1},i_m}$$

and

$$x_{i_1\,1\,i_3\cdots i_m}\,x_{i_2\,1\,i_4\cdots i_m i_1}\cdots x_{i_m\,1\,i_2\cdots i_{m-1}} =$$

$$\left[\left(\mathbf{1}_{n\times\cdots\times n} - \boldsymbol{\Delta}\right) \circ \left(\frac{\mathrm{Prod}_{\boldsymbol{\Delta}^{(1)}}\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right) - \mathrm{Prod}_{\boldsymbol{\Delta}^{(0)}}\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right)}{2}\right)\right]_{i_1,i_2\cdots,i_{m-1},i_m}.$$

Consequently the system has at most $2\left(\left\lceil\frac{2^m-2}{m}\right\rceil + 2\right)$ distinct equations in the 2^m variables associated with entries of $\mathbf{X}$. Taking these two observations together we conclude that the solution $\mathbf{X}$ obtained via Gauss-Jordan elimination to the orthogonalization procedure is made up of rows whose of two entries correspond to monomials in some free variables both of which are scaled by a common radical expression which depends only on the entries of $\mathbf{A}$. Finally, the desired hypermatrix is obtained by assigning to all free variables the value 1 and factoring out the common factor of each row.

3.3 Direct sums and Kronecker products of hypermatrices

Recall from linear algebra that the direct sum and the Kronecker product of square matrices $\mathbf{A} \in \mathbb{C}^{n_0\times n_0}$, $\mathbf{B} \in \mathbb{C}^{n_1\times n_1}$ can both be defined in terms of bilinear forms. For notational convenience we express here multilinear forms as general BM products.

$$\mathrm{Prod}_{\mathbf{A}\oplus\mathbf{B}}\left(\begin{pmatrix}\mathbf{x}_1\\ \mathbf{y}_1\end{pmatrix}^{T^1}, \begin{pmatrix}\mathbf{x}_0\\ \mathbf{y}_0\end{pmatrix}^{T^0}\right) := \mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}_1^{T^1}, \mathbf{x}_0^{T^0}\right) + \mathrm{Prod}_{\mathbf{B}}\left(\mathbf{y}_1^{T^1}, \mathbf{y}_0^{T^0}\right)$$

and

$$\mathrm{Prod}_{\mathbf{A}\otimes\mathbf{B}}\left(\left(\mathbf{x}_1\otimes\mathbf{y}_1\right)^{T^1}, \left(\mathbf{x}_0\otimes\mathbf{y}_0\right)^{T^0}\right) := \mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}_1^{T^1}, \mathbf{x}_0^{T^0}\right) \cdot \mathrm{Prod}_{\mathbf{B}}\left(\mathbf{y}_1^{T^1}, \mathbf{y}_0^{T^0}\right),$$

where $\{\mathbf{x}_0, \mathbf{x}_1\} \subset \mathbb{C}^{n_0\times 1}$ and $\{\mathbf{y}_0, \mathbf{y}_1\} \subset \mathbb{C}^{n_1\times 1}$. These definitions conveniently extend to cubic hypermatrices of all orders as illustrated below for third order hypermatrices:

$$\mathrm{Prod}_{\mathbf{A}\oplus\mathbf{B}}\left(\begin{pmatrix}\mathbf{x}_2\\ \mathbf{y}_2\end{pmatrix}^{T^2}, \begin{pmatrix}\mathbf{x}_1\\ \mathbf{y}_1\end{pmatrix}^{T^1}, \begin{pmatrix}\mathbf{x}_0\\ \mathbf{y}_0\end{pmatrix}^{T^0}\right) := \mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}_2^{T^2}, \mathbf{x}_1^{T^1}, \mathbf{x}_0^{T^0}\right) + \mathrm{Prod}_{\mathbf{B}}\left(\mathbf{y}_2^{T^2}, \mathbf{y}_1^{T^1}, \mathbf{y}_0^{T^0}\right),$$

and

$$\text{Prod}_{\mathbf{A}\otimes\mathbf{B}}\left(\left(\mathbf{x}_2\otimes\mathbf{y}_2\right)^{T^2},\left(\mathbf{x}_1\otimes\mathbf{y}_1\right)^{T^1},\left(\mathbf{x}_0\otimes\mathbf{y}_0\right)^{T^0}\right):=\text{Prod}_{\mathbf{A}}\left(\mathbf{x}_2^{T^2},\mathbf{x}_1^{T^1},\mathbf{x}_0^{T^0}\right)\cdot\text{Prod}_{\mathbf{B}}\left(\mathbf{y}_2^{T^2},\mathbf{y}_1^{T^1},\mathbf{y}_0^{T^0}\right),$$

where $\{\mathbf{x}_0,\,\mathbf{x}_1,\,\mathbf{x}_2\}\subset\mathbb{C}^{n_0\times1\times1}$ and $\{\mathbf{y}_0,\,\mathbf{y}_1,\,\mathbf{y}_2\}\subset\mathbb{C}^{n_1\times1\times1}$.

Lemma 3 : For any two arbitrary uncorrelated m-tuples of hypermatrices $\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right)$ and $\left(\mathbf{B}^{(1)},\cdots,\mathbf{B}^{(m)}\right)$ the following m-tuples

$$\left(\mathbf{A}^{(1)}\oplus\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(k)}\oplus\mathbf{B}^{(k)},\cdots,\mathbf{A}^{(m)}\oplus\mathbf{B}^{(m)}\right)$$

and

$$\left(\mathbf{A}^{(1)}\otimes\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(k)}\otimes\mathbf{B}^{(k)},\cdots,\mathbf{A}^{(m)}\otimes\mathbf{B}^{(m)}\right)$$

also form uncorrelated hypermatrix tuples.

Proof: The fact that the m-tuple of hypermatrices $\left(\mathbf{A}^{(1)}\oplus\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(k)}\oplus\mathbf{B}^{(k)},\cdots,\mathbf{A}^{(m)}\oplus\mathbf{B}^{(m)}\right)$ forms an uncorrelated tuple (assuming that the m-tuples $\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}\right)$ and $\left(\mathbf{B}^{(1)},\cdots,\mathbf{B}^{(m)}\right)$ form uncorrelated m-tuples) follows from the fact that the BM product is well behaved relative to conformable block hypermatrix partitions. We express this fact as follows:

$$\left[\text{Prod}\left(\mathbf{U}^{(1)},\,\cdots,\mathbf{U}^{(m)}\right)\right]_{i_1,\cdots,i_m}=\sum_{0\le j<k}\text{Prod}\left(\mathbf{U}^{(1)}_{i_1\,j\,i_3\cdots i_m}\cdots\,\mathbf{U}^{(t)}_{i_1\cdots i_t\,j\,i_{t+2}\cdots i_m}\cdots\,\mathbf{U}^{(m)}_{j\,i_2\cdots i_m}\right),$$

where $\left\{\mathbf{U}^{(t)}_{i_1\,i_1\,i_3\cdots i_m}\right\}_{i,j,k,t}$ denotes the conformable block partitions of the hypermatrices $\left\{\mathbf{U}^{(t)}\right\}_{1\le t\le m}$. Finally, the fact that the m-tuple of hypermatrices

$$\left(\mathbf{A}^{(1)}\otimes\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(k)}\otimes\mathbf{B}^{(k)},\cdots,\mathbf{A}^{(m)}\otimes\mathbf{B}^{(m)}\right)$$

also forms an uncorrelated m-tuple follows from the easily verifiable BM-product identity

$$\text{Prod}\left(\mathbf{A}^{(1)}\otimes\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(k)}\otimes\mathbf{B}^{(k)},\cdots,\mathbf{A}^{(m)}\otimes\mathbf{B}^{(m)}\right)=$$

$$\text{Prod}\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(k)},\cdots,\mathbf{A}^{(m)}\right)\otimes\text{Prod}\left(\mathbf{B}^{(1)},\cdots,\mathbf{B}^{(k)},\cdots,\mathbf{B}^{(m)}\right)\tag{9}$$

The identity 9 extends to hypermatrices the classical matrix identity

$$\left(\mathbf{A}^{(1)}\otimes\mathbf{B}^{(1)}\right)\cdot\left(\mathbf{A}^{(2)}\otimes\mathbf{B}^{(2)}\right)=\left(\mathbf{A}^{(1)}\cdot\mathbf{A}^{(2)}\right)\otimes\left(\mathbf{B}^{(1)}\cdot\mathbf{B}^{(2)}\right).$$

3.4 From matrix transformations to hypermatrix transformations

Recall from linear algebra that we associate with some matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ a matrix transformation acting on $\mathbb{C}^{n\times 1}$ defined by

$$\mathbf{A} \cdot \mathbf{x}$$

In order to extend to hypermatrices the notion of transformation acting on a vector space, we express matrix transformations as follows:

$$\mathcal{T}_{\mathbf{A}^T,\mathbf{A}} : \mathbb{C}^{n\times 1} \to \mathbb{C}^{n\times 1}$$

$$\mathbf{y} = \mathcal{T}_{\mathbf{A}^T,\mathbf{A}}(\mathbf{x}) \Leftrightarrow \forall\, 0 \le k < n,\ y_k = \sqrt{\mathrm{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^T, \mathbf{x}\right)},$$

where $\mathbf{P}_k = \mathrm{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{A}^T, \mathbf{A}\right)$. Consequently, up to sign,

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x}.$$

(That is, this equation holds if we identify two complex numbers differing only by sign.)
Note that linear transformations such as $\mathcal{T}_{\mathbf{A}^T,\mathbf{A}}$ are special cases of an equivalence classes of non-linear transforms associated with an arbitrary pair of $n \times n$ matrices $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ defined by

$$\mathcal{T}_{\mathbf{A}^{(1)},\mathbf{A}^{(2)}} : \mathbb{C}^{n\times 1} \to \mathbb{C}^{n\times 1}$$

$$\mathbf{y} = \mathcal{T}_{\mathbf{A}^{(1)},\mathbf{A}^{(2)}}(\mathbf{x}) \Leftrightarrow \forall\, 0 \le k < n, \quad y_k = \sqrt{\mathrm{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^T, \mathbf{x}\right)},$$

where $\mathbf{P}_k = \mathrm{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}\right)$. Such equivalence classes of transformations naturally extend to hypermatrices and are motivated by the general Parseval identity. We define for an arbitrary m-tuple of order m hypermatrices $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$ the equivalence class of transforms $\mathcal{T}_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}}$ whose action on the vector space $\mathbb{C}^{n\times 1\times\cdots\times 1}$ is defined by

$$\mathcal{T}_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}} : \mathbb{C}^{n\times 1\times\cdots\times 1} \to \mathbb{C}^{n\times 1\times\cdots\times 1}$$

such that

$$\mathbf{y} = \mathcal{T}_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}}(\mathbf{x})$$

$$\Leftrightarrow$$

$$\forall\, 0 \le k < n, \quad y_k = \sqrt[m]{\mathrm{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^{(m-1)}}, \mathbf{x}^{T^{(m-2)}}, \cdots, \mathbf{x}^{T^j}, \cdots, \mathbf{x}^{T^1}, \mathbf{x}^{T^0}\right)}$$

where $\mathbf{P}_k = \mathrm{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$. The equivalence class of transforms associated with m-th order hypermatrices is defined modulo multiplication of the each entry of the image vector $\mathbf{y}$ with an arbitrary m-th root of unity.

