The Tensor Rank Problem over the Quaternions

by

Yungang Liang

(YG Liang)

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Mathematics

University of Manitoba

Winnipeg

Abstract

Tensors are defined as multidimensional arrays and therefore generalize matrices to multiple dimensions. To put it simple, a 1-tensor is a vector. A 2-tensor is a matrix and a 3-tensor is a cube. Similarly, we can imagine a tensor with a higher order.

A K-way tensor T is said to be *rank-1* if it can be written as the out product of K vectors, i.e.,

$$T = \vec{a}_1 \otimes \vec{a}_2 \otimes \cdots \otimes \vec{a}_K,$$

where $\otimes$ is the vector outer product. It means that each entry of the tensor is the product of the corresponding vector entries:

$$T_{i_1 i_2 \cdots i_K} = a_{1,i_1} a_{2,i_2} \cdots a_{K,i_K} \quad \text{for all } 1 \leqslant i_j \leqslant N_k, \quad k = 1, 2, ..., K.$$

The rank of a tensor T is defined by the smallest positive integer n such that

$$T = T_1 + T_2 + \ldots + T_n$$

for some rank-1 tensors $T_1, T_2, \ldots, T_n$. That being said, the objective of rank determination is to answer the question, "How may rank-1 tensors are required to express the given tensor?"

In this thesis, we give the maximal rank (the best possible upper bound) of quaternionic tensors in the $n_1 \times n_2 \times n_3$ cases where $2 \leqslant n_i \leqslant 3$. Decomposition of a quaternionic tensor T into simple tensors in the some of these cases are discussed. We also give an example of a complex tensor that has different ranks over the complex field and the real quaternion algebra.

Moreover, we give the maximal rank and canonical forms of $m \times 2 \times n$ quaternion tensors for any $m, n \in \mathbb{N}^+$. We also discuss some upper bounds for general quaternion tensors by using a block tensor approach.

Acknowledgements

I would like to express my deepest appreciation to my supervisor, Dr. Yang Zhang, who has gone above and beyond what is expected of a supervisor, for his unconditional support, his generosity, his advice, and his encouragement throughout this project.

I would like to extend my sincere gratitude to Dr. Tommy Kucera and Dr. Derek Krepski, who have been one of the most influential professors in my journey of mathematics in the University of Manitoba, for enlightening me in my study, research, teaching, as well as other aspects. Special thanks to Dr. Sergio Da Silva for the previous collaborative work in tensor rank.

Lastly, I would like to acknowledge the staff and faculty in the Department of Mathematics at the University of Manitoba, as well as the financial support provided by the Department of Mathematics and the Faculty of Graduate Studies at the University of Manitoba.

Contributions of Authors

Chapter 2 and Sections 3.1 - 3.3 in Chapter 3 are a version of a journal article co-authored with Sergio Da Silva and Yang Zhang which has been published in [45]. I am the primary author.

Section 3.4 and 3.5 in Chapter 3 are a version of journal article co-authored with Yang Zhang to be submitted for publication . I am the primary author of the article.

Chapters 4 is a version of journal article co-authored with Yang Zhang to be submitted for publication. I am the primary author of the article.

Contents

List of Tables

List of Symbols

$\|\|q\|\|$	the norm of the real quaternion q
$\Re(q)$	the real part of the real quaternion q
$\mathbb{R}$	the real field
$\mathbb{C}$	the complex field
$\mathbb{H}$	the real quaternion algebra
$\text{rank}_{\mathbb{R}}(T)$	the rank the tensor T over $\mathbb{R}$
$\text{rank}_{\mathbb{C}}(T)$	the rank the tensor T over $\mathbb{C}$
$\text{rank}_{\mathbb{H}}(T)$	the rank the tensor T over $\mathbb{H}$
$\dim_{\mathbb{H}}({}_{\mathbb{H}}M)$	the dimension of the left $\mathbb{H}$-module M
$\dim_{\mathbb{H}}(M_{\mathbb{H}})$	the dimension of the right $\mathbb{H}$-module M
$\text{max.rank}_{\mathbb{R}}(n_1, n_2, \ldots, n_p)$	the maximal rank of $n_1 \times n_2 \times \ldots\ n_p$ tensors over $\mathbb{R}$
$\text{max.rank}_{\mathbb{C}}(n_1, n_2, \ldots, n_p)$	the maximal rank of $n_1 \times n_2 \times \ldots\ n_p$ tensors over $\mathbb{C}$
$\text{max.rank}_{\mathbb{H}}(n_1, n_2, \ldots, n_p)$	the maximal rank of $n_1 \times n_2 \times \ldots\ n_p$ tensors over $\mathbb{H}$
A^T	the transpose of the matrix A
$\det(A)$	the determinant of the matrix A
$\overline{A}$	the conjugate matrix of the matrix A
$\chi(A)$	the complex adjoint matrix of A
$O_{m\times n}$	the $m \times n$ zero matrix
I_k	the $k \times k$ identity matrix
$J_k(a)$	the $k \times k$ Jordan block with eigenvalue a

$T^{\leqslant m}$	the sub-tensor consisting the first m rows of the tensor T
${}^{m+1<}T$	the sub-tensor with the first m rows of the tensor T deleted
T^*	the conjugate transpose of the tensor T
$R_i(T)$	the i-th row of the tensor T
$R(T)$	the row space of the tensor T
$C_i(T)$	the i-th column of the tensor T
$C(T)$	the column space of the tensor T
$\mathrm{Proj}(m, i, T)$	the i-th m-slice of the tensor T
$\mathbb{H}^{N_1 \times \cdots \times N_p}$	the set of all $N_1 \times \cdots \times N_p$ tensors over $\mathbb{H}$
$M_{m\times n}(\mathbb{H}^{N_1 \times \cdots \times N_p})$	the set of all $m \times n$ matrices with entries in $\mathbb{H}^{N_1 \times \cdots \times N_p}$
$[A_{ij}]_{\otimes}$	a matrix filled with tensor entries A_{ij}
$A \star B$	the product of the matrices (with tensor entries) A and B
$T \times_m A_m$	m-mode product of the matrix A_m and the tensor T

Chapter 1

Introduction

Tensors, as generalizations of matrices to higher dimensions, have many applications in various settings. They have many applications in various aspects, such as aerospace engineering ([22]), signal processing ([11, 18, 17]), data mining ([62]), machine learning ([47]), computer vision ([54, 65, 66]), higher-order statistics ([10, 13]), pattern recognition ([33, 52]), chemometrics ([12, 55]), graph analysis ([37]), numerical linear algebra ([19, 20, 35]), numerical analysis (Part I in [43]), etc. They are one of the most important algebraic structures in mathematics. There has been active research on tensors for the decades with topics including tensor decompositions, tensor ranks, tensor determinant, tensor eigenvalue problems, and low-rank approximation.

The concepts of tensor decompositions first appeared in Hitchcock's book in 1927 ([28, 29]), and later was developed to a multiway model by Cattell in 1944 [9, 8]. These concepts got very little attention until Tucker's work ([63]) in the 1960s and Harshman's work ([7]) in 1970, which appeared in psychometrics literature. In 1981, tensor decompositions were first used in chemometrics by Appellof and Davidson ([1]). Since then, tensors have become fairly popular in that field (see [27, 56, 5]). There was even a book written ([55]) in 2004. Along the side of the developments in chemometrics and psychometrics, decompositions of bilinear forms also aroused lots of interest in the field of algebraic complexity, such as [34, Section 4.6.4] by Knuth and an interesting example by Strassen about matrix multiplication, which describe 2×2 matrix multiplication by using a decomposition of a 4×4×4 tensor (see [57, 39, 42]).

1.1 Tensor Ranks and Tensor Decompositions

There have been extensive studies about tensor ranks and tensor decompositions over the past few decades. They both start with the consideration of sum of *rank-1* tensors which are closely related to outer product of vectors. To illustrate, given two vectors $\vec{u} = (u_1, u_2, \ldots, u_m)^T$ and $\vec{v} = (v_1, v_2, \ldots, v_n)^T$, their outer product, denoted by $\vec{u} \otimes \vec{v}$, is defined as the $m \times n$ matrix obtained by multiplying each element in $\vec{u}$ by each element in $\vec{v}$ in the following way:

$$\vec{u} \otimes \vec{v} := \begin{bmatrix} u_1v_1 & u_1v_2 & \ldots & u_1v_n \\ u_2v_1 & u_2v_2 & \ldots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_mv_1 & u_mv_2 & \ldots & u_mv_n \end{bmatrix}.$$

Similarly, A K-way tensor $T \in \mathbb{F}^{N_1 \times N_2 \times \cdots \times N_K}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) is said to be *rank-1* if it can be written as the out product of K vectors:

$$T = \vec{a}_1 \otimes \vec{a}_2 \otimes \cdots \otimes \vec{a}_K,$$

where $\vec{a}_k \in \mathbb{F}^{N_k}$, $k = 1, 2, ..., K$. That is, each entry of the tensor is the product of the corresponding vector entries:

$$T_{i_1 i_2 \cdots i_K} = a_{1,i_1} a_{2,i_2} \cdots a_{K,i_K} \quad \text{for all } 1 \leqslant i_j \leqslant N_k, \quad k = 1, 2, ..., K.$$

The rank of a tensor T is defined by the smallest positive integer n such that

$$T = T_1 + T_2 + \ldots + T_n$$

for some rank-1 tensors $T_1, T_2, \ldots, T_n$. That being said, the objective of determining tensor rank is simply to answer the question, "To express the given tensor, how many rank-1 tensors are required?"

Although the definition of tensor rank is similar to the definition of matrix rank, their properties are pretty different (see [40]).

1. There are straight-forward algorithms to compute the rank of a matrix, but there isn't any algorithm for computing the rank of a given tensor. As a matter of fact, the tensor rank problem is NP-hard (see [24]).

2. Matrix rank doesn't depend on the base field, but tensor rank does. For a given real tensor, its rank could be different over the real field $\mathbb{R}$ and the complex field $\mathbb{C}$. For example, the tensor

$$T = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right)$$

has rank 3 over $\mathbb{R}$ but only rank 2 over $\mathbb{C}$. Its rank decomposition over $\mathbb{R}$ is

$$T = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} ; \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \right) + \left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) + \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} ; \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \right),$$

while the rank decomposition over $\mathbb{C}$ is

$$T = \frac{1}{2} \left(\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} ; \begin{bmatrix} -i & 1 \\ -1 & i \end{bmatrix} \right) + \frac{1}{2} \left(\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} ; \begin{bmatrix} i & 1 \\ -1 & -i \end{bmatrix} \right).$$

3. The maximal rank and typical rank is the same for a matrix, but can be different for a tensor. The maximal rank is defined to be the largest rank that is attainable, however, the typical rank is any possible rank that can occur with probability greater than zero (see [36, Section 3.1]). For example, Kruskal ([40]) discusses the $2 \times 2 \times 2$ case, where the tensors have typical ranks of 2 and 3 over the real field $\mathbb{R}$. As a matter of fact, Monte Carlo experiments, which randomly draw each of the entries of the tensor from a normal distribution with standard deviation one and mean zero, show that the set of all $2 \times 2 \times 2$ tensors with rank 2 fills about 79% of the space, while those with rank 3 fill 21%.

