# RELATIONS ENCODED IN MULTIWAY ARRAYS

by

**David Warren Katz**

**A Dissertation**

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

**Doctor of Philosophy**

Department of Mathematics

West Lafayette, Indiana

May 2022

**THE PURDUE UNIVERSITY GRADUATE SCHOOL**
**STATEMENT OF COMMITTEE APPROVAL**

**Dr. Saugata Basu, Chair**
Department of Mathematics

**Dr. Andrei Gabrielov**
Department of Mathematics

**Dr. Uli Walther**
Department of Mathematics

**Dr. Jaroslaw Wlodarczyk**
Department of Mathematics

**Approved by:**
Dr. Plamen Stefanov

# ACKNOWLEDGMENTS

Foremost, I would like to express my gratitude to my advisor Saugata Basu for his invaluable guidance and encouragement. I would also like to thank Brian Hogle and Ryan Vitale for many insightful discussions. Finally, I want to thank Zachary Selk and Brooke Hejnal for proofreading this thesis.

# TABLE OF CONTENTS

# LIST OF FIGURES

# ABSTRACT

Unlike matrix rank, hypermatrix rank is not lower semi-continuous. As a result, optimal low rank approximations of hypermatrices may not exist. Characterizing hypermatrices without optimal low rank approximations is an important step in implementing algorithms with hypermatrices. The main result of this thesis is an original coordinate-free proof that real $2 \times 2 \times 2$ tensors that are rank three do not have optimal rank two approximations with respect to the Frobenius norm. This result was previously only proved in coordinates. Our coordinate-free proof expands on prior results by developing a proof method that can be generalized more readily to higher dimensional tensor spaces. Our proof has the corollary that the nearest point of a rank three tensor to the second secant set of the Segre variety is a rank three tensor in the tangent space of the Segre variety. The relationship between the contraction maps of a tensor generalizes, in a coordinate-free way, the fundamental relationship between the rows and columns of a matrix to hypermatrices. Our proof method demonstrates geometrically the fundamental relationship between the contraction maps of a tensor. For example, we show that a regular real or complex tensor is tangent to the $2 \times 2 \times 2$ Segre variety if and only if the image of any of its contraction maps is tangent to the $2 \times 2$ Segre variety.

# 1. INTRODUCTION

The maximum number of non-redundant rows of a two-way array of real numbers is always equal to the maximum number of non-redundant columns. A column is redundant if it can be written in terms of the other columns. For example, the fourth and fifth columns of the $4 \times 5$ matrix

$$\begin{bmatrix} 1 & \frac{3}{2} & 0 & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 3 & 5 \end{bmatrix} \tag{1.1}$$

can be written in terms of the first three columns.

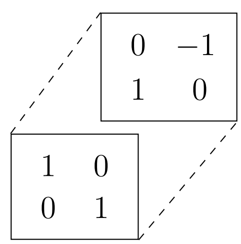
$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \end{bmatrix} = -\frac{3}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} \frac{3}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 5 \end{bmatrix} = 2\begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \end{bmatrix}$$

Each of the first three columns, however, cannot be written in terms of the other two, which can be observed by considering the entries carefully. For example, both columns two and three have a 1 in a position where the other two columns have 0's, and thus cannot be written in terms of the other two. By similar reasoning, the maximum number of non-redundant rows of matrix (1.1) is also three. For example, the first three rows cannot be written in terms of each other, but the fourth row can be written in terms of the first three.

$$\begin{bmatrix} 0 & 0 & 1 & 3 & 5 \end{bmatrix} = -2\begin{bmatrix} 1 & \frac{3}{2} & 0 & 0 & 0 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 & 1 \end{bmatrix} + 3\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Thus, the maximum number of non-redundant rows and the maximum number of non-redundant columns of matrix (1.1) is three. It is a remarkable observation that this is in fact always the case for matrices of real numbers. *All* real matrices have the same maximum number of non-redundant rows and non-redundant columns. We prove this for matrices over semirings in Theorem 5.

The equality of row rank and column rank makes it clear that a matrix is more than a collection of numbers. There is a fundamental relationship between the rows and columns of a matrix. Information encoded in this relationship is lost when data organized as a matrix is reorganized as a vector. Similarly, if data can more naturally be viewed in three separate but interconnected ways - as rows, columns, and pillars - then information is lost when the data is organized as a matrix or as a collection of matrices. In this situation, the data would be better organized as a 3-fold hypermatrix.




In Chapter 2, we show how matrix rank extends naturally to hypermatrices. The notion of hypermatrix rank has a long history that goes back at least to the work of Terracini in 1911 [1]. In its modern formulation, hypermatrix rank first appeared in the work of Hitchcock in 1927 [2] [3], and was popularized in the 1960s by Tucker [4] [5] [6]. Since then, hypermatrix rank has found many applications in machine learning [7] [8] [9] and phylogenetics [10] [11] [12].

Questions about the rank of a $d$-fold hypermatrix are much more difficult than their matrix analogues. For example, in [13] it is shown that computing the rank of a real 3-fold hypermatrix is NP-hard. This illustrates how remarkable it is that the same Gaussian elimination algorithm can compute the rank of a real or complex matrix of any dimensions $m \times n$ in polynomial time. However, it should be expected that there is no polynomial time algorithm that determines the rank of any $n_1 \times n_2 \times n_3$ hypermatrix, as the properties of the space of $3 \times 3 \times 3$ hypermatrices, for example, are fundamentally different than the properties of the space of hypermatrices of any other dimensions. Indeed, even the Gaussian elimination algorithm needs to be modified with respect to the dimensions $m \times n$ to determine the non-negative rank of a non-negative matrix.

Most practitioners working with hypermatrices choose to work with hypermatrices of dimensions $n_1 \times n_2 \times \cdots \times n_d$ because of some special property of those dimensions, rather than because their data can naturally be represented as an $n_1 \times n_2 \times \cdots \times n_d$ array. We believe this is a mistake. Just as the rank, eigenvalues, and singular values of an $n \times n$ matrix are lost when the same information is reorganized as a $n^2 \times 1$ vector, essential information encoded in a hypermatrix of dimensions $n_1 \times n_2 \times \cdots \times n_d$ is lost when the same information is encoded as a hypermatrix of any other dimensions. The dimensions of the space of hypermatrices chosen to store information should be determined by the information itself, rather than chosen to exploit some feature of those specific dimensions. In this thesis, we will study the special properties associated with spaces of hypermatrices of dimensions $2 \times 2 \times 2$.

Tensors are the coordinate-free counterparts of hypermatrices. Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, and $n_3$-dimensional vector spaces over a common field, and let $V^{1*}$, $V^{2*}$, and $V^{3*}$ denote their dual spaces. The tensor $\rho = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \otimes v_i^3$ in $V^1 \otimes V^2 \otimes V^3$ induces the three linear maps

$$\begin{array}{lll} V^{1*} \xrightarrow{\Pi_1(\rho)} V^2 \otimes V^3 & V^{2*} \xrightarrow{\Pi_2(\rho)} V^1 \otimes V^3 & V^{3*} \xrightarrow{\Pi_3(\rho)} V^1 \otimes V^2 \\ v^{1*} \mapsto \displaystyle\sum_{i=1}^{r} v^{1*}(v_i^1)\ v_i^2 \otimes v_i^3, & v^{2*} \mapsto \displaystyle\sum_{i=1}^{r} v^{2*}(v_i^2)\ v_i^1 \otimes v_i^3, \quad \text{and} & v^{3*} \mapsto \displaystyle\sum_{i=1}^{r} v^{3*}(v_i^3)\ v_i^1 \otimes v_i^2. \end{array}$$

These maps are called the mode-1, mode-2, and mode-3 contraction maps of $\rho$, respectively. In Theorem 6 of Chapter 2, we prove that the contraction maps are indeed well-defined. The relationship between the contraction maps of a tensor generalizes, in a coordinate-free way, the fundamental relationship between the rows and columns of a matrix to hypermatrices. The rank of a two-fold tensor $\rho$ in $V^1 \otimes V^2$ is equal to the dimension of the image of $\Pi_1(\rho)$, which is also equal to the dimension of the image of $\Pi_2(\rho)$.

For a 3-fold tensor $\rho$ in $V^1 \otimes V^2 \otimes V^3$, the rank of $\rho$ is equal to the minimum number $n$ such that the image of *any* of its mode-$i$ contraction maps is contained in the span of $n$ rank one tensors [14, p.68]. Although not often recognized as such, this theorem is actually a generalization of the equality of row rank and column rank to hypermatrices. This generalization as well as one other generalization of row rank equals column rank are discussed in Section 2.5. When restricted to regular tensors, this generalization of row rank equals column rank to $2 \times 2 \times 2$ tensors takes on a more geometric form. A tensor is regular if all of its contraction maps are full rank. In Section 3.4, we prove that the rank of a real or complex $2 \times 2 \times 2$ regular tensor is equal to the cardinally of the intersection of the projective image of any of its contraction maps with the projective Segre variety. The Segre variety is the variety of rank at most one tensors.

The rank of a tensor can best be understood geometrically in terms of the tangential and secant varieties of the Segre variety. Despite being well-studied, the ideals of the secant and tangential varieties of the Segre variety are only known in a few cases [15] [16]. In fact, the maximal possible rank of the $n_1 \times n_2 \times \cdots \times n_d$ tensor space and the dimensions of the secant and tangential varieties of the $n_1 \times n_2 \times \cdots \times n_d$ Segre variety are unknown in general. To illustrate how the special properties of the space of $n_1 \times n_2 \times \cdots \times n_d$ hypermatrices is best understood geometrically in terms of the curvature of the Segre variety, we give an example of the unique identifiability of rank one decompositions of $2 \times 2 \times 2$ regular tensors in Theorem 26.

In 1989, Kruskal brought attention to the remarkable fact that regular $2 \times 2 \times 2$ real hypermatrices could be written uniquely as the sum of rank one hypermatrices [17]. This result is known as Kruskal's Theorem, and was extended to regular $2 \times m \times m$ real hypermatrices by Ten Berge [18]. It was later realized that the $2 \times m \times m$ case was actually implied by the work of Weierstrass and Kronecker [19]. In a modern formulation, Kruskal's Theorem states that real $2 \times m \times m$ hypermatrices of rank $m$ have uniquely identifiable rank one decompositions. However, $n \times n$ matrices of rank $m$ do not have this property for any $m > 1$. Geometrically, this is because the $n \times n$ Segre variety is curved in such a way that every tensor on a secant line of the Segre variety is also on a tangent line. This is not true for the Segre variety of real $m \times m \times m$ hypermatrices for any $m \geq 2$. In Chapter 3, we illustrate this geometric perspective. In Theorem 17, we give a geometric proof of the existence of an open and dense set of rank two complex $2 \times 2 \times 2$ tensors.

In Theorem 19 of Chapter 3 we prove the original corollary that a regular tensor is tangent to the $2 \times 2 \times 2$ Segre variety if and only if the image of any of its contraction maps is tangent to the $2 \times 2$ Segre variety. This is a geometric example of the fundamental relationship between the contraction maps of a tensor. We also give a geometric proof that a regular $2 \times 2 \times 2$ real or complex tensor is rank two if and only if its contraction maps intersect the $2 \times 2$ Segre variety at two distinct points. This characterization of rank two regular tensors suggests that optimal rank two approximations cannot be regular. Indeed, we will prove this in Theorem 31 in Chapter 5.

The problem of characterizing when low rank approximations of $2 \times 2 \times 2$ real tensors exist will be the primary topic of this thesis. Small perturbations of the entries of a matrix will almost always result in a full rank matrix. Similarly, hypermatrices of real-world data will be of erroneously large rank due to noise in the data. Thus, there is a need to find optimal low rank approximations of hypermatrices.

Low rank approximations of hypermatrices are much less well-behaved than low rank approximations of matrices. For example, subtracting an optimal rank one approximation of a 3-fold hypermatrix may actually increase the rank of the hypermatrix [20]. In fact, the problem of finding optimal low rank approximations of hypermatrices is often ill-posed, for the set of real hypermatrices with no optimal low rank approximation often has positive Lebesgue measure [21]. One reason for the nonexistence of low rank approximations is the phenomenon of rank-jumping. A hypermatrix $(a_{ijk})$ is said to be rank-jumping if there is a sequence of rank $s$ hypermatrices that converges to $(a_{ijk})$, but the rank of $(a_{ijk})$ is greater than $s$. It is a classical observation dating back to at least Terracini [14, p.9] that hypermatrices in the form

$$\beta \;=\; x_2^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_1^2 \otimes x_2^3 \tag{1.2}$$

are limit points of the sequence of rank at most two hypermatrices

$$\rho_n \;=\; n\,(x_1^1 + \frac{1}{n}x_2^1) \otimes (x_1^2 + \frac{1}{n}x_2^2) \otimes (x_1^3 + \frac{1}{n}x_2^3) \;-\; n x_1^1 \otimes x_1^2 \otimes x_1^3,\; n \in \mathbb{N} \tag{1.3}$$

in the norm topology. Thus, when $\beta$ is rank three, it is an example of a rank-jumping hypermatrix without an optimal rank two approximation. In Chapter 4, we will show that the sequence (1.3) is best understood as a sequence of difference quotients that converges to the derivative of a smooth curve on the Segre variety. In [22], it is proved that all rank-jumping tensors over $\mathbb{C}$ that are limits of a sequence of rank three tensors are also derived from derivatives of smooth curves on the Segre variety. This suggests a fundamental relationship between rank-jumping tensors and derivatives. In Chapter 4, we show that a $2 \times 2 \times 2$ real hypermatrix is rank-jumping if and only if it is a regular element of the tangential variety of the Segre variety. We use this perspective to explain why rank-jumping never occurs for non-negative real hypermatrices.

The main result of this thesis is a coordinate-free proof in Theorem 36 that in fact *every* rank three $2 \times 2 \times 2$ tensor has no optimal rank two approximation with respect to the Frobenius norm, not just the rank-jumping tensors in the form of (1.2). General rank three $2 \times 2 \times 2$ real tensors are not rank-jumping. Rather, we will show that their failure to have optimal rank two approximations is due to the curvature of the Segre variety. Our proof has the corollary that the nearest point of a rank three tensor to the second secant set of the Segre variety is a rank three tensor in the tangent space of the Segre variety. Overall, this thesis develops a coordinate-free, geometric proof method that uses the contraction maps of a tensor to study the relations encoded in multiway arrays.

# 2. ROW RANK EQUALS COLUMN RANK FOR HYPERMATRICES

We define the rank of a hypermatrix and show that it is both a natural extension of matrix rank and a natural measurement of a hypermatrix's complexity. We introduce tensors and contraction maps, and show that the relationship between the contraction maps of a tensor generalizes, in a coordinate-free way, the fundamental relationship between the rows and columns of a matrix to hypermatrices. We then give two generalizations of the equality of row rank and column rank for hypermatrices.

## 2.1 The Rank of a Hypermatrix

Let $F$ be a fixed field.

**Definition 1.** *A d-fold hypermatrix over $F$ is a function from $(1, 2, \ldots, n_1) \times (1, 2, \ldots, n_2) \times \cdots \times (1, 2, \ldots, n_d)$ to $F$, where $n_1$, $n_2$, ..., $n_d$ are natural numbers. The set of all hypermatrices of dimensions $n_1 \times n_2 \times \cdots \times n_d$ forms a vector space with entry-wise addition and scalar multiplication. The space of all such hypermatrices is denoted $F^{n_1 \times n_2 \times \cdots \times n_d}$, and hypermatrices are written as $(a_{i_1 i_2 \ldots i_d})$, where each $a_{i_1 i_2 \ldots i_d} \in F$.*

We work with 3-fold hypermatrices of dimensions $2 \times 2 \times 2$, but all of our definitions easily extend to $d$-fold hypermatrices of arbitrary dimensions. The rank of a hypermatrix is an example of information that is lost when a hypermatrix is reorganized as a matrix. To define the rank of a hypermatrix, we first define the class of hypermatrices whose rank is less than or equal to one. We call such hypermatrices simple. As motivation, we first consider the case of $2 \times 2$ real matrices. If nonzero, the rank one matrix

$$\begin{pmatrix} x_1 y_1 & x_1 y_2 \\ x_2 y_1 & x_2 y_2 \end{pmatrix}$$

has a row space spanned by the vector $[y_1, y_2]$ and column space spanned by the vector $[x_1, x_2]$. Hence, when studying the row and column spaces of matrices, it is natural to consider the map

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1y_1 & x_1y_2 \\ x_2y_1 & x_2y_2 \end{pmatrix}, \tag{2.1}$$

which either sends two vectors to the zero matrix or to the rank one matrix with row and column spaces spanned by those two vectors. In fact, every rank less than or equal to one real $2 \times 2$ matrix is in the image of map (2.1). We use this characterization of rank less than or equal to one matrices to define rank less than or equal to one hypermatrices.

**Definition 2.** *The rank less than or equal to one hypermatrices in $F^{2\times2\times2}$ are the hypermatrices in the image of the coordinate tensor map $\otimes$, defined as*

$$F^2 \times F^2 \times F^2 \xrightarrow{\otimes} F^{2\times2\times2} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \times \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mapsto \begin{array}{c} \boxed{\begin{matrix} x_2y_1z_1 & x_2y_1z_2 \\ x_2y_2z_1 & x_2y_2z_2 \end{matrix}} \\ \boxed{\begin{matrix} x_1y_1z_1 & x_1y_1z_2 \\ x_1y_2z_1 & x_1y_2z_2 \end{matrix}} \end{array}.$$

An element in the image of the coordinate tensor map is called a simple hypermatrix, and is denoted as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \otimes \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} := \otimes\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \times \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right). \tag{2.2}$$

A simple hypermatrix is rank zero if and only if all of its entries are zero. This hypermatrix is called the zero hypermatrix. All nonzero simple hypermatrices are rank one. The vector $[x_1, x_2]$ spans the pillar space of rank one hypermatrix (2.2). The pillar space of a hypermatrix is the space spanned by its pillars. The pillars of hypermatrix (2.2) are circled in Figure 2.1. Similarly, the vector $[y_1, y_2]$ spans the column space of hypermatrix (2.2), and $[z_1, z_2]$ spans its row space. Hence, the rank one hypermatrices are precisely the hyperma-

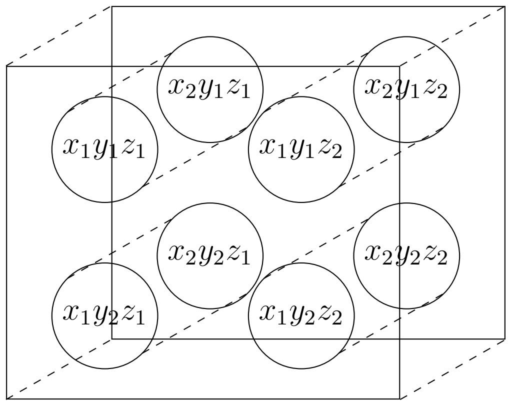



**Figure 2.1.** A visualization of rank one hypermatrix (2.2) with its four pillars circled.

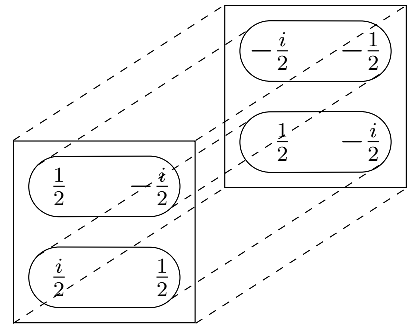



**Figure 2.2.** A visualization of a rank one complex $2 \times 2 \times 2$ hypermatrix with its two co-dimension one horizontal slices circled.

trices whose row, column, and pillar spaces are one dimensional. It is in this sense that the rank one hypermatrices are a natural class of simplest nontrivial hypermatrices, and, thus, the rank of a hypermatrix is a natural measurement of its complexity. Every hypermatrix is the sum of rank one hypermatrices, and the rank of a hypermatrix $B$, denoted $\text{rk}(B)$, is the minimum natural number $n$ such that $B$ is equal to the sum of $n$ simple hypermatrices.

**Definition 3.** *For $B \in F^{2\times 2\times 2}$, the rank of $B$ is defined as*

$$rk(B) = \min\{s \in \mathbb{N} \mid \ B \ = \sum_{i=1}^{s} \ v_i^1 \otimes v_i^2 \otimes v_i^3 \ \textit{for some } v_i^k \in F^2\}.$$

The row space, column space, and pillar space of a hypermatrix are all one dimensional if and only if the spaces of co-dimension one vertical, horizontal, and frontal slices are one dimensional. Co-dimension one slices are illustrated in Figure 2.2. This observation is our first example of the primary theme of this thesis. There are fundamental relationships between the slices of a hypermatrix, and the rank of a hypermatrix encodes information about these relationships.

## 2.2 How Rank Depends on Scalars

Real-valued hypermatrices have both a real rank and a complex rank. The real rank of real hypermatrix $C$ is the minimum number $n$ such that $C$ is equal to the sum of $n$ real-valued simple hypermatrices. The complex rank of real hypermatrix $C$, on the other hand, is the minimum number $n$ such that $C$ is equal to the sum of $n$ complex-valued simple hypermatrices. The complex rank may be strictly smaller than the real rank. For example, the hypermatrix

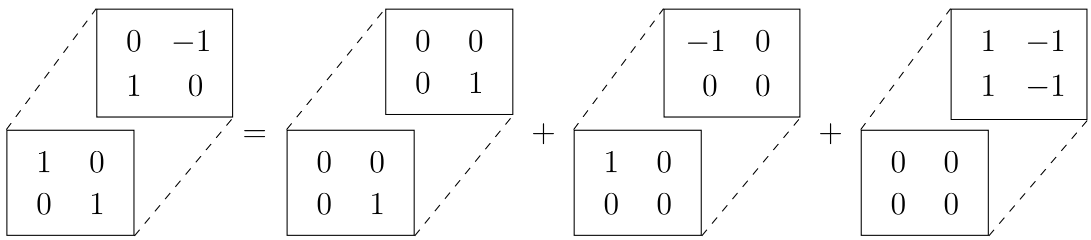
$$C = \begin{matrix} & \boxed{\begin{matrix}0 & -1\\ 1 & 0\end{matrix}} \\ \boxed{\begin{matrix}1 & 0\\ 0 & 1\end{matrix}} & \end{matrix}$$

has real rank three but complex rank two. Since the row space of $C$ is not one dimensional over $\mathbb{R}$ or $\mathbb{C}$, $C$ is not rank one over $\mathbb{R}$ or $\mathbb{C}$. It is, however, rank two over $\mathbb{C}$, as the following decomposition shows.

$$\begin{matrix} & \boxed{\begin{matrix}0 & -1\\ 1 & 0\end{matrix}} \\ \boxed{\begin{matrix}1 & 0\\ 0 & 1\end{matrix}} & \end{matrix} = \begin{matrix} & \boxed{\begin{matrix}-\frac{i}{2} & -\frac{1}{2}\\ \frac{1}{2} & -\frac{i}{2}\end{matrix}} \\ \boxed{\begin{matrix}\frac{1}{2} & -\frac{i}{2}\\ \frac{i}{2} & \frac{1}{2}\end{matrix}} & \end{matrix} + \begin{matrix} & \boxed{\begin{matrix}\frac{i}{2} & -\frac{1}{2}\\ \frac{1}{2} & \frac{i}{2}\end{matrix}} \\ \boxed{\begin{matrix}\frac{1}{2} & \frac{i}{2}\\ -\frac{i}{2} & \frac{1}{2}\end{matrix}} & \end{matrix}$$

$$= \begin{bmatrix}\frac{1}{2}\\ -\frac{i}{2}\end{bmatrix} \otimes \begin{bmatrix}1\\ i\end{bmatrix} \otimes \begin{bmatrix}1\\ -i\end{bmatrix} + \begin{bmatrix}\frac{1}{2}\\ \frac{i}{2}\end{bmatrix} \otimes \begin{bmatrix}1\\ -i\end{bmatrix} \otimes \begin{bmatrix}1\\ i\end{bmatrix}$$

Each of the rank one summands in the above decomposition contains complex numbers. It can be shown by contradiction that there is no rank two decomposition of $C$ with summands of entirely real entries. Hence, the real rank of $C$ is greater than two. The real rank of $C$ is in fact three, as demonstrated by the following decomposition.



$$= \begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} + \begin{bmatrix}1\\-1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\-1\end{bmatrix}$$

Expanding the set of scalars that a hypermatrix's entries can vary over allows for more possible rank one summands. This may decrease the rank of the hypermatrix, as it did for hypermatrix $C$ above. This phenomena of a hypermatrix's rank decreasing by considering rank one summands from a larger set of entries may seem surprising since it never occurs when considering the complex and real rank of real matrices. It is well-known that a real matrix is real rank $r$ if and only if it is complex rank $r$. This is equivalent to the fact that row operations with complex numbers are never needed to reduce a real matrix to row echelon form. Considering general hypermatrices thus reveals that this well-known fact is specific to the real and complex rank of 2-fold real hypermatrices.

Indeed, the equality of the real rank of a real matrix and the complex rank of a real matrix does not extend to non-negative matrices. The non-negative rank of non-negative real hypermatrix $D$ is the minimum number $n$ such that $D$ is equal to the sum of $n$ non-negative-valued simple hypermatrices. The non-negative rank of a non-negative real matrix is *not* in general equal to its real rank. For example, it can be shown by contradiction [23, p.153] that the non-negative matrix $D$ below is rank four over $\mathbb{R}_{\geq 0}$, but rank three over $\mathbb{R}$.

$$D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

### 2.3 Hypermatrices as Coordinatizations of Tensors

It is often convenient to view a matrix as a coordinatization of a linear map. Hypermatrices also have coordinate-free counterparts - tensors. If $\{\mathrm{e}_i^j\}_{i=1}^{n_j}$ is a basis of $F^{n_j}$ for $j = 1, 2, 3$, then $\{\mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3\}_{i,j,k=1}^{n_1,n_2,n_3}$ is a basis of $F^{n_1 \times n_2 \times n_3}$. Since a multilinear map on $F^{n_1} \times F^{n_2} \times F^{n_3}$ is determined by its values on $\{\mathrm{e}_i^1 \times \mathrm{e}_j^2 \times \mathrm{e}_k^3\}_{i,j,k=1}^{n_1,n_2,n_3}$, it follows that the coordinate tensor map $\otimes$ is a multilinear map with the following property: For every multilinear map $\phi : F^{n_1} \times F^{n_2} \times F^{n_3} \to F^m$ for some natural number $m$, there exists a unique linear map $\Phi : F^{n_1 \times n_2 \times n_3} \to F^m$ such that the following diagram commutes.

$$\begin{tikzcd}
F^{n_1} \times F^{n_2} \times F^{n_3} \arrow[r, "\otimes"] \arrow[dr, "\phi"'] & F^{n_1 \times n_2 \times n_3} \arrow[d, dashed, "\Phi"] \arrow[dl, phantom, "\circlearrowleft"] \\
& F^m
\end{tikzcd}$$

We use this characterization of the coordinate tensor map to define a coordinate-free version of hypermatrices.

**Definition 4.** *Let $V^i$ be finite $n_i$-dimensional vector spaces over a common field. A tensor product of $V^1 \times V^2 \times V^3$ is a multilinear mapping $\otimes : V^1 \times V^2 \times V^3 \to W$ to a vector space $W$ such that for any multilinear mapping $\phi : V^1 \times V^2 \times V^3 \to H$ to some vector space $H$, there exists a unique linear mapping $\Phi : W \to H$ such that $\phi = \Phi \circ \otimes$. That is, the following diagram commutes.*

$$\begin{tikzcd}
V^1 \times V^2 \times V^3 \arrow[r, "\otimes"] \arrow[dr, "\phi \text{ multilinear}"'] & W \arrow[d, dashed, "\Phi \text{ linear}"] \arrow[dl, phantom, "\circlearrowleft"] \\
& H
\end{tikzcd}$$

The above property is referred to the universal property of the tensor product [24, Ch.1]. The coordinate tensor map satisfies the universal property of the tensor product, which proves that a map with such a property indeed exists. It also follows from the universal property that any two tensor products of $V^1 \times V^2 \times V^3$ are unique up to a rank-preserving isomorphism. The vector space $W$ is denoted $V^1 \otimes V^2 \otimes V^3$ and its elements are called tensors. We define the rank of a tensor just as we did for hypermatrices. The tensors in the image of $\otimes$ are called simple tensors. The simple tensor $\otimes(v^1 \times v^2 \times v^3)$ is denoted $v^1 \otimes v^2 \otimes v^3$. Not every tensor is a simple tensor, but every tensor can be written as the sum of finitely many simple tensors. The rank of a tensor $\rho$ is the minimum number $n$ such that $\rho$ is the sum of $n$ simple tensors. That is,

$$\mathrm{rk}(\rho) = \min\{s \in \mathbb{N} \mid \ \rho \ = \sum_{i=1}^{s} v_i^1 \otimes v_i^2 \otimes v_i^3 \text{ for some } v_i^k \in V^k\}.$$

Once bases $\{\mathrm{e}_i^j\}_{i=1}^{n_j}$ of $V^j$ are chosen, a tensor in $V^1 \otimes V^2 \otimes V^3$ can be coordinatized as a hypermatrix in $F^{n_1 \times n_2 \times n_3}$ by the following linear map, defined on simple tensors as

$$\begin{array}{rcll} V^1 \otimes V^2 \otimes V^3 & \to & F^{n_1 \times n_2 \times n_3} & \\ v^1 \otimes v^2 \otimes v^3 & \longmapsto & (a_{i_1 i_2 i_3} = a_{i_1}^1 a_{i_2}^2 a_{i_3}^3) & \qquad \text{where } v^k = \sum_{i=1}^{n_k} a_i^k \ \mathrm{e}_i^k \text{ for } k = 1, 2, 3. \end{array}$$

For $2 \times 2 \times 2$ real tensors, the isomorphism from $V^1 \otimes V^2 \otimes V^3$ to $\mathbb{R}^{2\times 2\times 2}$ acts on simple tensors as

$$(a_1^1 \mathrm{e}_1^1 + a_2^1 \mathrm{e}_2^1) \otimes (a_1^2 \mathrm{e}_1^2 + a_2^2 \mathrm{e}_2^2) \otimes (a_1^3 \mathrm{e}_1^3 + a_2^3 \mathrm{e}_2^3) \quad \longmapsto \quad \boxed{\begin{array}{cc} a_2^1 a_1^2 a_1^3 & a_2^1 a_1^2 a_2^3 \\ a_2^1 a_2^2 a_1^3 & a_2^1 a_2^2 a_2^3 \end{array}} \quad \boxed{\begin{array}{cc} a_1^1 a_1^2 a_1^3 & a_1^1 a_1^2 a_2^3 \\ a_1^1 a_2^2 a_1^3 & a_1^1 a_2^2 a_2^3 \end{array}}$$

for some real constants $a_t^s$.

### 2.4 Proof of Row Rank Equals Column Rank Using Contraction Maps

Our definition of a hypermatrix and hypermatrix rank are well-defined in the general setting of hypermatrices over semirings, and this general setting allows us to discuss hypermatrices of complex, real, and non-negative real numbers simultaneously. A semiring is a generalization of a ring in which the requirement of the existence of an additive identity element and the requirement that each element must have an additive inverse are both dropped. This is also the appropriate setting for discussing hypermatrices because the relations encoded in multiway arrays exist in this generality. In this section, we present a proof of row rank equals column rank for matrices over semirings. The key idea of this generalized proof is based on the elegant proof in [25]. The proof in [25] does not work with semirings, but also does not require additive inverses. We then adapt this proof to the coordinate-free setting of contraction maps of tensors to demonstrate that contraction maps are an appropriate method of studying the relations that are encoded in the slices of a hypermatrix.