3.5 Hypermatrix Fourier transforms

Hypermatrix transforms also motivate a natural generalization of Fourier transforms. To emphasize the analogy between the hypermatrix Fourier transform and the matrix Fourier transform we briefly recall here a matrix variant of the Fourier transform. Given an inverse pair of $n \times n$ matrices (i.e. an uncorrelated pair of second order hypermatrices $\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}\right)$) their induced Fourier transform, denoted $\mathcal{T}_{\mathbf{A}^{(1)},\mathbf{A}^{(2)}}$, is defined as the map acting on the vector space $\mathbb{C}^{n\times 1}$ defined by

$$\mathcal{T}_{\mathbf{A}^{(1)},\mathbf{A}^{(2)}} : \mathbb{C}^{n\times 1} \rightarrow \mathbb{C}^{n\times 1}$$

such that

$$\mathbf{y} = \mathcal{T}_{\mathbf{A}^{(1)},\mathbf{A}^{(2)}}\left(\mathbf{x}\right) \Leftrightarrow \forall\, 0 \le k < n, \quad y_k = \sqrt{\text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^T, \mathbf{x}\right)},$$

where $\mathbf{P}_k = \text{Prod}_{\mathbf{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}\right)$. Although different choices of branches for the square root induce different transforms we consider all such transforms to belong to the equivalence class of transforms for which

$$\forall\, 0 \le k < n,\ \left(y_k\right)^2 = \text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^T, \mathbf{x}\right).$$

In linear algebra terms, we say that such maps are equivalent up to multiplication of the image vector $\mathbf{y}$ by a diagonal matrix whose diagonal entries are either -1 or 1. Furthermore, by Parseval's identity we know that the transform $\mathcal{T}_{\mathbf{A}^{(1)},\mathbf{A}^{(2)}}$ preserves the sum of squares of entries of the pre-image $\mathbf{x}$:

$$\mathbf{y} = \mathcal{T}_{\mathbf{A}^{(1)},\mathbf{A}^{(2)}}\left(\mathbf{x}\right) \Leftrightarrow \text{Prod}\left(\mathbf{y}^T, \mathbf{y}\right) = \text{Prod}\left(\mathbf{x}^T, \mathbf{x}\right)$$

Similarly, we associate with some arbitrary uncorrelated m-tuples of hypermatrices $\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$, each of order m and having side length n, a *hypermatrix Fourier transform* denoted $\mathcal{T}_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}}$ whose action on the vector space $\mathbb{C}^{n\times 1\times\cdots\times 1}$ is defined by

$$\mathcal{T}_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}} : \mathbb{C}^{n\times 1\times\cdots\times 1} \rightarrow \mathbb{C}^{n\times 1\times\cdots\times 1}$$

such that

$$\mathbf{y} = \mathcal{T}_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}}\left(\mathbf{x}\right)$$

$$\Leftrightarrow$$

$$\forall\, 0 \le k < n, \quad y_k = \sqrt[m]{\text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^{(m-1)}}, \mathbf{x}^{T^{(m-2)}}, \cdots, \mathbf{x}^{T^j}, \cdots, \mathbf{x}^{T^1}, \mathbf{x}^{T^0}\right)},$$

where $\mathbf{P}_k = \text{Prod}_{\mathbf{\Delta}^{(k)}}\left(\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}\right)$. Although different choices of branches for the m-th root induce different transforms, we consider all such transforms to belong to the equivalence class of transforms for which

$$\forall\, 0 \le k < n, \quad \left(y_k\right)^m = \text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^{(m-1)}}, \mathbf{x}^{T^{(m-2)}}, \cdots, \mathbf{x}^{T^j}, \cdots, \mathbf{x}^{T^1}, \mathbf{x}^{T^0}\right).$$

These transform are equivalent up to multiplication of each entry of the image vector $\mathbf{y}$ with an arbitrary m-th root of unity. By Proposition 1 it follows that the proposed transform preserves the sum of m-th powers of entries of $\mathbf{x}$:

$$\mathbf{y} = \mathcal{T}_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(m)}}\left(\mathbf{x}\right)$$

$$\Leftrightarrow$$

$$\mathrm{Prod}\left(\mathbf{x}^{T^{(m-1)}},\mathbf{x}^{T^{(m-2)}},\cdots,\mathbf{x}^{T^{j}},\cdots,\mathbf{x}^{T^{1}},\mathbf{x}^{T^{0}}\right)=\mathrm{Prod}\left(\mathbf{y}^{T^{(m-1)}},\mathbf{y}^{T^{(m-2)}},\cdots,\mathbf{y}^{T^{j}},\cdots,\mathbf{y}^{T^{1}},\mathbf{y}^{T^{0}}\right).$$

3.6 Third order DFT hypermatrices

We recall from matrix algebra that the inverse matrix pairs associated with the Discrete Fourier Transform (DFT) acting on the vector space $\mathbb{C}^{n\times 1}$ corresponds to $\mathcal{T}_{\mathbf{F},\overline{\mathbf{F}}^T}$ where the entries of the $n \times n$ matrix $\mathbf{F}$ are given by

$$[\mathbf{F}]_{u,v}=\frac{1}{\sqrt{n}}\exp\left\{i\,\frac{2\pi}{n}\,u\,v\right\}.$$

The definition crucially relies on the following geometric sum identity valid for every non zero integer n

$$\left(\frac{1}{n}\sum_{0\le t<n}\exp\left\{i\,\frac{2\pi}{n}\,u\,t-i\,\frac{2\pi}{n}\,t\,v\right\}\right)=\begin{cases}1 & \text{if } 0\le u=v<n\\ 0 & \text{otherwise}\end{cases}.$$

$$\Leftrightarrow$$

$$\sum_{0\le t<n}\left(\frac{\exp\left\{i\,\frac{2\pi}{n}\,u\,t\right\}}{\sqrt{n}}\right)\left(\frac{\exp\left\{-i\,\frac{2\pi}{n}\,t\,v\right\}}{\sqrt{n}}\right)=\begin{cases}1 & \text{if } 0\le u=v<n\\ 0 & \text{otherwise}\end{cases}. \tag{10}$$

The equality above expresses the fact that the $n \times n$ matrices $\mathbf{F}$ and $\overline{\mathbf{F}}^T$ are in fact inverse pairs (i.e. uncorrelated pair of second order hypermatrices). We therefore understand the DFT to be a special Fourier transform, in which the entries of the inverse matrix pairs are all roots of unity scaled by the normalizing factor $1/\sqrt{n}$. By Lemma 3 if $\mathcal{T}_{\mathbf{F},\overline{\mathbf{F}}^T}$ is a DFT then for every integer $k > 1$ the Fourier transform $\mathcal{T}_{\mathbf{F}^{\otimes^k},\overline{\mathbf{F}^{\otimes^k}}^T}$ is a also a DFT. Recall that $\otimes^k$ means k repeated Kronecker products.

There is a third order hypermatrix identity similar to the identity in 10, which is unfortunately only valid for appropriately chosen values of n. The third order DFT hypermatrix identity crucially relies on the following geometric sum identity

$$\left(\frac{1}{n}\sum_{0\le t<n}\exp\left\{i\frac{2\pi}{n}\left(u\sqrt{t}-\sqrt{t}\,w\right)^2+i\frac{2\pi}{n}\left(u\sqrt{t}-v\sqrt{t}\right)^2+i\frac{2\pi}{n}\left(\sqrt{t}\,v-\sqrt{t}\,w\right)^2\right\}\right)=$$

$$\begin{cases}1 & \text{if } 0\le u=v=w<n\\ 0 & \text{otherwise}\end{cases},$$

for appropriately chosen values of n. The identity above can be rewritten as

$$\left(\sum_{0\le t<n}\frac{\exp\left\{i\frac{2\pi}{n}\left(u\sqrt{t}-\sqrt{t}\,w\right)^2\right\}}{\sqrt[3]{n}}\,\frac{\exp\left\{i\frac{2\pi}{n}\left(u\sqrt{t}-v\sqrt{t}\right)^2\right\}}{\sqrt[3]{n}}\,\frac{\exp\left\{i\frac{2\pi}{n}\left(\sqrt{t}\,v-\sqrt{t}\,w\right)^2\right\}}{\sqrt[3]{n}}\right)=$$

$$\begin{cases} 1 & \text{if } 0 \leq u = v = w < n \\ 0 & \text{otherwise} \end{cases}.$$

The identity above expresses a BM product of the uncorrelated triple ($\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$). The entries of $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ are all n-th roots of unity scaled by the same normalizing factor $1/\sqrt[3]{n}$. The entries of $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ are thus given by

$$[\mathbf{F}]_{u,t,w} = \frac{\exp\left\{i\frac{2\pi}{n}t\,(u-w)^2\right\}}{\sqrt[3]{n}},\ [\mathbf{G}]_{u,v,t} = \frac{\exp\left\{i\frac{2\pi}{n}t\,(u-v)^2\right\}}{\sqrt[3]{n}},\ [\mathbf{H}]_{t,v,w} = \frac{\exp\left\{i\frac{2\pi}{n}t\,(v-w)^2\right\}}{\sqrt[3]{n}}. \tag{11}$$

As a result the transform $\mathcal{T}_{\mathbf{F},\mathbf{G},\mathbf{H}}$ is a hypermatrix DFT acting on the vector space $\mathbb{C}^{n\times 1\times 1}$. The smallest possible choice for the integer n is 5. By Lemma 3, if $\mathcal{T}_{\mathbf{F},\mathbf{G},\mathbf{H}}$ is a DFT over $\mathbb{C}^{n\times 1\times 1}$ then for every positive integer $k > 1$, $\mathcal{T}_{\mathbf{F}^{\otimes k},\mathbf{G}^{\otimes k},\mathbf{H}^{\otimes k}}$ is also a DFT over the vector space $\mathbb{C}^{n^k\times 1\times 1}$.
The following proposition determines necessary and sufficient condition on the positive integer n which ensures that the hypermatrices in 11 are uncorrelated.