In general, it is usually extremely difficult to determine the rank of a given tensor. There is an example of a particular $9 \times 9 \times 9$ tensor cited by Kruskal, whose rank is only known to be bounded between 18 and 23 (see [40]). In 2008, Comon et al. conjectured that the rank is 19 or 20 (see [14]).

1.2 Maximal Ranks of Complex and Real Tensors

Although it is generally difficult to determine the rank of a given tensor, there are extensive studies of upper bounds of ranks (maximal ranks) of tensors over the real field $\mathbb{R}$ and the complex field $\mathbb{C}$, which have been popular over the past few decades. Let $\mathbb{F}$ denote the real number field $\mathbb{R}$ or the complex field $\mathbb{C}$ and $\text{max.rank}_{\mathbb{F}}(n_1, n_2, \ldots, n_p)$ denote the maximal rank of $n_1 \times n_2 \times \ldots\ n_p$ tensors over $\mathbb{F}$. For example, in the matrix case, it is well-known that

$$\text{max.rank}_{\mathbb{F}}(n_1, n_2) = \max_{A \in M_{n_1 \times n_2}(\mathbb{F})} \text{rank}(A) = \min(n_1, n_2),$$

where $\mathbb{F}$ denotes the real number field $\mathbb{R}$ or the complex field $\mathbb{C}$.

In the tensor case, however, even for 3-tensors, $\text{max.rank}_{\mathbb{F}}(n_1, n_2, n_3)$ is unknown and difficult to determine. The following weak inequalities are known:

$$\max\{n_1, n_2, n_3\} \leqslant \text{max.rank}_{\mathbb{F}}(n_1, n_2, n_3) \leqslant \min\{n_1n_2, n_2n_3, n_3n_1\},$$

where $\mathbb{F}$ denotes the real number field $\mathbb{R}$ or the complex field $\mathbb{C}$.

As for better bounds for 3-tensors, Grigoryev [23] and Ja' Ja' [32] independently determine the rank of any $n_1 \times n_2 \times n_3$ tensor for $n_3 \leqslant 2$ in 1978 by using canonical forms. They showed that

$$\text{max.rank}_{\mathbb{F}}(m, n, 2) = \min\left\{n + \left\lfloor \frac{m}{2} \right\rfloor, m + \left\lfloor \frac{n}{2} \right\rfloor, 2m, 2n\right\},$$

where $\mathbb{F}$ denotes the real number field $\mathbb{R}$ or the complex field $\mathbb{C}$. In 1979, Atkinson and Stephens showed that

$$\text{max.rank}_{\mathbb{C}}(m, n, p) \leqslant m + \left\lfloor \frac{p}{2} \right\rfloor n$$

if $m \leqslant n$ (see [3]). In particular, when $m = n$, this becomes

$$\text{max.rank}_{\mathbb{C}}(n, n, p) \leqslant \begin{cases} \dfrac{(p+1)n}{2} & \text{if } p \text{ is odd} \\ \dfrac{(p+2)n}{2} & \text{if } p \text{ is even.} \end{cases}$$

Later in the same year, they improved the upper bound for the p even case (see [2]), showing that

$$\text{max.rank}_{\mathbb{C}}(n, n, p) \leqslant \frac{(p+1)n}{2}$$

for all p. Atkinson and Stephens also showed in [3, Theorem 2] that

$$\text{max.rank}_{\mathbb{C}}(m, n, mn-k) = mn - k^2 + \text{max.rank}_{\mathbb{C}}(k, k, k^2 - k).$$

As a result, the rank of all $3 \times 3 \times p$ tensor (except for $p = 5$) were obtained:

p	1	2	3	4	6	7	8	9
$\text{max.rank}_{\mathbb{C}}(3,3,p)$	3	4	5	6	7	8	8	9

Atkinson and Stephens also asserted that any $n \times n \times 3$ complex tensors have rank no more than $2n - 1$ (see [3]). In 2010, Sumi et al. proved Atkinson and Stephens' assertion and extended their result over the real number field for n odd case (see [59]). They also improved some other results from [2, 3]. In 2016, they summarized their results in [60] as follow:

(1) $\text{max.rank}_{\mathbb{C}}(n,n,3) \leqslant 2n-1$ and $\text{max.rank}_{\mathbb{R}}(n,n,3) \leqslant 2n$,

(2) If $m < n$, then $\text{max.rank}_{\mathbb{F}}(m,n,3) \leqslant m+n-1$,

(3) If n is not congruent to 0 modulo 4, then $\text{rank}_{\mathbb{R}}(n,n,3) \leqslant 2n-1$.

(4) $\text{rank}_{\mathbb{F}}(m,n,p) \leqslant (\varepsilon+1)m + \left\lfloor \frac{n(p-1-\varepsilon)}{2} \right\rfloor$, where $3 \leqslant m \leqslant n \leqslant p$ and $p = 2q + \varepsilon$ for an integer q and $\varepsilon = 0, 1$,

where $\mathbb{F}$ denotes the real number field $\mathbb{R}$ or the complex field $\mathbb{C}$. By extending Atkinson and Stephens' result in [3, Theorem 2] over the real numbers, they showed that

$$\text{max.rank}_{\mathbb{F}}(m,n,mn-k) = m(n-k) + \text{max.rank}_{\mathbb{F}}(m,k,mk-k),$$

and

$$\text{max.rank}_{\mathbb{F}}(m,n,mn-k) = mn - k^2 + \text{max.rank}_{\mathbb{F}}\left(k,k,k^2-k\right),$$

where $\mathbb{F}$ denotes the real number field $\mathbb{R}$ or the complex field $\mathbb{C}$. As a result, the rank of several types of tensors over the complex or reals were obtained:

(1) $\text{max.rank}_{\mathbb{F}}(3,3,3) \leqslant 5$;

(2) $\text{max.rank}_{\mathbb{C}}(4,4,3) \leqslant 7$;

(3) $\max.\mathrm{rank}_{\mathbb{F}}(5,5,3) \leqslant 9$;

(4) $\max.\mathrm{rank}_{\mathbb{C}}(6,6,3) \leqslant 11$,

where $\mathbb{F}$ denotes the real number field $\mathbb{R}$ or the complex field $\mathbb{C}$.

1.3 Canonical Forms

Apart from the maximal ranks, the canonical forms of 3-tensors with small sizes were also studied. In 2008, Vin De Silva and Lek-Heng Lim showed in [21] that any $2 \times 2 \times 2$ tensor over $\mathbb{R}$ can be reduced to one of the eight different canonical forms, i.e., for any $T \in \mathbb{R}^{2\times2\times2}$, there exist nonsingular matrices $A_1, A_2, A_3 \in M_2(\mathbb{R})$ such that $(A_1 \otimes A_2 \otimes A_3)(T)$ is a canonical form, where $A_1 \otimes A_2 \otimes A_3$ is defined by

$$\begin{aligned} A_1 \otimes A_2 \otimes A_3 : \mathbb{R}^{2\times2\times2} &\longrightarrow \mathbb{R}^{2\times2\times2}, \\ \vec{v}_1 \otimes \vec{v}_2 \otimes \vec{v}_3 &\mapsto A_1\vec{v}_1 \otimes A_2\vec{v}_2 \otimes A_3\vec{v}_3. \end{aligned}$$

Furthermore, each of the $2 \times 2 \times 2$ tensor over $\mathbb{R}$ can be transformed to a unique canonical form only, which gives a classification of all $2 \times 2 \times 2$ tensor over $\mathbb{R}$ into eight different orbits.

Orbit	**Canonical Form**	**Rank**
D_0	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	0
D_1	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	1
D_2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$	2

<table>
<tr><td>D_2'</td><td>$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$</td><td>2</td></tr>
<tr><td>D_2''</td><td>$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$</td><td>2</td></tr>
<tr><td>G_2</td><td>$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$</td><td>2</td></tr>
<tr><td>D_3</td><td>$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$</td><td>3</td></tr>
<tr><td>G_3</td><td>$\left[\begin{array}{cc|cc} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{array}\right]$</td><td>3</td></tr>
</table>

Table 1.2: Canonical Forms of $2 \times 2 \times 2$ Real Tensors

Later in 2009, Sumi et al. ([58, 60]) gave the canonical forms for complex 3-tensors with 2 frontal slices of arbitrary sizes: For any $T = (A; B)$, where $A, B \in M_{m \times n}(\mathbb{C})$, T is equivalent to a tensor of block diagonal form

$$\mathrm{Diag}\left((S_1; T_1), \ldots, (S_r; T_r)\right)$$

where each $(S_j; T_j)$ is one of the following:

1. zero tensor $(O_{k \times l}; O_{k \times l}) \in \mathbb{C}^{k \times 2 \times l}, k, l \geqslant 0$, $(k, l) \neq (0, 0)$;

2. $(J_k(a); I_k) \in \mathbb{C}^{k \times 2 \times l}$, $a \in \mathbb{C}$, $k \geqslant 1$;

3. $(I_k; J_k(0)) \in \mathbb{C}^{k \times 2 \times l}, k \geqslant 1$;

4. $([O_{k \times 1}, I_k]; [I_k, O_{k \times 1}]) \in \mathbb{C}^{k \times 2 \times (k+1)}, k \geqslant 1$;

5. $\left(\begin{bmatrix} O_{1\times k} \\ I_k \end{bmatrix}; \begin{bmatrix} I_k \\ O_{1\times k} \end{bmatrix}\right) \in \mathbb{C}^{(k+1)\times 2\times k}, k \geqslant 1,$

where I_k denotes the $k \times k$ identity matrix and $J_k(a)$ denotes the $k \times k$ Jordan block with eigenvalue a.

1.4 Quaternionic Tensors

In contrast to tensor decomposition and tensor rank problems over commutative algebras where there are some known results for specific cases over the complex and real numbers, the tensor decomposition and tensor rank problems over the real quaternion algebra

$$\mathbb{H} = \{a_0 + a_1 i + a_2 j + a_3 k | i^2 = j^2 = k^2 = ijk = -1; a_0, a_1, a_2, a_3 \in \mathbb{R}\}.$$

are at present far from fully developed and remain largely open questions.

Knowing how to check the rank of a given tensor and finding a minimal decomposition into simple tensors is not just of theoretical importance, but has many real life applications as well. Such results can be useful for questions arising in applied mathematics, engineering, physics and computer science. For example, Sylvester-type equations for tensors usually involves assumptions that depend on tensor ranks. Not all problems are commutative in nature however, so understanding tensors in the noncommutative case is also important.