**Theorem 5.** *Let $A = (a_{ij})$ be an $m \times n$ matrix over a semiring $K$. The size of a minimal spanning set of rows of $A$ is equal to the size of a minimal spanning set of columns of $A$.*

*Proof.* If $A = BC$ for some $m \times r$ matrix $B = (b_{ij})$ and some $r \times n$ matrix $C = (c_{ij})$,

$$\underset{m \times n}{\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}} = \underset{m \times r}{\begin{pmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1r} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2r} \\ \vdots & \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & b_{m3} & \dots & b_{mr} \end{pmatrix}} \underset{r \times n}{\begin{pmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ c_{r1} & c_{r2} & c_{r3} & \dots & c_{rn} \end{pmatrix}}$$

then the columns of $A$ are linear combinations of the columns of $B$.

$$\text{The } j^{th} \text{ column of } A = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = c_{1j} \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix} + c_{2j} \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{m2} \end{pmatrix} + \dots + c_{rj} \begin{pmatrix} b_{1r} \\ b_{2r} \\ \vdots \\ b_{mr} \end{pmatrix}$$

Thus, the size of a minimal spanning set of columns of $A$ is the minimal number $r$ such that $A = BC$ for some $m \times r$ matrix $B$ and some $r \times n$ matrix $C$. Additionally, the rows of $A$ are linear combinations of the rows of $C$.

$$\begin{aligned}
\text{The } j^{th} \text{ row of } A \; &= \begin{pmatrix} a_{j1} & a_{j2} & \dots & a_{jn} \end{pmatrix} \\
&= \; b_{j1}\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \end{pmatrix} \; + \; b_{j2}\begin{pmatrix} c_{21} & c_{22} & \dots & c_{2n} \end{pmatrix} \\
&\quad + \; \dots \; + \; b_{jr}\begin{pmatrix} c_{r1} & c_{r2} & \dots & c_{rn} \end{pmatrix}
\end{aligned}$$

Thus, the size of a minimal spanning set of rows of $A$ is *also* the minimal number $r$ such that $A = BC$ for some $m \times r$ matrix $B$ and some $r \times n$ matrix $C$. Hence, the size of a minimal spanning set of rows of $A$ is equal to the size of a minimal spanning set of columns of $A$. □

The proof of Theorem 5 does not require the existence of an additive identity or additive inverses, and thus indeed holds for matrices over semirings. Over fields, a set is a minimal spanning set if and only if it is a maximal independent set. Hence, for matrices over fields, we have shown that the size of a maximal independent set of columns is equal to the size of a maximal independent set of rows.

The connection between the theory of matrices and the theory of linear maps is sometimes emphasized to the point where matrices are *defined* as coordinatizations of linear maps. However, matrices can just as logically be thought of as coordinatizations of bilinear forms, and the over-identification of matrices with linear maps obscures the fact that matrices and linear maps are actually distinct structures. Matrices over semirings can no longer be identified with linear maps, but, as we have shown, there is still a fundamental relationship that exists between the rows and columns of a matrix over a semiring. Similarly, the theory of hypermatrices over semirings is more general than the theory of tensors over vector spaces. We return to working over fields now, but we will continue our study of hypermatrices over semirings when we discuss the absence of rank jumping in non-negative hypermatrices.

Though elegant, the proof of Theorem 5 is difficult to generalize to hypermatrices. We therefore now translate it into the coordinate-free setting of tensors, which lend themselves more readily to generalization. Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, and $n_3$-dimensional vector spaces over a common field, and let $V^{1*}$, $V^{2*}$, and $V^{3*}$ denote their dual spaces. The tensor $\rho = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \otimes v_i^3$ in $V^1 \otimes V^2 \otimes V^3$ induces the three linear maps

$$\begin{array}{lll} V^{1*} \xrightarrow{\Pi_1(\rho)} V^2 \otimes V^3 & V^{2*} \xrightarrow{\Pi_2(\rho)} V^1 \otimes V^3 & V^{3*} \xrightarrow{\Pi_3(\rho)} V^1 \otimes V^2 \\ v^{1*} \mapsto \displaystyle\sum_{i=1}^{r} v^{1*}(v_i^1)\ v_i^2 \otimes v_i^3, & v^{2*} \mapsto \displaystyle\sum_{i=1}^{r} v^{2*}(v_i^2)\ v_i^1 \otimes v_i^3, \quad \text{and} & v^{3*} \mapsto \displaystyle\sum_{i=1}^{r} v^{3*}(v_i^3)\ v_i^1 \otimes v_i^2. \end{array}$$

These maps are called the mode-1, mode-2, and mode-3 contraction maps of $\rho$, respectively.

**Theorem 6.** *The mode-$i$ contraction map $\Pi_i$ is well-defined.*

*Proof.* Suppose that $\sum_{t=1}^{r} v_t^1 \otimes v_t^2 \otimes v_t^3 = \sum_{t=1}^{s} w_t^1 \otimes w_t^2 \otimes w_t^3$ for vectors $v_t^i,\ w_t^i \in V^i$. We need to show that

$$\Pi_i \left( \sum_{t=1}^{r} v_t^1 \otimes v_t^2 \otimes v_t^3 \right) (v^{i*}) \ = \ \Pi_i \left( \sum_{t=1}^{s} w_t^1 \otimes w_t^2 \otimes w_t^3 \right) (v^{i*})$$

for every $v^{i*} \in V^{i*}$ for $i = 1, 2$ and 3. Choose bases $\{\mathrm{e}_t^i\}_{t=1}^{n_i}$ of $V^i$ with corresponding dual bases $\{\mathrm{e}_t^{i*}\}_{t=1}^{n_i}$ for each $i$. It is sufficient to show that

$$\Pi_1 \left( \sum_{t=1}^{r} v_t^1 \otimes v_t^2 \otimes v_t^3 \right) (\mathrm{e}_u^{1*}) \ = \ \Pi_1 \left( \sum_{t=1}^{s} w_t^1 \otimes w_t^2 \otimes w_t^3 \right) (\mathrm{e}_u^{1*})$$

for each $\mathrm{e}_u^{1*}$ in our dual basis. Let $a_{t,u}^i$ and $b_{t,u}^i$ be scalars such that $v_t^i = \sum_{u=1}^{n_i} a_{t,u}^i \mathrm{e}_u^i$ and $w_t^i = \sum_{u=1}^{n_i} b_{t,u}^i \mathrm{e}_u^i$ for all $i$ and $t$. It follows that $\sum_{t=1}^{r} v_t^1 \otimes v_t^2 \otimes v_t^3 = \sum_{t,i,j,k=1}^{r,n_1,n_2,n_3} a_{ti}^1 a_{tj}^2 a_{tk}^3\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$, which must equal $\sum_{t=1}^{s} w_t^1 \otimes w_t^2 \otimes w_t^3 = \sum_{t,i,j,k=1}^{s,n_1,n_2,n_3} b_{ti}^1 b_{tj}^2 b_{tk}^3\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$. Since a tensor is uniquely determined with respect to a basis, it follows that

$$\sum_{t=1}^{r} a_{ti}^1 a_{tj}^2 a_{tk}^3 \ = \ \sum_{t=1}^{s} b_{ti}^1 b_{tj}^2 b_{tk}^3 \text{ for all } i, j, k. \tag{2.3}$$

Hence, by the definition of the mode-1 contraction map

$$\begin{aligned} \Pi_1 \left( \sum_{t=1}^{r} v_t^1 \otimes v_t^2 \otimes v_t^3 \right) (\mathrm{e}_u^{1*}) \ &= \ \sum_{t,j,k=1}^{r,n_2,n_3} a_{tu}^1 a_{tj}^2 a_{tk}^3\ \mathrm{e}_j^2 \otimes \mathrm{e}_k^3 \\ &= \ \sum_{j,k=1}^{n_2,n_3} \left( \sum_{t=1}^{r} a_{tu}^1 a_{tj}^2 a_{tk}^3 \right)\ \mathrm{e}_j^2 \otimes \mathrm{e}_k^3. \end{aligned}$$

Equation (2.3) now implies that

$$
\begin{aligned}
\Pi_1\left(\sum_{t=1}^{r} v_t^1 \otimes v_t^2 \otimes v_t^3\right)(\mathrm{e}_u^{1*}) &= \sum_{j,k=1}^{n_2,n_3}\left(\sum_{t=1}^{s} b_{tu}^1 b_{tj}^2 b_{tk}^3\right)\ \mathrm{e}_j^2 \otimes \mathrm{e}_k^3 \\
&= \sum_{t,j,k=1}^{s,n_2,n_3} b_{tu}^1 b_{tj}^2 b_{tk}^3\ \mathrm{e}_j^2 \otimes \mathrm{e}_k^3 \\
&= \Pi_1\left(\sum_{t=1}^{s} w_t^1 \otimes w_t^2 \otimes w_t^3\right)(\mathrm{e}_u^{1*}).
\end{aligned}
$$

□

The rank of a two-fold tensor $\rho \in V^1 \otimes V^2$ is equal to the dimension of the image of $\Pi_1(\rho)$, which is also equal to the dimension of the image of $\Pi_2(\rho)$. Hence, Theorem 5 can be adapted to tensor spaces and contraction maps in the following way.

**Theorem 7.** *Let $V^1$ and $V^2$ be vectors spaces over a common field, and let $\rho \in V^1 \otimes V^2$. The dimension of the image of $\Pi_1(\rho)$ is equal to the dimension of the image of $\Pi_2(\rho)$.*

*Proof.* The dimension of the image of $\Pi_1(\rho)$ is equal to the minimum number $r$ such that there exists an $r$-dimensional vector space $U$ and linear maps $g$ and $h$ that make the following diagram commute.

$$
\begin{tikzcd}
V^{1*} \arrow[r, "\Pi_1(\rho)"] \arrow[dr, dashed, "g"'] & V^2 \\
& U \arrow[u, dashed, "h"'] \arrow[ul, phantom, "\circlearrowleft" near end, shift left=3]
\end{tikzcd}
\qquad \dim(U) = r
$$

Similarly, the dimension of the image of $\Pi_2(\rho)$ is equal to the minimum number $s$ such that there exists an $s$-dimensional vector space $W$ and linear maps $a$ and $b$ that make the following diagram commute.

$$
\begin{tikzcd}
V^{2*} \arrow[r, "\Pi_2(\rho)"] \arrow[dr, dashed, "a"'] & V^1 \\
& W \arrow[u, dashed, "b"'] \arrow[ul, phantom, "\circlearrowleft" near end, shift left=3]
\end{tikzcd}
\qquad \dim(W) = s
$$

However, the existence of an $r$-dimensional vector space $U$ and linear maps $g$ and $h$ that make the first diagram commute implies the existence of an $r$-dimensional vector space $U^*$, the dual space of $U$, and linear maps $\theta^{-1} \circ g^*$ and $h^*$ that make the second diagram commute, where

$$
\begin{aligned}
U^* &\xrightarrow{g^*} V^{1**} & V^{2*} &\xrightarrow{h^*} U^* & V^1 &\xrightarrow{\theta} V^{1**} \\
u^* &\mapsto \left(v^{1*} \mapsto u^* \circ g \circ v^{1*}\right), & v^{2*} &\mapsto \left(u \mapsto v^{2*} \circ h \circ u\right), \text{ and} & v^1 &\mapsto \left(v^{1*} \mapsto v^{1*}(v^1)\right).
\end{aligned}
$$

It follows that

$$
\begin{tikzcd}
V^{1*} \arrow[r, "\Pi_1(\rho)"] \arrow[dr, dashed, "g"'] & V^2 \arrow[d, phantom, "\circlearrowleft"' near start] & \Longrightarrow & V^{2*} \arrow[r, "\Pi_2(\rho)"] \arrow[dr, dashed, "h^*"'] & V^1 \arrow[d, phantom, "\circlearrowleft"' near start] \\
& U \arrow[u, dashed, "h"'] & & & U^* . \arrow[u, dashed, "\theta^{-1} \circ g^*"']
\end{tikzcd}
$$

Since the dimension of $U^*$ equals the dimension of $U$, the dimension of the image of $\Pi_2(\rho)$ must be less than or equal to the dimension of the image of $\Pi_1(\rho)$. Similarly,

$$
\begin{tikzcd}
V^{2*} \arrow[r, "\Pi_2(\rho)"] \arrow[dr, dashed, "a"'] & V^1 \arrow[d, phantom, "\circlearrowleft"' near start] & \Longrightarrow & V^{1*} \arrow[r, "\Pi_1(\rho)"] \arrow[dr, dashed, "b^*"'] & V^2 \arrow[d, phantom, "\circlearrowleft"' near start] \\
& W \arrow[u, dashed, "b"'] & & & W^* \arrow[u, dashed, "\Theta^{-1} \circ a^*"']
\end{tikzcd}
$$

where

$$
\begin{aligned}
W^{2*} &\xrightarrow{a^*} V^{2**} & V^{1*} &\xrightarrow{b^*} W^* & V^2 &\xrightarrow{\Theta} V^{2**} \\
w^* &\mapsto \left(v^{2*} \mapsto w^* \circ a \circ v^{2*}\right), & v^{1*} &\mapsto \left(w \mapsto v^{1*} \circ b \circ w\right), \text{ and} & v^2 &\mapsto \left(v^{2*} \mapsto v^{2*}(v^2)\right),
\end{aligned}
$$

which implies the theorem. $\square$

Unlike most proofs of row rank equals column rank [26, p.72] [27, p.37], our proof of Theorem 7 never invokes the rank nullity theorem. Our proof use the contraction maps of a tensor to derive relations between the slices of a hypermatrix. Thus, the relationship between the contraction maps of a tensor generalizes, in a coordinate-free way, the fundamental relationship between the rows and columns of a matrix to hypermatrices.

### 2.5 Two Generalizations of Row Rank Equals Column Rank to Hypermatrices

Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, and $n_3$-dimensional vector spaces, respectively, defined over a common field. In addition to the mode-1, 2, and 3 contraction maps, the tensor $\rho = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \otimes v_i^3$ in $V^1 \otimes V^2 \otimes V^3$ induces the linear map $\Pi^1(\rho)$, which is defined on simple tensors as

$$\begin{aligned} V^{2*} \otimes V^{3*} &\xrightarrow{\Pi^1(\rho)} V^1 \\ v^{2*} \otimes v^{3*} &\mapsto \sum_{i=1}^{r} \; v^{2*}(v_i^2) v^{3*}(v_i^3) \; v_i^1. \end{aligned}$$

The maps $\Pi^2(\rho)$ and $\Pi^3(\rho)$ are defined similarly. Note that these maps are denoted with a superscript rather than a subscript. In the case of 2-fold tensors, $\Pi^i(\rho)$ is the transpose of $\Pi_i(\rho)^T$ for $i = 1, 2$. In the 3-fold case and higher, equality does not hold. However, $\Pi^i(\rho)$ can still be identified isomorphically to the transpose of $\Pi_i(\rho)$. Hence, in the 3-fold case, the dimension of the image of $\Pi_i(\rho)$ is equal to the dimension of the image of $\Pi^i(\rho)$ for $i = 1, 2, 3$. In coordinates, this implies the following theorem.

**Theorem 8.** *Let $(a_{ijk})$ be a $n_1 \times n_2 \times n_3$ hypermatrix over some field.*

$$\begin{aligned} \dim\langle\, (a_{i::}) \,\rangle_{i=1}^{n_1} &= \dim\langle\, (a_{:jk}) \,\rangle_{j,k=1}^{n_2,n_3} \\ \dim\langle\, (a_{:j:}) \,\rangle_{j=1}^{n_2} &= \dim\langle\, (a_{i:k}) \,\rangle_{i,k=1}^{n_1,n_3} \\ \dim\langle\, (a_{::k}) \,\rangle_{k=1}^{n_3} &= \dim\langle\, (a_{ij:}) \,\rangle_{i,j=1}^{n_1,n_2}, \end{aligned}$$

*where $(a_{::t})$ is $n_1 n_2 \times 1$ vector of entries of $(a_{ijk})$ with the third coordinate fixed at $t$. That is, $(a_{::t}) = \begin{pmatrix} a_{11t} & a_{12t} & a_{13t} & \cdots & a_{21t} & a_{22t} & a_{23t} & \cdots & a_{n_1 n_2 t} \end{pmatrix}$. The $n_1 \times 1$ vector $(a_{:jk})$ is defined similarly.*

*Proof.* Let $\{(\mathrm{e}^i_j)\}_{j=1}^{n_i}$ be a basis of vector spaces $V^i$ for $i = 1, 2, 3$, and let $\{(\mathrm{e}^{i*}_j)\}_{j=1}^{n_i}$ denote the corresponding dual basis. Let $B$ denote the collection of these bases. For $\rho$ in $V^1 \otimes V^2 \otimes V^3$, there exists scalars $a_{ijk}$ such that

$$\rho \;=\; \sum_{i,j,k} a_{ijk}\ \mathrm{e}^1_i \otimes \mathrm{e}^2_j \otimes \mathrm{e}^3_k.$$

Let $[\rho]_B = (a_{ijk})$ be the hypermatrix defined by these scalars. The mode-3 contraction map of $\rho$ can also be coordinatized with respect to the same bases as

$$[\Pi_3(\rho)]_B = \begin{pmatrix} a_{111} & a_{121} & a_{131} & \cdots & a_{n_1n_21} \\ a_{112} & a_{122} & a_{122} & \cdots & a_{n_1n_22} \\ a_{113} & a_{123} & a_{133} & \cdots & a_{n_1n_23} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{11n_3} & a_{12n_3} & a_{13n_3} & \cdots & a_{n_1n_2n_3} \end{pmatrix}$$

$$= \begin{pmatrix} \rule[0.5ex]{2em}{0.4pt} & a_{::1} & \rule[0.5ex]{2em}{0.4pt} \\ \rule[0.5ex]{2em}{0.4pt} & a_{::2} & \rule[0.5ex]{2em}{0.4pt} \\ & \vdots & \\ \rule[0.5ex]{2em}{0.4pt} & a_{::n_3} & \rule[0.5ex]{2em}{0.4pt} \end{pmatrix} = \begin{pmatrix} \vert & \vert & & \vert \\ a_{11:} & a_{12:} & \cdots & a_{n_1n_2:} \\ \vert & \vert & & \vert \end{pmatrix}.$$

This demonstrates the desired relationship between the dimension-1 and co-dimension-1 slices of a hypermatrix. □

For 2-fold tensors, the equality of row rank and column rank of $n_1 \times n_2$ matrix $(a_{ijk})$ can be expressed as

$$\dim\langle\ (a_{i:})\ \rangle_{i=1}^{n_1} \;=\; \dim\langle\ (a_{:j})\ \rangle_{j=1}^{n_2}.$$

Thus, Theorem 8 is indeed a generalization of row rank equals column rank to 3-fold hypermatrices. For our second generalization, we introduce the notion of simple dimension. For a subset $W$ of tensor space $V^1 \otimes V^2 \otimes V^3$, the simple dimension of $W$, denoted sdim(W), is the minimum number of rank one tensors needed to span a space that contains $W$.

**Definition 9.** *Let $V^1$, $V^2$, and $V^3$ be finite dimensional vector spaces over a common field. For a subset $W$ of tensor space $V^1 \otimes V^2 \otimes V^3$, the simple dimension of $W$ is defined as*

$$sdim(W) = \min\{n \in \mathbb{N} \mid \; W \; \subseteq \; \langle v_1^1 \otimes v_1^2 \otimes v_1^3, \; v_2^1 \otimes v_2^2 \otimes v_2^3, \; \ldots, \; v_n^1 \otimes v_n^2 \otimes v_n^3 \; \rangle \quad \textit{for some } v_i^k \in V^k\}.$$

The simple dimension of a subspace may be strictly greater than its dimension as a vector space. For example, let $\{x_1^1, x_2^1\}$ and $\{y_1^1, y_2^1\}$ be linearly independent subsets of some vector space $V^1$, and let $\{x_1^2, x_2^2\}$ and $\{y_1^2, y_2^2\}$ be linearly independent subsets of some vector space $V^2$. The subspace $\langle x_1^1 \otimes x_1^2 + x_2^1 \otimes x_2^2 \rangle$ is a one-dimensional vector subspace of $V^1 \otimes V^2$, but has simple dimension two, since it is not generated by any rank one tensor. The rank of a tensor can be characterized in terms of the simple dimension of its contraction maps, which is a theorem from [14, p.68].

**Theorem 10.** *Let $V^1$, $V^2$, and $V^3$ be finite dimensional vector spaces over a common field, and let $\rho$ be a tensor in $V^1 \otimes V^2 \otimes V^3$. It follows that $rk(\rho) = sdim(\mathrm{im}(\Pi_i(\rho)))$ for any $i = 1$, $2$, or $3$.*

*Proof.* Without loss of generality, we prove the theorem for the mode-1 contraction. First, we show the inequality $\mathrm{sdim}(\mathrm{im}(\Pi_i(\rho))) \leq \mathrm{rk}(\rho)$. Suppose $\mathrm{rk}(\rho) = r$. There must then exist vectors $v_i^k$ such that $\rho = \sum_{i=1}^r v_i^1 \otimes v_i^2 \otimes v_i^3$. It follows that $\Pi_1(\rho)(v^{1*}) = \sum_{i=1}^r v^{1*}(v_i^1) \; v_i^2 \otimes v_i^3$ for every $v^{1*} \in V^{1*}$, so $\mathrm{im}(\Pi_1(\rho))$ is contained in $\langle \; v_1^2 \otimes v_1^3, \; v_2^2 \otimes v_2^3, \; \ldots, \; v_r^2 \otimes v_r^3 \; \rangle$. Hence, $\mathrm{sdim}(\mathrm{im}(\Pi_1(\rho)) \leq r$.

It remains to prove that $\mathrm{rk}(\rho) \leq \mathrm{sdim}(\mathrm{im}(\Pi_1(\rho)))$. Suppose that the simple dimension of $\mathrm{im}(\Pi_1(\rho))$ is $s$. It follows that there exists $w_i^k$ such that the image of $\Pi_1(\rho)$ is a subset of the span $\langle w_1^2 \otimes w_1^3, w_2^2 \otimes w_2^3, \ldots, w_s^2 \otimes w_s^3 \rangle$. Choose a basis $(\mathrm{e}_t^1)_{t=1}^{n_1}$ of $V^1$ with corresponding dual basis $(\mathrm{e}_t^{1*})_{t=1}^{n_1}$ of $V^{1*}$. For each $t \in \{1, 2, \ldots, n_1\}$, there must exist scalars $\{c_{i,t}\}_{i=1}^s$ such that $\Pi_1(\rho)(\mathrm{e}_t^{1*}) = \sum_{i=1}^s c_{i,t} \; w_i^2 \otimes w_i^3$. Note that

$$\Pi_1(\sum_{i,j=1}^{s,n_1} \mathrm{e}_j^1 \otimes c_{i,j} \; w_i^2 \otimes w_i^3)(\mathrm{e}_t^{1*}) = \sum_{i=1}^{s} c_{i,t} \; w_i^2 \otimes w_i^3 = \Pi_1(\rho)(\mathrm{e}_t^{1*})$$

for every $e_t^{1*}$, where $t = 1, 2, \ldots, n_1$. Since $\Pi_1$ is a vector space isomorphism, it follows that

$$\begin{aligned}
\rho &= \sum_{i,j=1}^{s,n_1} e_j^1 \otimes c_{i,j}\ w_i^2 \otimes w_i^3 \\
&= \sum_{i=1}^{s} (e_1^1 \otimes c_{i,1}\ w_i^2 \otimes w_i^3)\ +\ (e_2^1 \otimes c_{i,2}\ w_i^2 \otimes w_i^3)\ +\ \ldots\ +\ (e_{n_1}^1 \otimes c_{i,n_1}\ w_i^2 \otimes w_i^3) \\
&= \sum_{i=1}^{s} (c_{i,1}\ e_1^1 \otimes w_i^2 \otimes w_i^3)\ +\ (c_{i,2}\ e_2^1 \otimes w_i^2 \otimes w_i^3)\ +\ \ldots\ +\ (c_{i,n_1}\ e_{n_1}^1 \otimes w_i^2 \otimes w_i^3) \\
&= \sum_{i=1}^{s} (\sum_{j=1}^{n_1} c_{i,j} e_j^1) \otimes w_i^2 \otimes w_i^3, \text{ so } \mathrm{rk}(\rho) \leq s.
\end{aligned}$$

□

It follows that the simple dimensions of the row space, the column space, and the pillar space of a 3-fold hypermatrix are *all* equal. Furthermore, the simple dimension of all these spaces is equal to the rank of the hypermatrix. Since simple dimension and dimension are equivalent concepts for vector spaces, Theorem 10 is indeed a generalization of row rank equals column rank to hypermatrices.

# 3. COMPLEX GENERIC RANK

We give a proof of the existence of an open, dense set of constant rank complex $n_1 \times n_2 \times \cdots \times n_d$ tensors using Chevalley's theorem. We then contrast this with a more geometric proof of the same fact for complex $2\times2\times2$ tensors. Our geometric proof shows that a regular $2 \times 2 \times 2$ real or complex tensor is rank two if and only if its contraction maps intersect the $2\times2$ Segre variety at two distinct points. We further show that a regular tensor is tangent to the $2\times2\times2$ Segre variety if and only if the image of *any* of its contraction maps is tangent to the $2\times2$ Segre variety. This is a geometric example of the fundamental relationship between the contraction maps of a tensor.

## 3.1 Definition of Generic Rank

The observation that the number of solutions of a system of polynomial equations is often invariant with respect to perturbations of the coefficients of the polynomials is at the heart of algebraic geometry. We give an example of this phenomenon that can be visualized geometrically. Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, and $n_3$-dimensional complex vector spaces, respectively, and let $V^1 \otimes V^2 \otimes V^3$ denote the tensor product of these spaces. The tensors in the form $v^1 \otimes v^2 \otimes v^3$ for some vectors $v^1$, $v^2$, and $v^3$ are called simple tensors. The rank of a tensor $\rho$ is the minimum number $n$ such that $\rho$ is the sum of $n$ simple tensors.

It is a remarkable observation that there exists an open, dense set of constant rank tensors. That is, there exists a unique natural number $r$ such that there exists a set of rank $r$ tensors that is open and dense with respect to the norm topology. This $r$ is called the generic rank of the space $V^1 \otimes V^2 \otimes V^3$.

Characterizing a set as 'open' and 'dense' is a topological formalization that the set contains 'almost all' of the elements of the space. Intuitively, open sets reach in all directions from their interior points, and dense sets approach all of the points in the space. The rank of a tensor is a natural measurement of its complexity. Hence, the existence of an open and dense set of rank $r$ tensors implies that almost all tensors are of the same complexity. The existence of a generic rank is true for complex tensor spaces of arbitrary dimensions

$n_1 \times n_2 \times \cdots \times n_d$. However, it is particular to complex tensor spaces. There does *not* exist such an open, dense set of constant rank tensors in real tensor spaces in general.

Once bases of the vector spaces $V^1$, $V^2$, and $V^3$ are chosen, a tensor in $V^1 \otimes V^2 \otimes V^3$ can be coordinatized as a hypermatrix in $\mathbb{C}^{n_1 \times n_2 \times n_3}$. The generic rank of $\mathbb{C}^{2 \times n \times n}$ is $n$. For the $2 \times 2 \times 2$ case, the existence of a generic rank of two implies that there is an open, dense set of $2 \times 2 \times 2$ complex hypermatrices in the form

$$\begin{array}{l} \boxed{\begin{matrix} a_2^1 a_1^2 a_1^3 + b_2^1 b_1^2 b_1^3 & a_2^1 a_1^2 a_2^3 + b_2^1 b_1^2 b_2^3 \\ a_2^1 a_2^2 a_1^3 + b_2^1 b_2^2 b_1^3 & a_2^1 a_2^2 a_2^3 + b_2^1 b_2^2 b_2^3 \end{matrix}} \\ \qquad\qquad \boxed{\begin{matrix} a_1^1 a_1^2 a_1^3 + b_1^1 b_1^2 b_1^3 & a_1^1 a_1^2 a_2^3 + b_1^1 b_1^2 b_2^3 \\ a_1^1 a_2^2 a_1^3 + b_1^1 b_2^2 b_1^3 & a_1^1 a_2^2 a_2^3 + b_1^1 b_2^2 b_2^3 \end{matrix}} \end{array} \tag{3.1}$$

for some complex numbers $a_i^j$ and $b_i^j$. Choosing the entries of a matrix independently and 'randomly' will almost always result in a matrix of maximal possible rank. This seems intuitive, since choosing the entries of a matrix randomly and independently should almost never result in a relation among the rows or columns. It is less intuitive, however, why choosing the entries of a $2 \times n \times n$ hypermatrix randomly will almost always result in a hypermatrix of rank $n$, which is in general not the maximal rank of $2 \times n \times n$ hypermatrices.

The fact that almost all complex tensors have the same rank is often stated without proof in textbooks on tensors. We thus give a proof of this fact using Chevalley's theorem. We then give an alternative geometric proof that almost all $2 \times 2 \times 2$ complex hypermatrices are rank two. We show that a regular $2 \times 2 \times 2$ complex tensor is rank two if and only if the projective image of any of its mode-1 contraction maps intersect the $2 \times 2$ projective Segre variety at two distinct points. Since the $2 \times 2$ projective Segre variety is degree two, Bézout's theorem implies this occurs generically, thus explaining the existence of a complex generic rank. Our proof also illustrates an additional geometric relationship between the mode-1, 2, and 3 contraction maps of regular tensors.

## 3.2 The Sets of Constant Rank Complex Tensors are Constructible

We prove the existence of a unique complex generic rank by first showing that the sets of constant rank tensors are constructible and then invoking Chevalley's theorem on constructible sets. We present this proof to contrast it with our more geometric proof in later sections.

The Zariski closed sets are the sets of solutions of polynomial equations over the complex numbers. Zariski open sets are complements of Zariski closed sets. Constructible sets are the Boolean algebra generated by Zariski open sets and Zariski closed sets. That is, constructible sets are the closure of Zariski open and Zariski closed sets under finite union and complementation. A set is constructible if and only if it is open in its closure. It is a remarkable observation that constructible sets always contain an open, dense subset of their closure. We will prove a version of this statement specifically for tensor spaces below.