Proposition 4 : The $n \times n \times n$ hypermatrices $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ whose entries are specified by

$$[\mathbf{F}]_{u,t,w} = \frac{\exp\left\{i\frac{2\pi}{n}t\,(u-w)^2\right\}}{\sqrt[3]{n}},\ [\mathbf{G}]_{u,v,t} = \frac{\exp\left\{i\frac{2\pi}{n}t\,(u-v)^2\right\}}{\sqrt[3]{n}},\ [\mathbf{H}]_{t,v,w} = \frac{\exp\left\{i\frac{2\pi}{n}t\,(v-w)^2\right\}}{\sqrt[3]{n}}.$$

form an uncorrelated triple if and only if the equation

$$x^2 + 3\,y^2 \equiv 0 \mod n,$$

admits no solution other then the trivial solution $x \equiv 0 \mod n$ and $y \equiv 0 \mod n$.

Proof: The construction requires the following implication

$$\forall\, 0 \leq u, v, w < n, \quad u\left(e^{i\frac{2\pi}{3}}\right)^2 + v\left(e^{i\frac{2\pi}{3}}\right)^1 + w\left(e^{i\frac{2\pi}{3}}\right)^0 \neq 0 \Rightarrow (u-v)^2 + (v-w)^2 + (u-w)^2 \neq 0 \mod n.$$

Let $x = u - v$ and $y = v - w$, the implication becomes

$$\forall\, x, y \in \mathbb{N}, \quad x^2 + y^2 + (x+y)^2 \neq 0 \mod n.$$
$$\Rightarrow \forall\, x, y \in \mathbb{N}, \quad 2\left(x^2 + xy + y^2\right) \neq 0 \mod n.$$

If n is even then the choice $x = \frac{n}{2}$ and $y = 0$ always constitutes a counterexample. However if n is odd the constraints may be stated as follows: for all integers x, y not both zero mod n we require that

$$x^2 + xy + y^2 \neq 0 \mod n.$$
$$\Rightarrow \left(x + \frac{y}{2}\right)^2 + 3\left(\frac{y}{2}\right)^2 \neq 0 \mod n.$$

from which the sought after result follows.
In particular, when n is prime we need -3 to be a quadratic non-residue modulo n. An easy calculation shows that the primes of the forms $12m + 5$ and $12m + 11$ satisfy these conditions, and in particular there are infinitely many such n. We leave the case of composite n to the reader.

3.7 Hadamard hypermatrices

We discuss here Hadamard hypermatrices which are used to construct special DFT hypermatrices which have real entries. In fact we extend to hypermatrices Sylvester's classical Hadamard matrix construction. Recall from linear algebra that a matrix $\mathbf{H} \in \{-1,1\}^{n\times n}$ is a Hadamard matrix if

$$\left[\mathbf{H}\cdot\mathbf{H}^T\right]_{i,j} = \begin{cases} n & \text{if } 0 \le i = j < n \\ 0 & \text{otherwise} \end{cases}. \tag{12}$$

Hadamard matrices are of considerable importance in topics relating to combinatorial design and the analysis of boolean functions. They are also used to define the famous Hadamard–Rademacher–Walsh transform which plays an important role in signal processing. Hadamard matrices are also common occurrences in practical implementations of the Fast Fourier Transform. Furthermore, Hadamard matrices are well-known to be optimal matrices relative to the Hadamard determinant inequality

$$|\det \boldsymbol{\Theta}| \le \left(\sqrt{n}\right)^n,$$

valid over the set of all $n \times n$ matrices $\boldsymbol{\Theta}$ whose entries of are bounded in absolute value by 1. Equality is achieved in Hadamard's determinant inequality for Hadamard matrices. In 1867, James Joseph Sylvester proposed the classical construction of an infinite family of Hadamard matrices of size $2^n \times 2^n$ for any integer $n \ge 1$. Sylvester's construction starts with the 2×2 matrix

$$\left(\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right)$$

and considers the sequence of matrices

$$\left\{\left(\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right)^{\otimes^n} \in \mathbb{R}^{2^n\times 2^n}\right\}_{1<n<\infty}.$$

By Lemma 3 we know that every matrix in the sequence above will satisfy the Hadamard criterion 12. Having defined in Section 3.1 orthogonal hypermatrices, it is relatively straightforward to extend the Hadamard criterion 12 to hypermatrices of arbitrary orders, which can be used to extend to hypermatrices the Hadamard–Rademacher–Walsh transform. Formally, an order m hypermatrix $\mathbf{H} \in \{-1,1\}^{n\times\cdots\times n}$ is Hadamard if

$$\left[\text{Prod}\left(\mathbf{H}, \mathbf{H}^{T^{(m-1)}}, \cdots, \mathbf{H}^{T^k}, \cdots, \mathbf{H}^{T^2}, \mathbf{H}^T\right)\right]_{i_1,\cdots,i_m} = \begin{cases} n & \text{if } 0 \le i_1 = \cdots = i_m < n \\ 0 & \text{otherwise} \end{cases}. \tag{13}$$

The following theorem extends the scope of both Sylvester's constructions and the famous Hadamard matrix conjecture.

Theorem 5 : For every positive integer $n \ge 1$ and every positive integer m which is either odd or equal to 2, there exists an order m Hadamard hypermatrix of side length 2^n. In contrast, if m is an even integer larger than 2, then there is no order m Hadamard hypermatrix of side length 2.

Proof : By Lemma 3, it suffices to provide an explicit construction for odd order Hadamard hypermatrices of side length 2. For side length 2 hypermatrices of order $m > 2$, the Hadamard criterion 13 is expressed as follows

$$\forall \left(i_1, \cdots, i_m\right) \notin \left\{\left(0, 0, \cdots, 0\right), \left(1, 1, \cdots, 1\right)\right\}$$

$$\left(h_{i_1\,0\,i_3\cdots i_m}\, h_{i_2\,0\,i_4\cdots i_m i_1} \cdots h_{i_m\,0\,i_2\cdots i_{m-1}}\right) + \left(h_{i_1\,1\,i_3\cdots i_m}\, h_{i_2\,1\,i_4\cdots i_m i_1} \cdots h_{i_m\,1\,i_2\cdots i_{m-1}}\right) = 0$$

and

$$\forall\, 0 \le i < 2, \quad \left(h_{i\,0\,i\cdots i\,i}\right)^m + \left(h_{i\,1\,i\cdots i\,i}\right)^m = 2.$$

We can express the first set of constraints equivalently as

$$\forall \left(i_1, \cdots, i_m\right) \notin \left\{\left(0, 0, \cdots, 0\right), \left(1, 1, \cdots, 1\right)\right\}$$

$$\left(h_{i_1\,0\,i_3\cdots i_m}\, h_{i_2\,0\,i_4\cdots i_m i_1} \cdots h_{i_m\,0\,i_2\cdots i_{m-1}}\right) / \left(h_{i_1\,1\,i_3\cdots i_m}\, h_{i_2\,1\,i_4\cdots i_m i_1} \cdots h_{i_m\,1\,i_2\cdots i_{m-1}}\right) = -1$$

For ± 1 solutions, the second set of constraints just states that

$$\forall\, 0 \le i < 2, \quad \left(h_{i\,0\,i\cdots i\,i}\right)^m = \left(h_{i\,1\,i\cdots i\,i}\right)^m = 1.$$

For all $j_1, j_2, \ldots, j_{m-1}$,define $H_{j_1\,j_2\cdots j_{m-1}} = h_{j_1\,0\,j_2\cdots j_{m-1}} / h_{j_1\,1\,j_2\cdots j_{m-1}}$. The first set of constraints simplifies to

$$H_{i_1 i_2\cdots i_{m-1}} H_{i_2 i_3\cdots i_m} \cdots H_{i_m i_1\cdots i_{m-2}} = -1 \qquad \forall \left(i_1, \cdots, i_m\right) \notin \left\{\left(0, 0, \cdots, 0\right), \left(1, 1, \cdots, 1\right)\right\}$$

The second set of constraints states that

$$\forall\, 0 \le i < 2, \quad \left(H_{i\,i\cdots i\,i}\right)^m = 1.$$

Clearly the original constraints (in h) have a ± 1 solution iff the new constraints (in H) have a ± 1 solution.

We now show that if $m > 2$ is even then there are no solutions. Let $m = 2k$, and consider the constraint corresponding to $i_1 = 1, i_2 = \cdots = i_k = 0$,$i_{k+1} = 1, i_{k+2} = \cdots = i_m = 0$. This constraint states that

$$H^2_{10\cdots 0} H^2_{0\cdots 0} H^2_{0\cdots 01} \cdots H^2_{010\cdots 0} = -1,$$

which clearly has no ± 1 solution.

From now on, assume that $m > 1$ is odd. We immediately get that

$$H_{0\cdots 0} = H_{1\cdots 1} = 1.$$

Call a binary word of length m a *necklace* if it is lexicographically smaller than all its rotations. Since rotations of a word $i_1 \cdots i_m$ correspond to the same constraint, it is enough to consider constraints corresponding to necklaces. For each word $i_1 \cdots i_m$, a *window* consists of $m-1$ contiguous characters (where contiguity is cyclic). Thus there are m windows, some of which could be the same.

If a necklace is periodic with minimal period p, then each window will appear (at least) m/p times. The following lemma shows that periodicity is the only reason that a window repeats.

Lemma 5a : Suppose that a word $w_0 \ldots w_{m-1}$ satisfies $w_i = w_{i-p}$ for $i = 1, \ldots, m-1$ (but not necessarily for $i = 0$). Then w has a period π (possibly m) such that p is a multiple of π.

Proof of Lemma 5a : The proof is by induction on m. We can assume $0 \le p < m$. If $m = 1$ then there is nothing to prove. If p divides m then the constraints imply that p is a period of w, so again there is nothing to prove. Suppose therefore that $q = m \mod p > 0$. The constraints imply that w has the form

$$w_0 w_1 \ldots w_{p-1} w_0 w_1 \ldots w_{p-1} \cdots w_0 w_1 \ldots w_{q-1},$$

and furthermore $w_1 = w_{q+1}, \ldots, w_{p-1} = w_{p-1+q \mod p}$. That is, the word $w_0 \ldots w_{p-1}$satisfies the premise of the lemma with the shift q. By induction, $w_0 \ldots w_{p-1}$ has period π (which thus divides p) and q is a multiple of π. It follows that π divides m and so is a period of $w_0, \ldots, w_{m-1}$. This completes the proof of the lemma.

The lemma 5a implies that indeed if a necklace has minimal period p (possibly $p = m$) then each window appears m/p times, and so an odd number of times. We can thus restate the constraints as follows, for ± 1 solutions:

For each non-constant necklace $i_1 \ldots i_m$, the product of H-values corresponding to distinct non-constant windows of $i_1 \ldots i_m$ equals -1.