There are very few results involving tensor decompositions over $\mathbb{H}$, at least in comparison to known results over $\mathbb{C}$ and $\mathbb{R}$. In [58], upper bounds for the rank of tensors with size $2 \times \cdots \times 2$ over $\mathbb{R}$ or $\mathbb{C}$ are studied. For instance, the maximal rank of any real $2 \times 2 \times 2 \times 2$ tensor is 5, while the maximal rank for a complex tensor of the same size is 4 (these bounds were shown earlier in [38] and [6] respectively). In some recent work involving $\mathbb{H}$, a simultaneous diagonalization result in [26] produces solutions to a specific generalized Sylvester quaternion matrix equation, while expanded work in [68] provides solutions for a two-sided coupled Sylvester-type equation in a similar setting. These results however assume that the ranks of the tensors being used are known. It is therefore essential to find ways of determining the ranks of tensors over $\mathbb{H}$ if one hopes to utilize these results.

In the development of determining the ranks of other small real tensors, Kruskal first stated in 1989 ([40]) without proof that $3 \times 3 \times 3$ real tensors also have ranks at most 5. In 2013, a formal self-contained proof was given by Murray R. Bremner and Jiaxiong Hu ([4]). In 2016, a detailed simple proof, which mainly involves linear transformation and matrix diagonalization, was given by Toshio Sakata, Toshio Sumi and Mitsuhiro Miyazaki ([60]). In addition, they also used other simple evaluation methods to show that $2 \times 2 \times 2$ and $2 \times 2 \times 3$ real tensors have maximal rank 3, $2 \times 3 \times 3$ tensors have maximal rank 4 ([60]), and characterized the $2 \times 2 \times 2$ tensors having rank 3 by using the Cayley's hyperdeterminant ([61]).

In this thesis, we will start with showing that $2 \times 2 \times 2$ quaternion tensors have ranks no greater than 3. After that, we will bound the ranks of some other small sizes quaternionic tensors.

1. $\text{max.rank}_{\mathbb{H}}(2, 3, 2) = 3$;

2. $\text{max.rank}_{\mathbb{H}}(3, 2, 2) = 3$;

3. $\text{max.rank}_{\mathbb{H}}(3, 3, 2) = 4$;

4. $\text{max.rank}_{\mathbb{H}}(3, 2, 3) = 4$;

5. $\text{max.rank}_{\mathbb{H}}(3, 3, 3) = 5$.

In many of the results in this thesis, it is desirable to first simplify a tensor $T = (A_1; \ldots; A_k)$ by first applying column and row operations to the matrices A_i which preserve rank. Since we are working over a noncommutative division ring, we need to be careful with how $\mathbb{H}$ is acting in a column or row (see [70, Section 1.3.3] for an exposition on the difficulties of defining tensors over the quaternions).

If we however only act by $\mathbb{H}$ on the left when using row operations (i.e. horizontal slice operations), and by $\mathbb{H}$ on the right for column operations (i.e. lateral slice operations), then one can check that the rank is preserved. In particular, we are actually endowing $\mathbb{H}^{n_1\times\ldots\times n_k}$ with a bimodule structure where multiplication takes place on the left and right by nonsingular quaternionic matrices. Using real frontal slice operations is also allowed since $\mathbb{R}$ is the center of $\mathbb{H}$. We will call each of these **rank-preserving operations**. In the latter case for example, let A_i be an $N_i \times M_i$ matrix with entries in $\mathbb{R}$ for $i = 1, \ldots, p$. Consider the multilinear map defined by

$$A_1 \otimes \ldots \otimes A_p : \mathbb{H}^{N_1\times\ldots\times N_p} \longrightarrow \mathbb{H}^{M_1\times\ldots\times M_p},$$
$$\vec{v}_1 \otimes \ldots \otimes \vec{v}_p \mapsto A_1\vec{v}_1 \otimes \ldots \otimes A_p\vec{v}_p.$$

This map is well-defined since the action is linear in each component. When we apply $A_1\otimes\ldots\otimes A_p$ to T, we can bound the rank of the resulting image by $\operatorname{rank}(T)$. However, allowing all the A_i's to have entries in $\mathbb{H}$ would no longer define a multilinear map, even though using row operations with left $\mathbb{H}$ multiplication (or column operations with right $\mathbb{H}$ multiplication) coming from nonsingular matrices does preserve rank.

In the last chapter, we determine the maximal rank of $m \times 2 \times n$ quaternion tensors for any $m, n \in \mathbb{N}^+$ and their canonical forms (Theorems 4.3.7 and 4.3.4), and then we generalize some results in the complex 3-tensor case to the the quaternion tensor case with higher orders. For instance, in Theorem 4.4.4, we generalize the equality

$$\operatorname{max.rank}_{\mathbb{C}}(m, n, mn - k) = mn - k^2 + \operatorname{max.rank}_{\mathbb{C}}(k, k, k^2 - k),$$

which was given by Atkinson and Stephens ([3, Theorem 2]) in 1979, and in Theorem 4.2.3 we generalize a classical inequality by Frobenius about matrix ranks ([46, Theorem 5.66])

$$\operatorname{rank}(X) \geqslant \operatorname{rank}(UX) + \operatorname{rank}(XV) - \operatorname{rank}(UXV),$$

A similar result was also given by Kruskal ([39, Theorem 1]) in 1977 for complex 3-tensors.

Chapter 2

Preliminaries

In this chapter, we introduce some basic properties of quaternionic tensors.

2.1 Real Quaternions

The real quaternions were discovered by Hamilton in 1843. Since then, they have been widely used in many areas such as computer graphics, control theory, physics, robotics, image and signal processing, electromechanics, quantum mechanics, etc. We refer the readers to [44, 15, 41, 49, 51, 50, 53, 67, 69] for more information.

Denote the real field by $\mathbb{R}$ and the complex field by $\mathbb{C}$. The real quaternion algebra is defined by

$$\mathbb{H} = \left\{ a_0 + a_1 i + a_2 j + a_3 k \mid i^2 = j^2 = k^2 = ijk = -1; a_0, a_1, a_2, a_3 \in \mathbb{R} \right\}.$$

This is a 4-dimensional associative algebra over $\mathbb{R}$ with an ordered basis $1, i, j$ and k

with the multiplication table:

$\times$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

.

Table 2.1: Multiplication Table of the Real Quaternion Algebra

In other words, a real quaternion is a vector

$$x = x_0 + x_i i + x_2 j + x_3 k \in \mathbb{H}$$

with coefficients $x_0, x_1, x_2, x_3 \in \mathbb{R}$. Hence it has the usual addition and the scalar multiplication defined in the vector space $\mathbb{H}$ over $\mathbb{R}$,. Besides, the product of any two of the quaternions $1, i, j, k$ can be computed by using the table, i.e.

$$i^2 = j^2 = k^2 = -1$$
$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j.$$

Therefore, we can see that every element $x \in \mathbb{H}$ be can uniquely written in the form

$$x = x_0 + x_i i + x_2 j + x_3 k$$

for some $x_0, x_1, x_2, x_3 \in \mathbb{R}$, and thus real numbers and complex numbers can be thought of as quaternions in the natural way.

However, a major feature of quaternions is that multiplication of two quaternions is **noncommutative**. For example, in the given multiplication table, we have already seen that $ij = k$ but $ji = -k$, which shows that $ij \neq ji$. The following example even shows us a surprising fact.

Example 2.1.1. The equation $x^2 + 1 = 0$ has infinitely many solutions in $\mathbb{H}$.

Proof. Consider quaternions of the form

$$x = \cos\theta i + \sin\theta j$$

Taking the square, we have

$$x^2 = (\cos\theta i + \sin\theta j)^2 = \cos^2\theta i^2 + \sin^2\theta j^2 + \cos\theta\sin\theta(ij + ji).$$

The third term is zero, since $ij = -ji$. The first two terms sum to -1, so any number of the given form is a solution to $x^2 + 1 = 0$. Since there are uncountably many $\cos\theta \in [0, 1]$, there are uncountably many solutions. $\square$

2.2 Basics of Tensor Rank

Definition 2.2.1. A multiway array $T = (T_{i_1 i_2 \dots i_K})$ where $1 \leqslant i_1 \leqslant N_1, \dots, 1 \leqslant i_K \leqslant N_K$ is called a K-way tensor of size $(N_1, N_2, \dots, N_K)$. If each of the entries $T_{i_1 i_2 \dots i_K}$ are taken in $\mathbb{R}$, then we call T a real (or complex, quaternionic, respectively) tensor, denoted by $T \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_K}$ (or $T \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_K}$, $T \in \mathbb{H}^{N_1 \times N_2 \times \cdots \times N_K}$, respectively). We also say that T is an $N_1 \times N_2 \times \ldots \times N_K$ tensor.

When $K = 3$, we will use the convention that N_2 indicates the number of *frontal slices* of the array, so that the array consists of N_2 many $N_1 \times N_3$ matrices. Similarly, it will have N_1 *horizontal slices* and N_3 *lateral slices.* For example a $3 \times 2 \times 3$ tensor has 18 entries, which can be denoted by

$$T = \left(\begin{bmatrix} a_{111} & a_{121} & a_{131} \\ a_{211} & a_{221} & a_{231} \\ a_{311} & a_{321} & a_{331} \end{bmatrix} ; \begin{bmatrix} a_{112} & a_{122} & a_{132} \\ a_{212} & a_{222} & a_{232} \\ a_{312} & a_{322} & a_{332} \end{bmatrix} \right)$$

When $K > 3$, it is no longer convenient to use the same notation as above to describe tensors.

Note: We should mention that this notation is different from the one used by some authors cited in this thesis (for example, in [60], N_3 indicates the number of frontal slices).