Let $\otimes : \mathbb{C}^2 \times \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}^{2\times 2\times 2}$ denote the $2 \times 2 \times 2$ coordinate tensor map defined as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \times \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mapsto \begin{array}{c} \boxed{\begin{matrix} x_2y_1z_1 & x_2y_1z_2 \\ x_2y_2z_1 & x_2y_2z_2 \end{matrix}} \\ \boxed{\begin{matrix} x_1y_1z_1 & x_1y_1z_2 \\ x_1y_2z_1 & x_1y_2z_2 \end{matrix}} \end{array}.$$

The elements in the image of the coordinate tensor map are called simple hypermatrices and are denoted as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \otimes \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} := \otimes\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \times \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right). \tag{3.2}$$

The $n_1 \times n_2 \times \cdots \times n_d$ coordinate tensor map is defined in the same way. The hypermatrix (3.1) can thus be written as

$$\begin{pmatrix} a_1^1 \\ a_2^1 \end{pmatrix} \otimes \begin{pmatrix} a_1^2 \\ a_2^2 \end{pmatrix} \otimes \begin{pmatrix} a_1^2 \\ a_2^2 \end{pmatrix} + \begin{pmatrix} b_1^1 \\ b_2^1 \end{pmatrix} \otimes \begin{pmatrix} b_1^2 \\ b_2^2 \end{pmatrix} \otimes \begin{pmatrix} b_1^2 \\ b_2^2 \end{pmatrix}.$$

**Theorem 11.** *Let $V^1$, $V^2$, ..., $V^d$ be finite dimensional complex vector spaces. For each natural number $s$, the set $M_s^{\mathbb{C}} = \{\rho \ \in V^1 \otimes V^2 \otimes \cdots \otimes V^d \ \mid \ rk(\rho) = s\}$ is constructible with respect to the Zariski topology.*

*Proof.* We prove this theorem in coordinates using hypermatrices instead of tensors. Chevalley's theorem on constructible sets implies that the image of a polynomial map defined on a constructible set is also a constructible set [28, p.94]. It is thus sufficient to show that the sets of constant rank complex hypermatrices can be defined in terms of polynomial maps defined on constructible sets. We demonstrate this in the $2 \times 2 \times 2$ case for readability. Let $\psi_s^{\mathbb{C}}$ be the map from $(\mathbb{C}^2 \times \mathbb{C}^2 \times \mathbb{C}^2)^s \cong \mathbb{C}^{6s}$ to $\mathbb{C}^{2\times 2\times 2}$ that sends the Cartesian product of vectors

$$\left(\begin{pmatrix} a_{111}^1 \\ a_{211}^1 \end{pmatrix} \times \begin{pmatrix} a_{121}^1 \\ a_{221}^1 \end{pmatrix} \times \begin{pmatrix} a_{112}^1 \\ a_{212}^1 \end{pmatrix}\right) \times \ \dots \ \times \left(\begin{pmatrix} a_{111}^s \\ a_{211}^s \end{pmatrix} \times \begin{pmatrix} a_{121}^s \\ a_{222}^s \end{pmatrix} \times \begin{pmatrix} a_{112}^s \\ a_{212}^s \end{pmatrix}\right)$$

to the hypermatrix

$$\begin{pmatrix} a_{111}^1 \\ a_{211}^1 \end{pmatrix} \otimes \begin{pmatrix} a_{121}^1 \\ a_{221}^1 \end{pmatrix} \otimes \begin{pmatrix} a_{112}^1 \\ a_{212}^1 \end{pmatrix} \ + \ \dots \ + \begin{pmatrix} a_{111}^s \\ a_{211}^s \end{pmatrix} \otimes \begin{pmatrix} a_{121}^s \\ a_{221}^s \end{pmatrix} \otimes \begin{pmatrix} a_{112}^s \\ a_{212}^s \end{pmatrix},$$

which is equal to

$$\begin{array}{ll} \boxed{\begin{matrix} \sum_{i=1}^s a_{211}^i a_{121}^i a_{112}^i & \sum_{i=1}^s a_{211}^i a_{121}^i a_{212}^i \\ \sum_{i=1}^s a_{211}^i a_{221}^i a_{211}^i & \sum_{i=1}^s a_{211}^i a_{221}^i a_{212}^i \end{matrix}} & \\ & \boxed{\begin{matrix} \sum_{i=1}^s a_{111}^i a_{121}^i a_{112}^i & \sum_{i=1}^s a_{111}^i a_{121}^i a_{212}^i \\ \sum_{i=1}^s a_{111}^i a_{221}^i a_{211}^i & \sum_{i=1}^s a_{111}^i a_{221}^i a_{212}^i \end{matrix}} \end{array}. \tag{3.3}$$

Hypermatrix (3.3) has polynomial entries, so $\psi_s^{\mathbb{C}}$ is indeed a polynomial mapping. Hence, by Chevalley's theorem, the set $M_{\leq s}^{\mathbb{C}} = \{\rho \ \in V^1 \otimes V^2 \otimes V^3 \ \mid \ \mathrm{rk}(\rho) \leq s\}$ is constructible as it is the image of the polynomial map $\psi_s^{\mathbb{C}}$ defined on the constructible set $\mathbb{C}^{6s}$. Moreover, $M_s^{\mathbb{C}} = M_{\leq s}^{\mathbb{C}} \setminus M_{\leq s-1}^{\mathbb{C}}$ is constructible since constructible sets are closed under set complements. $\square$

We now show that Zariski constructible sets contain an open, dense subset of their Zariski closure. This in fact implies the existence of an open, dense set of constant rank tenors in the norm topology. Our proof modifies the proof in [29, Lem.2.1] to tensors spaces. We will also use the following two facts. Firstly, Zariski open sets are open and dense in the norm topology. Secondly, the tensor space $V^1 \otimes V^2 \otimes \cdots \otimes V^d$ is irreducible with respect to the Zariski topology. A set is irreducible if it cannot be written as the union of two proper closed sets. The tensor space $V^1 \otimes V^2 \otimes \cdots \otimes V^d$ is irreducible as it can be identified with the affine variety $\mathbb{C}^{n_1 n_2 \dots n_d}$ when $V^i$ is $n_i$-dimensional.

**Theorem 12.** *Let $V^1 \otimes V^2 \otimes \cdots \otimes V^d$ be a complex tensor space endowed with the norm topology. There exists a unique natural number $r$ such that there exists an open, dense set of rank $r$ tensors.*

*Proof.* Let $V = V^1 \otimes V^2 \otimes \cdots \otimes V^d$, and let $m$ be the maximal possible rank of tensors in $V$. Let $M_s^{\mathbb{C}}$ denote the set of rank $s$ tensors. Since $V = M_0^{\mathbb{C}} \cup M_2^{\mathbb{C}} \cup \ldots \cup M_m^{\mathbb{C}}$, it follows that

$$V = \widetilde{M_0^{\mathbb{C}}} \cup \widetilde{M_2^{\mathbb{C}}} \cup \ldots \cup \widetilde{M_m^{\mathbb{C}}},$$

where $\widetilde{M_i^{\mathbb{C}}}$ denotes the Zariski closure of $M_i^{\mathbb{C}}$. Since $V$ is irreducible with respect to the Zariski topology, there must thus exist a unique $r$ such that $\widetilde{M_r^{\mathbb{C}}} = V$. Furthermore, $M_r^{\mathbb{C}}$ is constructible by Theorem 11. There must thus exists sets $Y_1$, $Y_2$, ..., $Y_j$ that are each open in their closure such that $M_r^{\mathbb{C}}$ is equal to the union of all the $Y_i$'s. It follows that the set

$$V \setminus \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right) \tag{3.4}$$

is Zariski open, and thus open and dense in the norm topology. The set (3.4) is our open, dense set of rank $r$ tensors. It remains to show that (3.4) is in fact contained in $M_r^{\mathbb{C}}$, which follows from the following equalities and inclusion.

$$\begin{aligned}
V \setminus \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right) \;=\; \widetilde{M_r^{\mathbb{C}}} \setminus \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right) &= \widetilde{\left(\bigcup_{i=1}^{j} Y_i\right)} \setminus \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right) \\
&= \left(\bigcup_{i=1}^{j} \widetilde{Y_i}\right) \;\setminus\; \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right) \\
&= \bigcup_{i=1}^{j} \left(Y_i \;\cup\; \widetilde{Y_i} \setminus Y_i\right) \;\setminus\; \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right) \\
&= \left(\bigcup_{i=1}^{j} Y_i \;\cup\; \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right)\right) \;\setminus\; \bigcup_{i=1}^{j} \left(\widetilde{Y_i} \setminus Y_i\right) \\
&\subseteq \bigcup_{i=1}^{j} Y_i \;=\; M_r^{\mathbb{C}}
\end{aligned}$$

□

The proof of Theorem 11 is almost identical to the proof in [30, Thm.6.1] that the sets $M_s^{\mathbb{R}}$ are semialgebraic. The only difference is that Chevalley's theorem is replaced by the Tarski-Seidenberg theorem. An important difference, however, between constructible sets and semialgebraic sets is that semialgebraic sets do not generally contain an open, dense subset of their Zariski closure. This is what prevents us from extending the proof of Theorem 12 to real-valued hypermatrices. Indeed, there does not exist an open, dense set of constant rank real-valued hypermatrices in general. In the following sections, we give a geometric explanation for why no such set exists for real valued $2 \times 2 \times 2$ hypermatrices.

### 3.3 Computing the Generic Rank and the Maximal Rank of Tensor Spaces

Determining the generic rank of the complex $n_1 \times n_2 \times \cdots \times n_d$ tensor space is an open problem in general. The following intuitive argument is often used to calculate the expected generic rank (for example, [14, p.70]). Choose a basis $\{\mathrm{e}^i_j\}_{j=1}^{n_i}$ of $V^i$ for $i = 1, 2, 3$. A rank one tensor $v^1 \otimes v^2 \otimes v^3$ in $V^1 \otimes V^2 \otimes V^3$ can be written as

$$(a^1_1 \mathrm{e}^1_1 + \cdots + a^1_{n_1} \mathrm{e}^1_{n_1}) \otimes (a^2_1 \mathrm{e}^2_1 + \cdots + a^2_{n_2} \mathrm{e}^2_{n_2}) \otimes (a^3_1 \mathrm{e}^3_1 + \cdots + a^3_{n_3} \mathrm{e}^3_{n_3}) \tag{3.5}$$

for some parameters $a^j_i$. As a rank one tensor, each vector $v^1$, $v^2$, and $v^3$ must contain at least one nonzero coordinate. By multilinearity, one of the nonzero coordinates of each vector $v^1$, $v^2$, and $v^3$ can be normalized to 1. For example, if $a^1_1$, $a^2_1$, and $a^3_1$ are all nonzero, then (3.5) becomes

$$b\ (\mathrm{e}^1_1 + \frac{a^1_2}{a^1_1} \mathrm{e}^1_2 + \cdots + \frac{a^1_{n_1}}{a^1_1} \mathrm{e}^1_{n_1}) \otimes (\mathrm{e}^2_1 + \frac{a^2_2}{a^2_1} \mathrm{e}^2_2 + \cdots + \frac{a^2_{n_2}}{a^2_1} \mathrm{e}^2_{n_2}) \otimes (\mathrm{e}^3_1 + \frac{a^3_2}{a^3_1} \mathrm{e}^3_2 + \cdots + \frac{a^3_{n_3}}{a^3_1} \mathrm{e}^3_{n_3}),$$

where $b = a^1_1 a^2_1 a^3_1$. One thus expects the set of rank one tensors to be parameterized by

$$(n_1 - 1) + (n_2 - 1) + (n_3 - 1) + 1 \ = \ n_1 + n_2 + n_3 - 2$$

parameters. Similarly, one expects to need $s\,(n_1 + n_2 + n_3 - 2)$ parameters for the set of rank less than or equal to $s$ tensors $M^{\mathbb{C}}_{\leq s}$. Since the ascending chain

$$M^{\mathbb{C}}_{\leq 1} \ \subseteq \ M^{\mathbb{C}}_{\leq 2} \ \subseteq \ M^{\mathbb{C}}_{\leq 3} \ \subseteq \ M^{\mathbb{C}}_{\leq 4} \ \subseteq \ \ldots$$

eventually stabilizes, the generic rank must be the minimum $s$ such that the dimension of $M^{\mathbb{C}}_{\leq s}$ is equal to the dimension of $V^1 \otimes V^2 \otimes V^3$. Hence, one expects the generic rank to be equal to the minimum natural number $s$ such that $s(n_1 + n_2 + n_3 - 2) = n_1 n_2 n_3$. The expected complex generic rank of the space of $n_1 \times n_2 \times n_3$ complex tensors is thus

$$\lceil \frac{n_1 n_2 n_3}{n_1 + n_2 + n_3 - 2} \rceil.$$

The generic rank for complex $n \times n \times n$ tensors is indeed the expected rank for all $n$, except for $n = 3$. The expected rank for $3 \times 3 \times 3$ tensors is four, but the actual generic rank is five [14, p.70]. The maximal rank is, in general, much larger than the generic rank. We will compute the maximal rank of $2\times 2\times 2$ tensors now. We again let $V^i$ be two-dimensional vector spaces over a common field for $i=1,2,3$, and choose bases $\{\mathrm{e}^i_j\}_{j=1}^2$ for $i=1,2,3$. For a tensor $\rho$ in $V^1\otimes V^2\otimes V^3$, there are constants $a_{ijk}$ such that $\rho=\sum_{ijk} a_{ijk}\ \mathrm{e}^1_i\otimes\mathrm{e}^2_j\otimes\mathrm{e}^3_k$. The multilinearity of the tensor product implies the rank of a $2\times 2\times 2$ tensor is at most four, since

$$\begin{aligned}\rho \ &= \ \mathrm{e}^1_1\otimes\mathrm{e}^2_1\otimes(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2) \ + \ \mathrm{e}^1_1\otimes\mathrm{e}^2_2\otimes(a_{121}\mathrm{e}^3_1+a_{122}\mathrm{e}^3_2)\\ &\quad + \ \mathrm{e}^1_2\otimes\mathrm{e}^2_1\otimes(a_{211}\mathrm{e}^3_1+a_{212}\mathrm{e}^3_2) \ + \ \mathrm{e}^1_2\otimes\mathrm{e}^2_2\otimes(a_{221}\mathrm{e}^3_1+a_{222}\mathrm{e}^3_2).\end{aligned}$$

If the vectors $(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2)$ and $(a_{121}\mathrm{e}^3_1+a_{122}\mathrm{e}^3_2)$ are dependent, then there exists a constant $k$ such that

$$(a_{121}\mathrm{e}^3_1+a_{122}\mathrm{e}^3_2) \ = \ k(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2).$$

It follows that

$$\begin{aligned}\rho = \ &\mathrm{e}^1_1\otimes\mathrm{e}^2_1\otimes(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2) \ + \ \mathrm{e}^1_1\otimes\mathrm{e}^2_2\otimes k(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2)\\ &\quad + \ \mathrm{e}^1_2\otimes\mathrm{e}^2_1\otimes(a_{211}\mathrm{e}^3_1+a_{212}\mathrm{e}^3_2) \ + \ \mathrm{e}^1_2\otimes\mathrm{e}^2_2\otimes(a_{221}\mathrm{e}^3_1+a_{222}\mathrm{e}^3_2)\\ = \ &\mathrm{e}^1_1\otimes(\mathrm{e}^2_1+k\mathrm{e}^2_2)\otimes(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2) \ + \ \mathrm{e}^1_2\otimes\mathrm{e}^2_1\otimes(a_{211}\mathrm{e}^3_1+a_{212}\mathrm{e}^3_2)\\ &\quad + \ \mathrm{e}^1_2\otimes\mathrm{e}^2_2\otimes(a_{221}\mathrm{e}^3_1+a_{222}\mathrm{e}^3_2).\end{aligned}$$

Hence, $\rho$ is at most rank three in this case. If the vectors $(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2)$ and $(a_{121}\mathrm{e}^3_1+a_{122}\mathrm{e}^3_2)$ are independent, on the other hand, then they span $V^3$, so there exists constants $\alpha$, $\beta$, $\gamma$, $\delta$ such that

$$\begin{aligned}(a_{211}\mathrm{e}^3_1+a_{212}\mathrm{e}^3_2) \ &= \ \alpha(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2) \ + \ \beta(a_{121}\mathrm{e}^3_1+a_{122}\mathrm{e}^3_2), \text{and}\\ (a_{221}\mathrm{e}^3_1+a_{222}\mathrm{e}^3_2) \ &= \ \gamma(a_{111}\mathrm{e}^3_1+a_{112}\mathrm{e}^3_2) \ + \ \delta(a_{121}\mathrm{e}^3_1+a_{122}\mathrm{e}^3_2).\end{aligned}$$

It follows that $\rho$ is rank at most three by the not so obvious observation that

$$\begin{aligned}\rho \;=\; & (\mathrm{e}_1^1 + (\alpha-\beta)\mathrm{e}_2^1) \otimes \mathrm{e}_1^2 \otimes (a_{111}\mathrm{e}_1^3 + a_{112}\mathrm{e}_2^3) \\ & + \; (\mathrm{e}_1^1 + (\delta-\gamma)\mathrm{e}_2^1) \otimes \mathrm{e}_2^2 \otimes (a_{121}\mathrm{e}_1^3 + a_{122}\mathrm{e}_2^3) \\ & + \; \mathrm{e}_2^1 \otimes (\beta\mathrm{e}_1^2 + \gamma\mathrm{e}_2^2) \otimes \left((a_{111}\mathrm{e}_1^3 + a_{112}\mathrm{e}_2^3) + (a_{121}\mathrm{e}_1^3 + a_{122}\mathrm{e}_2^3)\right).\end{aligned} \tag{3.6}$$

We emphasize that relation (3.6) was not obvious, and the difficulty of finding such relations is a good illustration of why determining the maximal rank of tensor spaces is a difficult and open problem in general.

## 3.4 Characterizing Regular $2 \times 2 \times 2$ Tensors over $\mathbb{C}$ in terms of Contraction Maps

We will now give an alternative geometric proof that there is an open, dense set of rank two tensors in the $2 \times 2 \times 2$ complex tensor space. Let $V^1$, $V^2$, and $V^3$ be two-dimensional complex vector spaces, and let $V^{1*}$, $V^{2*}$, and $V^{3*}$ denote their dual spaces. Recall that the tensor $\rho = \sum_{i=1}^r v_i^1 \otimes v_i^2 \otimes v_i^3$ in $V^1 \otimes V^2 \otimes V^3$ induces the three linear maps

$$\begin{array}{ccc} V^{1*} \xrightarrow{\Pi_1(\rho)} V^2 \otimes V^3 & V^{2*} \xrightarrow{\Pi_2(\rho)} V^1 \otimes V^3 & V^{3*} \xrightarrow{\Pi_3(\rho)} V^1 \otimes V^2 \\ v^{1*} \mapsto \displaystyle\sum_{i=1}^r \; v^{1*}(v_i^1)\; v_i^2 \otimes v_i^3, & v^{2*} \mapsto \displaystyle\sum_{i=1}^r \; v^{2*}(v_i^2)\; v_i^1 \otimes v_i^3, \quad \text{and} & v^{3*} \mapsto \displaystyle\sum_{i=1}^r \; v^{3*}(v_i^3)\; v_i^1 \otimes v_i^2. \end{array}$$

These maps are called the mode-1, mode-2, and mode-3 contraction maps of $\rho$, respectively. A tensor is said to be regular if all of its contraction maps are full-rank. If the entries of a hypermatrix are selected by independent random variables that are absolutely continuous with respect to Lebesgue measure, then the set of hypermatrices with a nontrivial relation among its codimension-1 slices is zero, as expected. Regular tensors are precisely those tensors for which no such relations exist. Over $\mathbb{C}$, the set of regular tensors indeed is open and dense in the norm topology. Hence, in order to prove that there is an open, dense set of rank two tensors in $V^1 \otimes V^2 \otimes V^3$, it is sufficient to prove that the set of regular rank two tensors is open and dense in $V^1 \otimes V^2 \otimes V^3$.

The relationship between the contraction maps of a tensor generalizes, in a coordinate-free way, the fundamental relationship between the rows and columns of a matrix to hypermatrices. The rank of a two-fold tensor $\rho \in V^1 \otimes V^2$ is equal to the dimension of the image of $\Pi_1(\rho)$, which is also equal to the dimension of the image of $\Pi_2(\rho)$. This is generalizes to three-fold tensors in the following way [14, p.68].

**Theorem 13.** *Let $V^1$, $V^2$, and $V^3$ be finite dimensional vector spaces over a common field. For $\rho \in V^1 \otimes V^2 \otimes V^3$,*

$$\begin{aligned} rk(\rho) = \min\{n \in \mathbb{N} \mid image(\Pi_i(\rho)) \subseteq \ \langle v_1^1 \otimes v_1^2 \otimes v_1^3,\ v_2^1 \otimes v_2^2 \otimes v_2^3,\ \ldots,\ v_n^1 \otimes v_n^2 \otimes v_n^3 \ \rangle \\ \textit{for some } v_i^k \in V^k\} \end{aligned}$$

*for any $i = 1,\ 2,$ or $3$.*

For regular $2 \times 2 \times 2$ tensors, Theorem 13 becomes the geometric statement that the rank of regular tensor $\rho$ is equal to the cardinally of the intersection of the projective image of any of its contraction maps with the projective Segre variety. The Segre variety is the image of the map

$$\begin{aligned} V^1 \times V^2 \times V^3 \ &\rightarrow \ V^1 \otimes V^2 \otimes V^3 \\ v^1 \times v^2 \times v^3 \ &\mapsto \ v^1 \otimes v^2 \otimes v^3. \end{aligned}$$

The Segre variety of the $2 \times 2 \times 2$ complex tensor space is denoted $X_{\mathbb{C}}^{2\times2\times2}$ and the Segre variety of the $2 \times 2$ complex tensor space is denoted $X_{\mathbb{C}}^{2\times2}$. The $2 \times 2 \times 2$ Segre variety over an arbitrary field is denoted $X^{2\times2\times2}$, or simply as $X$ if the dimensions are clear from context. The tangent space of the Segre variety at a point $v$ is denoted as $T_v(X)$, and is characterized in [31] as in the following theorem.

**Theorem 14.** *Let $x_1^1 \otimes x_1^2 \otimes x_1^3$ be a rank one tensor in the Segre variety $X$. The tangent space of $X$ at $x_1^1 \otimes x_1^2 \otimes x_1^3$ is the space of all tensors in the form*

$$x_2^1 \otimes x_1^2 \otimes x_1^3 \ + \ x_1^1 \otimes x_2^2 \otimes x_1^3 \ + \ x_1^1 \otimes x_1^2 \otimes x_2^3 \tag{3.7}$$

*for some* $x_2^1 \in V^1$, $x_2^2 \in V^2$, *and* $x_2^3 \in V^3$.

The projective Segre variety is the the projective image of the rank one tensors. That is, it is the image of the injective Segre map defined as

$$\begin{aligned} \mathbb{P}(V^1) \times \mathbb{P}(V^2) \times \mathbb{P}(V^3) &\rightarrow \mathbb{P}(V^1 \otimes V^2 \otimes V^3) \\ [v^1] \times [v^2] \times [v^3] &\mapsto [v^1 \otimes v^2 \otimes v^3], \end{aligned}$$

where $[v]$ indicates the projectivization of vector $v$, and $\mathbb{P}(V^1)$ denotes the projectivization of vector space $V^1$. The $2 \times 2 \times 2$ complex projective Segre variety is denoted as $\mathbb{P}(X_{\mathbb{C}}^{2\times2\times2})$.

Regular $2 \times 2 \times 2$ complex tensors can be either rank two or three. We will show that a regular tensor $\rho$ is rank two if and only if $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ intersects $\mathbb{P}(X_{\mathbb{C}}^{2\times2})$ at two distinct points each with multiplicity one for any $j = 1, 2, 3$. Additionally, $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ intersects $\mathbb{P}(X_{\mathbb{C}}^{2\times2})$ at one point with multiplicity two for any $j = 1, 2, 3$ if and only if the rank of $\rho$ is three. This result is summarized in Figure 3.1 and proven in the following theorems.

**Lemma 15.** *Let* $V^1$, $V^2$, *and* $V^3$ *be two-dimensional vector spaces over a common field, and let* $\rho \in V^1 \otimes V^2 \otimes V^3$ *be a regular tensor. If* $\mathbb{P}(im\Pi_j(\rho))$ *intersects* $\mathbb{P}(X^{2\times2})$ *for any* $j$*, then* $\rho$ *is rank less than or equal to three.*

*Proof.* Suppose $\mathbb{P}(\mathrm{im}\Pi_1(\rho))$ intersects $\mathbb{P}(X^{2\times2})$ at some point $[x_1^2 \otimes x_1^3]$ for some vectors $x_1^2 \in V^2$ and $x_1^3 \in V^3$. Extend these vectors to a basis $\{x_1^2, x_2^2\}$ of $V^2$ and a basis $\{x_1^3, x_2^3\}$ of $V^3$. There must also be a basis $\{x_1^1, x_2^1\}$ of $V^1$ with corresponding dual basis $\{x_1^{1*}, x_2^{1*}\}$ such that

$$x_1^{1*} \mapsto x_1^2 \otimes x_1^3 \tag{3.8}$$

$$\begin{aligned} x_2^{1*} \mapsto\ & a\, x_1^2 \otimes x_1^3 + b\, x_1^2 \otimes x_2^3 + c\, x_2^2 \otimes x_1^3 + d\, x_2^2 \otimes x_2^3 \\ =\ & x_1^2 \otimes (a x_1^3 + b x_2^3) + x_2^2 \otimes (c x_1^3 + d x_2^3) \end{aligned} \tag{3.9}$$

| | | |
|---|---|---|
| $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ intersects $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ at two distinct points each with multiplicity one for any $j = 1, 2, 3$. | $\iff$ | $\mathrm{rk}(\rho) = 2$ |
| $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ intersects $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ at one point with multiplicity two for any $j = 1, 2, 3$. | $\iff$ | $\mathrm{rk}(\rho) = 3$ |

**Figure 3.1.** If $\rho$ is a regular complex $2 \times 2 \times 2$ tensor, then one of the above two conditions holds.

for some scalars $a, b, c, d$. Equation (3.8) follows from the fact that $x_1^2 \otimes x_1^3$ is in the image of $\Pi_1(\rho)$ by assumption, and equation (3.9) follows from the fact that $\{x_i^2 \otimes x_j^3\}_{i,j=1}^{2,2}$ is a basis of $V^2 \otimes V^3$. Hence,

$$\rho \;=\; x_1^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_1^2 \otimes (ax_1^3 + bx_2^3) \;+\; x_2^1 \otimes x_2^2 \otimes (cx_1^3 + dx_2^3),$$

as tensors are completely determined by the values of one of their contraction maps on a basis. Thus, the rank of $\rho$ is less than or equal to three. $\square$

**Lemma 16.** *Let $V^1$, $V^2$, and $V^3$ be two-dimensional vector spaces over a common field, and let $\rho$ be a tensor in $V^1 \otimes V^2 \otimes V^3$. If $\rho$ is rank two and regular, then for any rank two decomposition of $\rho$*

$$\rho \;=\; v_1^1 \otimes v_1^2 \otimes v_1^3 \;+\; v_2^1 \otimes v_2^2 \otimes v_2^3,$$

*the sets $\{v_1^i, v_2^i\}$ are linearly independent for $i = 1, 2, 3$.*

*Proof.* We will suppose for contradiction that the set $\{v_1^1, v_2^1\}$ is not independent. Suppose that there existed some scalar $k$ such that $v_2^1 = kv_1^1$. Extend the set $\{v_1^1\}$ to a basis $\{v_1^1, w_2^1\}$ of $V^1$ for some vector $w_2^1$, and let $\{v_1^{1*}, w_2^{1*}\}$ denote the corresponding dual basis. The action of the mode-1 contraction of $\rho$, $\Pi_1(\rho)$, on this basis is then

$$\begin{aligned} v_1^{1*} \;&\mapsto\; v_1^2 \otimes v_1^3 \;+\; kv_2^2 \otimes v_2^3 \\ w_2^{1*} \;&\mapsto\; 0. \end{aligned}$$

Thus, $\Pi_1(\rho)$ is not rank two, contradicting our assumption that $\rho$ is regular. $\square$

**Theorem 17.** *Let $V^1$, $V^2$, and $V^3$ be two-dimensional complex vector spaces. For a regular tensor $\rho$ in $V^1 \otimes V^2 \otimes V^3$, the following are equivalent:*

*(1) For some $j = 1,2,$ or $3$, $\mid \mathbb{P}(im\Pi_j(\rho)) \ \cap \ \mathbb{P}(X_{\mathbb{C}}^{2\times 2}) \mid = \ 2$*

*(2) $rk(\rho) = 2$*

*(3) For all $j = 1,2,$ or $3$, $\mid \mathbb{P}(im\Pi_j(\rho)) \ \cap \ \mathbb{P}(X_{\mathbb{C}}^{2\times 2}) \mid = \ 2$*

*Proof.* (1) $\Rightarrow$ (2) : Suppose without loss of generality that $\rho$ is a regular tensor such that $\mathbb{P}(\mathrm{im}\Pi_1(\rho))$ intersects $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ at the two distinct projective points $[v_1^2 \otimes v_1^3]$ and $[v_2^2 \otimes v_2^3]$ for some vectors $v_j^i$. The projective line $\mathbb{P}(\mathrm{im}\Pi_1(\rho))$ is then generated by these two points. It follows that there exists an independent subset $\{v_1^1, v_2^1\}$ of $V^1$ with dual basis $\{v_1^{1*}, v_2^{1*}\}$ such that $\Pi_1(v_1^{1*}) = v_1^2 \otimes v_1^3$ and $\Pi_1(v_2^{1*}) = v_2^2 \otimes v_2^3$. This implies that

$$\rho \ = \ v_1^1 \otimes v_1^2 \otimes v_1^3 \ + \ v_2^1 \otimes v_2^2 \otimes v_2^3.$$

Hence, $\rho$ is rank two.

(2) $\Rightarrow$ (3) : Suppose regular tensor $\rho$ is rank two. This implies that there exists vectors $v_j^i$ such that $\rho = v_1^1 \otimes v_1^2 \otimes v_1^3 + v_2^1 \otimes v_2^2 \otimes v_2^3$. By Lemma 16, the sets $\{v_1^i, v_2^i\}$ are a basis of $V^i$ for $i = 1, 2, 3$. Let $\{v_1^{i*}, v_2^{i*}\}$ denote the corresponding dual bases. It follows that

$$\begin{aligned}
&\{[v_1^2 \otimes v_1^3],\ [v_2^2 \otimes v_2^3]\} = \{[\Pi_1(\rho)(v_1^{1*})],\ [\Pi_1(\rho)(v_1^{1*})]\} \subseteq \mathbb{P}(\mathrm{im}\Pi_1(\rho)) \cap \mathbb{P}(X_{\mathbb{C}}^{2\times 2}),\\
&\{[v_1^1 \otimes v_1^3],\ [v_2^1 \otimes v_2^3]\} = \{[\Pi_2(\rho)(v_1^{2*})],\ [\Pi_2(\rho)(v_1^{2*})]\} \subseteq \mathbb{P}(\mathrm{im}\Pi_2(\rho)) \cap \mathbb{P}(X_{\mathbb{C}}^{2\times 2}),\ \text{and}\\
&\{[v_1^1 \otimes v_1^2],\ [v_2^1 \otimes v_2^2]\} = \{[\Pi_3(\rho)(v_1^{3*})],\ [\Pi_3(\rho)(v_1^{3*})]\} \subseteq \mathbb{P}(\mathrm{im}\Pi_3(\rho)) \cap \mathbb{P}(X_{\mathbb{C}}^{2\times 2}).
\end{aligned}$$

Thus, $\mathbb{P}(\mathrm{im}\Pi_i(\rho)) \cap \mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ contains at least two distinct points for each $i$. It remains to show that the intersection of $\mathbb{P}(\mathrm{im}\Pi_i(\rho))$ and $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ in fact contains only these two points. Suppose for contradiction that a projective point $[v_3^2 \otimes v_3^3]$ distinct from points $[v_1^2 \otimes v_1^3]$ and $[v_2^2 \otimes v_2^3]$ was contained in $\mathbb{P}(\mathrm{im}\Pi_1(\rho)) \cap \mathbb{P}(X_{\mathbb{C}}^{2\times 2})$. Since the line $\mathbb{P}(\mathrm{im}\Pi_1(\rho))$ is generated by any two distinct points that lie on it, it follows that there exists scalars $\alpha$ and $\beta$ such that

$$v_3^2 \otimes v_3^3 \ = \ \alpha \ v_1^2 \otimes v_1^3 \ + \ \beta \ v_2^2 \otimes v_2^3. \tag{3.10}$$

If either $\alpha$ or $\beta$ were zero, then $[v_3^2 \otimes v_3^3]$ is not distinct from $[v_1^2 \otimes v_1^3]$ and $[v_2^2 \otimes v_2^3]$ as projective points. Furthermore, if both $\alpha$ and $\beta$ were nonzero, then equation (3.10) is a contradiction, since a rank one tensor cannot be equal to a rank two tensor.