As an example, for $m = 5$ the non-constant necklaces are $00001, 00011, 00101, 00111, 01011, 01111$, and the corresponding constraints are

$$H_{0001} H_{0010} H_{0100} H_{1000} = -1$$

$$H_{0001} H_{0011} H_{0110} H_{1100} H_{1000} = -1$$

$$H_{0010} H_{0101} H_{1010} H_{0100} H_{1001} = -1$$

$$H_{0011} H_{0111} H_{1110} H_{1100} H_{1001} = -1$$

$$H_{0101}H_{1011}H_{0110}H_{1101}H_{1010} = -1$$

$$H_{0111}H_{1110}H_{1101}H_{1011} = -1$$

Consider now the graph whose vertex set consists of all non-constant necklaces, and edges connect two necklaces x, y if some rotations of x, y have Hamming distance 1. For example, 00101 and 01011 are connected since 01010 and 01011 differ in only one position. It is not hard to check that each non-constant window appears in exactly two constraints (corresponding to its two completions), and these constraints correspond to an edge. Continuing our example, the window 0101 appears in the constraints corresponding to the necklaces $00101, 01011$, and only there. (An edge can correspond to several windows: for example $(00001, 00011)$ corresponds to both 0001 and 1000). We will show that this graph contains a sub-graph in which all degrees are odd. If we set to -1 all variables corresponding to the chosen edges (one window per edge) and set to 1 all the other variables, then we obtain a solution to the set of constraints.

For example, the edges $\{(00001, 00011), (00101, 00111), (01011, 01111)\}$ constitute a matching in the graph, and so setting $H_{0001} = H_{1001} = H_{1011} = -1$ and setting all other variables to 1 yields a solution.

A well-known result states that a connected graph contains a sub-graph in which all degrees are odd iff it has an even number of vertices[3]. To complete the proof, it thus suffices to show that the number of necklaces (and so non-constant necklaces) is even. The classical formula for the number of necklaces (obtainable using the orbit-stabilizer theorem) states that the number of binary necklaces of length m is

$$\frac{1}{m}\sum_{k|m}\varphi(k)2^{m/k}.$$

Here φ is Euler's function. Since m is odd, it suffices to show that all summands are even. This is clear for all summands with $k < m$. When $k = m$, we use the easy fact that $\varphi(m)$ is even for all $m > 2$, which follows from the explicit formula for $\varphi(m)$ in terms of the factorization of m. This completes the proof of Theorem 5.

We close this section with an explicit example of a $2 \times 2 \times 2$ Hadamard hypermatrix:

$$\mathbf{H}[:,:,0] = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \mathbf{H}[:,:,1] = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

[3] Here is a proof of the hard direction, taken from Jukna's *Extremal combinatorics*: partition the graph into a list of pairs, and choose a path connecting each pair. Now take the XOR of all these paths.

4 Spectral decomposition of Kronecker products and direct sums of side length 2 hypermatrices

We describe here elementary methods for deriving generators for the spectral elimination ideals, which will be defined here. The method described here for computing generators for the first spectral elimination ideal applies to matrices of all sizes but is unfortunately restricted to side length 2 hypermatrices of order greater than 2.

4.1 The matrix case

We start by describing the derivation of generators for the *spectral elimination ideal* which we now define. Let $\mathbf{A} \in \mathbb{C}^{2\times 2}$ having distinct eigenvalues. Recall that the spectral decomposition equation is given by

$$\begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} = \left[\begin{pmatrix} u_{00} & u_{01} \\ u_{10} & u_{11} \end{pmatrix} \cdot \begin{pmatrix} \mu_0 & 0 \\ 0 & \mu_1 \end{pmatrix}\right] \cdot \left[\frac{\begin{pmatrix} u_{11} & -u_{10} \\ -u_{01} & u_{00} \end{pmatrix}}{u_{00}u_{11} - u_{01}u_{10}} \cdot \begin{pmatrix} \nu_0 & 0 \\ 0 & \nu_1 \end{pmatrix}\right]^T \tag{14}$$

The spectral constraints yield generators for the polynomial ideal $\mathcal{I}_{\mathbf{A}}$ in the polynomial ring $\mathbb{C}\left[u_{00}, u_{01}, u_{10}, u_{11}, \frac{u_{00}}{u_{00}u_{11}-u_{01}u_{10}}, \frac{u_{01}}{u_{00}u_{11}-u_{01}u_{10}}, \frac{u_{10}}{u_{00}u_{11}-u_{01}u_{10}}, \frac{u_{11}}{u_{00}u_{11}-u_{01}u_{10}}, \mu_0, \mu_1, \nu_0, \nu_1\right]$. The spectral elimination ideal is defined as

$$\mathcal{I}_{\mathbf{A}} \cap \mathbb{C}\left[\mu_0, \mu_1, \nu_0, \nu_1\right]$$

The spectral decomposition constraints can thus be rewritten as

$$\left(\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ \mu_0\nu_0 & \mu_1\nu_1 \end{pmatrix}\right) \cdot \begin{pmatrix} \frac{u_{00}\cdot u_{11}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{01}\cdot u_{10}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{00}\cdot u_{01}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{u_{01}\cdot u_{00}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{u_{10}\cdot u_{11}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{11}\cdot u_{10}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{10}\cdot u_{01}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{u_{11}\cdot u_{00}}{u_{00}u_{11}-u_{01}u_{10}} \end{pmatrix} = \begin{pmatrix} 1 \\ a_{00} \\ 0 \\ a_{01} \\ 0 \\ a_{10} \\ 1 \\ a_{11} \end{pmatrix},$$

from which it follows that

$$\begin{pmatrix} \frac{u_{00}\cdot u_{11}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{01}\cdot u_{10}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{00}\cdot u_{01}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{u_{01}\cdot u_{00}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{u_{10}\cdot u_{11}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{11}\cdot u_{10}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{-u_{10}\cdot u_{01}}{u_{00}u_{11}-u_{01}u_{10}} \\ \frac{u_{11}\cdot u_{00}}{u_{00}u_{11}-u_{01}u_{10}} \end{pmatrix} = \left(\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ \mu_0\nu_0 & \mu_1\nu_1 \end{pmatrix}^{-1}\right) \cdot \begin{pmatrix} 1 \\ a_{00} \\ 0 \\ a_{01} \\ 0 \\ a_{10} \\ 1 \\ a_{11} \end{pmatrix}.$$

Consequently, the entries of the vectors $\left(\begin{array}{c}\frac{u_{00}u_{11}}{u_{00}u_{11}-u_{01}u_{10}}\\ \frac{-u_{01}u_{10}}{u_{00}u_{11}-u_{01}u_{10}}\end{array}\right)$, $\left(\begin{array}{c}\frac{-u_{00}u_{01}}{u_{00}u_{11}-u_{01}u_{10}}\\ \frac{u_{01}u_{00}}{u_{00}u_{11}-u_{01}u_{10}}\end{array}\right)$, $\left(\begin{array}{c}\frac{u_{10}u_{11}}{u_{00}u_{11}-u_{01}u_{10}}\\ \frac{-u_{11}u_{10}}{u_{00}u_{11}-u_{01}u_{10}}\end{array}\right)$, $\left(\begin{array}{c}\frac{-u_{10}u_{01}}{u_{00}u_{11}-u_{01}u_{10}}\\ \frac{u_{11}u_{00}}{u_{00}u_{11}-u_{01}u_{10}}\end{array}\right)$

can be expressed as rational functions in the variables $\mu_0\nu_0$ and $\mu_1\nu_1$. The variables u_{00}, u_{01}, u_{10}, u_{11} are further eliminated via the algebraic relation

$$\left(\begin{array}{c}\frac{(u_{00}u_{11})(-u_{10}u_{01})}{(u_{00}u_{11}-u_{01}u_{10})^2}\\ \frac{(-u_{01}u_{10})(u_{11}u_{00})}{(u_{00}u_{11}-u_{01}u_{10})^2}\end{array}\right)=\left(\begin{array}{c}\frac{(-u_{00}u_{01})(u_{10}u_{11})}{(u_{00}u_{11}-u_{01}u_{10})^2}\\ \frac{(u_{01}u_{00})(-u_{11}u_{10})}{(u_{00}u_{11}-u_{01}u_{10})^2}\end{array}\right).$$

The algebraic relation above yields the characteristic polynomial

$$\left(\begin{array}{c}\frac{(\mu_1\nu_1-a_{00})(\mu_1\nu_1-a_{00})}{(\mu_1\nu_1-\mu_0\nu_0)^2}\\ \frac{(a_{00}-\mu_0\nu_0)(a_{11}-\mu_0\nu_0)}{(\mu_1\nu_1-\mu_0\nu_0)^2}\end{array}\right)=\left(\begin{array}{c}\frac{(-a_{01})(-a_{10})}{(\mu_1\nu_1-\mu_0\nu_0)^2}\\ \frac{a_{01}a_{10}}{(\mu_1\nu_1-\mu_0\nu_0)^2}\end{array}\right). \tag{15}$$

Once the determinant polynomial is derived, the generator of the spectral elimination ideal is more simply obtained by considering the polynomial

$$\det\left(\mathbf{A}-\mu\nu\,\mathbf{I}_n\right).$$

In particular, in the case $n=2$ we have

$$\det\left(\mathbf{A}-\mu\nu\,\mathbf{I}_n\right)=(\mu\nu)^2-\operatorname{Tr}\left(\mathbf{A}\right)(\mu\nu)+\det\left(\mathbf{A}\right).$$

We point out this well-known fact only to emphasize the close analogy with the hypermatrix case discussed in the next section.

Theorem 6 : Let $\mathbf{A}\in\mathbb{C}^{n\times n}$ be a matrix generated by arbitrary combinations of direct sums and Kronecker products of 2×2 matrices. Furthermore, assume that each 2×2 generator matrix admits a spectral decomposition. Then $\mathbf{A}$ admits a spectral decomposition of the form

$$\mathbf{A}=\left(\mathbf{U}\cdot\operatorname{diag}\left(\boldsymbol{\mu}\right)\right)\cdot\left(\left(\mathbf{U}^{-1}\right)^T\cdot\operatorname{diag}\left(\boldsymbol{\nu}\right)\right)^T$$

Proof: From the fact that each 2×2 generator matrix admits a spectral deccomposition, it follows that the spectral decomposition of $\mathbf{A}$ is obtained from the spectral decomposition of the generator matrices by repeated use of Lemma 3.