Definition 2.2.2. A nonzero real tensor (or complex, quaternionic, respectively) $T = (T_{i_1 i_2 \ldots i_K})$ is called a **simple tensor** if there exist real (or complex, quaternionic, respectively) vectors

$$\begin{aligned} \vec{a_1} &= (a_{11}, a_{12}, \ldots, a_{1N_1}), \\ \vec{a_2} &= (a_{21}, a_{22}, \ldots, a_{2N_2}), \\ &\vdots \\ \vec{a_K} &= (a_{K1}, a_{K2}, \ldots, a_{KN_K}), \end{aligned}$$

such that $T = (T_{i_1 i_2 \ldots i_K}) = (a_{1i_1} a_{2i_2} \ldots a_{Ki_K})$. We will also denote this by

$$T = \vec{a}_1 \otimes \vec{a}_2 \otimes \cdots \otimes \vec{a}_K.$$

Example 2.2.3. For the $2 \times 3 \times 2$ tensor

$$T = \left(\begin{bmatrix} 2 & 3 \\ 8 & 12 \end{bmatrix} ; \begin{bmatrix} -2 & -3 \\ -8 & -12 \end{bmatrix} ; \begin{bmatrix} 4 & 6 \\ 16 & 24 \end{bmatrix} \right),$$

there exist vectors

$$\vec{a} = (a_1, a_2) = (1, 4), \ \vec{b} = (b_1, b_2, b_3) = (1, -1, 2) \text{ and } \vec{c} = (c_1, c_2) = (2, 3)$$

such that

$$\begin{aligned} T &= \left(\begin{bmatrix} a_1b_1c_1 & a_1b_1c_2 \\ a_2b_1c_1 & a_2b_1c_2 \end{bmatrix}; \begin{bmatrix} a_1b_2c_1 & a_1b_2c_2 \\ a_2b_2c_1 & a_2b_2c_2 \end{bmatrix}; \begin{bmatrix} a_1b_3c_1 & a_1b_3c_2 \\ a_2b_3c_1 & a_2b_3c_2 \end{bmatrix} \right) \\ &= \vec{a} \otimes \vec{b} \otimes \vec{c}. \end{aligned}$$

Therefore, T is a simple tensor. □

Definition 2.2.4. Let T be a nonzero real (or complex, quaternionic, respectively) tensor. Then the rank of T is the smallest positive integer n such that

$$T = T_1 + T_2 + \ldots + T_n,$$

where $T_1, T_2, \ldots, T_n$ are simple real (or complex, quaternionic, respectively) tensors. We will say that T has rank n and denote this by $\text{rank}_{\mathbb{R}}(T) = n$ (or $\text{rank}_{\mathbb{C}}(T) = n$, $\text{rank}_{\mathbb{H}}(T) = n$, respectively).

Note: We will use simply the notation $\text{rank}(T)$ for general situations (including the real, complex, and quaternion cases) when the underlying base doesn't affect the statement.

An immediate consequence of this definition is that $\text{rank}(T + S) \leqslant \text{rank}(T) + \text{rank}(S)$. Any sum of simple tensors $T_1 + T_2 + \ldots + T_n = T$ is called a **tensor decomposition** for T, even if n is not minimal. We will highlight various nontrivial tensor decompositions in Chapter 3 where n is always equal to the minimal known bound on tensor rank. It should be noted that some authors refer to a tensor decomposition in the singular value decomposition sense, and not as a sum of simple tensors (see [25] for this type of tensor decomposition in the quaternion case).

For $A_i \in M_{M_i \times N_i}(\mathbb{R})$, $i = 2, ..., p-1$ and $A_j \in M_{M_i \times N_i}(\mathbb{H})$, $j = 1, p$, let $\bigotimes_{i=1}^{p} A_i = A_1 \otimes \cdots \otimes A_p$ be the multilinear map from $\mathbb{H}^{N_1 \times \cdots \times N_p}$ to $\mathbb{H}^{M_1 \times \cdots \times M_p}$ determined by

$$A_1 \otimes \cdots \otimes A_p : \mathbb{H}^{N_1 \times \cdots \times N_p} \longrightarrow \mathbb{H}^{M_1 \times \cdots \times M_p},$$
$$\vec{v}_1 \otimes \cdots \otimes \vec{v}_p \mapsto A_1 \vec{v}_1 \otimes A_2 \vec{v}_2 \otimes \cdots \otimes A_{p-1} \vec{v}_{p-1} \otimes \left(\vec{v}_p^{\,T} A_p^T \right)^T.$$

Since $A_2, ..., A_{p-1}$ are real matrices and the action is linear in each component, this map is well-defined by universal property.

For any $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$ and $m \in \{1, 2, ..., p\}$,

$$T \times_m A_m := \left(I_{N_1} \otimes \cdots \otimes I_{N_{m-1}} \otimes A_m \otimes I_{N_{m+1}} \otimes \cdots \otimes I_{N_p} \right)(T)$$

is called the m-mode product of A_m and T.

In [45], it was mentioned that one can use multilinear maps formed by real matrices as well as left and right module actions determined by quaternionic matrices to bound the rank of a given tensor. We now summarize it in the following lemma.

Lemma 2.2.5. *Let T be an $N_1 \times \ldots \times N_p$ quaternionic tensor and $A_1 \otimes \ldots \otimes A_p$ a multilinear map from $\mathbb{H}^{N_1 \times \ldots \times N_p}$ to $\mathbb{H}^{M_1 \times \ldots \times M_p}$ defined as above. Then*

$$\operatorname{rank}_{\mathbb{H}}((A_1 \otimes \ldots \otimes A_p)(T)) \leqslant \operatorname{rank}_{\mathbb{H}}(T).$$

Furthermore, if $A_1 \otimes \ldots \otimes A_p$ is invertible, then

$$\operatorname{rank}_{\mathbb{H}}((A_1 \otimes \ldots \otimes A_p)(T)) = \operatorname{rank}_{\mathbb{H}}(T).$$

Proof. Suppose $\operatorname{rank}_{\mathbb{H}}((T) = n$, and write T as a sum of simple tensors

$$T = \sum_{i=1}^{n} v_{i_1} \otimes \ldots \otimes v_{i_p}.$$

By multilinearity, we have

$$(A_1 \otimes \ldots \otimes A_p)(T) = \sum_{i=1}^{n} A_1 \vec{v}_{i_1} \otimes \ldots \otimes A_{p-1} \vec{v}_{i_{p-1}} \otimes \left(\vec{v_{i_p}}^T A_p^T \right)^T,$$

which implies $\operatorname{rank}_{\mathbb{H}}(A_1 \otimes \ldots \otimes A_p)(T)) \leqslant n$. □

A similar proof works if we replace $\mathbb{H}$ with $\mathbb{C}$ and meanwhile restrict all A_i's to be complex matrices. However, allowing all the A_i's to have entries in $\mathbb{H}$ would no longer define a multilinear map, even though using row operations with left $\mathbb{H}$ multiplication (or column operations with right $\mathbb{H}$ multiplication) coming from nonsingular matrices does preserve rank.

Now, we introduce the notation for the adjoint of a quaternion matrix. Given an $n \times n$ matrix A with entries in $\mathbb{H}$, we can uniquely write $A = A_1 + A_2 j$, where A_1 and A_2 are $n \times n$ matrices with entries in $\mathbb{C}$. The **complex adjoint matrix** of A (or simply the adjoint of A), is defined as the $2n \times 2n$ complex block matrix

$$\chi(A) = \begin{bmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{bmatrix}.$$

The adjoint matrix is very useful in converting a diagonalization problem over $\mathbb{H}$ into a diagonalization problem over $\mathbb{C}$. The following result is well-known ([49]).

Lemma 2.2.6. *An $n \times n$ matrix A with entries in $\mathbb{H}$ is diagonalizable if and only if $\chi(A)$ is diagonalizable.*

Moving on, we introduce two auxiliary lemmas which are important in the next chapter that follows.

Lemma 2.2.7. *Let $T = (A_1; A_2; \ldots; A_p)$ be an $m \times p \times n$ tensor. Then $\operatorname{rank}_{\mathbb{H}}(T) \leqslant r$ if and only if there are $r \times r$ diagonal matrices D_i, an $m \times r$ matrix P, and an $r \times n$ matrix Q such that $A_k = PD_kQ$, for $k = 1, \ldots, p$.*

Proof. The following argument is a slight alteration from the one found in [60, Proposition 2.1]. First suppose that $\operatorname{rank}_{\mathbb{H}}(T) \leqslant r$, and write T as a sum of simple tensors

$$T = \sum_{i=1}^{r} \vec{a_i} \otimes \vec{b_i} \otimes \vec{c_i}.$$

Let us write $\vec{b_i} = (b_{i1}, b_{i2}, \ldots, b_{ip})$ for $1 \leqslant i \leqslant r$, so that $A_k = \sum_{i=1}^{r} \vec{a_i} b_{ik} \vec{c_i}^T$. We will define P, Q and D_k by

$$P = [\vec{a_1}, \vec{a_2}, \ldots, \vec{a_r}], \quad Q = \begin{bmatrix} \vec{c_1}^T \\ \vec{c_2}^T \\ \vdots \\ \vec{c_r}^T \end{bmatrix}, \quad D_k = \operatorname{diag}(b_{1k}, b_{2k}, \ldots, b_{rk}).$$

Then

$$PD_kQ = [\vec{a_1}, \vec{a_2}, \ldots, \vec{a_r}] \begin{bmatrix} b_{1k} & & & \\ & b_{2k} & & \\ & & \ddots & \\ & & & b_{rk} \end{bmatrix} \begin{bmatrix} \vec{c_1}^T \\ \vec{c_2}^T \\ \vdots \\ \vec{c_r}^T \end{bmatrix} = \sum_{i=1}^{r} \vec{a_i} b_{ik} \vec{c_i}^T = A_k,$$

as required.

Proceeding similarly for the other direction, assume that there are $r \times r$ diagonal matrices D_k, an $m \times r$ matrix P, and an $r \times n$ matrix Q such that $A_k = PD_kQ$, for $k = 1, \ldots, p$. Then writing P, Q and D_k as above, we have:

$$A_k = PD_kQ = \sum_{i=1}^{r} \vec{a_i} b_{ik} \vec{c_i}^T \Longrightarrow T = \sum_{i=1}^{r} \vec{a_i} \otimes \vec{b_i} \otimes \vec{c_i}.$$

Hence $\text{rank}_{\mathbb{H}}(T) \leqslant r$, completing the proof. $\square$

Lemma 2.2.8. *Let $T = (A_1; A_2; \ldots; A_p)$ be an $n \times p \times n$ tensor, where A_1 is non-singular. Then $\text{rank}_{\mathbb{H}}(T) = n$ if and only if $\left\{A_j A_1^{-1} \mid j = 2, 3, \ldots, p\right\}$ can be simultaneously diagonalized.*

Proof. The following proof is similar to that found in [60, Proposition 2.5]. First suppose that $\text{rank}_{\mathbb{H}}(T) = n$. Then by Lemma 2.2.7, there exists an $n \times r$ matrix P, an $r \times n$ matrix Q and $r \times r$ diagonal matrices $D_1, D_2, \ldots, D_p$ such that

$$A_1 = PD_1Q, \;\; A_2 = PD_2Q, \;\; \ldots, \;\; A_p = PD_pQ.$$

Since A_1 is non-singular, D_1 must have rank n and thus $r = n$. This implies that P, Q and D_1 are each invertible. Therefore

$$A_j A_1^{-1} = PD_jQ(PD_1Q)^{-1} = PD_jQQ^{-1}D_1^{-1}P^{-1} = PD_jD_1^{-1}P^{-1}$$

for $j = 1, \ldots n$, and thus the $A_j A_1^{-1}$ can be simultaneously diagonalized.