Since the implication (3) $\Rightarrow$ (1) is clear, the theorem is proved. $\square$

Bézout's theorem states that the number of intersection points with multiplicity of $n$ homogeneous polynomials in $n+1$ variables with no common components is the product of the degrees of the polynomials in complex projective space. For a regular $2 \times 2 \times 2$ complex tensor $\rho$, $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ is a projective line generated by two degree one homogeneous polynomials in four variables. Additionally, the projective $2 \times 2$ Segre variety $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ is generated by the $2 \times 2$ matrix determinant, which is a degree two homogeneous polynomial in four variables. Hence, if $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ is not contained in $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$, Bézout's theorem implies that the intersection of $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ and $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ is generically two. This in turn implies that a regular tensor $\rho$ is generically rank two by Theorem 17. Thus, to show that there is an open and dense set of rank two complex $2 \times 2 \times 2$ tensors in the norm topology, it remains to prove the following theorem.

**Theorem 18.** *Let $V^1$, $V^2$, and $V^3$ be two-dimensional complex vector spaces. If $\rho \in V^1 \otimes V^2 \otimes V^3$ is a regular tensor, then im$\Pi_j(\rho) \not\subseteq X_{\mathbb{C}}^{2\times 2}$ for $j = 1, 2, 3$.*

*Proof.* Segre varieties contain numerous vector spaces in general. For example, the plane generated by the following two matrices as

$$\left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\rangle$$

is contained in the $X_{\mathbb{C}}^{2\times 2}$. We now show that this does not happen for the plane $\Pi_j(\rho)$ when $\rho$ is regular for any $j = 1, 2, 3$. Suppose for contradiction that the plane $\Pi_1(\rho)$ was contained in the variety $X_{\mathbb{C}}^{2\times 2}$. There would then exist a basis $\{\mathrm{e}_1^1, \mathrm{e}_2^1\}$ of $V^1$ with corresponding dual basis $\{\mathrm{e}_1^{1*}, \mathrm{e}_2^{1*}\}$ and some vectors $v_j^i$'s such that

$$\Pi_1(\rho)(\mathrm{e}_1^{1*}) \;=\; v_1^1 \otimes v_1^2 \quad \text{and} \quad \Pi_1(\rho)(\mathrm{e}_2^{1*}) \;=\; v_2^1 \otimes v_2^2.$$

It also follows that $v_1^1 \otimes v_1^2 + v_2^1 \otimes v_2^2 = x_1^2 \otimes x_1^3$ for some vectors $x_1^2$, $x_1^3$ since $v_1^1 \otimes v_1^2 + v_2^1 \otimes v_2^2 \in X_{\mathbb{C}}^{2\times 2}$. This implies that $\Pi_1(v_1^1 \otimes v_1^2 + v_2^1 \otimes v_2^2)$ is equal to $\Pi_1(x_1^2 \otimes x_1^3)$, and thus the set $\{v_1^2, v_2^2\}$ is linearly dependent. This, however, cannot occur when $\rho$ is regular by Lemma 16. $\square$

We could apply Bézout's theorem in the $2 \times 2 \times 2$ case as the ideal of Segre variety was just the $2 \times 2$ determinant. Finding the ideal of the $n_1 \times n_2 \times n_3$ Segre variety is a much more difficult problem in general.

### 3.5 Regular $2 \times 2 \times 2$ Tensors Tangent to the Segre Variety

If $\rho$ is a regular $2 \times 2 \times 2$ real or complex tensor, then each of the three contraction maps generate a projective line - the line that is the projectivization of the image of each contraction map. If *any* of these lines are tangent to the Segre variety, then *all* of the these lines are tangent to the Segre variety. Furthermore, these lines are tangent to the Segre variety if and only if $\rho$ itself is tangent to the Segre variety. This demonstrates the fundamental relationship between a tensor and its contraction maps geometrically, and is summarized in Figure 3.2.

**Theorem 19.** *Let $V^1$, $V^2$, and $V^3$ be two-dimensional complex or real vector spaces. For a regular tensor $\rho$ in $V^1 \otimes V^2 \otimes V^3$, the following are equivalent:*

*(1) For some $j = 1,2,$ or $3$, $\mathbb{P}(im\Pi_j(\rho)) \subseteq T_x(\mathbb{P}(X^{2\times 2}))$ for some $x \in \mathbb{P}(X^{2\times 2})$.*

*(2) $[\rho] \in T_x(\mathbb{P}(X^{2\times 2\times 2}))$ for some $x \in \mathbb{P}(X^{2\times 2\times 2})$.*

*(3) For all $j = 1,2,$ or $3$, $\mathbb{P}(im\Pi_j(\rho)) \subseteq T_x(\mathbb{P}(X^{2\times 2}))$ for some $x \in \mathbb{P}(X^{2\times 2})$.*

*Each of these conditions also implies that the rank of $\rho$ is three, and $|\ \mathbb{P}(im\Pi_j(\rho)) \cap \mathbb{P}(X^{2\times 2})\ | = 1$ for any $j = 1, 2, 3$. Furthermore, the intersection of the projective image of any mode-$j$ contraction of $\rho$ with the $2 \times 2$ projective Segre variety is one with multiplicity two.*

*Proof.* (1) $\Rightarrow$ (2) : Suppose $\mathbb{P}(\mathrm{im}\Pi_1(\rho)) \subset T_{[x_1^2 \otimes x_1^3]}(\mathbb{P}(X^{2\times 2}))$ for some $x_1^2 \otimes x_1^3 \in X^{2\times 2}$. Since $\rho$ is regular, there must be a basis $\{x_1^1, x_2^1\}$ of $V^1$ with dual basis $\{x_1^{1*}, x_2^{1*}\}$ such that

$$\Pi_1(\rho) : x_1^{1*} \mapsto \; x_1^2 \otimes x_1^3 \tag{3.11}$$

$$x_2^{1*} \mapsto \; x_2^2 \otimes x_1^3 \; + \; x_1^2 \otimes x_2^3 \quad \text{for some vectors } x_2^2, x_2^3. \tag{3.12}$$

Equation (3.11) follows from the fact that $x_1^2 \otimes x_1^3$ must be in $T_{x_1^2 \otimes x_1^3}(X^{2\times 2})$. Equation (3.12) follows from the fact that every element in $T_{x_1^2 \otimes x_1^3}(X^{2\times 2})$ is in the form $x_2^2 \otimes x_1^3 + x_1^2 \otimes x_2^3$ by Theorem 14. Hence,

$$[\rho] \;=\; [x_1^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_1^2 \otimes x_2^3] \in T_{[x_2^1 \otimes x_1^2 \otimes x_1^3]}(\mathbb{P}(X^{2\times 2\times 2})).$$

(2) $\Rightarrow$ (3) : Suppose $\rho$ is a regular tensor contained in $T_{[x_1^1 \otimes x_1^2 \otimes x_1^3]}(\mathbb{P}(X^{2\times 2\times 2}))$ for some simple tensor $x_1^1 \otimes x_1^2 \otimes x_1^3$. The tensor $\rho$ must be in the form

$$\rho \;=\; x_2^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_1^3 + \; x_1^1 \otimes x_1^2 \otimes x_2^3 \tag{3.13}$$

for some vectors $x_2^1$, $x_2^2$, and $x_2^3$ by Theorem 14. By Lemma 15, $\rho$ is rank three. Hence, the sets $\{x_1^i, x_2^i\}$ must be linearly independent for each $i$, or else the rank of $\rho$ would be less than three. Let $\{x_1^{i*}, x_2^{i*}\}$ be the corresponding dual spaces. It follows that

$$\Pi_1(\rho)(x_1^{1*}) = x_2^2 \otimes x_1^3 + x_1^2 \otimes x_2^3, \qquad \Pi_1(\rho)(x_2^{1*}) = x_1^2 \otimes x_1^3,$$
$$\Pi_2(\rho)(x_1^{2*}) = x_2^1 \otimes x_1^3 + x_1^1 \otimes x_2^3, \qquad \Pi_2(\rho)(x_2^{2*}) = x_1^1 \otimes x_1^3,$$
$$\Pi_3(\rho)(x_1^{3*}) = x_2^1 \otimes x_1^2 + x_1^1 \otimes x_2^2, \text{ and} \qquad \Pi_3(\rho)(x_2^{3*}) = x_1^1 \otimes x_1^2.$$

Since the image of each of these maps is two-dimensional, we conclude that

$$\mathbb{P}(\mathrm{im}\Pi_1(\rho)) = \langle [x_2^2 \otimes x_1^3 + x_1^2 \otimes x_2^3],\; [x_1^2 \otimes x_1^3] \rangle \subseteq T_{[x_1^2 \otimes x_1^3]}(\mathbb{P}(X^{2\times 2})), \tag{3.14}$$

$$\mathbb{P}(\mathrm{im}\Pi_2(\rho)) = \langle [x_2^1 \otimes x_1^3 + x_1^1 \otimes x_2^3],\; [x_1^1 \otimes x_1^3] \rangle \subseteq T_{[x_1^1 \otimes x_1^3]}(\mathbb{P}(X^{2\times 2})), \text{and} \tag{3.15}$$

$$\mathbb{P}(\mathrm{im}\Pi_3(\rho)) = \langle [x_2^1 \otimes x_1^2 + x_1^1 \otimes x_2^2],\; [x_1^1 \otimes x_1^2] \rangle \subseteq T_{[x_1^1 \otimes x_1^2]}(\mathbb{P}(X^{2\times 2})). \tag{3.16}$$

| | | |
|---|---|---|
| $\mathbb{P}(\mathrm{im}\Pi_j(\rho)) \subseteq T_x(\mathbb{P}(X_{\mathbb{C}}^{2\times 2}))$<br>for some $x \in (\mathbb{P}(X_{\mathbb{C}}^{2\times 2}))$<br>for any $j = 1, 2, 3$. | $\Longleftrightarrow$ | $[\rho] \in T_x(\mathbb{P}(X_{\mathbb{C}}^{2\times 2\times 2}))$<br>for some $x \in \mathbb{P}(X_{\mathbb{C}}^{2\times 2\times 2})$ |

**Figure 3.2.** If $\rho$ is a regular complex $2 \times 2 \times 2$ tensor, then the contraction maps of $\rho$ are tangent to the Segre variety if and only if $\rho$ is tangent to the Segre variety.

The set containments in Equations (3.14), (3.15), and (3.16) follow from the characterization of the tangent space of the Segre variety in Theorem 14. This implies that each projective line $\mathbb{P}(\mathrm{im}\Pi_i(\rho))$ is tangent to the projective Segre variety.

(3) $\Rightarrow$ (1) is obvious. Furthermore, by Bézout's theorem, there cannot be any additional intersections, since a line and a degree two variety can have at most two intersection in three-dimensional complex projective space. Hence, $\mid \mathbb{P}(\mathrm{im}\Pi_j(\rho)) \cap \mathbb{P}(X^{2\times 2}) \mid = 1$ for any $j = 1, 2, 3$.by the same reasoning, the intersection of the projective image of any mode-$j$ contraction of $\rho$ with the $2 \times 2$ projective Segre variety is one with multiplicity two. It can be shown by contradiction that any regular tensor in the form (3.13) is indeed rank three. Hence, the theorem is proven. $\square$

Let $\rho$ be a regular real $2 \times 2 \times 2$ tensor. The projective line generated by the image of the mode-1 contraction map of $\rho$ intersects $\mathbb{P}(X_{\mathbb{C}}^{2\times 2})$ at one point with multiplicity two or two points with multiplicity one. However, the projective line generate by $\mathrm{im}\Pi_1(\rho)$ could possibly not intersect the real $2 \times 2$ Segre variety $\mathbb{P}(X_{\mathbb{R}}^{2\times 2})$ at all. It is not possible for the the projective line to intersect the projective Segre variety one point of multiplicity one, since complex solutions to real polynomials come in complex pairs. Hence, over $\mathbb{R}$, there are three possible cases, which are summarized in Figure 3.3. This explains why there is no generic rank for $2 \times 2 \times 2$ real tensors, for non-intersection with the Segre variety is an open condition.

| | | |
|---|---|---|
| $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ intersects $\mathbb{P}(X_{\mathbb{R}}^{2\times 2})$ at two distinct points each with multiplicity one for any $j = 1, 2, 3$ | $\Longleftrightarrow$ | $\mathrm{rk}(\rho) = 2$ |
| $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ intersects $\mathbb{P}(X_{\mathbb{R}}^{2\times 2})$ at one point with multiplicity two for any $j = 1, 2, 3$ | $\Longleftrightarrow$ | $\mathrm{rk}(\rho) = 3$<br>$[\rho] \in T_x(\mathbb{P}(X_{\mathbb{R}}^{2\times 2\times 2}))$<br>for some $x \in \mathbb{P}(X_{\mathbb{R}}^{2\times 2\times 2})$ |
| $\mathbb{P}(\mathrm{im}\Pi_j(\rho))$ does not intersect $\mathbb{P}(X_{\mathbb{R}}^{2\times 2})$ for any $j = 1, 2, 3$ | $\Longleftrightarrow$ | $\mathrm{rk}(\rho) = 3$<br>$[\rho] \notin T_x(\mathbb{P}(X_{\mathbb{R}}^{2\times 2\times 2}))$<br>for some $x \in \mathbb{P}(X_{\mathbb{R}}^{2\times 2\times 2})$ |

**Figure 3.3.** If $\rho$ is a regular real $2 \times 2 \times 2$ tensor, then one of the above three conditions holds.

### 3.6 Quantifier Elimination and the Importance of a Geometric Perspective

It follows from the quadratic formula that the existential statement

$$(\exists x)(\exists y)(x^2 + bx + c = 0) \wedge (y^2 + by + c = 0) \wedge (x \neq y)$$

is true over $\mathbb{C}$ if and only if the quantifier-free statement

$$b^2 - 4c \neq 0$$

is true. In fact, every statement in the language of algebraically closed fields is equivalent to a statement without quantifiers. The statement '$\rho \in \mathbb{C}^{2\times2\times2}$ is a rank two tensor' can be formalized in the language of algebraically closed fields with the formula $\phi(a_{ijk})$

$$\begin{aligned}
&\exists u_1^1\ \exists u_1^2\ \exists v_1^1\ \exists v_1^2\ \exists w_1^1\ \exists w_1^2\ \exists u_2^1\ \exists u_2^2\ \exists v_2^1\ \exists v_2^2\ \exists w_2^1\ \exists w_2^2\ (a_{111} = u_1^1 v_1^1 w_1^1 + u_1^2 v_1^2 w_1^2) \wedge \\
&\quad (a_{112} = u_1^1 v_1^1 w_2^1 + u_1^2 v_1^2 w_2^2) \wedge (a_{121} = u_1^1 v_2^1 w_1^1 + u_1^2 v_2^2 w_1^2) \wedge (a_{122} = u_1^1 v_2^1 w_2^1 + u_1^2 v_2^2 w_2^2) \wedge \\
&\quad (a_{211} = u_2^1 v_1^1 w_1^1 + u_2^2 v_1^2 w_1^2) \wedge (a_{212} = u_2^1 v_1^1 w_2^1 + u_2^2 v_1^2 w_2^2) \wedge (a_{221} = u_2^1 v_2^1 w_1^1 + u_2^2 v_2^2 w_1^2) \wedge \\
&\quad (a_{222} = u_2^1 v_2^1 w_2^1 + u_2^2 v_2^2 w_2^2) \quad \wedge \\
&\neg\exists u_1^1\ \neg\exists u_1^2\ \neg\exists v_1^1\ \neg\exists v_1^2\ \neg\exists w_1^1\ \neg\exists w_1^2 (a_{111} = u_1^1 v_1^1 w_1^1) \wedge (a_{112} = u_1^1 v_1^1 w_2^1) \wedge (a_{121} = u_1^1 v_2^1 w_1^1) \\
&\quad \wedge (a_{122} = u_1^1 v_2^1 w_2^1) \wedge (a_{211} = u_2^1 v_1^1 w_1^1) \wedge (a_{212} = u_2^1 v_1^1 w_2^1 2) \\
&\quad \wedge (a_{221} = u_2^1 v_2^1 w_1^1) \wedge (a_{222} = u_2^1 v_2^1 w_2^1).
\end{aligned}$$

This is the entry-wise equivalent of the informal statement in the 'language of tensors' of

$$\exists \begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} \exists \begin{bmatrix} u_1^2 \\ u_2^2 \end{bmatrix} \exists \begin{bmatrix} v_1^1 \\ v_2^1 \end{bmatrix} \exists \begin{bmatrix} v_1^2 \\ v_2^2 \end{bmatrix} \exists \begin{bmatrix} w_1^1 \\ w_2^1 \end{bmatrix} \exists \begin{bmatrix} w_1^2 \\ w_2^2 \end{bmatrix} \Big(\rho = \begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} \otimes \begin{bmatrix} v_1^1 \\ v_2^1 \end{bmatrix} \otimes \begin{bmatrix} w_1^1 \\ w_2^1 \end{bmatrix} + \begin{bmatrix} u_1^2 \\ u_2^2 \end{bmatrix} \otimes \begin{bmatrix} v_1^2 \\ v_2^2 \end{bmatrix} \otimes \begin{bmatrix} w_2^2 \\ w_2^2 \end{bmatrix}\Big)$$
$$\wedge\ \neg\exists \begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} \neg\exists \begin{bmatrix} u_1^2 \\ u_2^2 \end{bmatrix} \neg\exists \begin{bmatrix} v_1^1 \\ v_2^1 \end{bmatrix} \Big(\rho = \begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} \otimes \begin{bmatrix} v_1^1 \\ v_2^1 \end{bmatrix} \otimes \begin{bmatrix} w_1^1 \\ w_2^1 \end{bmatrix}\Big).$$

We now construct a quantifier-free formula that is equivalent to $\phi(a_{ijk})$. Let $m_1(a_{ijk})$, $m_2(a_{ijk})$, and $m_3(a_{ijk})$ be the following quantifier-free formulas.

$$\begin{aligned} m_1(a_{ijk}) &= (a_{111}a_{212} - a_{211}a_{112} \neq 0) \\ &\vee (a_{111}a_{221} - a_{211}a_{121} \neq 0) \\ &\vee (a_{111}a_{221} - a_{211}a_{122} \neq 0) \\ &\vee (a_{112}a_{221} - a_{212}a_{121} \neq 0) \\ &\vee (a_{112}a_{222} - a_{212}a_{122} \neq 0) \\ &\vee (a_{121}a_{222} - a_{221}a_{122} \neq 0), \end{aligned} \qquad \begin{aligned} m_2(a_{ijk}) &= (a_{111}a_{122} - a_{121}a_{112} \neq 0) \\ &\vee (a_{111}a_{221} - a_{121}a_{211} \neq 0) \\ &\vee (a_{111}a_{222} - a_{121}a_{212} \neq 0) \\ &\vee (a_{112}a_{221} - a_{122}a_{211} \neq 0) \\ &\vee (a_{112}a_{222} - a_{122}a_{212} \neq 0) \\ &\vee (a_{211}a_{222} - a_{221}a_{212} \neq 0), \end{aligned}$$

$$\begin{aligned} \text{and } m_3(a_{ijk}) &= (a_{111}a_{122} - a_{112}a_{121} \neq 0) \\ &\vee (a_{111}a_{212} - a_{112}a_{211} \neq 0) \\ &\vee (a_{111}a_{222} - a_{112}a_{221} \neq 0) \\ &\vee (a_{121}a_{212} - a_{122}a_{211} \neq 0) \\ &\vee (a_{121}a_{222} - a_{122}a_{221} \neq 0) \\ &\vee (a_{211}a_{222} - a_{212}a_{221} \neq 0). \end{aligned}$$

Furthermore, let $\theta(a_{ijk})$ be the formula

$$\theta(a_{ijk}) = (m_1 \vee m_2 \vee \neg m_3) \ \vee \ (m_1 \vee m_3 \vee \neg m_2) \ \vee \ (m_2 \vee m_3 \vee \neg m_1),$$

and let $\Delta(a_{ijk})$ be the polynomial

$$\begin{aligned} \Delta(a_{ijk}) = (&a_{111}^2 a_{222}^2 + a_{112}^2 a_{221}^2 + a_{121}^2 a_{212}^2 + a_{122}^2 a_{211}^2) \\ &- 2(a_{111}a_{112}a_{221}a_{222} + a_{111}a_{121}a_{212}a_{222} + a_{111}a_{122}a_{211}a_{222}) \\ &- 2(a_{112}a_{121}a_{212}a_{221} + a_{112}a_{122}a_{221}a_{211} + a_{121}a_{122}a_{212}a_{211}) \\ &+ 4(a_{111}a_{122}a_{212}a_{221} + a_{112}a_{121}a_{211}a_{222}). \end{aligned}$$

The quantified formula $\phi(a_{ijk})$ is equivalent to the quantifier-free formula $(\Delta(a_{ijk}) \neq 0) \vee \theta(a_{ijk})$.

The rank of a hypermatrix can theoretically be determined algorithmically by computing and then evaluating the equivalent quantifier-free formulas. However, such formulas are incomprehensible and often infeasible for high dimensional tensor spaces, as the above example demonstrates. Hence, there is a need for a more geometric understanding of tensors. In this chapter, we have given a geometric explanation for the existence of an open, dense set of rank two tensors in the complex $2 \times 2 \times 2$ tensor space.

# 4. RANK-JUMPING

If a sequence of rank $r$ real matrices converges to a matrix $A$ in the norm topology, then the rank of $A$ is less than or equal to $r$. A sequence of real hypermatrices, however, can converge to a hypermatrix of greater rank. The limit points of such sequences are called rank-jumping hypermatrices. We show that a $2 \times 2 \times 2$ real hypermatrix is rank-jumping if and only if it is a regular element of the tangential variety of the Segre variety. We also explain why rank-jumping never occurs for non-negative real hypermatrices.

## 4.1 Rank-Jumping Hypermatrices

Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, and $n_3$-dimensional real vector spaces, respectively, and let $V^1 \otimes V^2 \otimes V^3$ denote the tensor product of these spaces. The tensors in the form $v^1 \otimes v^2 \otimes v^3$ for some vectors $v^1$, $v^2$, and $v^3$ are called simple tensors. The simple tensors form a variety, which is denoted $X$ and is called the Segre variety. The rank of a tensor $\rho$ is the minimum number $n$ such that $\rho$ is the sum of $n$ simple tensors. A tensor $\rho$ is said to be rank-jumping if the rank of $\rho$ is greater than $s$, and there exists a sequence of rank $s$ tensors that converge to $\rho$ in the norm topology.

We denote by $\tau(X)$ the set of all tensors that lie on lines tangent to the Segre variety. It is a classical observation that tensors of rank greater than two in $\tau(X)$ are rank-jumping. A tensor $\beta$ in $\tau(X)$ is in the form

$$\beta \;=\; x_2^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_1^2 \otimes x_2^3 \tag{4.1}$$

for some vectors $x_j^i$. This is an immediate calculation, but can also be found in [31, Prop.2.6] and [14, p.108]. The tensor $\beta$ is the limit of the sequence of rank at most two tensors

$$\rho_n \;=\; n\,(x_1^1 + \frac{1}{n}x_2^1) \otimes (x_1^2 + \frac{1}{n}x_2^2) \otimes (x_1^3 + \frac{1}{n}x_2^3) \;-\; n x_1^1 \otimes x_1^2 \otimes x_1^3,\;\; n \in \mathbb{N} \tag{4.2}$$

in the norm topology. It follows from the observation

$$\|\rho_n - \beta\| \;\leq\; \frac{1}{n}\left\|x_2^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_1^2 \otimes x_2^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_2^3\right\| \;+\; \frac{1}{n^2}\left\|x_2^1 \otimes x_2^2 \otimes x_2^3\right\|,$$

that the limit of $\rho_n$ is indeed equal to $\beta$. The rank-jumping of tensors in the form (4.1) is best understood in terms of secant lines of the Segre variety. A secant line of the Segre variety is a line that intersects the Segre variety at two points. A real tensor is rank at most two if and only if it is on a secant line of the Segre variety. We denote by $\sigma_2(X)$ the norm closure of the set of all tensors that lie on secant lines of the Segre variety. We define $\sigma_2(X)$ as the norm closure of the set of rank at most two tensors since the set of rank at most two tensors is not closed, as sequence (4.2) demonstrates.

Tensors in $\tau(X)$ are on tangent lines of the Segre variety, which are limits of secant lines. Hence, almost by definition, the rank-jumping tensors that are limits of rank two tensors should be the rank greater than two elements in $\tau(X)$. In this chapter, we will prove this characterization formally. We will prove that a $2 \times 2 \times 2$ real tensor is rank-jumping if and only if it is a regular element of $\tau(X)$.

Interestingly, rank jumping does not occur for non-negative real hypermatrices [32]. Recall that tensors can be coordinatized as hypermatrices. The $2 \times 2 \times 2$ hypermatrices in the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \otimes \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} := \begin{array}{c} \boxed{\begin{matrix} x_2y_1z_1 & x_2y_1z_2 \\ x_2y_2z_1 & x_2y_2z_2 \end{matrix}} \\ \boxed{\begin{matrix} x_1y_1z_1 & x_1y_1z_2 \\ x_1y_2z_1 & x_1y_2z_2 \end{matrix}} \end{array}$$

for some constants $x_i$, $y_i$, and $z_i$ are the simple hypermatrices, which are the coordinate counterparts of simple tensors.

A non-negative hypermatrix $(a_{ijk})$ has both a non-negative rank and a real rank. The non-negative rank of $(a_{ijk})$ is the minimum number $n$ such that $(a_{ijk})$ is equal to the sum of $n$ non-negative-valued simple hypermatrices. The real rank of $(a_{ijk})$, on the other hand,

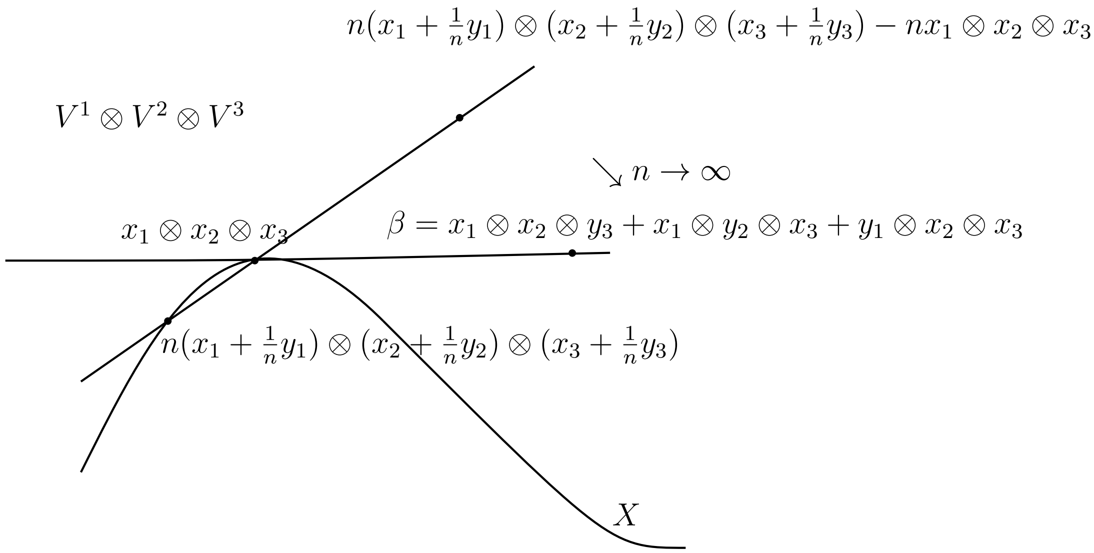



**Figure 4.1.** Tensor $\beta$ is a rank three element on a tangent line of the Segre variety that is the limit of rank two tensors on secant lines of the Segre variety. Hence, $\beta$ is a rank-jumping tensor. The sequence of rank two tensors that converge to $\beta$ contains a negative term. A negative term always appears in such sequences of rank two tensors that converge to rank three tensors, since rank-jumping never occurs for non-negative hypermatrices.

is the minimum number $n$ such that $(a_{ijk})$ is equal to the sum of $n$ real-valued simple hypermatrices. Tangent lines to the non-negative Segre variety still contain rank greater than two non-negative hypermatrices. However, over the non-negative real numbers, rank at most two tensors can no longer be identified with secant lines of the Segre variety. For example, the non-negative real hypermatrix

$$A = \boxed{\begin{matrix} \frac{23}{4} & 7 \\ 7 & 10 \end{matrix}} \quad \boxed{\begin{matrix} \frac{50}{4} & 16 \\ 16 & 22 \end{matrix}}$$

is on a secant line of the non-negative Segre variety. Indeed, $A$ is on the secant line that intersects the the non-negative Segre variety at the points

$$\begin{bmatrix} 3 \\ 6 \end{bmatrix} \otimes \begin{bmatrix} \frac{3}{2} \\ 2 \end{bmatrix} \otimes \begin{bmatrix} \frac{3}{2} \\ 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

since

$$A = \begin{bmatrix} 3 \\ 6 \end{bmatrix} \otimes \begin{bmatrix} \frac{3}{2} \\ 2 \end{bmatrix} \otimes \begin{bmatrix} \frac{3}{2} \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Non-negative hypermatrix $A$ is thus real rank two. It can be shown by contradiction, however, that $A$ cannot be written as the sum of two simple non-negative hypermatrices. Thus, the non-negative rank of $A$ is greater than two.

## 4.2 Rank Three Tensors On Tangent Lines Are Rank-Jumping

Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, and $n_3$-dimensional real vector spaces, and let $V^{1*}$, $V^{2*}$, and $V^{3*}$ denote their dual spaces. The tensor $\rho = \sum_{i=1}^{r} v_i^1 \otimes v_i^2 \otimes v_i^3$ in $V^1 \otimes V^2 \otimes V^3$ induces the three linear maps

$$\begin{array}{lll} V^{1*} \xrightarrow{\Pi_1(\rho)} V^2 \otimes V^3 & V^{2*} \xrightarrow{\Pi_2(\rho)} V^1 \otimes V^3 & V^{3*} \xrightarrow{\Pi_3(\rho)} V^1 \otimes V^2 \\ v^{1*} \mapsto \displaystyle\sum_{i=1}^{r} \; v^{1*}(v_i^1) \; v_i^2 \otimes v_i^3, & v^{2*} \mapsto \displaystyle\sum_{i=1}^{r} \; v^{2*}(v_i^2) \; v_i^1 \otimes v_i^3, \quad \text{and} & v^{3*} \mapsto \displaystyle\sum_{i=1}^{r} \; v^{3*}(v_i^3) \; v_i^1 \otimes v_i^2. \end{array}$$

These maps are called the mode-1, mode-2, and mode-3 contraction maps of $\rho$, respectively. The relationship between the contraction maps of a tensor generalizes, in a coordinate-free way, the fundamental relationship between the rows and columns of a matrix to hypermatrices. The rank of a two-fold tensor $\rho \in V^1 \otimes V^2$ is equal to the dimension of the image of $\Pi_1(\rho)$, which is also equal to the dimension of the image of $\Pi_2(\rho)$. We now use contraction maps to show when tensors in the form (4.1) are rank three. First, we need the following lemma.