4.2 The hypermatrix case

The spectral decomposition of a hypermatrix $\mathbf{A}\in\mathbb{C}^{2\times2\times2}$ is expressed in terms of uncorrelated an triple $\mathbf{U}$,$\mathbf{V}$,$\mathbf{W}$ The $2\times1\times2$ hypermatrix column slices $\{\mathbf{U}[:,k,:]\,,\ \mathbf{V}[:,k,:]\,,\ \mathbf{W}[:,k,:]\}_{0\le k<2}$ collect the "*eigenmatrices*" of $\mathbf{A}$. We recall from [GER11] that the spectral decomposition is expressed as

$$\mathbf{A}=\operatorname{Prod}\left(\operatorname{Prod}\left(\mathbf{U},\mathbf{D}_0,\mathbf{D}_0^T\right),\operatorname{Prod}\left(\mathbf{V},\mathbf{D}_1,\mathbf{D}_1^T\right)^{T^2},\operatorname{Prod}\left(\mathbf{W},\mathbf{D}_2,\mathbf{D}_2^T\right)^T\right), \tag{16}$$

where the $2 \times 2 \times 2$ hypermatrices $\mathbf{D}_0$, $\mathbf{D}_1$, and $\mathbf{D}_2$ are third-order analogs of the diagonal matrices

$$\begin{pmatrix} \mu_0 & 0 \\ 0 & \mu_1 \end{pmatrix}, \begin{pmatrix} \nu_0 & 0 \\ 0 & \nu_1 \end{pmatrix}$$

used in 14. The entries of the hypermatrices $\mathbf{D}_0$, $\mathbf{D}_1$, and $\mathbf{D}_2$ are respectively given by

$$\mathbf{D}_0\left[:,:,0\right] = \begin{pmatrix} \mu_{00} & 0 \\ \mu_{01} & 0 \end{pmatrix}, \ \mathbf{D}_0\left[:,:,1\right] = \begin{pmatrix} 0 & \mu_{01} \\ 0 & \mu_{11} \end{pmatrix},$$

$$\mathbf{D}_1\left[:,:,0\right] = \begin{pmatrix} \nu_{00} & 0 \\ \nu_{01} & 0 \end{pmatrix}, \ \mathbf{D}_1\left[:,:,1\right] = \begin{pmatrix} 0 & \nu_{01} \\ 0 & \nu_{11} \end{pmatrix},$$

$$\mathbf{D}_2\left[:,:,0\right] = \begin{pmatrix} \omega_{00} & 0 \\ \omega_{01} & 0 \end{pmatrix}, \ \mathbf{D}_2\left[:,:,1\right] = \begin{pmatrix} 0 & \omega_{01} \\ 0 & \omega_{11} \end{pmatrix}.$$

The spectral constraints yield generators for the polynomial ideal $\mathcal{I}_{\mathbf{A}}$ in the polynomial ring $\mathbb{C}\left[u_{000}, \cdots, u_{111}, \ v_{000}, \cdots, v_{111}, \ w_{000}, \cdots, w_{111}, \ \mu_{00}, \mu_{01}, \mu_{11}, \ \nu_{00}, \nu_{01}, \nu_{11}, \ \omega_{00}, \omega_{01}, \omega_{11}\right]$.
By analogy to the matrix derivation, generators for the spectral elimination ideal are generators for the polynomial ideal

$$\mathcal{I}_{\mathbf{A}} \cap \mathbb{C}\left[\mu_{00}, \mu_{01}, \mu_{11}, \ \nu_{00}, \nu_{01}, \nu_{11}, \ \omega_{00}, \omega_{01}, \omega_{11}\right].$$

The generators will therefore yield expressions for the $2 \times 2 \times 2$ determinant as well as for the corresponding characteristic polynomial. We rewrite the hypermatrix spectral decomposition constraints 16 as follows:

$$\left[\bigoplus_{0 \le i,j,k < 2} \left(\mathbf{I}_2 \otimes \begin{pmatrix} 1 & 1 \\ \mu_{0i}\mu_{0k}\nu_{0j}\nu_{0i}\omega_{0k}\omega_{0j} & \mu_{i1}\mu_{k1}\nu_{j1}\nu_{i1}\omega_{k1}\omega_{j1} \end{pmatrix}\right)\right] \begin{pmatrix} u_{00k} \cdot v_{000} \cdot w_{000} \\ u_{010} \cdot v_{010} \cdot w_{010} \\ \vdots \\ u_{i0k} \cdot v_{j0i} \cdot w_{k0j} \\ u_{i1k} \cdot v_{j1i} \cdot w_{k1j} \\ \vdots \\ u_{101} \cdot v_{101} \cdot w_{101} \\ u_{111} \cdot v_{111} \cdot w_{111} \end{pmatrix} = \begin{pmatrix} 1 \\ a_{000} \\ 0 \\ a_{001} \\ 0 \\ a_{010} \\ 0 \\ a_{011} \\ 0 \\ a_{100} \\ 0 \\ a_{101} \\ 0 \\ a_{110} \\ 1 \\ a_{111} \end{pmatrix}.$$

It therefore follows from the equality above that

$$\begin{pmatrix} u_{00k}\cdot v_{000}\cdot w_{000} \\ u_{010}\cdot v_{010}\cdot w_{010} \\ \vdots \\ u_{i0k}\cdot v_{j0i}\cdot w_{k0j} \\ u_{i1k}\cdot v_{j1i}\cdot w_{k1j} \\ \vdots \\ u_{101}\cdot v_{101}\cdot w_{101} \\ u_{111}\cdot v_{111}\cdot w_{111} \end{pmatrix} = \left[\bigoplus_{0\le i,j,k<2} \left(\mathbf{I}_2 \otimes \begin{pmatrix} 1 & 1 \\ \mu_{0i}\mu_{0k}\nu_{0j}\nu_{0i}\omega_{0k}\omega_{0j} & \mu_{i1}\mu_{k1}\nu_{j1}\nu_{i1}\omega_{k1}\omega_{j1} \end{pmatrix} \right) \right]^{-1} \cdot \begin{pmatrix} 1 \\ a_{000} \\ 0 \\ a_{001} \\ 0 \\ a_{010} \\ 0 \\ a_{011} \\ 0 \\ a_{100} \\ 0 \\ a_{101} \\ 0 \\ a_{110} \\ 1 \\ a_{111} \end{pmatrix},$$

implicitly assuming that

$$0 \neq \prod_{0\le i,j,k<2} \left(\mu_{0i}\mu_{0k}\nu_{0j}\nu_{0i}\omega_{0k}\omega_{0j} - \mu_{i1}\mu_{k1}\nu_{j1}\nu_{i1}\omega_{k1}\omega_{j1} \right).$$

Consequently the entries of the vectors $\left\{ \begin{pmatrix} u_{i0k}\cdot v_{j0i}\cdot w_{k0j} \\ u_{i1k}\cdot v_{j1i}\cdot w_{k1j} \end{pmatrix} \right\}_{0\le i,j,k<2}$ are rational functions of the variables $\mu_{00}, \mu_{01}, \mu_{11},\ \nu_{00}, \nu_{01}, \nu_{11},\ \omega_{00}, \omega_{01}, \omega_{11}$. The variables $u_{000}, \cdots, u_{111},\ v_{000}, \cdots, v_{111},\ w_{000}, \cdots, w_{111}$ are thus eliminated via the relation

$$\begin{pmatrix} (u_{00k}v_{000}w_{000})\cdot(u_{001}v_{100}w_{101})\cdot(u_{101}v_{001}w_{100})\cdot(u_{100}v_{101}w_{001}) \\ (u_{01k}v_{010}w_{010})\cdot(u_{011}v_{110}w_{111})\cdot(u_{111}v_{011}w_{110})\cdot(u_{100}v_{101}w_{001}) \end{pmatrix}$$
$$=$$
$$\begin{pmatrix} (u_{001}v_{000}w_{100})\cdot(u_{000}v_{100}w_{001})\cdot(u_{100}v_{001}w_{000})\cdot(u_{101}v_{101}w_{101}) \\ (u_{011}v_{010}w_{110})\cdot(u_{010}v_{110}w_{011})\cdot(u_{110}v_{011}w_{010})\cdot(u_{101}v_{101}w_{101}) \end{pmatrix},$$

which yields the third order characterisitic polynomial

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \prod_{0\le i,j,k<2} \left(\mu_{0i}\mu_{0k}\nu_{0j}\nu_{0i}\omega_{0k}\omega_{0j} - \mu_{i1}\mu_{k1}\nu_{j1}\nu_{i1}\omega_{k1}\omega_{j1} \right)^{-2} \times$$

$$\begin{pmatrix} a_{001}a_{010}a_{100}\left(\mu_{11}\nu_{11}\omega_{11}\right)^2 - a_{011}a_{101}a_{110}\left(\mu_{01}\nu_{01}\omega_{01}\right)^2 + a_{000}a_{011}a_{101}a_{110} - a_{001}a_{010}a_{100}a_{111} \\ a_{001}a_{010}a_{100}\left(\mu_{01}\nu_{01}\omega_{01}\right)^2 - a_{011}a_{101}a_{110}\left(\mu_{00}\nu_{00}\omega_{00}\right)^2 + a_{000}a_{011}a_{101}a_{110} - a_{001}a_{010}a_{100}a_{111} \end{pmatrix}.$$

The generators for the spectral elimination ideal correspond to generators for the polynomial ideals

$$\mathcal{I}_{\mathbf{A}} \cap \mathbb{C}\left[\mu_{00}, \mu_{01},\ \nu_{00}, \nu_{01},\ \omega_{00}, \omega_{01}\right] \ \text{ and } \mathcal{I}_{\mathbf{A}} \cap \mathbb{C}\left[\mu_{00}, \mu_{01},\ \nu_{00}, \nu_{01},\ \omega_{00}, \omega_{01}\right]$$

respectively given by

$$a_{001}a_{010}a_{100}\left(\mu_{01}\nu_{01}\omega_{01}\right)^{2}-a_{011}a_{101}a_{110}\left(\mu_{00}\nu_{00}\omega_{00}\right)^{2}+a_{000}a_{011}a_{101}a_{110}-a_{001}a_{010}a_{100}a_{111}$$

and

$$a_{001}a_{010}a_{100}\left(\mu_{11}\nu_{11}\omega_{11}\right)^{2}-a_{011}a_{101}a_{110}\left(\mu_{01}\nu_{01}\omega_{01}\right)^{2}+a_{000}a_{011}a_{101}a_{110}-a_{001}a_{010}a_{100}a_{111}.$$

The derivation suggests that the $2\times2\times2$ hypermatrix characteristic polynomial is the polynomial

$$p\left(\mu_0\nu_0\omega_0,\ \mu_1\nu_1\omega_1\right)=a_{001}a_{010}a_{100}\left(\mu_1\nu_1\omega_1\right)^{2}-a_{011}a_{101}a_{110}\left(\mu_0\nu_0\omega_0\right)^{2}+\left(a_{000}a_{011}a_{101}a_{110}-a_{001}a_{010}a_{100}a_{111}\right)$$

whose constant term $a_{000}a_{011}a_{101}a_{110}-a_{001}a_{010}a_{100}a_{111}$ corresponds to the $2\times2\times2$ hypermatrix analog of the determinant polynomial. Note that the determinant polynomial is linear in the row, column and depth slices. Furthermore, the $2\times2\times2$ hyperdeterminant changes sign with a row, column or depth slice exchange.