For the other direction, suppose that there exists an $n \times n$ matrix P where

$$D_j = P^{-1}A_jA_1^{-1}P, \;\; j = 2, 3, \ldots, p$$

are diagonal matrices. Consider the tensor

$$T' = P^{-1}TA_1^{-1}P = (I_n; D_2; \ldots; D_p).$$

It is easy to check that $\text{rank}_{\mathbb{H}}(T') = n$, and since multiplication by invertible matrices is rank preserving, $\text{rank}_{\mathbb{H}}(T) = n$ as required. $\square$

Definition 2.2.9. Let R_i be the $\mathbb{H}$-homomorphism induced by

$$R_i : \mathbb{H}^{N_1 \times \ldots \times N_p} \longrightarrow \mathbb{H}^{N_2 \times \ldots \times N_p},$$
$$\vec{v}_1 \otimes \ldots \otimes \vec{v}_p \mapsto v_{1i} \left(\vec{v}_2 \otimes \ldots \otimes \vec{v}_p\right),$$

where $v_{1i} \in \mathbb{H}$ is the i-th coordinate of the vector $\vec{v}_1 \in \mathbb{H}^{N_1}$, $i = 1, 2, ..., N_1$. The left $\mathbb{H}$-module generated by $\{R_i(T) \mid i = 1, 2, \ldots, N_1\}$ is denoted by $R(T)$. The dimension of $R(T)$ is denoted by $\dim_{\mathbb{H}} R(T)$.

Definition 2.2.10. Let C_j be the $\mathbb{H}$-homomorphism induced by

$$C_j : \mathbb{H}^{N_1 \times \ldots \times N_p} \longrightarrow \mathbb{H}^{N_1 \times \ldots \times N_{p-1}},$$
$$\vec{v}_1 \otimes \ldots \otimes \vec{v}_p \mapsto \left(\vec{v}_1 \otimes \ldots \otimes \vec{v}_{p-1}\right) v_{pj},$$

where $v_{pj} \in \mathbb{H}$ is the j-th coordinate of the vector $\vec{v}_p \in \mathbb{H}^{N_p}$, $j = 1, 2, ..., N_p$. The right $\mathbb{H}$-module generated by $\{C_j(T) \mid j = 1, 2, \ldots, N_p\}$ is denoted by $C(T)$. The dimension of $C(T)$ is denoted by $\dim_{\mathbb{H}} C(T)$.

Definition 2.2.11. Let T be an $N_1 \times N_2 \times \cdots \times N_p$ tensor over the quaternions. The **conjugate transpose** of T is an $N_p \times N_{p-1} \times \cdots \times N_1$ tensor T^* defined via the $(\mathbb{H}, \mathbb{H})$-bimodule anti-isomorphism induced by

$$* : \mathbb{H}^{N_1 \times \cdots \times N_p} \longrightarrow \mathbb{H}^{N_p \times \cdots \times N_1},$$
$$\vec{v}_1 \otimes \cdots \otimes \vec{v}_p \mapsto (\vec{v_p}^*)^T \otimes \cdots \otimes (\vec{v_1}^*)^T.$$

Therefore, for any $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, T has the same rank as T^*. This gives us a useful tool to bound the ranks of quaternionic tensors.

Proposition 2.2.12. *The maximal rank of $N_1 \times N_2 \times \cdots \times N_p$ quaternionic tensors equals the maximal rank of $N_p \times N_{p-1} \times \cdots \times N_1$ quaternionic tensors.*

Proof. Since the conjugate transpose $*$ is a bijective multilinear map, by lemma 2.2.5 this is obvious. $\square$

Chapter 3

Small Quaternionic 3-Tensors

In this chapter, we give the ranks of some other small quaternionic tensors:

1. $\text{max.rank}_{\mathbb{H}}(2, 2, 2) = 3$;
2. $\text{max.rank}_{\mathbb{H}}(2, 3, 2) = 3$;
3. $\text{max.rank}_{\mathbb{H}}(3, 2, 2) = 3$;
4. $\text{max.rank}_{\mathbb{H}}(3, 3, 2) = 4$;
5. $\text{max.rank}_{\mathbb{H}}(3, 2, 3) = 4$;
6. $\text{max.rank}_{\mathbb{H}}(3, 3, 3) = 5$.

3.1 The $2 \times 2 \times 2$ case

We will show in this section that $2 \times 2 \times 2$ quaternion tensors have a rank no greater than 3.

Proposition 3.1.1. *Let*

$$T = \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}; \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \right)$$

be a $2 \times 2 \times 2$ *quaternion tensor. If* $A_{11}B_{11} \neq 0$*, then* $T = T_1 + T_2 + T_3$ *where:*

$$T_1 = \left(\begin{bmatrix} A_{11} & A_{11}(A_{11}^{-1}A_{12}) \\ A_{21} & A_{21}(A_{11}^{-1}A_{12}) \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} B_{11} & B_{11}(B_{11}^{-1}B_{12}) \\ B_{21} & B_{21}(B_{11}^{-1}B_{12}) \end{bmatrix} \right),$$

$$T_3 = \left(\begin{bmatrix} 0 & 0 \\ 0 & A_{22} - A_{21}(A_{11}^{-1}A_{12}) \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & B_{22} - B_{21}(B_{11}^{-1}B_{12}) \end{bmatrix} \right).$$

One could ask whether this tensor decomposition is unique up to rescaling or permutation indeterminacy (see [47] for details about uniqueness of tensor decompositions). While it is not difficult to find examples where a decomposition does not appear to be unique, we will not consider uniqueness questions for the purposes of this thesis. We can however bound the rank of any quaternion tensor T of size $2 \times 2 \times 2$ using the explicit decomposition into at most 3 simple tensors from the proposition. To ensure that we can apply this result, we first need to ensure that we can write T in the desired form, and this requires us to use row and column operations.

Theorem 3.1.2. *Let* T *be a* $2 \times 2 \times 2$ *quaternion tensor. Then* $\mathrm{rank}_{\mathbb{H}}(T) \leqslant 3$.

Proof. Let us write the tensor as

$$T = \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}; \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \right).$$

Since $\mathbb{R}$ is the center of $\mathbb{H}$, by Lemma 2.2.5 we can perform real elementary operations on T (including frontal slice, horizontal slice and lateral slice operations). Except for the trivial cases with too many zero entries (in which case the tensor rank is obvious), we may assume that A_{11} and B_{11} are not zero. Then $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 3$ follows from Proposition 3.1.1 since the rank of each T_i is no more than 1. $\square$

It is not difficult to show that the tensor

$$T = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right)$$

has rank 3. In fact, it follows immediately from Lemma 2.2.8. Therefore, the bound on the rank for $2 \times 2 \times 2$ quaternion tensors is the best possible.

3.2 The $2 \times 2 \times 3$ and $2 \times 3 \times 2$ cases

When working over a commutative ring, there is no difference between the $2 \times 2 \times 3$, $2 \times 3 \times 2$, and $3 \times 2 \times 2$ tensor cases. Over the quaternions however, the $2 \times 3 \times 2$ tensor differs from the other two cases.

We will start our discussion by working over the complex numbers. While it is known that the rank of a complex tensor of this size is bounded by 3, as far as we are aware, an explicit decomposition like the one below has not been made readily available.

Theorem 3.2.1. *Let T be a complex $2 \times 3 \times 2$ tensor*

$$T = \left(\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} ; \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} ; \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} \right)$$

such that the two matrices M and N are invertible:

$$M = \begin{bmatrix} A_2 & A_3 & A_4 \\ B_2 & B_3 & B_4 \\ C_2 & C_3 & C_4 \end{bmatrix}, \quad N = \begin{bmatrix} A_1 & A_2 & A_4 \\ B_1 & B_2 & B_4 \\ C_1 & C_2 & C_4 \end{bmatrix}.$$

Then T has a decomposition as the sum of the 3 simple tensors defined below.

Proof. Define S_1, S_2, T_1 and T_2 by

$$S_1 = A_2B_3C_4 - A_2B_4C_3 - A_3B_2C_4 + A_3B_4C_2 + A_4B_2C_3 - A_4B_3C_2,$$

$$S_2 = A_1B_2C_4 - A_1B_4C_2 - A_2B_1C_4 + A_2B_4C_1 + A_4B_1C_2 - A_4B_2C_1,$$

$$T_1 = A_1B_3C_4 - A_1B_4C_3 - A_3B_1C_4 + A_3B_4C_1 + A_4B_1C_3 - A_4B_3C_1,$$

$$T_2 = A_1B_2C_3 - A_1B_3C_2 - A_2B_1C_3 + A_2B_3C_1 + A_3B_1C_2 - A_3B_2C_1.$$

We can check by inspection that

$$A_2T_1 - A_1S_1 + A_4T_2 = A_3S_2,$$

$$B_2T_1 - B_1S_1 + B_4T_2 = B_3S_2,$$

$$C_2T_1 - C_1S_1 + C_4T_2 = C_3S_2.$$

Since $S_1 = \det(M) \neq 0$ and $S_2 = \det(N) \neq 0$, we can write $T = T_1 + T_2 + T_3$ where:

$$T_1 = \left(\begin{bmatrix} A_2T_1S_1^{-1} & A_2 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} B_2T_1S_1^{-1} & B_2 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} C_2T_1S_1^{-1} & C_2 \\ 0 & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} 0 & 0 \\ A_4T_2S_2^{-1} & A_4 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ B_4T_2S_2^{-1} & B_4 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ C_4T_2S_2^{-1} & C_4 \end{bmatrix} \right),$$

$$T_3 = (S_1; S_2; S_3)\,,$$

where $S_1 = \begin{bmatrix} (A_1S_1 - A_2T_1)S_1^{-1} & 0 \\ -(A_1S_1 - A_2T_1)S_2^{-1} & 0 \end{bmatrix}$, $\quad S_2 = \begin{bmatrix} (B_1S_1 - B_2T_1)S_1^{-1} & 0 \\ -(B_1S_1 - B_2T_1)S_2^{-1} & 0 \end{bmatrix}$,

and $S_3 = \begin{bmatrix} (C_1S_1 - C_2T_1)S_1^{-1} & 0 \\ -(C_1S_1 - C_2T_1)S_2^{-1} & 0 \end{bmatrix}$. □

If we try and generalize Theorem 3.1.2 or Theorem 3.2.1 to write an explicit decomposition for the $2 \times 2 \times 3$ quaternion case, we immediately run into problems with the operations needed to write T in an appropriate form. In Theorem 3.1.2 for example, we only needed real row and column operations. As mentioned in remark after Lemma 2.2.5, we cannot freely apply row and column operations using quaternions, and can only use row operations where $\mathbb{H}$ is acting on the left, and column operations where $\mathbb{H}$ is acting on the right (that is, horizontal and lateral slice operations). Nonetheless, we can provide a basic bound on the rank even if we cannot write down an explicit decomposition. Let us start with a motivational example.