**Lemma 20.** *Let $V^1$ and $V^2$ be finite dimensional real vector spaces, let $\{x_i^1\}_{i=1}^r$ and $\{y_i^1\}_{i=1}^s$ be linearly independent subsets of $V^1$, and let $\{x_i^2\}_{i=1}^r$ and $\{y_i^2\}_{i=1}^s$ be linearly independent subsets of $V^2$. If $\sum_{i=1}^r x_i^1 \otimes x_i^2 = \sum_{i=1}^s y_i^1 \otimes y_i^2$, then $r = s$.*

*Proof.* Since $\{x_i^1\}_{i=1}^r$ is independent, we can choose $\{x_j^{1*}\}_{j=1}^r \subseteq V^{1*}$ such that $x_j^{1*}(x_i^1) = \delta_{ij}$. By taking the mode-1 contraction of the tensor with respect to each of the two representations, it follows that

$$\sum_{i=1}^{r} \; x_j^{1*}(x_i^1) \; x_i^2 \;\; = \;\; \sum_{i=1}^{s} x_j^{1*}(y_i^1) \; y_i^2 \quad \text{for } j = 1, 2, \ldots, r.$$

This implies $x_j^2 = \sum_{i=1}^s x_j^{1*}(y_i^1) \; y_i^2$ for $j = 1, 2, \ldots, r$, so each $x_j^2$ is in the span of $y_1^2$, $y_2^2$, ...$y_s^2$, which we denote by $\langle y_i^2 \rangle_{i=1}^s$. Thus, we conclude that the linear space $\langle x_j^2 \rangle_{j=1}^r$ is a subset of $\langle y_i^2 \rangle_{i=1}^s$. Similarly, $\langle y_i^2 \rangle_{i=1}^s \subseteq \langle x_j^2 \rangle_{j=1}^r$, so, by the independence of each set, $r = s$. $\square$

**Theorem 21.** *Let $V^1$, $V^2$, and $V^3$ be finite dimensional real vector spaces. If the sets $\{x_1^1, x_2^1\}$, $\{x_1^2, x_2^2\}$, and $\{x_1^3, x_2^3\}$ in $V^1$, $V^2$, and $V^3$, respectively, are linearly independent, then the tensor*

$$\beta \;=\; x_2^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_1^2 \otimes x_2^3$$

*is rank three.*

*Proof.* Suppose for contradiction that $\beta$ were rank less than three. That is, suppose that

$$\beta \;=\; y_1^1 \otimes y_1^2 \otimes y_1^3 \;+\; y_2^1 \otimes y_2^2 \otimes y_2^3$$

for some vectors $y_j^i$. Since the image of $\Pi_i(\beta)$ is equal to $\langle x_1^i, x_2^i \rangle$ and $\langle y_1^i, y_2^i \rangle$ for each $i$, it follows that the sets $\{y_1^i, y_2^i\}$ are independent for each $i$. Let $x_1^{1*}$ and $x_2^{1*}$ be the dual vectors such that $x_i^{1*}(x_j^1) = \delta_{ij}$ for $i, j = 1, 2$. By considering the mode-1 contraction of both representations of $\beta$, it follows that

$$\begin{aligned}\Pi_1(\beta)(x_2^{1*}) &= x_2^{1*}(x_2^1)x_1^2 \otimes x_1^3 \;+\; x_2^{1*}(x_1^1)x_2^2 \otimes x_1^3 \;+\; x_2^{1*}(x_1^1)x_1^2 \otimes x_2^3, \text{ and}\\ \Pi_1(\beta)(x_2^{1*}) &= \; x_2^{1*}(y_1^1)y_1^2 \otimes y_1^3 \;+\; x_2^{1*}(y_2^1)y_2^2 \otimes y_2^3.\end{aligned}$$

This implies that

$$x_1^2 \otimes x_1^3 \;=\; x_2^{1*}(y_1^1)y_1^2 \otimes y_1^3 \;+\; x_2^{1*}(y_2^1)y_2^2 \otimes y_2^3.$$

Since the sets $\{y_1^i, y_2^i\}$ are independent, Lemma 20 implies that either $x_2^{1*}(y_1^1) = 0$ or $x_2^{1*}(y_2^1) = 0$. We consider the case when $x_2^{1*}(y_2^1) = 0$, and leave the remaining similar case to the reader. It follows that

$$\operatorname{im} \Pi_2(\; \Pi_1(\beta)(x_2^{1*}) \;) = \operatorname{im} \Pi_2(x_1^2 \otimes x_1^3) = \operatorname{im} \Pi_2(\; x_2^{1*}(y_1^1)y_1^2 \otimes y_1^3 \;),$$

which implies that $\langle x_1^2 \rangle = \langle y_1^2 \rangle$. Let $k$ be a scalar such that $y_1^2 = kx_1^2$. Similarly,

$$\operatorname{im}\Pi_1(\ \Pi_1(\beta)(x_2^{1*})\ ) = \operatorname{im}\Pi_1(x_1^2 \otimes x_1^3) = \operatorname{im}\Pi_1(\ x_2^{1*}(y_1^1)y_1^2 \otimes y_1^3\ ),$$

so $\langle x_1^3 \rangle = \langle y_1^3 \rangle$. Finally, we derive a contradiction by considering the mode-2 contraction of both representations of $\beta$.

$$\begin{aligned}
\Pi_2(\beta)(x_2^{2*}) = \ x_1^1 \otimes x_1^3 \ &= \ x_2^{2*}(y_1^2)y_1^1 \otimes y_1^3 \ + \ x_2^{2*}(y_2^2)y_1^2 \otimes y_2^3 \\
&= \ x_2^{2*}(kx_1^2)y_1^1 \otimes y_1^3 \ + \ x_2^{2*}(y_2^2)y_1^2 \otimes y_2^3 \\
&= \ x_2^{2*}(y_2^2)y_1^2 \otimes y_2^3.
\end{aligned}$$

However, this implies that $y_2^3 \in \langle x_1^3 \rangle$. This is a contradiction, since we have already shown that $y_1^3 \in \langle x_1^3 \rangle$ and the set $\{y_1^3, y_2^3\}$ is independent. We leave it to the reader to check the similar case of $x_2^{1*}(y_1^1) = 0$. $\square$

We have now shown that when $V^1$, $V^2$, and $V^3$ are finite dimensional real vector spaces and the sets $\{x_1^i, x_2^i\} \subseteq V^i$ are linearly independent, then the tensor

$$\beta = \ x_2^1 \otimes x_1^2 \otimes x_1^3 \ + \ x_1^1 \otimes x_2^2 \otimes x_1^3 \ + \ x_1^1 \otimes x_1^2 \otimes x_2^3$$

is rank three. However, it is also the limit of the sequence of rank at most two tensors

$$\rho_n = \ n\ (x_1^1 + \frac{1}{n}x_2^1) \otimes (x_1^2 + \frac{1}{n}x_2^2) \otimes (x_1^3 + \frac{1}{n}x_2^3) \ - \ nx_1^1 \otimes x_1^2 \otimes x_1^3,\ n \in \mathbb{N}$$

in the norm topology. Hence, such tensors are indeed rank-jumping.

### 4.3 Continuous Curves of Rank Two Hypermatrices

In the coordinate-free setting of tensors, the rank one summands of a $2 \times 2 \times 2$ regular tensor can be determined by considering the contraction maps of the tensors. A tensor is said to be regular if all of its contraction maps are full-rank. If $\rho$ is regular and rank two, then

$$\rho \;=\; v_1^1 \otimes v_1^2 \otimes v_1^3 \;+\; v_2^1 \otimes v_2^2 \otimes v_2^3$$

for some vectors $v_j^i$, and the sets $\{v_1^1, v_2^1\}$, $\{v_1^2, v_2^2\}$, and $\{v_1^3, v_2^3\}$ are linearly independent. Let $\{v_1^{1*}, v_2^{1*}\}$ denote the dual basis of $\{v_1^1, v_2^1\}$. By considering the values of $\Pi_1(\rho)$ on the dual vectors $v_1^{1*}$ and $v_2^{1*}$, we concluded that

$$\operatorname{im}\Pi_1(\rho) \;=\; \langle v_1^2 \otimes v_1^3,\; v_2^2 \otimes v_2^3\rangle. \tag{4.3}$$

By similar reasoning, we can conclude that the vector $v_1^3$ is in the image of the map $\Pi_1(v_1^2 \otimes v_1^3)$ and the vector $v_1^2$ is in the image of the map $\Pi_2(v_1^2 \otimes v_1^3)$, where the contractions are now taken with respect to two-fold tensors. We now translate these observations into the coordinate setting of hypermatrices. Let $B$ denote the collection of bases of $\{\mathrm{e}_1^i, \mathrm{e}_2^i\}$ of $V^i$ for $i = 1$, 2, 3, and let $\{\mathrm{e}_1^{i*}, \mathrm{e}_2^{i*}\}$ denote their corresponding dual bases. The tensor $\rho = \sum_{ijk} a_{ijk}\mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$ is identified with the hypermatrix

$$(a_{ijk}) = \begin{array}{cc} & \boxed{\begin{matrix} a_{211} & a_{212} \\ a_{221} & a_{222} \end{matrix}} \\ \boxed{\begin{matrix} a_{111} & a_{112} \\ a_{121} & a_{122} \end{matrix}} & \end{array}$$

with respect to bases $B$. The mode-$i$ contractions maps of $\rho$ can also be coordinitized as $4 \times 2$ matrix unfoldings of $(a_{ijk})$.

$$\Pi_1(a_{ijk}) = \begin{pmatrix} a_{111} & a_{211} \\ a_{112} & a_{212} \\ a_{121} & a_{221} \\ a_{122} & a_{222} \end{pmatrix} \qquad \Pi_2(a_{ijk}) = \begin{pmatrix} a_{111} & a_{121} \\ a_{112} & a_{122} \\ a_{211} & a_{221} \\ a_{212} & a_{222} \end{pmatrix} \qquad \Pi_3(a_{ijk}) = \begin{pmatrix} a_{111} & a_{112} \\ a_{121} & a_{122} \\ a_{211} & a_{212} \\ a_{221} & a_{222} \end{pmatrix}$$

The matrices $\Pi_1(a_{ijk})$, $\Pi_2(a_{ijk})$, and $\Pi_3(a_{ijk})$ are called the mode-1, mode-2, and mode-3 coordinate contractions of $(a_{ijk})$, respectively. Hypermatrix $(a_{ijk})$ is simple if and only if all of the mode-$i$ coordinate contraction maps are rank less than or equal to one. Hence, hypermatrix $(a_{ijk})$ is simple if and only if all the $2 \times 2$ minors of its mode-$i$ coordinate contraction maps are zero. This shows that the set of simple hypermatrices is indeed a variety. It is in fact sufficient to show that two of the three mode-$i$ coordinate contraction maps are simple, for this implies the third one is simple as well.

The columns of the contraction maps $\Pi_i(a_{ijk})$ can also be unfolded into $2 \times 2$ matrices, which are denoted as follows.

$$a_{1::} = \begin{pmatrix} a_{111} & a_{112} \\ a_{121} & a_{122} \end{pmatrix} \qquad a_{2::} = \begin{pmatrix} a_{211} & a_{212} \\ a_{221} & a_{222} \end{pmatrix}$$
$$a_{:1:} = \begin{pmatrix} a_{111} & a_{112} \\ a_{211} & a_{212} \end{pmatrix} \qquad a_{:2:} = \begin{pmatrix} a_{121} & a_{122} \\ a_{221} & a_{222} \end{pmatrix}$$
$$a_{::1} = \begin{pmatrix} a_{111} & a_{121} \\ a_{211} & a_{221} \end{pmatrix} \qquad a_{::2} = \begin{pmatrix} a_{112} & a_{122} \\ a_{212} & a_{222} \end{pmatrix}$$

These matrices are referred to as the co-dimension one slices of hypermatrix $(a_{ijk})$. Hypermatrix $(a_{ijk})$ is said to be regular if it is the coordinatization of a regular tensor. Thus, a hypermatrix is regular if and only if all of its mode-$i$ coordinate contraction maps are full rank. Thus, a $2 \times 2 \times 2$ hypermatrix is regular if and only if $\Pi_1(a_{ijk})$, $\Pi_2(a_{ijk})$, and $\Pi_3(a_{ijk})$ are all rank two. The $4 \times 2$ matrix $\Pi_1(a_{ijk})$ is rank two if and only if the set of $2 \times 2$ matrices $\{a_{1::}, a_{2::}\}$ is linearly independent. Hence, a $2 \times 2 \times 2$ hypermatrix is regular if and only if there are no linear relations among its co-dimension one slices. Intuitively, generating the entries of a $2 \times 2 \times 2$ hypermatrix randomly and independently should result in a regular hypermatrix. Indeed, the set of $2 \times 2 \times 2$ hypermatrices form an open and dense set in $\mathbb{R}^{2\times2\times2}$ with respect to the norm topology.

We prove that if regular $2 \times 2 \times 2$ real hypermatrices are parameterized continuously, then the rank one summands can also be parameterized continuously and semi-algebraically. We will need the following generalization of Cramer's rule.

**Theorem 22.** *Let* $(0,1] \mapsto A(t)$ *be a continuous curve of* $m \times n$ *matrices, and let* $(0,1] \mapsto b(t)$ *be a continuous curve of* $m \times 1$ *vectors. If there exists a* $t_0$ *in the interval* $(0,1)$ *such that the linear system* $A(t_0)x = b(t_0)$ *has a solution, then there exists a neighborhood* $(a,b)$ *of* $t_0$ *and a rational function* $p(t)$ *from* $(a,b)$ *to* $\mathbb{R}^n$ *such that* $A(t)p(t) = b(t)$ *on* $(a,b)$.

*Proof.* First, recall that if $A$ is a square matrix, and $b$ be is a vector such that the linear system $Ax = b$ has a solution, then the solution $x$ must be completely determined by the coefficients of $A$ and $b$. Cramer's rule is the classical observation that the vector with $i^{th}$ entry $x_i$ equal to

$$x_i \;=\; \frac{\det(A_i)}{\det A}$$

is the solution to $Ax = b$, where $A_i$ is the matrix obtained by replacing the $i^{th}$ column of $A$ with the column vector $b$. Cramer's formula thus demonstrates that the solution of a square system of equations $Ax = b$ is in fact determined by a ratio of polynomial functions in the coefficients of $A$ and $b$. If a polynomial is zero at a point, then it is zero on a neighborhood of the point. Hence, if $A(t)$ is a continuous curve of $n \times n$ matrices defined on $(0,1]$ and $b(t)$ is a continuous curve of $n \times 1$ vectors defined on $(0,1]$, then if $A(t_0)x = b(t_0)$ has a solution for some $t_0 \in (0,1]$, there must exist a neighborhood of $(a,b)$ of $t_0$ on which $A(t)x = b(t)$ has a solution. Let $(a,b)$ be the maximal neighborhood of $t_0$ such that $\det A(t)$ is nonzero. It then follows from Cramer's rule that

$$t \mapsto \left( \frac{\det(A_i(t))}{\det A(t)} \right)$$

is the desired rational function $p(t)$. Furthermore, Cramer's rule can be extended to rectangular matrices. The solutions of rectangular systems of linear equations can still be determined by rational functions [33]. This implies the theorem. □

**Theorem 23.** *Let* $\gamma : (0,1] \to \mathbb{R}^{2\times 2\times 2}$ *be a continuous curve of regular rank two hypermatrices. There exists* $(0,1] \xrightarrow{v_j^i} \mathbb{R}^2$ *continuous and semi-algebraic for all* $i,j$ *such that*

$$\gamma(t) \;=\; v_1^1(t) \otimes v_1^2(t) \otimes v_1^3(t) \;+\; v_2^1(t) \otimes v_2^2(t) \otimes v_2^3(t)$$

*for all* $t \in (0,1]$.

*Proof.* Since $(a_{ijk}(t))$ is a regular rank two hypermatrix for every $t$, we know that

$$(a_{ijk}(t)) = \begin{pmatrix} x^1_{11}(t) \\ x^1_{12}(t) \end{pmatrix} \otimes \begin{pmatrix} x^2_{11}(t) \\ x^2_{12}(t) \end{pmatrix} \otimes \begin{pmatrix} x^3_{11}(t) \\ x^3_{12}(t) \end{pmatrix} + \begin{pmatrix} x^1_{21}(t) \\ x^1_{22}(t) \end{pmatrix} \otimes \begin{pmatrix} x^2_{21}(t) \\ x^2_{22}(t) \end{pmatrix} \otimes \begin{pmatrix} x^3_{21}(t) \\ x^3_{22}(t) \end{pmatrix}$$

for some constants $x^i_{jk}(t)$ for every $t$ in $(0,1]$. By translating Equation (4.3) into the language of hypermatrices, we conclude that there exists two ordered pairs $(\alpha_1(t), \beta_1(t))$ and $(\alpha_2(t), \beta_2(t))$ for every $t$ such that the following equalities hold.

$$\begin{pmatrix} x^2_{11}(t) \\ x^2_{12}(t) \end{pmatrix} \otimes \begin{pmatrix} x^3_{11}(t) \\ x^3_{12}(t) \end{pmatrix} = \alpha_1(t)a_{1::}(t) + \beta_1(t)a_{2::}(t) \tag{4.4}$$

$$\begin{pmatrix} x^2_{21}(t) \\ x^2_{22}(t) \end{pmatrix} \otimes \begin{pmatrix} x^3_{21}(t) \\ x^3_{22}(t) \end{pmatrix} = \alpha_2(t)a_{1::}(t) + \beta_2(t)a_{2::}(t) \tag{4.5}$$

These pairs $(\alpha_1(t), \beta_1(t))$ and $(\alpha_2(t), \beta_2(t))$ are precisely the solutions to the equation $\det(xa_{1::}(t) + ya_{2::}(t)) = 0$. By similarly considering the mode-3 contraction maps, we conclude that there also exists ordered pairs $(\alpha_3(t), \beta_3(t))$ and $(\alpha_4(t), \beta_4(t))$ for every $t$ such that the following equalities hold.

$$\begin{pmatrix} x^1_{11}(t) \\ x^1_{12}(t) \end{pmatrix} \otimes \begin{pmatrix} x^2_{11}(t) \\ x^2_{12}(t) \end{pmatrix} = \alpha_3(t)a_{::1}(t) + \beta_3(t)a_{::3}(t) \tag{4.6}$$

$$\begin{pmatrix} x^1_{21}(t) \\ x^1_{22}(t) \end{pmatrix} \otimes \begin{pmatrix} x^2_{21}(t) \\ x^2_{22}(t) \end{pmatrix} = \alpha_4(t)a_{::1}(t) + \beta_4(t)a_{::2}(t) \tag{4.7}$$

Choose a $t_0$ such that $a_{1::}(t_0)$, $a_{2::}(t_0)$, $a_{::1}(t_0)$, and $a_{::2}(t_0)$ are rank all rank two. Thus, on a neighborhood $(c,d)$ of $t_0$, we can conclude that the solutions of (4.4), (4.5), (4.6), and (4.7) do not have zero coordinates. The multilinearity of the determinant function then implies that for each $t$ in $(c,d)$, there exists $\delta_1(t)$, $\delta_2(t)$, $\delta_3(t)$, and $\delta_4(t)$ on $(c,d)$ such that the following equations hold.

$$\det\left(a_{1::}(t) + \delta_1(t)\ a_{2::}(t)\right) = 0 \qquad \det\left(a_{::1}(t) + \delta_3(t)\ a_{::2}(t)\right) = 0$$
$$\det\left(a_{1::}(t) + \delta_2(t)\ a_{2::}(t)\right) = 0 \qquad \det\left(a_{::1}(t) + \delta_4(t)\ a_{::2}(t)\right) = 0$$

Note that $\det\left(a_{1::}(t) + \delta_1(t)\ a_{2::}(t)\right) = 0$ implies $\delta_1(t)^2\ A(t) + \delta_1(t)\ B(t)\ +\ C(t)\ =\ 0,$

where
$$A(t) = \det\left(a_{2::}(t)\right)$$
$$B(t) = \left(\mathrm{tr}(a_{1::}(t))\mathrm{tr}(a_{2::}(t)) - \mathrm{tr}\left(a_{1::}(t)a_{2::}(t)\right)\right)$$
$$C(t) = \det\left(a_{1::}(t)\right).$$

It follows that
$$\delta_{1,2}(t)\ \ =\ \ \frac{-B(t) \pm \sqrt{B(t)^2 - 4A(t)C(t)}}{2A(t)},$$

so $\delta_1(t)$ and $\delta_2(t)$ are continuous and semialgebraic. Note that $\delta_1(t)$ and $\delta_2(t)$ are defined on $(c,d)$ since $a_{2::}(t)$ and $a_{::2}(t)$ are rank two on this interval. The functions $\delta_3(t)$ and $\delta_4(t)$ are continuous and semialgebraic on $(c,d)$ by the same reasoning. Define the continuous and semialgebraic functions $D_1(t)$, $D_2(t)$, $E_1(t)$, and $E_2(t)$ as follows.

$$D_1(t) = a_{1::}(t)\ +\ \delta_1(t)\ a_{2::}(t) \qquad E_1(t) = a_{::1}(t)\ +\ \delta_3(t)\ a_{::2}(t)$$
$$D_2(t) = a_{1::}(t)\ +\ \delta_2(t)\ a_{2::}(t) \qquad E_2(t) = a_{:1}(t)\ +\ \delta_4(t)\ a_{::2}(t)$$

We can thus compute the vector components of the rank one summands of $(a_{ijk})$ as the continuous and semialgebraic functions $v^i_j(t)$ defined as follows.

$$
\begin{aligned}
v^1_1(t) &= \begin{pmatrix} v^1_{11}(t) \\ v^1_{12}(t) \end{pmatrix} = E_2(t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\qquad v^1_2(t) &= \begin{pmatrix} v^1_{21}(t) \\ v^1_{22}(t) \end{pmatrix} = E_1(t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
v^2_1(t) &= \begin{pmatrix} v^2_{11}(t) \\ v^2_{12}(t) \end{pmatrix} = D_2(t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\qquad v^2_2(t) &= \begin{pmatrix} v^2_{21}(t) \\ v^2_{22}(t) \end{pmatrix} = D_1(t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
v^3_1(t) &= \begin{pmatrix} v^3_{11}(t) \\ v^3_{12}(t) \end{pmatrix} = \left( \begin{pmatrix} 1 & 0 \end{pmatrix} D_2(t) \right)^T &\qquad v^3_2(t) &= \begin{pmatrix} v^3_{21}(t) \\ v^3_{22}(t) \end{pmatrix} = \left( \begin{pmatrix} 1 & 0 \end{pmatrix} D_1(t) \right)^T
\end{aligned}
$$

There must thus be some constant $p_1(t)$ and $p_2(t)$ for every $t$ on our interval $(c, d)$ such that

$$(a_{ijk}) \;=\; p_1(t)\; v^1_1(t) \otimes v^2_1(t) \otimes v^3_1(t) \;+\; p_2(t)\; v^1_2(t) \otimes v^2_2(t) \otimes v^3_2(t).$$

By Theorem 22, there exists such $p_1(t)$ and $p_2(t)$ continuous and semialgebraic on our interval $(c, d)$. We leave it to the reader to check the similar case when one of the slabs of the hypermatrix $(a_{ijk})$ is rank one, and to confirm that it is sufficient to find a continuous and semialgebraic parameterization on a subinterval. $\square$

### 4.4 The Sets of Constant Rank Real Hypermatrices are Semialgebraic

Let $V^1$, $V^2$, and $V^3$ be two-dimensional real vector spaces. We denote the set of rank $r$ tensors by $M^{\mathbb{R}}_r$, and the set of rank at most $r$ tensors as $M^{\mathbb{R}}_{\leq r}$.

$$
\begin{aligned}
M^{\mathbb{R}}_r \;&=\; \{\ \rho\ \in V^1 \otimes V^2 \otimes V^3 \ \mid\ (\rho) = r\ \} \\
M^{\mathbb{R}}_{\leq r} \;&=\; \{\ \rho\ \in V^1 \otimes V^2 \otimes V^3 \ \mid\ (\rho) \leq r\ \}
\end{aligned}
$$

It thus follows that in our $2 \times 2 \times 2$ case, the tensor space $V^1 \otimes V^2 \otimes V^3$ is thus equal to the disjoint union $M^{\mathbb{R}}_{\leq 1} \sqcup M^{\mathbb{R}}_2 \sqcup M^{\mathbb{R}}_3$. The set $M^{\mathbb{R}}_{\leq 1}$ is closed in the norm topology, since a tensor is simple if and only if all of its linear contraction maps are rank less than or equal than one. It is well-known that set of linear maps that are rank less than or to one is closed in the norm topology. Hence, the only real $2 \times 2 \times 2$ rank-jumping tensors are rank three tensors that are limit points of a sequence of rank two tensors. That is, tensor $\beta$ is a $2 \times 2 \times 2$ real rank-jumping tensor if and only if $\beta$ is rank three and there exists a sequence of tensors $(\rho_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} \rho_n = \ \beta \ \text{ and } \ (\rho_n)_{n\in\mathbb{N}} \subseteq M_2^{\mathbb{R}}. \tag{4.8}$$

This is in general weaker than the statement that there exists a continuous mapping $\gamma$ such that

$$\gamma : \ [0,1] \subseteq \mathbb{R} \to V^1 \otimes V^2 \otimes V^3 \ \text{ and } \ \gamma(0) = \beta, \gamma(0,1] \subseteq M_2^{\mathbb{R}}. \tag{4.9}$$

However, (4.8) does imply (4.9) if $M_2^{\mathbb{R}}$ is semi-algebraic by the Curve Selection Lemma. In the following theorem, which is modified from [30, Thm.6.1], we show that the sets $M_r^{\mathbb{R}}$ are indeed semialgebraic by considering the map $\phi_r^{\mathbb{R}}$ from $(\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2)^r \cong \mathbb{R}^{6r}$ to $\mathbb{R}^{2\times 2\times 2}$ that sends

$$(\begin{pmatrix} a_{1,1}^1 \\ a_{1,2}^1 \end{pmatrix} \times \begin{pmatrix} a_{2,1}^1 \\ a_{2,2}^1 \end{pmatrix} \times \begin{pmatrix} a_{n_d,1}^1 \\ a_{n_d,2}^1 \end{pmatrix}) \ \times \ \ldots \ \times \ (\begin{pmatrix} a_{1,1}^r \\ a_{1,2}^r \end{pmatrix} \times \begin{pmatrix} a_{2,1}^r \\ a_{2,2}^1 \end{pmatrix} \times \begin{pmatrix} a_{n_d,1}^r \\ a_{n_d,2}^r \end{pmatrix})$$

to the tensor

$$\begin{pmatrix} a_{1,1}^1 \\ a_{1,2}^1 \end{pmatrix} \otimes \begin{pmatrix} a_{2,1}^1 \\ a_{2,2}^1 \end{pmatrix} \otimes \begin{pmatrix} a_{n_d,1}^1 \\ a_{n_d,2}^1 \end{pmatrix} \ + \ \ldots \ + \ \begin{pmatrix} a_{1,1}^r \\ a_{1,2}^r \end{pmatrix} \otimes \begin{pmatrix} a_{2,1}^r \\ a_{2,2}^1 \end{pmatrix} \otimes \begin{pmatrix} a_{n_d,1}^r \\ a_{n_d,2}^r \end{pmatrix}.$$

**Theorem 24.** *For any natural number $r$, the set $M_r^{\mathbb{R}}$ is semialgebraic.*

*Proof.* In coordinates, by the isomorphism (5.1), $\phi_r^{\mathbb{R}}$ sends

$$(\begin{pmatrix} a_{1,1}^1 \\ a_{1,2}^1 \end{pmatrix} \times \begin{pmatrix} a_{2,1}^1 \\ a_{2,2}^1 \end{pmatrix} \times \begin{pmatrix} a_{n_d,1}^1 \\ a_{n_d,2}^1 \end{pmatrix}) \ \times \ \ldots \ \times \ (\begin{pmatrix} a_{1,1}^r \\ a_{1,2}^r \end{pmatrix} \times \begin{pmatrix} a_{2,1}^r \\ a_{2,2}^1 \end{pmatrix} \times \begin{pmatrix} a_{n_d,1}^r \\ a_{n_d,2}^r \end{pmatrix})$$

to the hypermatrix

$$\begin{array}{|cc|}\hline \sum_{i=1}^r a_{211}^i a_{121}^i a_{112}^i & \sum_{i=1}^r a_{211}^i a_{121}^i a_{212}^i \\ \sum_{i=1}^r a_{211}^i a_{221}^i a_{211}^i & \sum_{i=1}^r a_{211}^i a_{221}^i a_{212}^i \\ \hline \end{array} \quad \begin{array}{|cc|}\hline \sum_{i=1}^r a_{111}^i a_{121}^i a_{112}^i & \sum_{i=1}^r a_{111}^i a_{121}^i a_{212}^i \\ \sum_{i=1}^r a_{111}^i a_{221}^i a_{211}^i & \sum_{i=1}^r a_{111}^i a_{221}^i a_{212}^i \\ \hline \end{array}. \tag{4.10}$$

Hypermatrix (4.10) has polynomial entries, so $\psi_2^{\mathbb{R}}$ is indeed a polynomial mapping. The Tarski-Seidenberg theorem implies that polynomial maps over $\mathbb{R}$ preserve semi-algebraic sets [34]. Hence, the set $M_{\leq r}^{\mathbb{R}}$ is semialgebraic as it is the image of $\psi_2^{\mathbb{R}}$, which has a semialgebraic domain. Moreover, $M_2^{\mathbb{R}} = M_{\leq 2}^{\mathbb{R}} \setminus M_{\leq 1}^{\mathbb{R}}$ is semialgebraic as semialgebraic sets are closed under set complements. □

### 4.5 A Coordinate-Free Characterization of Real $2 \times 2 \times 2$ Rank-Jumping Tensors

We have shown that rank three tensors in $\tau(X)$ are rank-jumping. We now show that a $2 \times 2 \times 2$ tensor is rank-jumping if and only if it is a rank three element in $\tau(X)$.