Theorem 7 : Let $\mathbf{A}\in\mathbb{C}^{n\times n\times n}$ be a hypermatrix generated by some arbitrary combination of direct sums and Kronecker products of $2\times2\times2$ hypermatrices. Furthermore, assume that each $2\times2\times2$ generator admits a spectral decomposition. Then $\mathbf{A}$ admits a spectral decomposition of the form

$$\mathbf{A}=\text{Prod}\left(\text{Prod}\left(\mathbf{U},\mathbf{D}_0,\mathbf{D}_0^{T}\right),\text{Prod}\left(\mathbf{V},\mathbf{D}_1,\mathbf{D}_1^{T}\right)^{T^2},\text{Prod}\left(\mathbf{W},\mathbf{D}_2,\mathbf{D}_2^{T}\right)^{T}\right),$$

subject to

$$\text{Prod}\left(\mathbf{U},\mathbf{V}^{T^2},\mathbf{W}^{T}\right)=\boldsymbol{\Delta}$$

$$\left[\mathbf{D}_0\right]_{ijk}=\begin{cases}\mu_{jk}=\mu_{kj} & \text{if } 0\le i=k<n\\ 0 & \text{otherwise}\end{cases},$$

$$\left[\mathbf{D}_1\right]_{ijk}=\begin{cases}\nu_{jk}=\nu_{kj} & \text{if } 0\le i=k<n\\ 0 & \text{otherwise}\end{cases},$$

$$\left[\mathbf{D}_2\right]_{ijk}=\begin{cases}\omega_{jk}=\omega_{kj} & \text{if } 0\le i=k<n\\ 0 & \text{otherwise}\end{cases}.$$

Proof: From the fact that each $2\times2\times2$ generator admits a spectral decomposition, It follows that the spectral decomposition of $\mathbf{A}$ is derived from the spectral decomposition of the generators by repeated use of Lemma 3.

As in the matrix case, the characteristic polynomial can be obtained directly from the hyperdeterminant of the cubic hypermatrix $\mathbf{B}$ whose entries are given by

$$\left[\mathbf{B}\right]_{ijk}=a_{ijk}-\sum_{0\le t<n}\left(\mu_i u_{itk}\mu_k\right)\left(\nu_j v_{jti}\nu_i\right)\left(\omega_k w_{ktj}\omega_j\right)$$

$$\Rightarrow [\mathbf{B}]_{ijk} = a_{ijk} - \mu_i \mu_k \nu_j \nu_i \omega_k \omega_j \sum_{0 \le t < n} u_{itk} v_{jti} w_{ktj}.$$

From the fact that $\mathrm{Prod}\left(\mathbf{U}, \mathbf{V}, \mathbf{W}\right) = \boldsymbol{\Delta}$ it follows that

$$[\mathbf{B}]_{ijk} = \begin{cases} a_{iii} - \left(\mu_i \nu_i \omega_i\right)^2 & \text{if } 0 \le i = j = k < n \\ a_{ijk} & \text{otherwise} \end{cases}.$$

In particular, for $n = 2$ the characteristic polynomial is obtained by computing the hyperdeterminant of $\mathbf{B}$ given by

$$\mathrm{Hyperdeterminant}\left(\mathbf{B}\right) =$$

$$a_{001} a_{010} a_{100} \left(\mu_1 \nu_1 \omega_1\right)^2 - a_{011} a_{101} a_{110} \left(\mu_0 \nu_0 \omega_0\right)^2 + \left(a_{000} a_{011} a_{101} a_{110} - a_{001} a_{010} a_{100} a_{111}\right). \tag{17}$$

The m-th order hyperdeterminant is derived from the family of spectral elimination ideals and given by

$$\mathrm{Hyperdeterminant}\left(\mathbf{A}\right) =$$

$$\left(\prod_{\substack{\mathbf{j} \in \{0,1\}^{1\times m} \\ \|\mathbf{j}\|_{\ell_1} \equiv 0 \mod 2}} a_{\mathbf{j}}\right) - \left(\prod_{\substack{\mathbf{j} \in \{0,1\}^{1\times m} \\ \|\mathbf{j}\|_{\ell_1} \equiv 1 \mod 2}} a_{\mathbf{j}}\right), \tag{18}$$

for an order m hypermatrix $\mathbf{A}$ of side length 2.

4.3 Spectra of adjacency hypermatrices of groups

As an illustration of naturally occurring hypermatrices we consider the adjacency hypermatrices of finite groups. To an arbitrary finite group G of order n, one associates an $n \times n \times n$ adjacency hypermatrix $\mathbf{A}_G$ with binary entries specified as follows:

$$\forall\, i, j, k \in G, \quad a_{ijk} = \begin{cases} 1 & \text{if } i \cdot j = k \text{ in } G \\ 0 & \text{otherwise} \end{cases}.$$

As illustration, we consider here adjacency hypermatrices associated with the family of groups of the form $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \cdots \times \mathbb{Z}/2\mathbb{Z}$. Note that by definition

$$\mathbf{A}_{\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \cdots \times \mathbb{Z}/2\mathbb{Z}} = \mathbf{A}_{\mathbb{Z}/2\mathbb{Z}} \otimes \mathbf{A}_{\mathbb{Z}/2\mathbb{Z}} \otimes \cdots \otimes \mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}.$$

Consequently, the spectral decomposition of the adjacency hypermatrix $\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}\times\cdots\times\mathbb{Z}/2\mathbb{Z}}$ is determined by the spectral decomposition of the $2\times 2\times 2$ hypermatrix $\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}$. The entries of the hypermatrix $\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}$ are given by

$$\forall\, i,j,k \in \mathbb{Z}/2\mathbb{Z}, \quad \left[\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}\right]_{i,j,k} = \begin{cases} 1 & \text{if } i+j\equiv k \mod 2 \\ 0 & \text{otherwise} \end{cases},$$

$$\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}\left[:,:,0\right] = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right), \quad \mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}\left[:,:,1\right] = \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right).$$

By symmetry the hypermatrix $\mathbf{A}$ admits a spectral decomposition of the form

$$\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}} = \text{Prod}\left(\text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right), \text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T^2}, \text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T}\right),$$

where the the hypermatrix $\mathbf{D}$ is of the form

$$\mathbf{D}_0\left[:,:,0\right] = \left(\begin{array}{cc} \lambda_{00} & 0 \\ \lambda_{01} & 0 \end{array}\right), \ \mathbf{D}_0\left[:,:,1\right] = \left(\begin{array}{cc} 0 & \lambda_{01} \\ 0 & \lambda_{11} \end{array}\right),$$

and the hypermatrix $\mathbf{Q}$ is subject to the orthogonality constraints expressed by

$$\text{Prod}\left(\mathbf{Q},\mathbf{Q}^{T^2},\mathbf{Q}^{T}\right) = \mathbf{\Delta}.$$

The spectrum of $\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}$ is determined by the following parametrization of orthogonal hypermatrices

$$\mathbf{Q}\left[:,:,0\right] = \left(\begin{array}{cc} \left(x^3+1\right)^{-\frac{1}{3}} & \left(\frac{1}{x^3}+1\right)^{-\frac{1}{3}} \\ -x & 1 \end{array}\right), \mathbf{Q}\left[:,:,1\right] = \left(\begin{array}{cc} 1 & 1 \\ \left(x^3+1\right)^{-\frac{1}{3}} & \left(\frac{1}{x^3}+1\right)^{-\frac{1}{3}} \end{array}\right),$$

as well as by the following parametrization for the hypermatrix $\mathbf{D}$:

$$\mathbf{D}\left[:,:,0\right] = \left(\begin{array}{cc} \left(-x^3\right)^{\frac{1}{12}} & 0 \\ \left(-x^3\right)^{\frac{1}{6}} & 0 \end{array}\right), \ \mathbf{D}\left[:,:,1\right] = \left(\begin{array}{cc} 0 & \left(-x^3\right)^{\frac{1}{6}} \\ 0 & 1 \end{array}\right).$$

The parametrization above ensures that

$$\forall\ (i,j,k) \in \left\{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\right\},$$

$$\left[\text{Prod}\left(\text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right), \text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T^2}, \text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T}\right)\right]_{i,j,k} = 0.$$

Finally, by symmetry, the spectral decomposition of $\mathbf{A}$ is obtained by solving for the parameter x in the equation

$$\left[\text{Prod}\left(\text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right), \text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T^2}, \text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T}\right)\right]_{0,0,0} =$$

$$\left[\text{Prod}\left(\text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right),\text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T^{2}},\text{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T}\right)\right]_{0,1,1},$$

which yields the equation

$$\frac{(-1)^{\frac{5}{6}}\,x^{\frac{7}{2}}-(-1)^{\frac{1}{3}}\,x^{2}}{\left(x^{2}-x+1\right)^{\frac{1}{3}}\,(x+1)^{\frac{1}{3}}}-\frac{x^{6}+x\sqrt{-x}}{x^{3}+1}=0,$$

for which the existence of complex roots follows immediately from the fundamental theorem of algebra. Consequently, by Lemma 3 the spectral decomposition of the m-th order adjacency hypermatrix of the group $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}\times\cdots\times\mathbb{Z}/2\mathbb{Z}$ is expressed as

$$\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}\times\cdots\times\mathbb{Z}/2\mathbb{Z}}=\left(\mathbf{A}_{\mathbb{Z}/2\mathbb{Z}}\right)^{\otimes^{n}}=$$

$$\text{Prod}\left(\text{Prod}\left(\mathbf{Q}^{\otimes^{n}},\mathbf{D}^{\otimes^{n}},\left(\mathbf{D}^{\otimes^{n}}\right)^{T}\right),\text{Prod}\left(\mathbf{Q}^{\otimes^{n}},\mathbf{D}^{\otimes^{n}},\left(\mathbf{D}^{\otimes^{n}}\right)^{T}\right)^{T^{2}},\text{Prod}\left(\mathbf{Q}^{\otimes^{n}},\mathbf{D}^{\otimes^{n}},\left(\mathbf{D}^{\otimes^{n}}\right)^{T}\right)^{T}\right).$$

5 General matrix and hypermatrix Rayleigh quotient

The Rayleigh quotient is central to many applications of the spectral decomposition of matrices. We prove here a slight generalization of the matrix Rayleigh quotient, more specifically we do not assume the matrices to be Hermitian. We then proceed to extend the formulation to hypermatrices.