Example 3.2.2. Consider the quaternion $2 \times 2 \times 3$ tensor

$$T = \left(\begin{bmatrix} 1 & i & 0 \\ 0 & -j & 1+i \end{bmatrix}; \begin{bmatrix} 0 & 1+j & 0 \\ 0 & i+k & 1+j \end{bmatrix} \right).$$

We can apply rank-preserving row and column operations:

(a) $C_2 \to C_2 j$,

(b) $C_3 \to C_3 - C_2(1+i) + C_1(j+k)$,

(c) $C_2 \to 3C_2 - C_3(-2+i-k)$

to reduce T to a form that is easier to decompose into 3 simple tensors.

$$T \longrightarrow \left(\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 1+i \end{bmatrix}; \begin{bmatrix} 0 & -1+j & 0 \\ 0 & -i+k & 1+j \end{bmatrix} \right) \quad \text{using (a)}$$

$$\longrightarrow \left(\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \end{bmatrix}; \begin{bmatrix} 0 & -1+j & 1+i-j+k \\ 0 & -i+k & i-k \end{bmatrix} \right) \quad \text{using (b)}$$

$$\longrightarrow \left(\begin{bmatrix} 1 & 3k & 0 \\ 0 & 3 & 0 \end{bmatrix}; \begin{bmatrix} 0 & -1-j+2k & 1+i-j+k \\ 0 & 2-i+k & i-k \end{bmatrix} \right) \quad \text{using (c)}$$

Then the reduced tensor

$$T' := \left(\begin{bmatrix} 1 & 3k & 0 \\ 0 & 3 & 0 \end{bmatrix}; \begin{bmatrix} 0 & -1-j+2k & 1+i-j+k \\ 0 & 2-i+k & i-k \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} 1 & 3k & 0 \\ 0 & 3 & 0 \end{bmatrix}; \begin{bmatrix} 0 & k(2-i+k) & (-i-j)(i-k) \\ 0 & 2-i+k & i-k \end{bmatrix} \right)$$

can be decompose into $T' = T_1 + T_2 + T_3$ where:

$$T_1 = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & (-i-j)(i-k) \\ 0 & 0 & i-k \end{bmatrix} \right),$$

$$T_3 = \left(\begin{bmatrix} 0 & k \cdot 3 & 0 \\ 0 & 3 & 0 \end{bmatrix}; \begin{bmatrix} 0 & k(2-i+k) & 0 \\ 0 & 2-i+k & 0 \end{bmatrix} \right).$$

Therefore, $\mathrm{rank}_{\mathbb{H}}(T) \leqslant 3$. $\square$

Theorem 3.2.3. *Let T be a $2 \times 2 \times 3$ or a $3 \times 2 \times 2$ quaternion tensor. Then* $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 3$.

Proof. Let $T = (A; B)$ where

$$A = \begin{bmatrix} \vec{a} & \vec{b} & \vec{c} \end{bmatrix}, \quad B = \begin{bmatrix} \vec{d} & \vec{e} & \vec{f} \end{bmatrix},$$

and $\vec{a}, \ldots, \vec{f}$ are 2-dimensional column vectors. If either A or B has rank no greater than 1, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \leqslant \operatorname{rank}_{\mathbb{H}}(A; 0)) + \operatorname{rank}_{\mathbb{H}}(0; B)) \leqslant 1 + 2 = 3.$$

Therefore, let us assume that $\operatorname{rank}_{\mathbb{H}}(A) = \operatorname{rank}_{\mathbb{H}}(B) = 2$. Without loss of generality, we can assume that $\operatorname{rank}_{\mathbb{H}}(\vec{a}, \vec{b}) = 2$. Then, by a column operation, we have

$$T = \left(\begin{bmatrix} \vec{a} & \vec{b} & \vec{c} \end{bmatrix}; \begin{bmatrix} \vec{d} & \vec{e} & \vec{f} \end{bmatrix} \right) \longrightarrow \left(\begin{bmatrix} \vec{a} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} \vec{d} & \vec{e} & \vec{g} \end{bmatrix} \right).$$

If $\vec{g} = 0$, then $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 3$ by Theorem 3.1.2. Let us assume then that $\vec{g} \neq 0$, and without loss of generality, that $\operatorname{rank}_{\mathbb{H}}(\vec{e}, \vec{g}) = 2$ (since B has full rank). By further column operations, we have

$$\left(\begin{bmatrix} \vec{a} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} \vec{d} & \vec{e} & \vec{g} \end{bmatrix} \right) \longrightarrow \left(\begin{bmatrix} \vec{h} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} 0 & \vec{e} & \vec{g} \end{bmatrix} \right).$$

Since $\operatorname{rank}_{\mathbb{H}}(\vec{h}, \vec{b}) = 2$, we can write

$$\vec{e} = \vec{h}c_1 + \vec{b}c_2 \ , \quad \vec{g} = \vec{h}c_3 + \vec{b}c_4.$$

If $c_3 \neq 0$, then

$$\left(\begin{bmatrix} \vec{h} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} 0 & \vec{e} & \vec{g} \end{bmatrix} \right) = \left(\begin{bmatrix} \vec{h} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} 0 & \vec{h}c_1 + \vec{b}c_2 & \vec{h}c_3 + \vec{b}c_4 \end{bmatrix} \right)$$

$$\longrightarrow \left(\begin{bmatrix} \vec{h} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} 0 & \vec{b}c_5 & \vec{h}c_3 + \vec{b}c_4 \end{bmatrix}\right),$$

which has rank at most 3 (we can write it as the sum of 3 simple tensors defined using the 2×2 slices in the lateral direction). If $c_3 = 0$, then $c_4 \neq 0$ and we have

$$\left(\begin{bmatrix} \vec{h} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} 0 & \vec{e} & \vec{g} \end{bmatrix}\right) = \left(\begin{bmatrix} \vec{h} & \vec{b} & 0 \end{bmatrix}; \begin{bmatrix} 0 & \vec{h}c_1 + \vec{b}c_2 & \vec{b}c_4 \end{bmatrix}\right)$$

$$\longrightarrow \left(\begin{bmatrix} \vec{h} & \vec{h}c_6 + \vec{b}c_7 & 0 \end{bmatrix}; \begin{bmatrix} 0 & (\vec{h}c_6 + \vec{b}c_7)c_8 & \vec{b}c_4 \end{bmatrix}\right),$$

using the column operation $C_2 \to C_2 + C_1k_1 + C_3k_3$ and choosing k_1, k_3 appropriately (depending on whether $c_1 = 0$ or $c_2 = 0$ for example). Then the rank is at most 3 by the same argument as above.

Finally, the $3 \times 2 \times 2$ case can be proven by using proposition 2.2.12. □

Theorem 3.2.4. *Let T be a $2 \times 3 \times 2$ quaternion tensor. Then* $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 3$.

Proof. Let

$$T = (A; B; C) = \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}\right).$$

If either A, B or C is singular (and if so, we may assume that it's A), then by rank-preserving row and column operations, the tensor can be reduced to

$$T = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} b'_{11} & b'_{12} \\ b'_{21} & b'_{22} \end{bmatrix}; \begin{bmatrix} c'_{11} & c'_{12} \\ c'_{21} & c'_{22} \end{bmatrix}\right).$$

If $b'_{22} = c'_{22} = 0$, then T is the sum of 3 simple tensors defined using the 3 nonzero vectors in the lateral direction ($\vec{v} = (1, b'_{11}, c'_{11})$ is one such vector). Otherwise if either $b'_{22} \neq 0$ or $c'_{22} \neq 0$, then by adding a frontal slice to another we can assume that $b'_{22}c'_{22} \neq 0$, and we have the following decomposition into simple tensors $T = T_1 + T_2 + T_3$:

$$T_1 = \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} b'_{12}(b'_{22})^{-1}b'_{21} & b'_{12} \\ b'_{22}(b'_{22})^{-1}b'_{21} & b'_{22} \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} c'_{12}(c'_{22})^{-1}c'_{21} & c'_{12} \\ c'_{22}(c'_{22})^{-1}c'_{21} & c'_{22} \end{bmatrix} \right),$$

$$T_3 = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} b'_{11} - b'_{12}(b'_{22})^{-1}b'_{21} & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} c'_{11} - c'_{12}(c'_{22})^{-1}c'_{21} & 0 \\ 0 & 0 \end{bmatrix} \right).$$

In both cases we see that $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 3$.

Therefore, we may assume that A, B and C are nonsingular. By performing row and column operations, we can assume that

$$T = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \right).$$

If either the second or the third frontal slice of T is diagonalizable (without loss of generality assume the second one is diagonalizable), then T can be further reduced to

$$S = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} b'_{11} & 0 \\ 0 & b'_{22} \end{bmatrix}; \begin{bmatrix} c'_{11} & c'_{12} \\ c'_{21} & c'_{22} \end{bmatrix} \right), \tag{3.2.1}$$

which has the following decomposition $S = S_1 + S_2 + S_3$ with $\operatorname{rank}_{\mathbb{H}}(S_i) \leqslant 1$:

$$S_1 = \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} 1 & c'_{12} \\ c'_{21} & c'_{21}c'_{12} \end{bmatrix} \right),$$

$$S_2 = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} b'_{11} & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} c'_{11} - 1 & 0 \\ 0 & 0 \end{bmatrix} \right),$$

$$S_3 = \left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & b'_{22} \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & c'_{22} - c'_{21}c'_{12} \end{bmatrix} \right).$$

Otherwise, neither of them is diagonalizable. Then by [71, Theorem 6.3] we can apply unitary triangularization for the second frontal slice to reduce T to

$$S = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} b'_{11} & 0 \\ b'_{21} & b'_{22} \end{bmatrix}; \begin{bmatrix} c'_{11} & c'_{12} \\ c'_{21} & c'_{22} \end{bmatrix} \right).$$

Since the second frontal slice of S is not diagonalizable, b'_{11} and b'_{22} must be equivalent. So we can assume by similarity that $b'_{11} = b'_{22} \in \mathbb{C}$. By adding a real multiple of the first frontal slice to the second, we can assume further that $b'_{11} = b'_{22} = ai$ for some $a \in \mathbb{R}$. Since the second slice is not singular, we have $a \neq 0$. Then we can assume by rescaling the second slice that

$$S = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ b'_{21} & i \end{bmatrix}; \begin{bmatrix} c'_{11} & c'_{12} \\ c'_{21} & c'_{22} \end{bmatrix} \right).$$

Denote $b'_{21} = q_0 + q_1 i + q_2 j + q_3 k$ for some $q_0, q_1, q_2, q_3 \in \mathbb{R}$. We have

$$\begin{bmatrix} 1 & 0 \\ \frac{q_3 j - q_2 k}{2} & 1 \end{bmatrix} \begin{bmatrix} i & 0 \\ b'_{21} & i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{q_3 j - q_2 k}{2} & 1 \end{bmatrix} = \begin{bmatrix} i & 0 \\ q_0 + q_1 i & i \end{bmatrix}.$$

Since the second slice is not diagonalizable by the assumption, $q_0 + q_1 i \neq 0$. So we have

$$\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{q_0+q_1 i} \end{bmatrix} \begin{bmatrix} i & 0 \\ q_0 + q_1 i & i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & q_0 + q_1 i \end{bmatrix} = \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix}.$$

If we apply the same operations as above to the tensor S, we get the resulting tensor

$$S^{(1)} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix}; \begin{bmatrix} c''_{11} & c''_{12} \\ c''_{21} & c''_{22} \end{bmatrix} \right).$$

Similarly, since the third frontal slice is not diagonalizable by the assumption, it can also be written as

$$\begin{bmatrix} c''_{11} & c''_{12} \\ c''_{21} & c''_{22} \end{bmatrix} = P \begin{bmatrix} m & 1 \\ 0 & m \end{bmatrix} P^{-1}$$

for some invertible matrix $P \in M_n(\mathbb{H})$ and some $m \in \mathbb{C}$. Therefore, by adding a multiple of the first frontal slice to the third as well as multiplying the third frontal by a real constant, $S^{(1)}$ can be further reduced to

$$S^{(2)} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix}; P \begin{bmatrix} i & 1 \\ 0 & i \end{bmatrix} P^{-1} \right).$$

Denote $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and decompose $P = P_1^{-1} P_2$ according to the following situations.