**Theorem 25.** *Let $V^1$, $V^2$, and $V^3$ be two-dimensional real vector spaces. A tensor $\beta$ in $V^1 \otimes V^2 \otimes V^3$ is rank-jumping and the limit of rank two tensors if and only if it is an element of $\tau(X)$ of rank greater than two.*

*Proof.* Let $\beta$ be a rank-jumping tensor that is the limit of rank two tensors. Since the set of rank two tensors is semialgebraic by Theorem 24, the Curve Selection Theorem [35, p.38] implies there exists a continuous mapping $\gamma$ such that

$$\begin{aligned} \gamma :\ [0,1] \subseteq \mathbb{R} \to V^1 \otimes V^2 \otimes V^3 \quad \text{and} \quad & \gamma(0) = \beta \\ & \gamma(0,1] \subseteq M_2^{\mathbb{R}}. \end{aligned} \tag{4.11}$$

By Theorem 23, we can further conclude that there exists continuous and semialgebraic functions $\gamma_1(t)$ and $\gamma_2(t)$ such that

$$\gamma(t) \;=\; \gamma_1(t) \;+\; \gamma_2(t) \;\text{ for all }\; t \in (0,1],$$

where $\gamma_1(t)$ and $\gamma_2(t)$ are rank one tensors for all $t \in (0,1]$. We switch to working in the projective space, since this space is compact. Denote the projectivization of a vector $v$ as $[v]$. The functions $t \mapsto [\gamma_1(t)]$ and $t \mapsto [\gamma_2(t)]$ are thus continuous and semialgebraic functions whose image is in a compact space. We can thus extend these functions continuously to functions $\widetilde{\gamma_1}$ and $\widetilde{\gamma_2}$ that are defined on the entire closed interval $[0,1]$. Similarly, let $\lambda(t)$ be the function from the half open interval $(0,1]$ to the space of lines in $\mathbb{P}(V^1 \otimes V^2 \otimes V^3)$ defined as

$$\lambda(t) \;=\; \langle\; [\gamma_1(t)],\; [\gamma_2(t)] \;\rangle,$$

where $\langle[\gamma_1(t)], [\gamma_2(t)]\rangle$ denotes the projective line spanned by $[\gamma_1(t)]$ and $[\gamma_2(t)]$. Theorem 23 also implies that $\lambda(t)$ is continuous and semialgebraic, so it can also be continuously extended to a function $\widetilde{\lambda}(t)$ defined on the interval $[0,1]$. By the continuity of this function, $[\beta]$, $\widetilde{\gamma_1}(0)$, and $\widetilde{\gamma_2}(0)$ are all on the projective line $\widetilde{\lambda}(0)$. Note that if $\widetilde{\gamma_1}(0) \neq \widetilde{\gamma_2}(0)$, then these two points would span the line $\widetilde{\lambda}(0)$, which contains $[\beta]$. This would imply that $\beta$ is rank at most two, contradicting our assumption that $\beta$ is rank greater than two. Hence, $\widetilde{\lambda}(0)$ is the limit of secant lines that is tangent to the Segre variety at the point $\widetilde{\gamma_1}(0) = \widetilde{\gamma_2}(0)$. Thus, $[\beta]$ is contained in the tangent line $\widetilde{\lambda}(0)$. $\square$

### 4.6 Uniqueness of Low Rank Decompositions

Our characterization of rank two tensors in Theorem 17 in fact implies that rank two $2{\times}2{\times}2$ tensors have unique decompositions as sums of rank one tensors. Such decompositions are not unique in general for rank three $2 \times 2 \times 2$ tensors. For example, let $V^1$, $V^2$, $V^3$ be two-dimensional vector spaces over a common field, and let $\{x_1^i, x_2^i\} \subset V^i$ for $i = 1, 2, 3$. It can be shown that the tensor

$$\rho \;=\; x_1^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_2^3 \;+\; x_2^1 \otimes x_1^2 \otimes x_2^3$$

is rank three, and can be decomposed as a sum of rank one tensors in the two following distinct ways:

$$\begin{aligned}
\rho \;&=\; (x_1^1 + x_2^1) \otimes (x_1^2 + x_2^2) \otimes x_1^3 \;+\; (-x_1^1) \otimes (x_1^2 + x_2^2) \otimes (x_1^3 - x_2^3) \\
&\qquad\qquad\qquad +\; (x_1^1 - x_2^1) \otimes x_1^2 \otimes (x_1^3 - x_2^3), \text{ and} \\
\rho \;&=\; (x_1^1 + x_2^1) \otimes (x_1^2 + x_2^2) \otimes x_1^3 \;+\; x_2^1 \otimes (x_1^2 + x_2^2) \otimes (x_2^3 - x_1^3) \\
&\qquad\qquad\qquad +\; (x_1^1 - x_2^1) \otimes x_1^2 \otimes (x_2^3 - x_1^3).
\end{aligned}$$

Hence, the decomposition of $\rho$ as three rank one summands is not unique. This does not happen for rank two tensors.

**Theorem 26.** *Let $V^1$, $V^2$, and $V^3$ be two-dimensional real or complex vector spaces. If $\rho$ is a regular tensor in $V^1 \otimes V^2 \otimes V^3$ such that*

$$\begin{aligned}
\rho \;&=\; v_1^1 \otimes v_1^2 \otimes v_1^3 \;+\; v_2^1 \otimes v_2^2 \otimes v_2^3, \quad \textit{and} \\
\rho \;&=\; \widehat{v_1^1} \otimes \widehat{v_1^2} \otimes \widehat{v_1^3} \;+\; \widehat{v_2^1} \otimes \widehat{v_2^2} \otimes \widehat{v_2^3},
\end{aligned}$$

*then there exists some permutation $\sigma \in S_2$ such that*

$$[v_j^i] = [\widehat{v_{\sigma j}^i}] \quad \textit{for all } i, j.$$

*That is, the set of rank one summands of $\rho$, $\{[v_1^1 \otimes v_1^2 \otimes v_1^3], [v_2^1 \otimes v_2^2 \otimes v_2^3]\}$, is unique as a set of projective points.*

*Proof.* Since $\rho$ is regular and rank two, by Theorem 17

$$\begin{aligned}
\mathbb{P}(\mathrm{im}\Pi_1(\rho)) \cap X^{2\times 2} &= \{\ [v_1^2 \otimes v_1^3],\ [v_2^2 \otimes v_2^3]\ \} \ = \ \{\ [\widehat{v_1^2} \otimes \widehat{v_1^3}],\ [\widehat{v_2^2} \otimes \widehat{v_2^3}]\ \},\\
\mathbb{P}(\mathrm{im}\Pi_2(\rho)) \cap X^{2\times 2} &= \{\ [v_1^1 \otimes v_1^3],\ [v_2^1 \otimes v_2^3]\ \} \ = \ \{\ [\widehat{v_1^2} \otimes \widehat{v_1^3}],\ [\widehat{v_2^2} \otimes \widehat{v_2^3}]\ \},\ \text{and}\\
\mathbb{P}(\mathrm{im}\Pi_3(\rho)) \cap X^{2\times 2} &= \{\ [v_1^1 \otimes v_1^2],\ [v_2^1 \otimes v_2^2]\ \} \ = \ \{\ [\widehat{v_1^2} \otimes \widehat{v_1^3}],\ [\widehat{v_2^2} \otimes \widehat{v_2^3}]\ \}.
\end{aligned}$$

Each of the two elements of $\mathbb{P}(\mathrm{im}\Pi_3(\rho)) \cap X^{2\times 2}$ induces a linear map from $V^{2*}$ to $V^1$ with a well-defined and one dimensional kernel. Let $f_1$ and $f_2$ be these two linear maps such that

$$\ker(f_1) = \langle v_1^1 \rangle \text{ and } \ker(f_2) = \langle v_2^1 \rangle.$$

$\widehat{v_1^1}$ must be in the kernel of one of these maps. Hence, $[\widehat{v_1^1}] = [v_1^1]$ or $[\widehat{v_1^1}] = [v_2^1]$. The other equalities follow similarly. $\square$

The uniqueness of low rank decompositions is an example of the usefulness of representing data as a hypermatrix as opposed to a matrix. Consider the following set of data organized as matrix $A = (a_{ij})$, where $a_{ij}$ represents the $i^{\text{th}}$ person's score on test $j$. This example is taken from [36]. What would it mean if the above matrix consistently had a good rank $r$ approximation for some $r$? If $A$ were a rank one matrix, it would be reasonable to conjecture that every person's test performances are the same. Similarly, if $A$ had a good rank $r$ approximation, that is if $a_{ij} \approx \alpha_i^1 \beta_j^1 + \alpha_i^2 \beta_j^2 + \ldots + \alpha_i^r \beta_j^r$ for all $i$,$j$ for some $\alpha_i^t$'s and $\beta_j^t$'s, then it would be reasonable to conjecture that the sampled exams are testing $r$ abilities, $\alpha_i^t$ represents the $i^{\text{th}}$ person's $t$ ability, and $\beta_j^t$ represents the degree to which exam $j$ tests ability $t$. It is a remarkable observation that many data sets do indeed consistently have good low rank matrix approximations. However, one immediate obstacle to interpreting matrix rank as above is that low rank matrix decompositions are not unique, so there is no way to recover the $\alpha_i^t$'s and $\beta_j^t$'s. To see this, note that if $m \times m$ matrix $A = \sum_{i=1}^r u^i (v^i)^t$, where $u^i, v^i \in \mathbb{C}^m$, then

$$A = \sum_{i=1}^{r} \begin{pmatrix} | \\ u^i \\ | \end{pmatrix} \begin{pmatrix} \text{——} & (v^i)^t & \text{——} \end{pmatrix} = \begin{pmatrix} | & | & & | \\ u^1 & u^2 & \cdots & u^r \\ | & | & & | \end{pmatrix} \begin{pmatrix} \text{——} & (v^1)^t & \text{——} \\ \text{——} & (v^2)^t & \text{——} \\ & \vdots & \\ \text{——} & (v^r)^t & \text{——} \end{pmatrix}$$

$$= UV^t \qquad \text{for } m \times r \text{ matrices } U = \begin{pmatrix} | & | & & | \\ u^1 & u^2 & \cdots & u^r \\ | & | & & | \end{pmatrix}, V = \begin{pmatrix} | & | & & | \\ v^1 & v^2 & \cdots & v^r \\ | & | & & | \end{pmatrix}$$

$$= UCC^{-1}V^t \ \text{ for any invertible } r \times r \quad C = \begin{pmatrix} | & | & & | \\ c^1 & c^2 & \cdots & c^r \\ | & | & & | \end{pmatrix}, \ C^{-1} = \begin{pmatrix} | & | & & | \\ d^1 & d^2 & \cdots & d^r \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & | & & | \\ Uc^1 & Uc^2 & \cdots & Uc^r \\ | & | & & | \end{pmatrix} \begin{pmatrix} \text{——} & (d^1)^t V^t & \text{——} \\ \text{——} & (d^2)^t V^t & \text{——} \\ & \vdots & \\ \text{——} & (d^r)^t V^t & \text{——} \end{pmatrix} = \sum_{i=1}^{r} \begin{pmatrix} | \\ Uc^i \\ | \end{pmatrix} \begin{pmatrix} \text{——} & (d^i)^t V^t & \text{——} \end{pmatrix}$$

So $A$ can also be written as $\sum_{i=1}^r Uc^i (Vd^i)^t$.

We have shown that hypermatrices often do not have this problem. If $B \in \mathbb{R}^{2 \times n \times 2}$ is a generic rank $n$ hypermatrix such that $B = \sum_{i=1}^r \ u^i \otimes v^i \otimes w^i$ with $\{u^i\}_{i=1}^r$, $\{v^i\}_{i=1}^r$, $\{w^i\}_{i=1}^r$ linearly independent, then the $u^i$'s, $v^i$'s, and $w^i$'s are uniquely determined up to scalar multiple.

### 4.7 Non-Negative Real Hypermatrices

The author's of [30, p.21] cite a private correspondence with Landsberg that offers an explanation of how the classical observation (4.1) could have been predicted. Let $\omega(t)$ be the continuous curve of rank one tensors

$$\omega(t) \;=\; (x_1^1 + tx_2^1) \otimes (x_1^2 + tx_2^2) \otimes (x_1^3 + tx_2^3), \;\; t \in \mathbb{R}.$$

It follows from the Leibniz rule on derivatives of products that

$$\begin{aligned} \frac{d\omega(t)}{dt} \;&=\; \frac{d(x_1^1 + tx_2^1)}{dt} \otimes (x_1^2 + tx_2^2) \otimes (x_1^3 + tx_2^3) \\ &\qquad + \;(x_1^1 + tx_2^1) \otimes \frac{d(x_1^2 + tx_2^2)}{dt} \otimes (x_1^3 + tx_2^3) \\ &\qquad + \;(x_1^1 + tx_2^1) \otimes (x_1^2 + tx_2^2) \otimes \frac{d(x_1^3 + tx_2^3)}{dt} \\ &=\; x_2^1 \otimes (x_1^2 + tx_2^2) \otimes (x_1^3 + tx_2^3) \\ &\qquad + \;(x_1^1 + tx_2^1) \otimes x_2^2 \otimes (x_1^3 + tx_2^3) \\ &\qquad + \;(x_1^1 + tx_2^1) \otimes (x_1^2 + tx_2^2) \otimes x_2^3. \end{aligned} \tag{4.12}$$

By evaluating (4.12) at 0, we conclude

$$\left.\frac{d\omega(t)}{dt}\right|_{t=0} \;=\; x_2^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_1^2 \otimes x_2^3 \;=\; \beta. \tag{4.13}$$

On the other hand, by the definition of the derivative

$$\begin{aligned} \left.\frac{d\omega(t)}{dt}\right|_{t=0} \;&=\; \lim_{t\to 0} \frac{\omega(t) - \omega(0)}{t - 0} \\ &=\; \lim_{t\to 0} \frac{(x_1^1 + tx_2^1) \otimes (x_1^2 + tx_2^2) \otimes (x_1^3 + tx_2^3) \;-\; x_1^1 \otimes x_1^2 \otimes x_1^3}{t}. \end{aligned} \tag{4.14}$$

We can write the limit (4.14) in terms of the natural numbers $n$ as

$$
\begin{aligned}
&= \lim_{\frac{1}{n}\to 0} \frac{(x_1^1+\frac{1}{n}x_2^1)\otimes(x_1^2+\frac{1}{n}x_2^2)\otimes(x_1^3+\frac{1}{n}x_2^3) \;-\; x_1^1\otimes x_1^2\otimes x_1^3}{\frac{1}{n}},\ n\in\mathbb{N} \\
&= \lim_{n\to\infty} n(x_1^1+\frac{1}{n}x_2^1)\otimes(x_1^2+\frac{1}{n}x_2^2)\otimes(x_1^3+\frac{1}{n}x_2^3) \;-\; nx_1^1\otimes x_1^2\otimes x_1^3. \qquad (4.15)
\end{aligned}
$$

In fact, in [22], it is proved that all rank-jumping tensors over $\mathbb{C}$ are limits of a sequence of rank three tensors also come from derivatives. That is, it is shown that if tensor $\beta$ in the complex tensor space $V^1\otimes V^2\otimes\cdots\otimes V^d$ of arbitrary dimensions is the limit of rank three tensors and rank greater than three, then $\beta$ is in one of the following forms:

1. $\beta \;=\; x'(0) \;+\; y$, for some smooth curve $x(t)$ contained in $X$ and some $y\in X$.

2. $\beta \;=\; x'(0) \;+\; y'(0)$, for some smooth curves $x(t)$ and $y(t)$ contained in $X$.

3. $\beta \;=\; x'(0) \;+\; x''(0)$, for some smooth curve $x(t)$ in $X$.

This suggests a fundamental relationship between rank-jumping tensors and derivatives, and perhaps explains why rank-jumping never occurs for non-negative real hypermatrices. Like derivatives, rank-jumping tensors are limits of sequences of difference quotients, and it is not possible to form difference quotients over the non-negative reals.

We summarize the formal proof in [32] that rank-jumping never occurs for non-negative hypermatrices now. Choose a norm $\|\cdot\|$ on the tensor space $V^1\otimes V^2\otimes V^3$. The components vectors of a tensor in $V^1\otimes V^2\otimes V^3$ are not bounded, in general. For example,

$$\begin{bmatrix}1\\1\end{bmatrix}\otimes\begin{bmatrix}1\\1\end{bmatrix} \;=\; \begin{bmatrix}n\\n\end{bmatrix}\otimes\begin{bmatrix}\frac{1}{n}\\\frac{1}{n}\end{bmatrix}$$

for any real number $n$. Using the multilinearity of the tensor product, however, a rank $r$ tensor can always be written in the form

$$\sum_{i=1}^{r}\ \sigma_i\ v_i^1\otimes v_i^2\otimes v_i^3$$

for some constants $\sigma_i$, where $\|v_i^j\| = 1$ for all $i, j$. By imposing more conditions of the component vectors, we can often make the values $\sigma_i$ unique. The singular value decomposition of matrices is a well-known example of this. Imposing the restriction on non-negativity is sufficient to show that the non-negative rank one summands of non-negative hypermatrix cannot be arbitrarily large. Hence, if the sequence of rank less than or equal to $r$ non-negative hypermatrices

$$\rho_n \; = \; \sum_{i=1}^{r} \; v_{n,i}^1 \otimes v_{n,i}^2 \otimes v_{n,i}^3$$

converges to some non-negative hypermatrix $\rho$ in the norm topology, then the sets $\{v_{n,i}^k\}_n$ must be bounded. Hence, each sequence $(v_{n,i}^k)_n$ contains a subsequential limit $v_i^k$. This implies that

$$\sum_{i=1}^{r} \; v_{n,i}^1 \otimes v_{n,i}^2 \otimes v_{n,i}^3 \; \to \; \sum_{i=1}^{r} \; v_i^1 \otimes v_i^2 \otimes v_i^3 \; = \; \rho,$$

and thus $\rho$ must be rank less than or equal to $r$. As a result, rank jumping never occurs for non-negative hypermatrices.

Another difference between real and non-negative real rank is the possible maximal rank. Real $2 \times 2 \times 2$ hypermatrices have a maximal real rank of three. Non-negative $2 \times 2 \times 2$ hypermatrices, however, have a maximal non-negative rank of four. Let $\{\mathrm{e}_1^i, \mathrm{e}_2^i\}$ be the standard basis of $\mathbb{R}^2$ for $i = 1, 2, 3$. Any tensor $\rho$ can be written in terms of the induced basis on $V^1 \otimes V^2 \otimes V^3$ as

$$\begin{aligned}
\rho \; &= a_{111} \; \mathrm{e}_1^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_1^3 \; + \; a_{112} \; \mathrm{e}_1^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_2^3 \; + \; a_{121} \; \mathrm{e}_1^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_1^3 \; + \; a_{122} \; \mathrm{e}_1^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_2^3 \\
&\quad + \; a_{211} \; \mathrm{e}_2^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_1^3 \; + \; a_{212} \; \mathrm{e}_2^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_2^3 \; + \; a_{221} \; \mathrm{e}_2^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_1^3 \; + \; a_{222} \; \mathrm{e}_2^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_2^3
\end{aligned}$$

for some scalars $a_{ijk}$. By the multilinearity of the tensor product,

$$\begin{aligned}
\rho \; &= \mathrm{e}_1^1 \otimes \mathrm{e}_1^2 \otimes (a_{111} \; \mathrm{e}_1^3 + a_{112} \; \mathrm{e}_2^3) \; + \; \mathrm{e}_1^1 \otimes \mathrm{e}_2^2 \otimes (a_{121} \; \mathrm{e}_1^3 + a_{122} \; \mathrm{e}_2^3) \\
&\quad + \; \mathrm{e}_2^1 \otimes \mathrm{e}_1^2 \otimes (a_{211} \; \mathrm{e}_1^3 + a_{212} \; \mathrm{e}_2^3) \; + \; \mathrm{e}_2^1 \otimes \mathrm{e}_2^2 \otimes (a_{221} \; \mathrm{e}_1^3 + a_{222} \; \mathrm{e}_2^3).
\end{aligned}$$

Hence, the rank of a $2 \times 2 \times 2$ tensor is at most four. We now prove that four is indeed the maximum possible rank in $\mathbb{R}_{\geq 0}^{2\times 2\times 2}$, which shows that rank of a $2 \times 2 \times 2$ tensor cannot be further bounded in general.

**Theorem 27.** *Let* $\{\mathrm{e}_1^i, \mathrm{e}_2^i\}$ *denote the standard basis of* $\mathbb{R}^2$ *for* $i = 1, 2, 3$. *The non-negative real* $2 \times 2 \times 2$ *tensor which equals*

$$\rho \;=\; \mathrm{e}_1^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_1^3 \;+\; \mathrm{e}_1^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_2^3 \;+\; \mathrm{e}_2^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_2^3 \;+\; \mathrm{e}_2^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_1^3$$

*is non-negative real rank four.*

*Proof.* Suppose for contradiction that $\rho$ were rank less than or equal to three. There would then exists vectors $u_j^i$ such that

$$\rho \;=\; u_1^1 \otimes u_1^2 \otimes u_1^3 \;+\; u_2^1 \otimes u_2^2 \otimes u_2^3 \;+\; u_3^1 \otimes u_3^2 \otimes u_3^3.$$

Furthermore, there must be scalars $\{a_{j,t}^i\}_{t,i,j=1}^{2,3,3}$ such that $u_i^j = a_{i,1}^j \mathrm{e}_1^j + a_{i,2}^j \mathrm{e}_2^j$ for all $i, j$ since the sets $\{\mathrm{e}_1^i, \mathrm{e}_2^i\}$ are bases. It then follows from multilinearity that

$$\begin{aligned}
\rho &= u_1^1 \otimes u_1^2 \otimes u_1^3 \;+\; u_2^1 \otimes u_2^2 \otimes u_2^3 \;+\; u_3^1 \otimes u_3^2 \otimes u_3^3 \\
&= (a_{11}^1 \mathrm{e}_1^1 + a_{12}^1 \mathrm{e}_2^1) \otimes (a_{11}^2 \mathrm{e}_1^2 + a_{12}^2 \mathrm{e}_2^2) \otimes (a_{11}^3 \mathrm{e}_1^3 + a_{12}^3 \mathrm{e}_2^3) \\
&\qquad + (a_{21}^1 \mathrm{e}_1^1 + a_{22}^1 \mathrm{e}_2^1) \otimes (a_{21}^2 \mathrm{e}_1^2 + a_{22}^2 \mathrm{e}_2^2) \otimes (a_{21}^3 \mathrm{e}_1^3 + a_{22}^3 \mathrm{e}_2^3) \\
&\qquad + \; (a_{31}^1 \mathrm{e}_1^1 + a_{32}^1 \mathrm{e}_2^1) \otimes (a_{31}^2 \mathrm{e}_1^2 + a_{32}^2 \mathrm{e}_2^2) \otimes (a_{31}^3 \mathrm{e}_1^3 + a_{32}^3 \mathrm{e}_2^3) \\
&= \sum_{i,j,k=1}^{2} \left( \sum_{t=1}^{3} a_{ti}^1 \; a_{tj}^2 \; a_{tk}^3 \right) \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3.
\end{aligned}$$

Since $T$ also equals $\mathrm{e}_1^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_1^3 \ + \ \mathrm{e}_1^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_2^3 \ + \ \mathrm{e}_2^1 \otimes \mathrm{e}_1^2 \otimes \mathrm{e}_2^3 \ + \ \mathrm{e}_2^1 \otimes \mathrm{e}_2^2 \otimes \mathrm{e}_1^3$, the following eight equations must hold.

$$a_{11}^1 a_{11}^2 a_{11}^3 \ + \ a_{21}^1 a_{21}^2 a_{21}^3 \ + \ a_{31}^1 a_{31}^2 a_{31}^3 \ = \ 1 \tag{4.16}$$
$$a_{11}^1 a_{11}^2 a_{12}^3 \ + \ a_{21}^1 a_{21}^2 a_{22}^3 \ + \ a_{31}^1 a_{31}^2 a_{32}^3 \ = \ 0 \tag{4.17}$$
$$a_{11}^1 a_{12}^2 a_{11}^3 \ + \ a_{21}^1 a_{22}^2 a_{21}^3 \ + \ a_{31}^1 a_{32}^2 a_{31}^3 \ = \ 0 \tag{4.18}$$
$$a_{11}^1 a_{12}^2 a_{12}^3 \ + \ a_{21}^1 a_{22}^2 a_{22}^3 \ + \ a_{31}^1 a_{32}^2 a_{32}^3 \ = \ 1 \tag{4.19}$$
$$a_{12}^1 a_{11}^2 a_{11}^3 \ + \ a_{22}^1 a_{21}^2 a_{21}^3 \ + \ a_{32}^1 a_{31}^2 a_{31}^3 \ = \ 0 \tag{4.20}$$
$$a_{12}^1 a_{11}^2 a_{12}^3 \ + \ a_{22}^1 a_{21}^2 a_{22}^3 \ + \ a_{32}^1 a_{31}^2 a_{32}^3 \ = \ 1 \tag{4.21}$$
$$a_{12}^1 a_{12}^2 a_{11}^3 \ + \ a_{22}^1 a_{22}^2 a_{21}^3 \ + \ a_{32}^1 a_{32}^2 a_{31}^3 \ = \ 1 \tag{4.22}$$
$$a_{12}^1 a_{12}^2 a_{12}^3 \ + \ a_{22}^1 a_{22}^2 a_{22}^3 \ + \ a_{32}^1 a_{32}^2 a_{32}^3 \ = \ 0 \tag{4.23}$$

It remains to show that these eight equations cannot hold simultaneously. Equation (4.16) implies that at least one element in the set $\{a_{11}^1 a_{11}^2 a_{11}^3,\ a_{21}^1 a_{21}^2 a_{21}^3,\ a_{31}^1 a_{31}^2 a_{31}^3\}$ does not equal zero. Suppose $a_{11}^1 a_{11}^2 a_{11}^3 \neq 0$, which would imply that $a_{11}^1 \neq 0$, $a_{11}^2 \neq 0$ and $a_{11}^3 \neq 0$. $a_{11}^1, a_{11}^2 \neq 0$ and (4.17) implies $a_{12}^3 = 0$. $a_{11}^2, a_{11}^3 \neq 0$ and (4.18) implies $a_{12}^2 = 0$. $a_{11}^3, a_{11}^2 \neq 0$ and (4.20) implies $a_{12}^1 = 0$. Hence,

$$a_{12}^1, a_{12}^2, a_{12}^3 = 0. \tag{4.24}$$

Equations (4.24) and (4.19) further implies that one element in the set $\{a_{21}^1 a_{22}^2 a_{22}^3, a_{31}^1 a_{32}^2 a_{32}^3\}$ does not equal zero. Suppose $a_{21}^1 a_{22}^2 a_{22}^3 \neq 0$, which would imply that $a_{21}^1 \neq 0$, $a_{22}^2 \neq 0$ and $a_{22}^3 \neq 0$. $a_{21}^1, a_{22}^3 \neq 0$ and (4.17) implies $a_{21}^2 = 0$. $a_{21}^1, a_{22}^2 \neq 0$ and (4.18) implies $a_{21}^3 = 0$. $a_{22}^2, a_{22}^3 \neq 0$ and (4.23) implies $a_{22}^1 = 0$. Hence,

$$a_{22}^1, a_{21}^2, a_{21}^3 = 0. \tag{4.25}$$

Equations (4.24), (4.25), and (4.21) now implies that $a^1_{32}a^2_{31}a^3_{32}$ does not equal zero, which implies that $a^1_{32} \neq 0$, $a^2_{31} \neq 0$ and $a^3_{32} \neq 0$. $a^2_{21}, a^3_{21} \neq 0$ and (4.17) implies $a^1_{32} = 0$. $a^1_{32}, a^2_{31} \neq 0$ and (4.20) implies $a^3_{31} = 0$. $a^1_{32}, a^3_{32} \neq 0$ and (4.23) implies $a^2_{32} = 0$. Hence,

$$a^1_{32}, a^2_{32}, a^3_{31} = 0. \tag{4.26}$$

However, equations (4.24), (4.25), and (4.26) contradict equation (4.22). We leave it to the reader to check that the remaining cases result in a similar contradiction. $\square$

# 5. LOW RANK APPROXIMATIONS

We provide a coordinate-free proof that real $2\times2\times2$ rank three tensors do not have optimal rank two approximations with respect to the Frobenius norm. This result was first proved in [30, Thm. 8.1] by considering the $\mathrm{GL}(V^1)\times\mathrm{GL}(V^2)\times\mathrm{GL}(V^3)$ orbit classes of $V^1\otimes V^2\otimes V^3$ and the $2\times2\times2$ hyperdeterminant. Our coordinate-free proof expands on the result in [30] by developing a proof method that can be generalized more readily to higher dimensional $n_1\times n_2\times n_3$ tensor spaces.

## 5.1 Optimal Low Rank Approximations May Not Exist

Let $V^1$, $V^2$, and $V^3$ be two-dimensional real vector spaces, respectively, and let $V^1\otimes V^2\otimes V^3$ denote the tensor product of these spaces. The tensors in the form $v^1\otimes v^2\otimes v^3$ for some vectors $v^1$, $v^2$, and $v^3$ are called simple tensors. Every tensor can be written as the sum of finitely many simple tensors. The rank of a tensor $\rho$ is the minimum number $n$ such that $\rho$ is the sum of $n$ simple tensors. Once a basis $\{\mathrm{e}^s_t\}_{t=1}^2$ of $V^s$ is chosen for $s=1,2,3$, a tensor in $V^1\otimes V^2\otimes V^3$ can be coordinatized as a hypermatrix in $\mathbb{R}^{2\times2\times2}$ by the isomorphism from $V^1\otimes V^2\otimes V^3$ to $\mathbb{R}^{2\times2\times2}$ defined on simple tensors as

$$(a^1_1\mathrm{e}^1_1+a^1_2\mathrm{e}^1_2)\otimes(a^2_1\mathrm{e}^2_1+a^2_2\mathrm{e}^2_2)\otimes(a^3_1\mathrm{e}^3_1+a^3_2\mathrm{e}^3_2)\quad\longmapsto\quad\begin{array}{c}\boxed{\begin{matrix}a^1_2a^2_1a^3_1 & a^1_2a^2_1a^3_2\\ a^1_2a^2_2a^3_1 & a^1_2a^2_2a^3_2\end{matrix}}\\ \boxed{\begin{matrix}a^1_1a^2_1a^3_1 & a^1_1a^2_1a^3_2\\ a^1_1a^2_2a^3_1 & a^1_1a^2_2a^3_2\end{matrix}}\end{array}\tag{5.1}$$

for some real constants $a^s_t$. A lot of information about the rank of the $2\times2\times2$ real hypermatrix

$$(a_{ijk})\;=\;\begin{array}{c}\boxed{\begin{matrix}a_{211} & a_{212}\\ a_{221} & a_{222}\end{matrix}}\\ \boxed{\begin{matrix}a_{111} & a_{112}\\ a_{121} & a_{122}\end{matrix}}\end{array}$$

can be inferred from the sign of the polynomial

$$\begin{aligned}\Delta(a_{ijk}) = (&a_{111}^2a_{222}^2 + a_{112}^2a_{221}^2 + a_{121}^2a_{212}^2 + a_{122}^2a_{211}^2) \\ &- 2(a_{111}a_{112}a_{221}a_{222} + a_{111}a_{121}a_{212}a_{222} + a_{111}a_{122}a_{211}a_{222}) \\ &- 2(a_{112}a_{121}a_{212}a_{221} + a_{112}a_{122}a_{221}a_{211} + a_{121}a_{122}a_{212}a_{211}) \\ &+ 4(a_{111}a_{122}a_{212}a_{221} + a_{112}a_{121}a_{211}a_{222}).\end{aligned}$$

The polynomial $\Delta$ is called the $2\times2\times2$ hyperdeterminant. $\Delta(a_{ijk}) > 0$ implies hypermatrix $(a_{ijk})$ is rank two, and $\Delta(a_{ijk}) < 0$ implies $(a_{ijk})$ is rank three. The existence of such a polynomial is not particular to the $2 \times 2 \times 2$ case. The sets of constant rank real hypermatrices are semialgebraic for any dimensions $n_1 \times n_2 \times \cdots \times n_d$ by Theorem 24. Hence, the rank of a real hypermatrix can always be computed by evaluating a finite number of polynomial equalities and inequalities, where the variables of the polynomials are the entries of the hypermatrix. Because of this, real tensors are usually studied in coordinates by representing tensors as hypermatrices and then exploiting properties of these polynomials. The polynomials that define the sets of constant rank $2 \times 2 \times 2$ tensors are different, however, from the polynomials that define the sets of constant rank $n_1 \times n_2 \times n_3$ tensors for any other dimensions $n_1 \times n_2 \times n_3$. As a result, coordinate proofs about $2\times2\times2$ tensors that rely on these polynomials cannot be readily generalized to the arbitrary $n_1 \times n_2 \times n_3$ case. In this chapter, we take an alternative coordinate-free approach in our study of $2 \times 2 \times 2$ tensors that lends itself more readily to generalization. We prove that real $2 \times 2 \times 2$ rank three tensors do not have optimal rank two approximations with respect to the Frobenius norm without relying on hypermatrices or the hyperdeterminant.