Theorem 8 : Let $\mathbf{A}\in\mathbb{C}^{n\times n}$ having non-negative eigenvalues. Let the spectral decomposition of $\mathbf{A}$ be given by

$$\mathbf{A}=\mathbf{U}\cdot\text{diag}\begin{pmatrix}\lambda_{0}\\ \vdots\\ \lambda_{n-1}\end{pmatrix}\cdot\mathbf{V}^{T},\text{ subject to }\quad\mathbf{I}_{n}=\mathbf{U}\cdot\mathbf{V}^{T},$$

Let $\mathbf{P}_{k}=\text{Prod}_{\boldsymbol{\Delta}^{(k)}}\left(\mathbf{U},\mathbf{V}\right)$ and $S_{k}\subset\mathbb{C}^{n\times1}\times\mathbb{C}^{n\times1}$ be such that $\forall\,(\mathbf{x},\mathbf{y})\in S_{k}$, $\text{Prod}_{\mathbf{P}_{k}}\left(\mathbf{x}^{T},\mathbf{y}\right)\geq0$, then

$$\forall\,(\mathbf{x},\mathbf{y})\in\bigcap_{0\leq k<n}S_{k},\quad\min_{0\leq t<n}\left(\lambda_{t}\right)\leq\frac{\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T},\mathbf{y}\right)}{\text{Prod}\left(\mathbf{x}^{T},\mathbf{y}\right)}\leq\max_{0\leq t<n}\left(\lambda_{t}\right).$$

Proof : The proof follows from the inequality

$$\sum_{0\leq k<n}\lambda_{0}\,\text{Prod}_{\mathbf{P}_{k}}\left(\mathbf{x}^{T},\mathbf{y}\right)\leq\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T},\mathbf{y}\right)\leq\sum_{0\leq k<n}\lambda_{n-1}\,\text{Prod}_{\mathbf{P}_{k}}\left(\mathbf{x}^{T},\mathbf{y}\right),$$

$$\Rightarrow\lambda_{0}\sum_{0\leq k<n}\text{Prod}_{\mathbf{P}_{k}}\left(\mathbf{x}^{T},\mathbf{y}\right)\leq\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T},\mathbf{y}\right)\leq\lambda_{n-1}\sum_{0\leq k<n}\text{Prod}_{\mathbf{P}_{k}}\left(\mathbf{x}^{T},\mathbf{y}\right),$$

which follows from the fact that $\forall\, 0 \leq k < n,\ \mathrm{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^T, \mathbf{y}\right) \geq 0$. By the Parseval identity

$$\mathrm{Prod}\left(\mathbf{x}^T, \mathbf{y}\right) = \sum_{0 \leq k < n} \mathrm{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^T, \mathbf{y}\right),$$

we have

$$\lambda_0\, \mathrm{Prod}\left(\mathbf{x}^T, \mathbf{y}\right) \leq \mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}^T, \mathbf{y}\right) \leq \lambda_{n-1}\, \mathrm{Prod}\left(\mathbf{x}^T, \mathbf{y}\right),$$

and the sought after result follows

$$\lambda_0 \leq \frac{\mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}^T, \mathbf{y}\right)}{\mathrm{Prod}\left(\mathbf{x}^T, \mathbf{y}\right)} \leq \lambda_{n-1}.\square$$

It is then easily verified that the bounds are attained for the choices

$$\mathbf{x} = \mathbf{V}\left[:, 0\right], \quad \mathbf{y} = \mathbf{U}\left[:, 0\right]$$

and

$$\mathbf{x} = \mathbf{V}\left[:, n-1\right], \quad \mathbf{y} = \mathbf{U}\left[:, n-1\right]$$

For practical uses of the general Rayleigh quotient it is useful to provide some explicit description for vectors in the set

$$\forall\, (\mathbf{x}, \mathbf{y}) \in \bigcap_{0 \leq k < n} S_k \Leftrightarrow \mathbf{x} = \sum_{0 \leq i < n} \alpha_i\, \mathbf{V}\left[:, i\right],\ \mathbf{y} = \sum_{0 \leq j < n} \beta_j\, \mathbf{U}\left[:, j\right] \text{ s.t. } \left\{\alpha_k\, \beta_k\right\}_{0 \leq k < n} \subset \mathbb{R}_{\geq 0}.$$

Having reviewed the matrix formulation of the general Rayleigh quotient, we now discuss the hypermatrix formulation of the Rayleigh quotient. For notational convenience we restrict the discussion to third order hypermatrices, but the formulation extends to hypermatrices of all orders.

Theorem 9 : Let $\mathbf{A} \in \mathbb{C}^{n \times n \times n}$, whose spectral decomposition is given by

$$\mathbf{A} = \mathrm{Prod}\left(\mathrm{Prod}\left(\mathbf{U}, \mathbf{D}_0, \mathbf{D}_0^T\right), \mathrm{Prod}\left(\mathbf{V}, \mathbf{D}_1, \mathbf{D}_1^T\right)^{T^2}, \mathrm{Prod}\left(\mathbf{W}, \mathbf{D}_2, \mathbf{D}_2^T\right)^T\right)$$

subject to

$$\mathbf{U}, \mathbf{V}, \mathbf{W} \in \mathbb{C}^{n \times n \times n} \text{ and } \left[\mathrm{Prod}\left(\mathbf{U}, \mathbf{V}^{T^2}, \mathbf{W}^T\right)\right]_{i,j,k} = \begin{cases} 1 & \text{if } 0 \leq i = j = k < n \\ 0 & \text{otherwise} \end{cases},$$

where

$$\left[\mathbf{D}_0\right]_{ijk} = \begin{cases} \mu_{jk} = \mu_{kj} \geq 0 & \text{if } 0 \leq i = k < n \\ 0 & \text{otherwise} \end{cases},$$

$$\left[\mathbf{D}_1\right]_{ijk} = \begin{cases} \nu_{jk} = \nu_{kj} \geq 0 & \text{if } 0 \leq i = k < n \\ 0 & \text{otherwise} \end{cases},$$

$$[\mathbf{D}_2]_{ijk} = \begin{cases} \omega_{jk} = \omega_{kj} \geq 0 & \text{if } 0 \leq i = k < n \\ 0 & \text{otherwise} \end{cases}.$$

Let $\mathbf{P}_k = \text{Prod}_{\mathbf{\Delta}^{(k)}}\left(\mathbf{U}, \mathbf{V}^{T^2}, \mathbf{W}^T\right)$ and $S_k \subset \mathbb{C}^{n\times1\times1} \times \mathbb{C}^{n\times1\times1} \times \mathbb{C}^{n\times1\times1}$ be such that

$$\forall\ (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in S_k, \quad \text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) \geq 0.$$

Then $\forall\ (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \bigcap_{0\leq k<n} S_k$,

$$\min_{0\leq i,j,k,t<n} \left(\mu_{it}\mu_{tk}\, \nu_{jt}\nu_{ti}\, \omega_{kt}\omega_{tj}\right) \leq \frac{\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right)}{\text{Prod}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right)} \leq \max_{0\leq i,j,k,t<n} \left(\omega_{it}\omega_{kt}\, \nu_{jt}\nu_{it}\, \mu_{kt}\mu_{jt}\right)$$

Proof : The proof argument is similar to the matrix case. Recall from the general Parseval identity

$$\forall\ (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in\ \mathbb{C}^{n\times1\times1} \times \mathbb{C}^{n\times1\times1} \times \mathbb{C}^{n\times1\times1},$$

$$\text{Prod}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) = \sum_{0\leq k<n} \text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right),$$

and from the inequalities expressed by

$$\forall\ (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \bigcap_{0\leq k<n} S_k,$$

$$\sum_{0\leq t<n} \min_{0\leq i,j,k<n} \{\mu_{it}\mu_{tk}\, \nu_{jt}\nu_{ti}\, \omega_{kt}\omega_{tj}\}\ \text{Prod}_{\mathbf{P}_t}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) \leq \text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right)$$

and

$$\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) \leq \sum_{0\leq t<n} \max_{0\leq i,j,k<n} \{\mu_{it}\mu_{tk}\, \nu_{jt}\nu_{ti}\, \omega_{kt}\omega_{tj}\}\ \text{Prod}_{\mathbf{P}_t}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right)$$

hence

$$\min_{0\leq i,j,k,t<n} \{\mu_{it}\mu_{tk}\, \nu_{jt}\nu_{ti}\, \omega_{kt}\omega_{tj}\} \sum_{0\leq t<n} \text{Prod}_{\mathbf{P}_t}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) \leq \text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right)$$

and

$$\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) \leq \max_{0\leq i,j,k,t<n} \{\mu_{it}\mu_{tk}\, \nu_{jt}\nu_{ti}\, \omega_{kt}\omega_{tj}\} \sum_{0\leq t<n} \text{Prod}_{\mathbf{P}_t}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right),$$

which follows from the fact that $\forall\ (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in S_k$, $\text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) \geq 0$. By the Parseval identity

$$\text{Prod}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right) = \sum_{0\leq k<n} \text{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^2}, \mathbf{y}^T, \mathbf{z}\right),$$

we have

$$\min_{0\le i,j,k,t<n} \{\mu_{it}\mu_{tk}\,\nu_{jt}\nu_{ti}\,\omega_{kt}\omega_{tj}\}\ \mathrm{Prod}\left(\mathbf{x}^{T^2},\mathbf{y}^{T},\mathbf{z}\right) \le \mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2},\mathbf{y}^{T},\mathbf{z}\right)$$

and

$$\mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2},\mathbf{y}^{T},\mathbf{z}\right) \le \max_{0\le i,j,k,t<n} \{\mu_{it}\mu_{tk}\,\nu_{jt}\nu_{ti}\,\omega_{kt}\omega_{tj}\}\ \mathrm{Prod}\left(\mathbf{x}^{T^2},\mathbf{y}^{T},\mathbf{z}\right).$$

from which we obtain the sought after result

$$\min_{0\le i,j,k,t<n} \{\mu_{it}\mu_{tk}\,\nu_{jt}\nu_{ti}\,\omega_{kt}\omega_{tj}\} \le \frac{\mathrm{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2},\mathbf{y}^{T},\mathbf{z}\right)}{\mathrm{Prod}\left(\mathbf{x}^{T^2},\mathbf{y}^{T},\mathbf{z}\right)} \le \max_{0\le i,j,k,t<n} \{\omega_{it}\omega_{kt}\,\nu_{jt}\nu_{it}\,\mu_{kt}\mu_{jt}\}.$$

For practical uses of the hypermatrix formulation of the Rayleigh quotient it is useful to provide some explicit description for vectors in the set

$$\bigcap_{0\le k<n} S_k.$$

We provide here such a characterization. Let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ denote uncorrelated third order hypermatrix triple. Let $\mathbf{P}_k$ denote the outer product

$$\mathbf{P}_k = \mathrm{Prod}_{\mathbf{\Delta}^{(k)}}\left(\mathbf{U},\mathbf{V}^{T^2},\mathbf{W}^{T}\right).$$