(1) If $a = 0$, then define $P_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $P_2 = \begin{bmatrix} c & d \\ 0 & b \end{bmatrix}$.

(2) If $a \neq 0$, then define $P_1 = \begin{bmatrix} 1 & 0 \\ -ca^{-1} & 1 \end{bmatrix}$ and $P_2 = \begin{bmatrix} a & b \\ 0 & d - ca^{-1}b \end{bmatrix}$.

Then in either case, we have

$$S^{(3)} = P_1 S^{(2)} P_1^{-1} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; P_1 \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix} P_1^{-1}; P_2 \begin{bmatrix} i & 1 \\ 0 & i \end{bmatrix} P_2^{-1} \right).$$

For (1), we can easily see that both the second and the third frontal slices of $S^{(3)}$ are upper triangular. So, $S^{(3)}$ is the sum of 3 simple tensors defined using the 3 nonzero vectors in the lateral direction.

For (2), we can compute

$$P_1 \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix} P_1^{-1} = \begin{bmatrix} 1 & 0 \\ -ca^{-1} & 1 \end{bmatrix} \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ ca^{-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 - ca^{-1}i + ica^{-1} & 1 \end{bmatrix}$$

and

$$P_2 \begin{bmatrix} i & 1 \\ 0 & i \end{bmatrix} P_2^{-1} = \begin{bmatrix} a & b \\ 0 & d - ca^{-1}b \end{bmatrix} \begin{bmatrix} i & 1 \\ 0 & i \end{bmatrix} \begin{bmatrix} a^{-1} & -a^{-1}b(d-ca^{-1}b)^{-1} \\ 0 & (d-ca^{-1}b)^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} aia^{-1} & -aia^{-1}b(d-ca^{-1}b)^{-1} + (a+bi)(d-ca^{-1}b)^{-1} \\ 0 & (d-ca^{-1}b)i(d-ca^{-1}b)^{-1} \end{bmatrix}.$$

Denote $c_1 = 1 - ca^{-1}i + ica^{-1}$, $b_1 = -aia^{-1}b(d-ca^{-1}b)^{-1} + (a+bi)(d-ca^{-1}b)^{-1}$ and $d_1 = (d - ca^{-1}b)$. We can write

$$S^{(3)} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ c_1 & i \end{bmatrix}; \begin{bmatrix} aia^{-1} & b_1 \\ 0 & d_1 i d_1{}^{-1} \end{bmatrix} \right).$$

Note that $c_1 \neq 0$ by the assumption. Let $P_3 = \begin{bmatrix} 1 & 0 \\ 0 & c_1 \end{bmatrix}$ and we have

$$S^{(4)} = P_3 S^{(3)} P_3^{-1} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ 1 & c_1 i c_1^{-1} \end{bmatrix}; \begin{bmatrix} aia^{-1} & b_1 c_1^{-1} \\ 0 & (c_1 d_1) i (c_1 d_1)^{-1} \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ 1 & i_1 \end{bmatrix}; \begin{bmatrix} i_2 & b_1 c_1^{-1} \\ 0 & i_3 \end{bmatrix} \right),$$

where $i_1 = c_1 i c_1^{-1}$, $i_2 = aia^{-1}$ and $i_3 = (c_1 d_1) i (c_1 d_1)^{-1}$ are all equivalent to i. Since the second frontal slice of $S^{(4)}$ is not diagonalizable, we have $i_1 \neq -i$. So there exists $0 \neq t \in \mathbb{R}$ such that

$$\alpha := -(i + t i_2) \neq (i_1 + t i_3) := \beta.$$

Add the second frontal slice to t times the third frontal slice, then $S^{(4)}$ is equivalent to

$$S^{(5)} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ 1 & i_1 \end{bmatrix}; \begin{bmatrix} i + t i_2 & t b_1 c_1^{-1} \\ 1 & i_1 + t i_3 \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} i & 0 \\ 1 & i_1 \end{bmatrix}; \begin{bmatrix} -\alpha & \gamma \\ 1 & \beta \end{bmatrix} \right),$$

where $\gamma = t b_1 c_1^{-1}$. We will show that the third frontal slice of $S^{(5)}$ is diagonalizable, so that $S^{(5)}$ reduces to the situation of Equality 3.2.1. Consider the following quaternion equation

$$x^2 + \alpha x + x\beta - \gamma = 0.$$

Since $\alpha \neq \beta$, according to [31, Theorem 2.3.1] there exists a solution x with $\Re(x) \neq 0$. Therefore, we have

$$\begin{bmatrix} 1 & -x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\alpha & \gamma \\ 1 & \beta \end{bmatrix} \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\alpha - x & 0 \\ 1 & \beta + x \end{bmatrix}.$$

Note that both α and β are pure imaginary numbers and $\Re(x) \neq 0$, we have

$$\Re(-\alpha - x) \neq \Re(\beta + x),$$

which implies that $-\alpha - x$ and $\beta + x$ must be non-equivalent. It follows that $\begin{bmatrix} -\alpha - x & 0 \\ 1 & \beta + x \end{bmatrix}$ is diagonalizable, and thus so is $\begin{bmatrix} -\alpha & \gamma \\ 1 & \beta \end{bmatrix}$ by similarity.

□

Since we can view a $2 \times 2 \times 2$ quaternion tensor as a special case of the $2 \times 2 \times 3$ case, it is clear that a bound of 3 is the best possible bound since there already exist examples of such tensors that have rank 3.

3.3 The $2 \times 3 \times 3$ and $3 \times 2 \times 3$ cases

We wish to bound the rank of a quaternion tensor with size $2 \times 3 \times 3$ (and hence the $3 \times 3 \times 2$ case) using the results of the previous sections. In order to do this, we will first need the following auxiliary lemma.

Lemma 3.3.1. *Let $A, B \in M_n(\mathbb{H})$ be $n \times n$ matrices with entries in $\mathbb{H}$. If A is invertible, then there exists an $x_0 \in \mathbb{H}$ such that $x_0 A + B$ is singular.*

Proof. By [71, Theorem 5.3], every $n \times n$ quaternion matrix has at least one left eigenvalue in $\mathbb{H}$, which means that we can always choose $x_0 \in \mathbb{H}$ such that $x_0 I + BA^{-1}$ is singular. Therefore, $x_0 A + B$ is also singular. □

Theorem 3.3.2. *Let T be a $2 \times 3 \times 3$ or a $3 \times 3 \times 2$ quaternion tensor. Then $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4$.*

Proof. Let

$$T = \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \end{bmatrix}; \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}; \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \right).$$

We can write $T = T_1 + T_2$, where

$$T_1 = \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ b_{11} & b_{12} & b_{13} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \right).$$

Consider the matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}.$$

Notice that $\mathrm{rank}_{\mathbb{H}}(T_1) = \mathrm{rank}_{\mathbb{H}}(A)$ and $\mathrm{rank}_{\mathbb{H}}(T_2) = \mathrm{rank}_{\mathbb{H}}(B)$. If both A and B have rank at most 2, then

$$\mathrm{rank}_{\mathbb{H}}(T) \leqslant \mathrm{rank}_{\mathbb{H}}(T_1) + \mathrm{rank}_{\mathbb{H}}(T_2) \leqslant 2 + 2 = 4.$$

Therefore, without loss of generality, let us assume that $\mathrm{rank}_{\mathbb{H}}(A) = 3$. By Lemma 3.3.1, there exists $x_0 \in \mathbb{H}$ such that

$$C = x_0 A + B = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

is singular. Therefore $\operatorname{rank}_{\mathbb{H}} C) \leqslant 2$ and we can assume that

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \longrightarrow \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & 0 \end{bmatrix}.$$

By rank-preserving row and column operations, we can write

$$\begin{aligned} T &= \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \end{bmatrix}; \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}; \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \right), \\ &\longrightarrow \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ c_{11} & c_{12} & c_{13} \end{bmatrix}; \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}; \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \right), \\ &\longrightarrow \left(\begin{bmatrix} a_{11} & a_{12} & a'_{13} \\ c_{11} & c_{12} & 0 \end{bmatrix}; \begin{bmatrix} a_{21} & a_{22} & a'_{23} \\ c_{21} & c_{22} & 0 \end{bmatrix}; \begin{bmatrix} a_{31} & a_{32} & a'_{33} \\ c_{31} & c_{32} & 0 \end{bmatrix} \right). \end{aligned}$$

If $a'_{13} = a'_{23} = a'_{33} = 0$, then we immediately have $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4$. Otherwise, we may that assume $a'_{13} \neq 0$. By column operations, the tensor can be reduced to

$$\left(\begin{bmatrix} 0 & 0 & a'_{13} \\ c_{11} & c_{12} & 0 \end{bmatrix}; \begin{bmatrix} a'_{21} & a'_{22} & a'_{23} \\ c_{21} & c_{22} & 0 \end{bmatrix}; \begin{bmatrix} a'_{31} & a'_{32} & a'_{33} \\ c_{31} & c_{32} & 0 \end{bmatrix} \right).$$

If $c_{11} = c_{12} = 0$, then T is the sum of a $2 \times 2 \times 3$ tensor and a simple tensor, so we again have $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4$ by Theorem 3.2.3. Therefore, by switching columns if necessary, we can assume that $c_{12} \neq 0$. After performing further column operations we can reduce T to

$$\left(\begin{bmatrix} 0 & 0 & a'_{13} \\ 0 & c_{12} & 0 \end{bmatrix}; \begin{bmatrix} a''_{21} & a'_{22} & a'_{23} \\ c'_{21} & c_{22} & 0 \end{bmatrix}; \begin{bmatrix} a''_{31} & a'_{32} & a'_{33} \\ c'_{31} & c_{32} & 0 \end{bmatrix} \right).$$

If $a''_{31} = a''_{21} = 0$, then T is the sum of 4 simple tensors defined from the 4 nonzero 1×3 vectors in the lateral direction, and we immediately have $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4$. We can therefore assume that $a''_{31}, a''_{21} \neq 0$ (by adding one matrix to the other if only one was nonzero). Now the remaining tensor can be decomposed as the sum of T_1, T_2, T_3 and T_4 defined below, showing that $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4$:

$$T_1 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & c_{12} & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_{22} - c'_{21}(a''_{21})^{-1}a'_{22} & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_{32} - c'_{31}(a''_{31})^{-1}a'_{32} & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} a''_{31} & a''_{31}(a''_{31})^{-1}a'_{32} & 0 \\ c'_{31} & c'_{31}(a''_{31})^{-1}a'_{32} & 0 \end{bmatrix} \right),$$

$$T_3 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} a''_{21} & a''_{21}(a''_{21})^{-1}a'_{22} & 0 \\ c'_{21} & c'_{21}(a''_{21})^{-1}a'_{22} & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_4 = \left(\begin{bmatrix} 0 & 0 & a'_{13} \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & a'_{23} \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & a'_{33} \\ 0 & 0 & 0 \end{bmatrix} \right).$$

Finally, we can apply Proposition 2.2.12 for the $3 \times 3 \times 2$ case. □

In order to show that the bound of 4 is the best possible for the $2 \times 3 \times 3$ and $3 \times 3 \times 2$ cases, we introduce the next theorem.