Matrix rank is lower semi-continuous for matrices of real numbers. That is, if a sequence of rank $r$ real matrices converges to a matrix $A$ in the norm topology, then the rank of $A$ is less than or equal to $r$. This guarantees the existence of optimal low rank approximations of real matrices. A sequence of real hypermatrices, however, can converge to a hypermatrix of greater rank. As a result, optimal low rank approximations may not exist for real hypermatrices. In fact, the set of real hypermatrices with no optimal low rank approximation often has positive Lebesgue measure [21], so characterizing such real hypermatrices is an important step in implementing algorithms with real hypermatrices.

Let $\rho \in V^1 \otimes V^2 \otimes V^3$ be a tensor of rank $r$, and let $s \leq r$. An optimal rank $s$ approximation of $\rho$ with respect to norm $\|\cdot\|$ is a tensor $\upsilon$ such that

$$\|\rho - \upsilon\| \;=\; \inf_{\substack{\upsilon' \in V^1 \otimes V^2 \otimes V^3 \\ \mathrm{rk}(\upsilon') \leq s}} \|\rho - \upsilon'\|.$$

It is a classical observation that tensors in the form

$$\beta \;=\; x_2^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_1^2 \otimes x_2^3 \tag{5.2}$$

are limit points of the sequence of rank at most two tensors

$$\rho_n \;=\; n\,(x_1^1 + \frac{1}{n}x_2^1) \otimes (x_1^2 + \frac{1}{n}x_2^2) \otimes (x_1^3 + \frac{1}{n}x_2^3) \;-\; n x_1^1 \otimes x_1^2 \otimes x_1^3,\; n \in \mathbb{N} \tag{5.3}$$

in the norm topology. It follows from the triangle inequality and the multilinearity of the tensor product that

$$\|\rho_n - \beta\| \;\leq\; \frac{1}{n}\left\|x_2^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_1^2 \otimes x_2^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_2^3\right\| \;+\; \frac{1}{n^2}\left\|x_2^1 \otimes x_2^2 \otimes x_2^3\right\|.$$

Hence, $\lim_{n\to\infty} \|\rho_n - \beta\| = 0$, so $\lim_{n\to\infty} \rho_n$ indeed equals $\beta$. In Section 4.2, we gave a coordinate-free proof that tensors in the form of (5.2) are in fact rank three when the sets $\{x_1^1, x_2^1\}$, $\{x_1^2, x_2^2\}$, and $\{x_1^3, x_2^3\}$ are independent. Thus, when $\{x_1^1, x_2^1\}$, $\{x_1^2, x_2^2\}$, and $\{x_1^3, x_2^3\}$ are independent, $\beta$ is an example of a rank three tensor without an optimal rank two approximation. The main result of this chapter is a coordinate-free proof that in fact *every* rank three $2 \times 2 \times 2$ tensor has no optimal rank two approximation with respect to the Frobenius norm, not just the tensors in the form of (5.2). General rank three $2 \times 2 \times 2$ real tensors are not limit points of sequences of rank two tensors. Rather, we will show that their failure to have optimal rank two approximations is due to the curvature of the Segre variety.

This result was first proved in coordinates in [30, Thm. 8.1]. The argument in [30] can be summarized as follows. Suppose for contradiction that the $2 \times 2 \times 2$ real hypermatrix

$B$ is an optimal rank two approximation of rank three $2 \times 2 \times 2$ real hypermatrix $A$ with respect to the Frobenius norm. By considering properties of the polynomial $\Delta$, it follows that $\Delta(B) = 0$. The polynomial $\Delta$ is invariant on the $\mathrm{GL}(V^1) \times \mathrm{GL}(V^2) \times \mathrm{GL}(V^3)$ orbit classes of $V^1 \otimes V^2 \otimes V^3$, and only three of the eight orbit classes are zero on $\Delta$. Hypermatrices in these three orbit classes are equivalent up to an orthogonal change of coordinates to hypermatrices in the form

$$\begin{array}{cc} & \boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}} \\ \boxed{\begin{matrix} \lambda & 0 \\ 0 & \mu \end{matrix}} & \end{array} \tag{5.4}$$

for some real $\lambda$ and $\mu$. Since $B$ is rank two, we may thus assume $B$ is in form (5.4) with both $\lambda$ and $\mu$ nonzero. Finally, it is shown that if $H$ is a $2 \times 2 \times 2$ hypermatrix such that

$$\Delta(B + \epsilon H) = 0 \tag{5.5}$$

for all real $\epsilon$, then $A - B$ is orthogonal to $H$. By considering various hypermatrices $H$ that satisfy (5.5), the authors of [30] then conclude that

$$A - B = \begin{array}{cc} & \boxed{\begin{matrix} a\mu & 0 \\ 0 & -a\lambda \end{matrix}} \\ \boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}} & \end{array},$$

for some constant $a$. This implies that

$$A = \begin{array}{cc} & \boxed{\begin{matrix} a\mu & 0 \\ 0 & -a\lambda \end{matrix}} \\ \boxed{\begin{matrix} \lambda & 0 \\ 0 & \mu \end{matrix}} & \end{array}.$$

However, this hypermatrix is rank two, contradicting that $A$ is rank three.

Our coordinate-free proof expands on the result in [30] by developing a proof method that can be generalized more readily to higher dimensional tensor spaces. Our proof is also a proof by contradiction. We suppose for contradiction that rank two tensor $\upsilon$ in $V^1 \otimes V^2 \otimes V^3$ is an optimal rank two approximation of rank three tensor $\rho$. By considering the relationship between the mode-1 contraction maps of $\rho$ and $\upsilon$, we also derive the contradiction that $\rho$ is rank two. Our proof has the interesting geometric corollary that the nearest point of a rank three tensor to the second secant set of the Segre variety is a rank three tensor in the tangent space of the Segre variety.

## 5.2 A Characterization of Optimal Rank Two Approximations

The set of simple tensors is a variety, and it is called the Segre variety. It is the image of the map

$$\begin{aligned} V^1 \times V^2 \times V^3 &\rightarrow V^1 \otimes V^2 \otimes V^3 \\ \upsilon^1 \times \upsilon^2 \times \upsilon^3 &\mapsto \upsilon^1 \otimes \upsilon^2 \otimes \upsilon^3. \end{aligned}$$

The Segre variety of the $n_1 \times n_2 \times n_3$ tensor space $V^1 \otimes V^2 \otimes V^3$ is denoted $X^{n_1 \times n_2 \times n_3}$ or simply $X$ when the dimensions are clear from context. The tangent space of the Segre variety at a point $\upsilon$ is denoted as $T_\upsilon(X)$, and is characterized in [31] as in the following theorem.

**Theorem 28.** *Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, $n_3$-dimensional real vector spaces, respectively, and let $x_1^1 \otimes x_1^2 \otimes x_1^3$ be a rank one tensor in the Segre variety $X$. The tangent space of $X$ at $x_1^1 \otimes x_1^2 \otimes x_1^3$ is the space of all tensors in the form*

$$x_2^1 \otimes x_1^2 \otimes x_1^3 \; + \; x_1^1 \otimes x_2^2 \otimes x_1^3 \; + \; x_1^1 \otimes x_1^2 \otimes x_2^3 \tag{5.6}$$

*for some $x_2^1 \in V^1$, $x_2^2 \in V^2$, and $x_2^3 \in V^3$.*

We have seen that the set of rank at most two tensors is not closed with respect to the norm topology, which motivates the following definition of the $2^{nd}$ secant set of the Segre variety.

**Definition 29.** *The $2^{nd}$ secant set of the Segre variety, denoted $\sigma_2(X)$, is the norm closure of the set of all tensors that lie on secant lines of the Segre variety $X$.*

If $\upsilon$ is an optimal rank two approximation of $\rho$ with respect to an inner product norm, then $\upsilon - \rho$ must be orthogonal to the tangent space of $\sigma_2(X)$ at $\upsilon$, which we characterize in the following theorem using Theorem 28 and Terracini's Lemma [37].

**Theorem 30.** *Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, $n_3$-dimensional real vector spaces, respectively, and let $\upsilon = x_1^1 \otimes x_1^2 \otimes x_1^3 + x_2^1 \otimes x_2^2 \otimes x_2^3$ be a rank two tensor in $V^1 \otimes V^2 \otimes V^3$. The tangent space of $\sigma_2(X)$ at $\upsilon$ is the space of all tensors in the form*

$$\begin{aligned} & x_3^1 \otimes x_1^2 \otimes x_1^3 \ + \ x_1^1 \otimes x_3^2 \otimes x_1^3 \ + \ x_1^1 \otimes x_1^2 \otimes x_3^3 \\ & \quad + \ x_4^1 \otimes x_2^2 \otimes x_2^3 \ + \ x_2^1 \otimes x_4^2 \otimes x_2^3 \ + \ x_2^1 \otimes x_2^2 \otimes x_4^3 \end{aligned} \tag{5.7}$$

*for some $x_3^1, x_4^1 \in V^1$, $x_3^2, x_4^2 \in V^2$, and $x_3^3, x_4^3 \in V^3$.*

Until now, we have worked with $n_1 \times n_2 \times n_3$ tensors. The next theorem, however, is our first statement that must be restricted to $2 \times 2 \times 2$ tensors.

**Theorem 31.** *Let $V^1$, $V^2$, and $V^3$ be 2-dimensional real vector spaces, and let $\rho \in V^1 \otimes V^2 \otimes V^3$ be of rank greater than two. If $\upsilon \in V^1 \otimes V^2 \otimes V^3$ is an optimal rank two approximation of $\rho$ with respect to an inner product norm, then $\operatorname{im} \Pi_i(\upsilon)$ is not dimension two for some $i = 1, 2,$ or 3.*

*Proof.* Suppose for contradiction that $\operatorname{im} \Pi_i(\upsilon)$ were dimension two for $i = 1, 2,$ and 3. Then, there would exist three linearly independent sets $\{x_1^i, x_2^i\}$ in $V^i$ for $i = 1, 2,$ and 3, such that

$$\upsilon \ = \ x_1^1 \otimes x_1^2 \otimes x_1^3 \ + \ x_2^1 \otimes x_2^2 \otimes x_2^3.$$

It follows that the set $\{x_i^1 \otimes x_j^2 \otimes x_k^3\}_{i,j,k=1}^{2,2,2}$ is a basis of $V^1 \otimes V^2 \otimes V^3$, so there must exist constants $c_{ijk}$ such that $\rho = \sum_{i,j,k=1}^{2,2,2} c_{ijk}\ x_i^1 \otimes x_j^2 \otimes x_k^3$. By multilinearity,

$$\begin{aligned}\rho = (c_{111}x_1^1 + c_{211}x_2^1) \otimes x_1^2 \otimes x_1^3 \ + \ x_1^1 \otimes c_{121}x_2^2 \otimes x_1^3 \ + \ x_1^1 \otimes x_1^2 \otimes c_{112}x_2^3 \\ + \ c_{122}x_1^1 \otimes x_2^2 \otimes x_2^3 \ + \ x_2^1 \otimes c_{212}x_1^2 \otimes x_2^3 \ + \ x_2^1 \otimes x_2^2 \otimes (c_{221}x_1^3 + c_{222}x_2^3).\end{aligned}$$

Hence, $\rho$ is in the form of (5.7), and, thus, it is in the tangent space of $\sigma_2(X)$ at $\upsilon$. This implies that $\upsilon - \rho$ is both in $T_\upsilon\sigma_2(X)$ and orthogonal to $T_\upsilon\sigma_2(X)$, which implies that $\rho = \upsilon$. However, this is a contradiction, since the rank of $\rho$ is not equal to the rank of $\upsilon$ by hypothesis. $\square$

## 5.3 P-Norms

Once bases of vector spaces $V^1$, $V^2$, and $V^3$ are chosen, we can define explicit norms on the tensor space $V^1 \otimes V^2 \otimes V^3$. Choose basis $\{\mathrm{e}_j^i\}_{j=1}^{n_i}$ of $V^i$ for $i = 1$, 2, and 3, and denote the corresponding dual basis as $\{\mathrm{e}_j^{i*}\}_{j=1}^{n_i}$ also for $i = 1$, 2, and 3. Let $B$ denote the collection of these bases. For $\rho \in V^1 \otimes V^2 \otimes V^3$, we define the following class of norms for any positive integer $p$.

$$\|\rho\|_{B,p} \ = \ \left( \sum_{i,j,k=1,1,1}^{n_1,n_2,n_3} \|a_{ijk}\|^p \right)^{\frac{1}{p}},$$

where $\rho = \sum_{i,j,k} a_{ijk}\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$. Similarly, for $\kappa \in V^r \otimes V^s$ and $v^r \in V^r$ for $1 \le r, s \le 3$, we define

$$\|\kappa\|_{B,p} \ = \ \left( \sum_{i,j=1,1}^{n_r,n_s} \|a_{ij}\|^p \right)^{\frac{1}{p}}, \text{ and } \ \|v^r\|_{B,p} \ = \ \left( \sum_{i=1}^{n_r} \|a_i\|^p \right)^{\frac{1}{p}},$$

where $\kappa = \sum_{i,j} a_{ij}\ \mathrm{e}_i^r \otimes \mathrm{e}_j^s$ and $v^r = \sum_i a_i\ \mathrm{e}_i^r$. These norms are a convenient choice for working in tensor spaces as they work well with contraction maps.

**Theorem 32.** *Let $V^1$, $V^2$, and $V^3$ be finite $n_1$, $n_2$, and $n_3$-dimensional real vector spaces, and let $B$ be the collection of bases of $V^1$, $V^2$, and $V^3$ defined above. For $\rho \in V^1 \otimes V^2 \otimes V^3$,*

$$\|\rho\|_{B,p}^p \ = \ \sum_{t=1}^{n_i} \ \|\Pi_i(\rho)(\mathrm{e}_t^{i*})\|_{B,p}^p$$

*for any mode-$i$ contraction $\Pi_i$ for $i = 1, 2, 3$.*

*Proof.* Without loss of generality, we prove the theorem for the the mode-1 contraction. Suppose $\rho = \sum_{s=1}^{r} v_s^1 \otimes v_s^2 \otimes v_s^3$ for some vectors $v_s^i$, and let $a_{s,t}^i$ be scalars such that $v_s^i = \sum_{t=1}^{n_i} a_{s,t}^i \mathrm{e}_t^i$ for all $i$ and $s$. It follows that

$$\sum_{s=1}^{r} v_s^1 \otimes v_s^2 \otimes v_s^3 \;=\; \sum_{i,j,k=1}^{n_1,n_2,n_3} \left( \sum_{s=1}^{r} a_{si}^1 a_{sj}^2 a_{sk}^3 \right) \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3.$$

Hence,

$$\begin{aligned} \|\rho\|_{B,p}^p &= \sum_{i,j,k=1}^{n_1,n_2,n_3} \left\| \sum_{s=1}^{r} a_{si}^1 a_{sj}^2 a_{sk}^3 \right\|^p \\ &= \sum_{i=1}^{n_1} \sum_{j,k=1}^{n_2,n_3} \left\| \sum_{s=1}^{r} a_{si}^1 a_{sj}^2 a_{sk}^3 \right\|^p \;=\; \sum_{i=1}^{n_1} \|\Pi_1(\rho)(\mathrm{e}_i^{1*})\|_{B,p}^p. \end{aligned}$$

□

Furthermore, when $p = 2$, the norm $\|\cdot\|_{B,2}$ is induced by the inner product

$$\langle \rho, \upsilon \rangle_{B,2} = \sum_{i,j,k} a_{ijk} b_{ijk},$$

where $\rho = \sum_{ijk} a_{ijk}\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$ and $\upsilon = \sum_{ijk} b_{ijk}\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$. The norm $\|\cdot\|_{B,2}$ is called the Frobenius norm with respect to bases $B$.

**Theorem 33.** *The Frobenius inner product has the following property on rank one tensors.*

$$\langle v_1^1 \otimes v_1^2 \otimes v_1^3 \mid v_2^1 \otimes v_2^2 \otimes v_2^3 \rangle_{B,2} \;=\; \langle v_1^1 \mid v_2^1 \rangle_{B,2}\ \langle v_1^2 \mid v_2^2 \rangle_{B,2}\ \langle v_1^3 \mid v_2^3 \rangle_{B,2}$$

*for any rank one tensors $v_1^1 \otimes v_1^2 \otimes v_1^3,\ v_2^1 \otimes v_2^2 \otimes v_2^3 \in V^1 \otimes V^2 \otimes V^3$.*

*Proof.* Let $a_{t,s}^i$ be scalars such that $v_t^i = \sum_{t=1}^{n_i} a_{t,s}^i \mathrm{e}_s^i$ for $i = 1, 2, 3$ and $t = 1, 2$. It follows that

$$\begin{aligned} v_1^1 \otimes v_1^2 \otimes v_1^3 &= \sum_{i,j,k=1}^{n_1,n_2,n_3} a_{1i}^1 a_{1j}^2 a_{1k}^3\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3, \text{ and} \\ v_2^1 \otimes v_2^2 \otimes v_2^3 &= \sum_{i,j,k=1}^{n_1,n_2,n_3} a_{2i}^1 a_{2j}^2 a_{2k}^3\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3. \end{aligned}$$

Hence,

$$\begin{aligned}
\langle v_1^1 \otimes v_1^2 \otimes v_1^3 \mid v_2^1 \otimes v_2^2 \otimes v_2^3 \rangle_{B,2} &= \sum_{i,j,k=1}^{n_1,n_2,n_3} a_{1i}^1 a_{1j}^2 a_{1k}^3 a_{2i}^1 a_{2j}^2 a_{2k}^3 \\
&= \left( \sum_{i=1}^{n_1} a_{1i}^1 a_{2i}^1 \right) \left( \sum_{j=1}^{n_2} a_{1j}^2 a_{2j}^2 \right) \left( \sum_{k=1}^{n_3} a_{1k}^3 a_{2k}^3 \right) \\
&= \langle v_1^1 \mid v_2^1 \rangle_{B,2} \ \langle v_1^2 \mid v_2^2 \rangle_{B,2} \ \langle v_1^3 \mid v_2^3 \rangle_{B,2}.
\end{aligned}$$

□

An optimal rank $r$ approximation could be rank strictly less than $r$ by our definition. We now show that an optimal rank $r$ approximations with respect to $p$-norms must be rank $r$. This theorem is modified from [30, Lemma 8.2].

**Theorem 34.** *Let $V^i$ be finite $n_i$-dimensional real vector spaces, and let $\rho \in V^1 \otimes V^2 \otimes V^3$ have rank greater than $r$. If $\upsilon$ is an optimal rank $r$ approximation of $\rho$ with respect to $\|\cdot\|_{B,p}$, then $\upsilon$ must be rank $r$.*

*Proof.* Suppose for contradiction that there existed an optimal rank $r$ approximation $\upsilon$ with rank strictly less than $r$. Let $a_{ijk}$ and $b_{ijk}$ be scalars such that $\rho = \sum_{ijk} a_{ijk}\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$ and $\upsilon = \sum_{ijk} b_{ijk}\ \mathrm{e}_i^1 \otimes \mathrm{e}_j^2 \otimes \mathrm{e}_k^3$. Since $\rho$ and $\upsilon$ have different ranks, there must exists some triple $(\alpha, \beta, \gamma)$ such that $a_{\alpha\beta\gamma} \neq b_{\alpha\beta\gamma}$. The tensor $\upsilon' = \upsilon + (a_{\alpha\beta\gamma} - b_{\alpha\beta\gamma})\mathrm{e}_\alpha^1 \otimes \mathrm{e}_\beta^2 \otimes \mathrm{e}_\gamma^3$ is rank less than or equal to $r$ by construction. It follows that

$$\begin{aligned}
\|\rho - \upsilon'\|_{B,p} &= \|\rho - \left(\upsilon + (a_{\alpha\beta\gamma} - b_{\alpha\beta\gamma})\mathrm{e}_\alpha^1 \otimes \mathrm{e}_\beta^2 \otimes \mathrm{e}_\gamma^3\right)\|_{B,p} \\
&= \left( \sum_{ijk \neq \alpha\beta\gamma} \|a_{ijk} - b_{ijk}\|^p \right)^{\frac{1}{p}} \\
&< \left( \sum_{ijk} \|a_{ijk} - b_{ijk}\|^p \right)^{\frac{1}{p}} = \|\rho - \upsilon\|_{B,p},
\end{aligned}$$

which contradicts that $\upsilon$ is an optimal rank $r$ approximation of $\rho$. □

If $\upsilon$ is an optimal rank two approximation of $\rho$ with respect to the Frobenius norm, then the contractions of $\upsilon$ must be related to the contractions of $\rho$ in the following way.

**Theorem 35.** *Let $V^1$, $V^2$, and $V^3$ be $n_1$, $n_2$, and $n_3$-dimensional real vector spaces, respectively. Let $B$ denote the collection of bases $\{\mathrm{e}^i_j\}_{j=1}^{n_i}$ of $V^i$ for $i = 1$, 2, and 3. Furthermore, denote the corresponding dual bases as $\{\mathrm{e}^{i*}_j\}_{j=1}^{n_i}$ for $i = 1$, 2, and 3. Let $\rho \in V^1 \otimes V^2 \otimes V^3$ be of rank greater than two, and let $\upsilon \in V^1 \otimes V^2 \otimes V^3$ be an optimal rank two approximation of $\rho$ with respect to the Frobenius norm $\|\cdot\|_{B,2}$. Let $P_{\operatorname{im}\Pi_i(\upsilon)}$ denote the projection onto the image of the mode-$i$ contraction of $\upsilon$. It follows that*

$$P_{\operatorname{im}\Pi_i(\upsilon)}(\Pi_i(\rho)(\mathrm{e}^{i*}_j)) \;=\; \Pi_i(\upsilon)(\mathrm{e}^{i*}_j) \textit{ for any } i, j.$$

*Proof.* As $\upsilon$ is rank two, there must exist vectors $x^i_j$ such that

$$\upsilon = \; x^1_1 \otimes x^2_1 \otimes x^3_1 \;+\; x^1_2 \otimes x^2_2 \otimes x^3_2.$$

Without loss of generality, we prove the theorem for $i$ and $j$ both equal to 1. First, note that the image of $\Pi_1(\upsilon)$ contains vectors in the form $ax^2_1 \otimes x^3_1 + bx^2_2 \otimes x^3_2$ for some constants $a$ and $b$. If the set $\{x^1_1, x^1_2\}$ is independent, it can be extended to a basis $\{x^1_j\}_{j=1}^{n_1}$ with dual basis $\{x^{1*}_j\}_{j=1}^{n_1}$. It then follows that the image of $\Pi_1(\upsilon)$ is the span of $\Pi_1(\upsilon)(x^{1*}_1) = x^2_1 \otimes x^3_1$ and $\Pi_1(\upsilon)(x^{1*}_2) = x^2_2 \otimes x^3_2$, so every element in the image of $\Pi_1(\upsilon)$ can indeed be written as $ax^2_1 \otimes x^3_1 + bx^2_2 \otimes x^3_2$ for some constants $a$ and $b$. On the other hand, if $x^1_2 = kx^1_1$ for some constant $k$, then the set $\{x^1_1\}$ can similarly be extended to a basis $\{x^1_1, y^1_j\}_{j=2}^{n_1}$ of $V^1$ with the corresponding dual basis $\{x^{1*}_1, y^{1*}_j\}_{j=2}^{n_1}$. It follows that the image of $\Pi_1(\upsilon)$ is the span of $\Pi_1(\upsilon)(x^{1*}_1) = x^2_1 \otimes x^3_1 + kx^2_2 \otimes x^3_2$. Hence, in this case, it is also true that every element in the image of $\Pi_1(\upsilon)$ can be written as $ax^2_1 \otimes x^3_1 + bx^2_2 \otimes x^3_2$ for some constants $a$ and $b$.

Suppose for contradiction that $P_{\operatorname{im}\Pi_1(\upsilon)}(\Pi_1(\rho)(\mathrm{e}^{1*}_1)) \neq \Pi_1(\upsilon)(\mathrm{e}^{1*}_1)$. Let $\alpha$ and $\beta$ be scalars such that

$$P_{\operatorname{im}\Pi_1(\upsilon)}(\Pi_1(\rho)(\mathrm{e}^{1*}_1)) \;=\; \alpha\; x^2_1 \otimes x^3_1 \;+\; \beta\; x^2_2 \otimes x^3_2.$$

Furthermore, let $a_i$ and $b_i$ be scalars such that

$$\Pi_1(\upsilon)(\mathrm{e}_i^{1*}) \;=\; a_i\; x_1^2 \otimes x_1^3 \;+\; b_i\; x_2^2 \otimes x_2^3$$

for $i = 2, \ldots, n_1$. Define $\upsilon' \in V^1 \otimes V^2 \otimes V^3$ as the unique tensor with the following mode-1 contraction:

$$\begin{aligned}
\Pi_1(\upsilon')(\mathrm{e}_1^{1*}) \;&=\; \alpha\; x_1^2 \otimes x_1^3 \;+\; \beta\; x_2^2 \otimes x_2^3 \;=\; P_{\mathrm{im}\,\Pi_1(\upsilon)}(\Pi_1(\rho)(\mathrm{e}_1^{1*})),\\
\Pi_1(\upsilon')(\mathrm{e}_i^{1*}) \;&=\; a_i\; x_1^2 \otimes x_1^3 \;+\; b_i\; x_2^2 \otimes x_2^3 \;=\; \Pi_1(\upsilon)(\mathrm{e}_i^{1*}) \;\; \text{for } i = 2, \ldots, n_1.
\end{aligned}$$

It follows that

$$\upsilon' \;=\; (\alpha \mathrm{e}_1^1 + \sum_{j=2}^{n_1} a_j \mathrm{e}_j^1) \otimes x_1^2 \otimes x_1^3 \;+\; (\beta \mathrm{e}_1^1 + \sum_{j=2}^{n_1} b_j \mathrm{e}_j^1) \otimes x_2^2 \otimes x_2^3,$$

so $\upsilon'$ is rank $\leq 2$. Furthermore,

$$\begin{aligned}
\|\rho - \upsilon'\|_{B,2}^2 \;&=\; \sum_{i=1}^{n_1} \|\Pi_1(\rho)(\mathrm{e}_i^{1*}) - \Pi_1(\upsilon')(\mathrm{e}_i^{1*})\|_{B,2}^2 && (5.8)\\
&=\; \|\Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon')(\mathrm{e}_1^{1*})\|_{B,2}^2 \;+\; \sum_{i=2}^{n_1} \|\Pi_1(\rho)(\mathrm{e}_i^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_i^{1*})\|_{B,2}^2 && (5.9)\\
&<\; \|\Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_1^{1*})\|_{B,2}^2 \;+\; \sum_{i=2}^{n_1} \|\Pi_1(\rho)(\mathrm{e}_i^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_i^{1*})\|_{B,2}^2 && (5.10)\\
&=\; \|\rho - \upsilon\|_{B,2}^2.
\end{aligned}$$

Equation (5.8) follows from Theorem 32, and equation (5.9) follows from the fact that $\Pi_1(\upsilon')(\mathrm{e}_i^{1*}) = \Pi_1(\upsilon)(\mathrm{e}_i^{1*})$ for $i = 2, 3, \ldots, n_1$. Equation (5.10) follows from our hypothesis that $\alpha x_1^2 \otimes x_1^3 + \beta x_2^2 \otimes x_2^3$ is a better approximation of $\Pi_1(\rho)(\mathrm{e}_1^{1*})$ with respect to $\|\cdot\|_{B,2}$ than $\Pi_1(\upsilon)(\mathrm{e}_1^{1*})$. Hence, $\|\rho - \upsilon'\|_{B,2} < \|\rho - \upsilon\|_{B,2}$, contradicting that $\upsilon$ is a best rank two approximation of $\rho$. $\square$

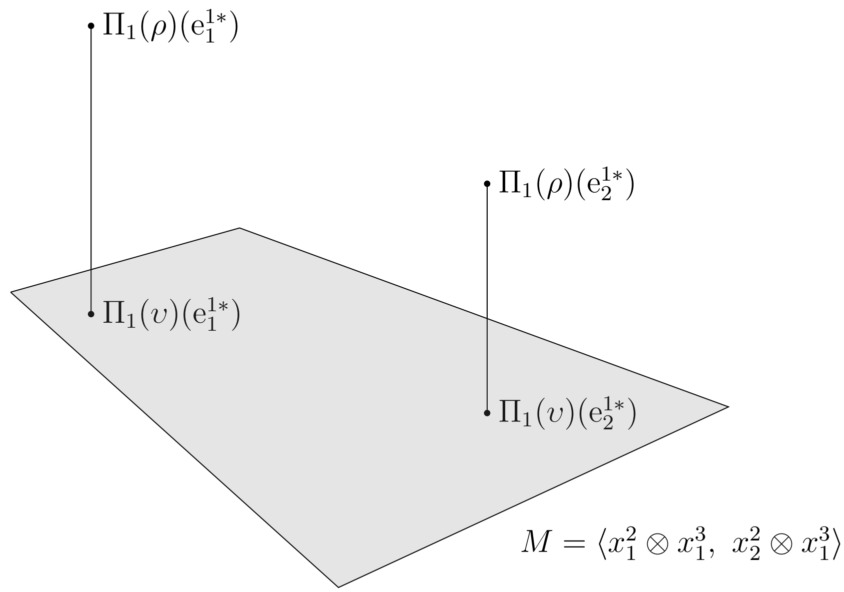



**Figure 5.1.** In this figure, we consider the $2 \times 2 \times 2$ case. Let $\upsilon = x_1^1 \otimes x_1^2 \otimes x_1^3 + x_2^1 \otimes x_2^2 \otimes x_1^3$ be an optimal rank two approximation of rank three tensor $\rho \in V^1 \otimes V^2 \otimes V^3$. In this case, $\operatorname{im} \Pi_1(\upsilon)$ is a plane in $V^2 \otimes V^3$ that contains all the rank one tensors in the span of $x_1^2 \otimes x_1^3$ and $x_2^2 \otimes x_1^3$. This plane is denoted as $M$ in the figure. Theorem 35 shows that the projection of $\Pi_1(\rho)(\mathrm{e}_i^{1*})$ onto the plane $M$ must be $\Pi_1(\upsilon)(\mathrm{e}_i^{1*})$ for $i = 1, 2$.

## 5.4 Optimal Rank Two Approximations Do Not Exist With Respect to the Frobenius Norm

In this section, we show that rank three $2 \times 2 \times 2$ tensors over $\mathbb{R}$ do not have optimal rank two approximations with respect to the Frobenius norm. Given any rank two tensor $\upsilon$ and rank three tensor $\rho$, we construct a rank two tensor $\upsilon'$ that is a better approximation of $\rho$ than $\upsilon$ with respect to the Frobenius norm. Let $V^i$ be two-dimensional real vector spaces for $i = 1, 2, 3$. As in the previous section, let $B$ denote a collection of bases $\{\mathrm{e}^i_j\}_{j=1}^2$ of $V^i$ for $i =$ 1, 2, 3. Furthermore, denote the corresponding dual bases as $\{\mathrm{e}^{i*}_j\}_{j=1}^2$ for $i =$ 1, 2, 3.