We first observe that for each $\mathbf{P}_k$ there is a unique matrix $\mathbf{M}_k(\mathbf{z})$ for which the following equality holds

$$\forall\, 0\le k<2,\quad \mathrm{Prod}_{\mathbf{P}_k}\left(\mathbf{x}^{T^2},\mathbf{y}^{T},\mathbf{z}\right) = \mathbf{x}^{T}\cdot\mathbf{M}_k(\mathbf{z})\cdot\mathbf{y}$$

consequently the vector $\mathbf{z}$ must be chosen if at all possible to ensure that the $n\times n$ matrix $\mathbf{M}_k(\mathbf{z})$ is diagonalizable with positive eigenvalues for all $0\le k<n$. In the special case $n=2$ both of these requirement are met when

$$\forall\, 0\le k<2,$$

$$\mathrm{Tr}\left(\mathbf{M}_k(\mathbf{z})\right)^2 - 4\det\left(\mathbf{M}_k(\mathbf{z})\right) > 0,\ \mathrm{Tr}\left(\mathbf{M}_k(\mathbf{z})\right)\ge 0\ \text{ and }\ \det\left(\mathbf{M}_k(\mathbf{z})\right)\ge 0.$$

or

$$\mathrm{Tr}\left(\mathbf{M}_k(\mathbf{z})\right)^2 - 4\det\left(\mathbf{M}_k(\mathbf{z})\right) = 0,\ \mathrm{Tr}\left(\mathbf{M}_k(\mathbf{z})\right)\ge 0\ \text{ and }\ \mathbf{M}_k(\mathbf{z}) = \mathbf{M}_k^{T}(\mathbf{z}).$$

Finally, provided that $\mathbf{z}$ is chosen such that $\mathbf{M}_k(\mathbf{z})$ is diagonalizable with positive eigenvalues for all $0\le k<n$, Theorem 7 provides a complete characterization for the possible vectors $\mathbf{x}$ and $\mathbf{y}$ for each of the sets S_k.

Furthermore, for a symmetric $n\times n\times n$ hypermatrix $\mathbf{A}$ whose spectral decomposition is expressed by

$$\mathbf{A} = \mathrm{Prod}\left(\mathrm{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right),\mathrm{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T^2},\mathrm{Prod}\left(\mathbf{Q},\mathbf{D},\mathbf{D}^{T}\right)^{T}\right),$$

such that

$$\text{Prod}\left(\mathbf{Q},\mathbf{Q}^{T^2},\mathbf{Q}^{T}\right)=\boldsymbol{\Delta},$$

$$[\mathbf{D}]_{ijk}=\begin{cases}\lambda_{jk}=\lambda_{kj}>0 & \text{if } 0\le i=k<n\\ 0 & \text{otherwise}\end{cases},$$

where $\mathbf{Q}\in\mathbb{R}^{n\times n\times n}$ and $\mathbf{D}\in\mathbb{R}_{\ge 0}^{n\times n\times n}$, then

$$\forall\, 0\le k<n,\quad \left(\mathbf{M}_k\left(\mathbf{z}\right)\right)^T=\mathbf{M}_k\left(\mathbf{z}\right).$$

Consequently each $\mathbf{M}_k\left(\mathbf{z}\right)$ is diagonalizable and has real eigenvalues for all choices of the vector $\mathbf{z}$. In particular, for $n=2$ it suffices to choose $\mathbf{z}$ such that

$$\forall\, 0\le k<2,\quad \text{Tr}\left(\mathbf{M}_k\left(\mathbf{z}\right)\right)\ge 0,\quad \det\left(\mathbf{M}_k\left(\mathbf{z}\right)\right)\ge 0,$$

which asserts that

$$q_{000}^3 z_0+q_{001}q_{100}q_{101}z_0+q_{000}q_{001}q_{100}z_1+q_{101}^3 z_1\ge 0,$$
$$\left(q_{000}^3 z_0+q_{000}q_{001}q_{100}z_1\right)\left(q_{001}q_{100}q_{101}z_0+q_{101}^3 z_1\right)-\left(q_{000}q_{001}q_{100}z_0+q_{001}q_{100}q_{101}z_1\right)^2\ge 0,$$
$$q_{010}^3 z_0+q_{011}q_{110}q_{111}z_0+q_{010}q_{011}q_{110}z_1+q_{111}^3 z_1\ge 0,$$
$$\left(q_{010}^3 z_0+q_{010}q_{011}q_{110}z_1\right)\left(q_{011}q_{110}q_{111}z_0+q_{111}^3 z_1\right)-\left(q_{010}q_{011}q_{110}z_0+q_{011}q_{110}q_{111}z_1\right)^2\ge 0.$$

For all $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ chosen as indicated above we have

$$\min_{0\le i,j,k,t<n}\left\{\lambda_{it}^2\,\lambda_{jt}^2\,\lambda_{kt}^2\right\}\le\frac{\text{Prod}_{\mathbf{A}}\left(\mathbf{x}^{T^2},\mathbf{y}^T,\mathbf{z}\right)}{\text{Prod}\left(\mathbf{x}^{T^2},\mathbf{y}^T,\mathbf{z}\right)}\le\max_{0\le i,j,k,t<n}\left\{\lambda_{it}^2\,\lambda_{jt}^2\,\lambda_{kt}^2\right\}.$$

By Lemma 3, this characterization of the domain of application of the hypermatrix Rayleigh quotient bound extends to hypermatrices generated by arbitrary combinations of Kronecker product and direct sums of symmetric $2\times 2\times 2$ hypermatrices.

6 Some related algorithmic problems

6.1 Logarithmic least square

Let $\mathbf{A}\in\mathbb{C}^{m\times n}$, $\mathbf{b}\in\mathbb{C}^{m\times 1}$ and consider the monomial constraints in the unknown $\mathbf{x}$ of size $n\times 1$ vector

$$\left\{b_i=\prod_{0\le j<n}x_j^{a_{ij}}\right\}_{0\le i<m}. \tag{19}$$

The logarithmic least square solution to 19 is obtained by solving for $\mathbf{x}$ in the modified system

$$\left\{\prod_{0\le t<m}b_t^{\overline{a_{ti}}}=\prod_{0\le j<n}\left(\prod_{0\le t<m}x_j^{\overline{a_{ti}}\,a_{tj}}\right)\right\}_{1\le i\le n}.$$

By the least square argument the modified system is known to always admit a solution vector $\mathbf{x}$ which minimizes

$$\sum_{0\le i<m}\left|\ln\left(\prod_{0\le k<n} x_j^{a_{ij}}\right)\right|^2.$$

Such a solution is called the logarithmic least square solution of the system 19 and can be obtained via Gauss-Jordan elimination.

6.2 Logarithmic least square BM-rank one approximation

A solution to the general BM-rank ρ approximation of an order m hypermatrix $\mathbf{H} \in \mathbb{C}^{n\times n\times\cdots\times n}$ (for some positive integer $0 < \rho < n$) is obtained by solving for an m-tuple of hypermatrices $\left(\mathbf{X}^{(i)}\right)_{1\le i\le m}$ of the same size as $\mathbf{H}$ which minimize the norm

$$\left\|\mathbf{H} - \sum_{0\le t<\rho} \text{Prod}_{\mathbf{\Delta}^{(t)}}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(m)}\right)\right\|.$$

Recall that

$$\left[\mathbf{\Delta}^{(t)}\right]_{i_1,\cdots,i_m} = \begin{cases} 1 & \text{if } 0 \le t = i_1 = \cdots = i_m < n \\ 0 & \text{otherwise} \end{cases}.$$

Consequently, the BM-rank of $\mathbf{H}$ over $\mathbb{C}$ is the smallest positive integer ρ for which we can solve for the m-tuple $\left(\mathbf{X}^{(i)}\right)_{1\le i\le m}$ subject to

$$\left\|\mathbf{H} - \sum_{0\le t<\rho} \text{Prod}_{\mathbf{\Delta}^{(t)}}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(m)}\right)\right\| = 0.$$

It is easy to see that for all order m hypermatrix $\mathbf{H} \in \mathbb{C}^{n\times n\times\cdots\times n}$

$$0 \le \text{BM-rank}\left(\mathbf{H}\right) \le n.$$

In particular the necessary and sufficient condition for a side length 2 hypermatrix $\mathbf{A}$ to have BM-rank 2 is given by

$$\text{Hyperdeterminant}\left(\mathbf{A}\right) \neq 0,$$

where the hyperdeterminant refers to the polynomial described in 18.

Finally BM-rank 1 approximation constraints are monomial constraints of same type as the ones in 19. The corresponding system admits no solution if BM-rank($\mathbf{H}$) > 1. The proposed BM-rank 1 approximation of $\mathbf{H}$ is thus obtained by solving the corresponding system in the logarithmic least square sense.

6.3 Logarithmic least square direct sum and Kronecker product approximation

Let $\mathbf{A}$ denote a cubic order m hypermatrix of side length n such that

$$\mathbf{A} = \bigoplus_{0<j\leq\beta} \mathbf{A}^{(j)},$$

where $\mathbf{A}^{(j)} \in \mathbb{C}^{2^j\times 2^j\times\cdots\times 2^j}$. A direct sum and Kronecker product approximation of $\mathbf{A}$ is obtained by solving for entries of a hypermatrix $\mathbf{B}$ subject to two constraints. The first of which asserts that $\mathbf{B}$ must be generated by some a arbitrary combinations of Kronecker products and direct sums of cubic side length 2 hypermatrices. The second constraint asserts that $\mathbf{B}$ should be chosen so as to minimize the norm $\|\mathbf{A} - \mathbf{B}\|$. The problem reduces to a system of the same form as 19 given by

$$\forall\, 0 \leq j < \beta, \quad \mathbf{A}^{(j)} = \bigotimes_{0\leq i<j} \mathbf{X}^{(i,j)},$$

where $\mathbf{X}^{(i,j)} \in \mathbb{C}^{2\times 2\times\cdots\times 2}$. Consequently the system admits no solution if if $\mathbf{A}$ is not generated by a combination of Kronecker product and direct sums of side length 2 hypermatrices. The proposed direct sum and Kronecker product approximation of $\mathbf{A}$ is obtained by solving the corresponding system in the logarithmic least square sense.

References

[Bha95] P. Bhattacharya, *A new 3-d transform using a ternary product*, Signal Processing, IEEE Transactions on **43** (1995), no. 12, 3081–3084.

[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.

[Gna13] Edinah K. Gnang, *Computational aspects of the combinatorial nullstellensatz method via a polynomial approach to matrix and hypermatrix algebra*, Ph.D. thesis, Rutgers New Brunswick, 2013.

[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 6.7)*, The Sage Development Team, 2015, `http://www.sagemath.org`.