**Theorem 36.** *If $\rho \in V^1 \otimes V^2 \otimes V^3$ is rank three, then there does not exist an optimal rank two approximation of $\rho$ with respect to $\|\cdot\|_{B,2}$.*

*Proof.* Suppose for contradiction that there existed a tensor $\upsilon$ that was an optimal rank two approximation of $\rho$. By Theorem 31, we may assume

$$\upsilon \;=\; x^1_1 \otimes x^2_1 \otimes x^3_1 \;+\; x^1_2 \otimes x^2_2 \otimes x^3_1$$

for some vectors $x^i_j$. Since $\{\mathrm{e}^1_1, \mathrm{e}^1_2\}$ is a basis of $V^1$, there exists scalars $a$, $b$, $c$, and $d$ such that

$$x^1_1 = a\mathrm{e}^1_1 + c\mathrm{e}^1_2 \;\text{ and }\; x^1_2 = b\mathrm{e}^1_1 + d\mathrm{e}^1_2.$$

It follows that

$$\upsilon \;=\; \mathrm{e}^1_1 \otimes (ax^2_1 + bx^2_2) \otimes x^3_1 \;+\; \mathrm{e}^1_2 \otimes (cx^2_1 + dx^2_2) \otimes x^3_1.$$

If the set $\{x^1_1, x^1_2\}$ were linearly dependent, then $\upsilon$ would be rank one by multilinearity, which would contradict Theorem 34. Hence, $\{x^1_1, x^1_2\}$ is independent, and is thus a basis of $V^1$. Let $\{x^{1*}_1, x^{1*}_2\}$ denote its dual basis. The image of $\Pi_1(\upsilon)$ is the span of $\Pi_1(\upsilon)(x^{1*}_1) = x^2_1 \otimes x^3_1$ and $\Pi_1(\upsilon)(x^{1*}_2) = x^2_2 \otimes x^3_1$. These two tensors are linearly independent in $V^2 \otimes V^3$ since the set $\{x^2_1, x^2_2\}$ is independent, which also follows from the fact that $\upsilon$ is rank two. Hence, the image of $\Pi_1(\upsilon)$ is the plane spanned by the tensors $x_1^2 \otimes x_1^3$ and $x_2^2 \otimes x_1^3$. Let $M$ denote this plane and let $P_M$ denote the projection onto this plane. By Theorem 35,

$$P_M(\Pi_1(\rho)(\mathrm{e}_1^{1*})) = \Pi_1(\upsilon)(\mathrm{e}_1^{1*}) = (ax_1^2 + bx_2^2) \otimes x_1^3, \text{ and}$$
$$P_M(\Pi_1(\rho)(\mathrm{e}_2^{1*})) = \Pi_1(\upsilon)(\mathrm{e}_2^{1*}) = (cx_1^2 + dx_2^2) \otimes x_1^3.$$

Let $x_2^3 \in V^3$ be a vector such that the set $\{x_1^3, x_2^3\}$ is orthogonal with respect to $\langle \cdot \mid \cdot \rangle_{B,2}$. The set $\{x_1^2 \otimes x_2^3,\ x_2^2 \otimes x_2^3\}$ spans the orthogonal complement of the plane $M$ in $V^2 \otimes V^3$ since

$$\langle x_i^2 \otimes x_2^3 \mid x_j^2 \otimes x_1^3 \rangle_{B,2} = \langle x_i^2 \mid x_j^2 \rangle_{B,2}\ \langle x_2^3 \mid x_1^3 \rangle_{B,2} = 0 \text{ for } i,j \in \{1,2\},$$

by property (33). Thus, there must exist some constants $r$, $s$, $p$, and $q$ such that

$$\Pi_1(\rho)(\mathrm{e}_1^{1*}) = (ax_1^2 + bx_2^2) \otimes x_1^3 + (rx_1^2 + sx_2^2) \otimes x_2^3, \text{ and}$$
$$\Pi_1(\rho)(\mathrm{e}_2^{1*}) = (cx_1^2 + dx_2^2) \otimes x_1^3 + (px_1^2 + qx_2^2) \otimes x_2^3.$$

It follows that $\rho$ can be written in the form

$$\begin{aligned}\rho \ =\ & \mathrm{e}_1^1 \otimes (ax_1^2 + bx_2^2) \otimes x_1^3 \ +\ \mathrm{e}_1^1 \otimes (rx_1^2 + sx_2^2) \otimes x_2^3 \\ & +\ \mathrm{e}_2^1 \otimes (cx_1^2 + dx_2^2) \otimes x_1^3 \ +\ \mathrm{e}_2^1 \otimes (px_1^2 + qx_2^2) \otimes x_2^3.\end{aligned}$$

If we can show that

$$\langle\ (cx_1^2 + dx_2^2) \mid (rx_1^2 + sx_2^2)\ \rangle_{B,2} = 0, \tag{5.11}$$
$$\langle\ (ax_1^2 + bx_2^2) \mid (px_1^2 + qx_2^2)\ \rangle_{B,2} = 0, \text{ and} \tag{5.12}$$
$$\langle\ (ax_1^2 + bx_2^2) \mid (rx_1^2 + sx_2^2)\ \rangle_{B,2} = 0, \tag{5.13}$$

then this would imply that

$$(px_1^2 + qx_2^2) = k_1(rx_1^2 + sx_2^2) \text{ and } (cx_1^2 + dx_2^2) = k_2(ax_1^2 + bx_2^2)$$

for some constants $k_1$, $k_2$, since these vectors are in a two-dimensional space. It then follows by the multilinearity of the tensor product that

$$\begin{aligned}\rho \;&=\; \mathrm{e}_1^1 \otimes (ax_1^2 + bx_2^2) \otimes x_1^3 \;+\; \mathrm{e}_1^1 \otimes (rx_1^2 + sx_2^2) \otimes x_2^3 \\ &\qquad +\; \mathrm{e}_2^1 \otimes k_2(ax_1^2 + bx_2^2) \otimes x_1^3 \;+\; \mathrm{e}_2^1 \otimes k_1(rx_1^2 + sx_2^2) \otimes x_2^3 \\ &=\; (\mathrm{e}_1^1 + k_2\mathrm{e}_2^1) \otimes (ax_1^2 + bx_2^2) \otimes x_1^3 \;+\; (\mathrm{e}_1^1 + k_1\mathrm{e}_2^1) \otimes (rx_1^2 + sx_2^2) \otimes x_2^3,\end{aligned}$$

contradicting that $\rho$ is rank three.

We first prove equality (5.11) by considering the tensor

$$\upsilon_1(\epsilon) = \mathrm{e}_2^1 \otimes (cx_1^2 + dx_2^2) \otimes x_1^3 \;+\; \mathrm{e}_1^1 \otimes (ax_1^2 + bx_2^2) \otimes x_1^3 \;+\; \mathrm{e}_1^1 \otimes (cx_1^2 + dx_2^2) \otimes \epsilon x_2^3.$$

Suppose for contradiction that (5.11) were not true. It would then follow that

$$\langle\; (rx_1^2 + sx_2^2) \otimes x_2^3 \mid (cx_1^2 + dx_2^2) \otimes x_2^3 \;\rangle_{B,2} = \langle (rx_1^2 + sx_2^2) \mid (cx_1^2 + dx_2^2) \rangle_{B,2}\; \|x_2^3\|_{B,2}^2 \neq 0.$$

We can always choose a real $\epsilon$ small enough in absolute value such that

$$B_1(\epsilon) = -2\epsilon\; \langle\; (rx_1^2 + sx_2^2) \otimes x_2^3 \mid (cx_1^2 + dx_2^2) \otimes x_2^3 \;\rangle_{B,2} \;+\; \epsilon^2\; \|(cx_1^2 + dx_2^2) \otimes x_2^3\|_{B,2}^2$$

is negative. For example, if $\langle\; (rx_1^2 + sx_2^2) \otimes x_2^3 \mid (cx_1^2 + dx_2^2) \otimes x_2^3 \;\rangle_{B,2} < 0$, an $\epsilon < 0$ small enough in absolute value would result in a $B_1(\epsilon)$ negative. Suppose such an $\epsilon$ is chosen. Observe that

$$\begin{aligned}&\Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_1^{1*}) = (ax_1^2 + bx_2^2) \otimes x_1^3 \;+\; \epsilon(cx_1^2 + dx_2^2) \otimes x_2^3, \text{ and} \\ &\Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_2^{1*}) = (cx_1^2 + dx_2^2) \otimes x_1^3 = \Pi_1(\upsilon)(\mathrm{e}_2^{1*}).\end{aligned}$$

It thus follows that

$$\|\rho - \upsilon_1(\epsilon)\|^2_{B,2} = \|\Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_1^{1*})\|^2_{B,2} + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_2^{1*})\|^2_{B,2} \tag{5.14}$$

$$= \|\Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_1^{1*})\|^2_{B,2} + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2}. \tag{5.15}$$

Equation (5.14) follows from Theorem 32, and equation (5.15) follows from the fact that $\Pi_1(\upsilon)(\mathrm{e}_2^{1*}) = \Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_2^{1*})$. By the multilinearity of the tensor product and the multilinearity of the inner product, we conclude that

$$\begin{aligned}
\|\rho - \upsilon_1(\epsilon)\|^2_{B,2} &= \langle\ \Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_1^{1*})\ \mid\ \Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon_1(\epsilon))(\mathrm{e}_1^{1*})\ \rangle_{B,2} \\
&\qquad + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2} \\
&= \langle\ (rx_1^2 + sx_2^2)\otimes x_2^3 - \epsilon(cx_1^2 + dx_2^2)\otimes x_2^3\ \mid\ (rx_1^2 + sx_2^2)\otimes x_2^3 - \epsilon(cx_1^2 + dx_2^2)\otimes x_2^3\ \rangle_{B,2} \\
&\qquad + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2} \\
&= \|(rx_1^2 + sx_2^2)\otimes x_2^3\|^2_{B,2} - 2\epsilon\ \langle\ (rx_1^2 + sx_2^2)\otimes x_2^3\ \mid\ (cx_1^2 + dx_2^2)\otimes x_2^3\ \rangle_{B,2} \\
&\qquad + \epsilon^2\ \|(cx_1^2 + dx_2^2)\otimes x_2^3\|^2_{B,2} + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2} \\
&= \|(rx_1^2 + sx_2^2)\otimes x_2^3\|^2_{B,2} + B_1(\epsilon) + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2}.
\end{aligned}$$

Since $\epsilon$ was chosen specifically to make $B_1(\epsilon)$ negative, we can conclude that

$$\begin{aligned}
\|\rho - \upsilon_1(\epsilon)\|^2_{B,2} &= \|(rx_1^2 + sx_2^2)\otimes x_2^3\|^2_{B,2} + B_1(\epsilon) + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2} \\
&< \|(rx_1^2 + sx_2^2)\otimes x_2^3\|^2_{B,2} + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2} \\
&= \|\Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_1^{1*})\|^2_{B,2} + \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|^2_{B,2} \\
&= \|\rho - \upsilon\|^2_{B,2}.
\end{aligned}$$

Hence, $\upsilon_1(\epsilon)$ is a better approximation of $\rho$ than $\upsilon$. The tensor $\upsilon_1(\epsilon)$ is also in the tangent space of the Segre variety at $\mathrm{e}_1^1 \otimes (cx_1^2 + dx_2^2) \otimes x_1^3$. Hence, there exists a sequence of rank two tensors that converges to $\upsilon_1(\epsilon)$. Thus, there must be some rank two tensor in the sequence that is better approximation to $\rho$ than $\upsilon$, contradicting that $\upsilon$ is an optimal rank two approximation.

Equality (5.12) can be proven in the same way by considering the tensor

$$\upsilon_2(\epsilon) \;=\; \mathrm{e}^1_1 \otimes (ax^2_1 + bx^2_2) \otimes x^3_1 \;+\; \mathrm{e}^1_2 \otimes (cx^2_1 + dx^2_2) \otimes x^3_1 \;+\; \mathrm{e}^1_2 \otimes (ax^2_1 + bx^2_2) \otimes \epsilon x^3_2,$$

which is in the tangent space of the Segre variety at $\mathrm{e}^1_2 \otimes (ax^2_1 + bx^2_2) \otimes x^3_1$.

It thus remains to show equality (5.13), which we prove by considering the rank two tensor

$$\upsilon_3(\epsilon) \;=\; \mathrm{e}^1_1 \otimes (ax^2_1 + bx^2_2) \otimes (x^3_1 + \epsilon(x^3_1 + x^3_2)) \;+\; \mathrm{e}^1_2 \otimes (cx^2_1 + dx^2_2) \otimes x^3_1.$$

It follows that the mode-1 contractions of $\upsilon_3(\epsilon)$ are as follows.

$$\begin{aligned}
\Pi_1(\upsilon_3(\epsilon))(\mathrm{e}^{1*}_1) \;&=\; (ax^2_1 + bx^2_2) \otimes x^3_1 \;+\; \epsilon(ax^2_1 + bx^2_2) \otimes (x^3_1 + x^3_2), \text{ and}\\
\Pi_1(\upsilon_3(\epsilon))(\mathrm{e}^{1*}_2) \;&=\; (cx^2_1 + dx^2_2) \otimes x^3_1 \;=\; \Pi_1(\upsilon)(\mathrm{e}^{1*}_2).
\end{aligned}$$

Suppose for contradiction that (5.13) were nonzero. It would then follow that

$$\langle\, (rx^2_1 + sx^2_2) \otimes x^3_2 \mid (ax^2_1 + bx^2_2) \otimes (x^3_1 + x^3_2) \,\rangle_{B,2} = \langle\, (rx^2_1 + sx^2_2) \mid (ax^2_1 + bx^2_2) \,\rangle_{B,2}\; \|x^3_2\|^2_{B,2} \neq 0.$$

We could then choose an $\epsilon$ small enough in absolute value such that $D(\epsilon) < 0$, where

$$D(\epsilon) = \; -2\epsilon\, \langle (rx^2_1 + sx^2_2) \otimes x^3_2 \mid (ax^2_1 + bx^2_2) \otimes (x^3_1 + x^3_2)\rangle_{B,2} \;+\; \epsilon^2\, \|(ax^2_1 + bx^2_2) \otimes (x^3_1 + x^3_2)\|^2_{B,2}.$$

For example, if $\langle\, (rx^2_1 + sx^2_2) \otimes x^3_2 \mid (ax^2_1 + bx^2_2) \otimes (x^3_1 + x^3_2) \,\rangle_{B,2} < 0$, an $\epsilon < 0$ small enough in absolute value will yield a negative $D(\epsilon)$. Suppose such an $\epsilon$ is chosen. Then,

$$\begin{aligned}
\|\rho - \upsilon_3(\epsilon)\|^2_{B,2} \;&=\; \|\Pi_1(\rho)(\mathrm{e}^{1*}_1) - \Pi_1(\upsilon_3(\epsilon))(\mathrm{e}^{1*}_1)\|^2_{B,2} \;+\; \|\Pi_1(\rho)(\mathrm{e}^{1*}_2) - \Pi_1(\upsilon_3(\epsilon))(\mathrm{e}^{1*}_2)\|^2_{B,2}\\
&=\; \|\Pi_1(\rho)(\mathrm{e}^{1*}_1) - \Pi_1(\upsilon_3(\epsilon))(\mathrm{e}^{1*}_1)\|^2_{B,2} \;+\; \|\Pi_1(\rho)(\mathrm{e}^{1*}_2) - \Pi_1(\upsilon)(\mathrm{e}^{1*}_2)\|^2_{B,2}.
\end{aligned}$$

Observe that

$$\|\Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon_3(\epsilon))(\mathrm{e}_1^{1*})\|_{B,2}^2 \;=\; \|(rx_1^2 + sx_2^2)\otimes x_2^3 - \epsilon(ax_1^2 + bx_2^2)\otimes(x_1^3 + x_2^3)\|_{B,2}^2.$$

Writing this as an inner product, we see that

$$\begin{aligned}
\|\rho - \upsilon_3(\epsilon)\|_{B,2}^2 \;&=\; \|(rx_1^2 + sx_2^2)\otimes x_2^3\|_{B,2}^2 \;-\; 2\epsilon\,\langle (rx_1^2 + sx_2^2)\otimes x_2^3 \mid (ax_1^2 + bx_2^2)\otimes(x_1^3 + x_2^3)\rangle_{B,2} \\
&\qquad +\; \epsilon^2\,\|(ax_1^2 + bx_2^2)\otimes(x_1^3 + x_2^3)\|_{B,2}^2 \;+\; \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|_{B,2}^2 \\
&=\; \|(rx_1^2 + sx_2^2)\otimes x_2^3\|_{B,2}^2 \;+\; D(\epsilon) \;+\; \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|_{B,2}^2.
\end{aligned}$$

Finally, since $\epsilon$ was chosen so that $D(\epsilon)$ would be negative, it follows that

$$\begin{aligned}
\|\rho - \upsilon_3(\epsilon)\|_{B,2}^2 \;&<\; \|(rx_1^2 + sx_2^2)\otimes x_2^3\|_{B,2}^2 \;+\; \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|_{B,2}^2 \\
&=\; \|\Pi_1(\rho)(\mathrm{e}_1^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_1^{1*})\|_{B,2}^2 \;+\; \|\Pi_1(\rho)(\mathrm{e}_2^{1*}) - \Pi_1(\upsilon)(\mathrm{e}_2^{1*})\|_{B,2}^2 \\
&=\; \|\rho - \upsilon\|_{B,2}^2.
\end{aligned}$$

Hence, the rank two tensor $\upsilon_3(\epsilon)$ is a better approximation of $\rho$ than $\upsilon$, contradicting that $\upsilon$ is an optimal rank two approximation. $\square$

We have thus shown that rank three $2\times 2\times 2$ real tensors have no optimal rank two approximations with respect to the Frobenius norm. This implies that the nearest point of a rank three $2\times 2\times 2$ tensor $\rho$ to the second secant set $\sigma_2(X)$ of the Segre variety $X$ with respect to the Frobenius norm must be rank three. In fact, our proof above demonstrates that the nearest point of $\rho$ to $\sigma_2(X)$ is in fact on the tangential set $\tau(X)$ of the Segre variety, which is the set of all tensors contained in the tangent space of the Segre variety at some point $u \in X$.

$$\tau(X) \;=\; \{\rho \in V^1\otimes V^2\otimes V^3 \mid \rho \in T_u(X) \text{ for some } u \in X\}$$

From Theorem (28), we know that

$$\tau(X) \;=\; \{\rho \in V^1 \otimes V^2 \otimes V^3 \;\mid\; \rho \;=\; x_2^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_2^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_1^2 \otimes x_2^3 \quad \text{for some } x_j^i \in V^i\}.$$

There is an open, dense subset of $\tau(X)$ of rank three tensors. However, there are rank two elements in $\tau(X)$, and the fact that these rank two tensors can never be the the nearest point on $\tau(X)$ of any rank three tensor implies there is an interesting curvature of $\tau(X)$ at these points. We now consider an example of such a rank two tensor in $\tau(X)$. Let $\{x_1^i, x_2^i\}$ be independent vectors in $V^i$ for $i = 1, 2, 3$. The tensor

$$\nu \;=\; x_1^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_2^2 \otimes x_1^3$$

is a rank two tensor in $\tau(X)$. For every $\epsilon \neq 0$, the tensor

$$\rho(\epsilon) \;=\; x_1^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_2^1 \otimes x_2^2 \otimes x_1^3 \;+\; \epsilon(x_1^1 + x_2^1) \otimes (x_1^2 + x_2^2) \otimes x_2^3$$

is rank three. Clearly, $\lim_{\epsilon \to 0} \rho(\epsilon) = \nu$. However, the nearest point to $\rho(\epsilon)$ on $\tau(X)$ with respect to the Frobenius norm is never $\nu$, even when $\epsilon$ is infinitesimally small, since the nearest point to $\rho(\epsilon)$ on $\tau(X)$ must be rank three. The tangential variety of the $2 \times 2 \times 2$ Segre variety must thus have significant curvature at its rank two points, which is already suggested by the fact that $\nu$ is tangent to the Segre variety at any tensors in the form $x_2^1 \otimes x_1^2 \otimes y_1^3$ and $x_1^1 \otimes x_2^2 \otimes y_1^3$ for some $y_1^3 \in V^3$. In contrast, the rank three tensors in $\tau(X)$ are tangent to distinct point on the Segre variety up to a multiplicative constant.

**Theorem 37.** *Let $V^1$, $V^2$, and $V^3$ be two-dimensional real vector spaces. Let $X$ denote the Segre variety of simple tensors in $V^1 \otimes V^2 \otimes V^3$, and let $\tau(X)$ denote the tangential variety of $X$. If $\rho \in \tau(X)$ is rank three, then it is tangent to a unique point of $X$ up to a multiplicative constant.*

*Proof.* Suppose $\rho$ is tangent to both the tensors $x_1^1 \otimes x_1^2 \otimes x_1^3$ and $x_2^1 \otimes x_2^2 \otimes x_2^3$. It then follows from Theorem 28 that

$$\rho \;=\; y_1^1 \otimes x_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes y_1^2 \otimes x_1^3 \;+\; x_1^1 \otimes x_1^2 \otimes y_1^3, \text{ and} \tag{5.16}$$

$$\rho \;=\; y_2^1 \otimes x_2^2 \otimes x_2^3 \;+\; x_2^1 \otimes y_2^2 \otimes x_2^3 \;+\; x_2^1 \otimes x_2^2 \otimes y_2^3, \tag{5.17}$$

for some vectors $y_j^i$. Since $\rho$ is rank three, the sets $\{y_j^i, x_j^i\}$ must be independent for all $i, j$. Let $\{y_j^{i*}, x_j^{i*}\}$ be the corresponding dual bases. Considering the contraction maps of $\rho$ with respect to both of these representation (5.16) and (5.17), we conclude that

$$\Pi_1(\rho)(y_1^{1*}) \;=\; x_1^2 \otimes x_1^3 \;=\; x_2^2 \otimes \left(y_1^{1*}(y_2^1)x_2^3 + y_1^{1*}(x_2^1)y_2^3\right) \;+\; y_1^{1*}(x_2^1)y_2^2 \otimes x_2^3.$$

By Lemma 20, it follows that $y_1^{1*}(x_2^1) = 0$, so $x_1^2 \otimes x_1^3 \;=\; y_1^{1*}(y_2^1)\; x_2^2 \otimes x_2^3$. The fact that $y_1^{1*}(x_2^1) = 0$ implies that $x_2^1 = k_1 x_1^1$ for some constant $k_1$. Furthermore, $\operatorname{im}\Pi_2(x_1^2 \otimes x_1^3) = \operatorname{im}\Pi_2(y_1^{1*}(y_2^1)x_2^2 \otimes x_2^3)$ implies that $x_1^2 = k_2 x_2^2$ for some constant $k_2$, and $\operatorname{im}\Pi_1(x_1^2 \otimes x_1^3) = \operatorname{im}\Pi_1(y_1^{1*}(y_2^1)x_2^2 \otimes x_2^3)$ implies that $x_2^3 = k_3 x_1^3$ for some constant $k_3$. Hence, $x_2^1 \otimes x_2^2 \otimes x_2^3$ $= k_1 k_2 k_3\; x_1^1 \otimes x_1^2 \otimes x_1^3$. □

Tangential sets with the property that a general point lies on a unique tangent line is said to be strongly nondegenerate [38, p.2]. We have thus shown the $\tau(X)$ is strongly nondegenerate.

# 6. FUTURE WORK

We have shown that the nearest point of a rank three $2 \times 2 \times 2$ real tensor $\rho$ to the second secant set $\sigma_2(X)$ of the Segre variety $X$ with respect to the Frobenius norm is in fact on the tangential variety $\tau(X)$. We conjecture that the nearest point of a rank greater than three $3 \times 3 \times 3$ real tensor to the third secant set $\sigma_3(X)$ of the Segre variety $X$ is on the second tangential set $\tau_2(X)$. The third secant set $\sigma_3(X)$ is the norm closure of all rank three tensors, and the second tangential set $\tau_2(X)$ is the norm closure of all the tensors $\rho$ such that $\rho = \kappa'(0) + \kappa''(0)$ for some smooth curve $\kappa$ contained in the Segre variety. Notice that such a tensor $\rho$ is indeed the limit of rank at most three tensors, since

$$\begin{aligned}
\rho &= \left.\frac{d\kappa(t)}{dt}\right|_{t=0} + \left.\frac{d^2\kappa(t)}{dt^2}\right|_{t=0} \\
&= \lim_{t\to 0}\frac{\kappa(t)-\kappa(0)}{t-0} + \lim_{t\to 0}\frac{\kappa(t)-2\kappa(0)+\kappa(-t)}{(t-0)^2} \\
&= \lim_{n\to\infty}\frac{\kappa(\frac{1}{n})-\kappa(0)}{\frac{1}{n}} + \lim_{n\to\infty}\frac{\kappa(\frac{1}{n})-2\kappa(0)+\kappa(-\frac{1}{n})}{\frac{1}{n^2}} \\
&= \lim_{n\to\infty}\ (n+n^2)\kappa(\frac{1}{n}) - (2n^2-n)\kappa(0) + n^2\kappa(-\frac{1}{n}).
\end{aligned} \tag{6.1}$$

Equation (6.1) follows from the finite difference formula for the second derivative. Our proof method is well-suited to be generalized to this $3 \times 3 \times 3$ case since it does not rely on the specific polynomials that define the tangential and secant sets of the $2 \times 2 \times 2$ of the Segre variety. If true, this conjecture would suggest interesting curvature of $\tau_2(X)$ that would need to be studied further.

# REFERENCES


[1] A. Terracini, “Sulle $V_k$ per cui la varietà degli $S_h(h+1)$ seganti ha dimensione minore dell’ordinario,” *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, vol. 31, no. 1, pp. 392–396, 1911.

[2] F. L. Hitchcock, “The expression of a tensor or a polyadic as a sum of products,” *Journal of Mathematics and Physics*, vol. 6, no. 1-4, pp. 164–189, 1927.

[3] F. L. Hitchcock, “Multiple invariants and generalized rank of a p-way matrix or tensor,” *Journal of Mathematics and Physics*, vol. 7, no. 1-4, pp. 39–79, 1928.

[4] L. R. Tucker, “Implications of factor analysis of three-way matrices for measurement of change,” *Problems in measuring change*, vol. 15, no. 122-137, p. 3, 1963.

[5] L. R. Tucker *et al.*, “The extension of factor analysis to three-dimensional matrices,” *Contributions to mathematical psychology*, vol. 110119, 1964.

[6] L. R. Tucker, “Some mathematical notes on three-mode factor analysis,” *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[7] S. Rabanser, O. Shchur, and S. Günnemann, “Introduction to tensor decompositions and their applications in machine learning,” *arXiv preprint arXiv:1711.10781*, 2017.

[8] X. Li, S. Lin, S. Yan, and D. Xu, “Discriminant locally linear embedding with high-order tensor data,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 2, pp. 342–352, 2008.

[9] A. Anandkumar, R. Ge, D. Hsu, S. M. Kakade, and M. Telgarsky, “Tensor decompositions for learning latent variable models,” *Journal of machine learning research*, vol. 15, pp. 2773–2832, 2014.

[10] E. S. Allman, P. D. Jarvis, J. A. Rhodes, and J. G. Sumner, “Tensor rank, invariants, inequalities, and applications,” *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 1014–1045, 2013.

[11] J. A. Rhodes and S. Sullivant, “Identifiability of large phylogenetic mixture models,” *Bulletin of mathematical biology*, vol. 74, no. 1, pp. 212–231, 2012.

[12] J. G. Sumner, M. A. Charleston, L. S. Jermiin, and P. D. Jarvis, “Markov invariants, plethysms, and phylogenetics,” *Journal of theoretical biology*, vol. 253, no. 3, pp. 601–615, 2008.

[13] C. J. Hillar and L.-H. Lim, "Most tensor problems are $NP$-hard," *Journal of the ACM (JACM)*, vol. 60, no. 6, pp. 1–39, 2013.

[14] J. M. Landsberg, *Tensors: geometry and applications*. American Mathematical Society, 2012, vol. 128.

[15] J. M. Landsberg and L. Manivel, "On the ideals of secant varieties of segre varieties," *Foundations of Computational Mathematics*, vol. 4, no. 4, pp. 397–422, 2004.

[16] C. Raicu, "Secant varieties of segre–veronese varieties," *Algebra & Number Theory*, vol. 6, no. 8, pp. 1817–1868, 2012.

[17] J. B. Kruskal, "Rank, decomposition, and uniqueness for 3-way and n-way arrays," in *Multiway data analysis*, 1989, pp. 7–18.

[18] J. M. Ten Berge, "Kruskal's polynomial for 2× 2× 2 arrays and a generalization to 2× n× n arrays," *Psychometrika*, vol. 56, no. 4, pp. 631–636, 1991.

[19] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, *Algebraic complexity theory*. Springer Science & Business Media, 2013, vol. 315.

[20] A. Stegeman and P. Comon, "Subtracting a best rank-1 approximation may increase tensor rank," *Linear Algebra and its Applications*, vol. 433, no. 7, pp. 1276–1300, 2010.

[21] A. Stegeman, "Nonexistence of best low-rank approximations for real-valued three-way arrays and what to do about it," in *21st International Symposium on Mathematical Theory of Networks and Systems (MTNS2014)*, 2014, pp. 954–957.

[22] J. Buczyński and J. M. Landsberg, "On the third secant variety," *Journal of Algebraic Combinatorics*, vol. 40, no. 2, pp. 475–502, 2014.

[23] J. E. Cohen and U. G. Rothblum, "Nonnegative ranks, decompositions, and factorizations of nonnegative matrices," *Linear Algebra and its Applications*, vol. 190, pp. 149–168, 1993.

[24] W. Greub, *Multilinear Algebra*. Springer-Verlag, New York-Heidelberg, 1978.

[25] W. P. Wardlaw, "Row rank equals column rank," *Mathematics Magazine*, vol. 78, no. 5, pp. 316–318, 2005.

[26] K. Hoffman and R. Kunze, *Linear Algebra*, second. Prentice-Hall, 1971.

[27] P. D. Lax, *Linear Algebra and its Applications*, second. John Wiley & Sons, 2007.

[28] R. Hartshorne, *Algebraic geometry*. Springer Science & Business Media, 2013, vol. 52.

[29] J. An, "Rigid geometric structures, isometric actions, and algebraic quotients," *Geometriae Dedicata*, vol. 157, no. 1, pp. 153–185, 2012.

[30] V. De Silva and L.-H. Lim, "Tensor rank and the ill-posedness of the best low-rank approximation problem," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1084–1127, 2008.

[31] G. Ottaviani, "Introduction to the hyperdeterminant and to the rank of multidimensional matrices," in *Commutative algebra*, Springer, 2013, pp. 609–638.

[32] L.-H. Lim and P. Comon, "Nonnegative approximations of nonnegative tensors," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 23, no. 7-8, pp. 432–441, 2009.

[33] A. Akhtyamov, M. Amram, M. Dagan, and A. Mouftahkov, "Cramer's rule for nonsingular m$\times$ n matrices.," *Teaching of Mathematics*, vol. 20, no. 1, 2017.

[34] S. Basu, R. Pollack, and M.-F. Roy, *Algorithms in Real Algebraic Geometry*. Springer, 2006.

[35] J. Bochnak, M. Coste, and M.-F. Roy, *Real algebraic geometry*. Springer Science & Business Media, 2013, vol. 36.

[36] T. Roughgarden and G. Valiant, *Lecture 10: Tensors, and low-rank tensor recovery*. [Online]. Available: timroughgarden.org/notes.html.

[37] F. L. Zak, *Tangents and secants of algebraic varieties*. American Mathematical Soc., 1993, vol. 127.

[38] J. Landsberg and J. Weyman, "On tangential varieties of rational homogeneous varieties," *Journal of the London Mathematical Society*, vol. 76, no. 2, pp. 513–530, 2007.