Theorem 3.3.3. *Let $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, $\mathcal{B} = \{B_j \mid j = 1, 2, \ldots, m+1\}$ be a basis of $R(T)$ (or $C(T)$ respectively), and M be the left (or right respectively) $\mathbb{H}$-module generated by $\mathcal{B}_1 = \mathcal{B} - \{B_1\}$. Then we have*

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}} (B_1 + A) + \dim_{\mathbb{H}} R(T) - 1 \geqslant \dim_{\mathbb{H}} R(T),$$

(or respectively)

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}}(B_1 + A) + \dim_{\mathbb{H}} C(T) - 1 \geqslant \dim_{\mathbb{H}} C(T).$$

Proof. Suppose $\operatorname{rank}_{\mathbb{H}}(T) = r$, then $R(T)$ is contained in a left $\mathbb{H}$-module M generated a set of r simple tensors $\mathcal{S} = \{S_1, S_2, ..., S_r\}$. Since $\mathcal{B}$ is a basis of $R(T)$, we have

$$\mathcal{B} \subseteq {}_{\mathbb{H}}R(T) \leqslant {}_{\mathbb{H}}M.$$

Denote $m = \dim_{\mathbb{H}} R(T) - 1$. By Replacement Theorem, we can select m elements in $\mathcal{B}$ to replace some of those in $\mathcal{S}$ such that the resulting set is equivalent to $\mathcal{S}$. Without loss of generality, we may assume that

$$\mathcal{S}_1 = \{S_j \mid j = 1, 2, ..., r - m\} \cup \{B_j \mid j = 2, ..., m + 1\}.$$

is equivalent to $\mathcal{S}$. Then we can express B_1 as a left $\mathbb{H}$-linear combination of elements in $\mathcal{S}_1$. We have

$$B_1 = \sum_{j=1}^{r-m} c_j S_j + \sum_{j=2}^{m+1} d_j B_j,$$

which implies

$$B_1 - \sum_{j=2}^{m+1} d_j B_j = \sum_{j=1}^{r-m} c_j S_j.$$

Denote $A_1 = -\sum_{j=2}^{m+1} d_j B_j$. Since $B_1, B_2, ..., B_{m+1}$ are left $\mathbb{H}$-linearly independent, we have $B + A_1 \neq 0$. Therefore, we have

$$r - m \geqslant \operatorname{rank}_{\mathbb{H}}(B_1 + A_1) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}}(B_1 + A),$$

which implies

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}}(B_1 + A) + (\dim_{\mathbb{H}} R(T) - 1).$$

Noticing $\min_{A \in M} \operatorname{rank}_{\mathbb{H}}(B_1 + A)$ is at least 1, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}}(B_1 + A) + \dim_{\mathbb{H}} R(T) - 1 \geqslant \dim_{\mathbb{H}} R(T).$$

Proceeding in a similar manner, we also have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}}(B_1 + A) + \dim_{\mathbb{H}} C(T) - 1 \geqslant \dim_{\mathbb{H}} C(T).$$

□

Note that a similar proof still works even if we replace $\mathbb{H}$ with $\mathbb{C}$ or $\mathbb{R}$ in theorem 3.3.3.

Example 3.3.4. Let T be the $2 \times 3 \times 3$ tensor determined by the horizontal slices:

$$R_1(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ R_2(T) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since $R_1(T)$ and $R_2(T)$ are left $\mathbb{H}$-linear independent, we have $\dim_{\mathbb{H}} R(T) = 2$. Denote the left $\mathbb{H}$-module generated by $\{R_2(T)\}$ by M. Then by Theorem 3.3.3, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}}(R_1(T) + A) + \dim_{\mathbb{H}} R(T) - 1 = 3 + 2 - 1 = 4.$$

The proof of the $3 \times 2 \times 3$ case however will require a more delicate argument than the one used in the $2 \times 3 \times 3$ case.

Theorem 3.3.5. *Let T be a $3 \times 2 \times 3$ quaternion tensor. Then* $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4$.

Proof. Let

$$T = (A; B) = \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} ; \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \right).$$

If either A or B is singular (and without loss of generality assume that it's A), then we can use row operations to reduce the tensor to

$$\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \right),$$

which is the sum of a $2 \times 2 \times 3$ tensor (using the first 2 horizontal slices) and a simple tensor defined by the row vector (b_{31}, b_{32}, b_{33}). Then by Theorem 3.2.3, $\text{rank}_{\mathbb{H}}(T) \leqslant 3 + 1 \leqslant 4$. For the same reason, if any horizontal or lateral slice of the tensor does not have maximal rank, then $\text{rank}_{\mathbb{H}}(T) \leqslant 4$. For example, the first lateral slice of T is the matrix

$$\begin{bmatrix} a_{11} & b_{11} \\ a_{21} & b_{21} \\ a_{31} & b_{31} \end{bmatrix}.$$

If it has rank 1, then T can be written as the sum of a $3 \times 2 \times 2$ tensor and a simple tensor, proving that $\text{rank}_{\mathbb{H}}(T) \leqslant 4$. The maximal rank of any horizontal or lateral slice is 2.

Let us therefore assume that both A and B are nonsingular, and that any horizontal and lateral slice has rank 2. This means that there is at least one nonzero entry in every row and column of A and B. It also means that we cannot have more than one row or column of zeros in any horizontal or lateral slice.

Start by performing rank-preserving row operations to simplify the first column.

Note that we cannot have equal first column vectors for A and B since this would contradict the rank assumption on the first lateral slice:

$$T = \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} ; \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \right) \longrightarrow \left(\begin{bmatrix} 1 & a'_{12} & a'_{13} \\ 0 & a'_{22} & a'_{23} \\ 0 & a'_{32} & a'_{33} \end{bmatrix} ; \begin{bmatrix} 0 & b'_{12} & b'_{13} \\ 1 & b'_{22} & b'_{23} \\ 0 & b'_{32} & b'_{33} \end{bmatrix} \right).$$

Similarly let us perform rank-preserving column operations to simplify the third row. We can also make $a'_{12} = 0$ using the first column:

$$\longrightarrow \left(\begin{bmatrix} 1 & a''_{12} & a''_{13} \\ 0 & a''_{22} & a''_{23} \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & b''_{12} & b''_{13} \\ 1 & b''_{22} & b''_{23} \\ 0 & 1 & 0 \end{bmatrix} \right) \longrightarrow \left(\begin{bmatrix} 1 & 0 & a''_{13} \\ 0 & a''_{22} & a''_{23} \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & b''_{12} & b''_{13} \\ 1 & b'''_{22} & b''_{23} \\ 0 & 1 & 0 \end{bmatrix} \right).$$

Next we can use row operations to make $a''_{23} = b''_{12} = 0$:

$$\longrightarrow \left(\begin{bmatrix} 1 & 0 & a''_{13} \\ 0 & a''_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & b''_{12} & b''_{13} \\ 1 & b^{(4)}_{22} & b''_{23} \\ 0 & 1 & 0 \end{bmatrix} \right) \longrightarrow \left(\begin{bmatrix} 1 & 0 & a'''_{13} \\ 0 & a''_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & b''_{13} \\ 1 & b^{(4)}_{22} & b''_{23} \\ 0 & 1 & 0 \end{bmatrix} \right).$$

One final column operation can be used to make $a'''_{13} = 0$. We can also scale the second row since $a''_{22} \neq 0$ (by the rank assumptions):

$$\longrightarrow \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & a''_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & b''_{13} \\ 1 & b^{(4)}_{22} & b'''_{23} \\ 0 & 1 & 0 \end{bmatrix} \right)$$

$$\longrightarrow \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & b''_{13} \\ (a''_{22})^{-1} & (a''_{22})^{-1} b^{(4)}_{22} & (a''_{22})^{-1} b'''_{23} \\ 0 & 1 & 0 \end{bmatrix} \right).$$

Let us relabel the entries by w, x, y and z to write the resulting tensor as

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & w \\ x & y & z \\ 0 & 1 & 0 \end{bmatrix}\right).$$

Notice that if the second matrix (with entries w, x, y and z) were diagonalizable, then the resulting tensor would have rank 3. This means that if we can find a vector $\vec{v} = (e, f, g)$ such that the matrix

$$\begin{bmatrix} 0 & 0 & w \\ x+e & y+f & z+g \\ 0 & 1 & 0 \end{bmatrix}$$

is diagonalizable, then the tensor

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & w \\ x+e & y+f & z+g \\ 0 & 1 & 0 \end{bmatrix}\right)$$

would have rank 3, and T could be written as the sum of a simple tensor (defined by $\vec{v}$) and a rank 3 tensor, showing again that $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4$. Therefore, the problem reduces to choosing quaternions x, y and z such that the matrix

$$M = \begin{bmatrix} 0 & 0 & w \\ x & y & z \\ 0 & 1 & 0 \end{bmatrix}$$

(where $w \neq 0$ by the rank assumptions) is diagonalizable.

Set $w = a + bi + cj + dk$, and let us choose $x = u + vi$ to be a nonzero complex number and $y = z = 0$. By Lemma 2.2.6, the resulting matrix M is diagonalizable if and only if its complex adjoint $\chi(M)$ is. The adjoint of M is the complex matrix

$$\chi(M) = \begin{bmatrix} 0 & 0 & a+bi & 0 & 0 & c+di \\ u+vi & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -c+di & 0 & 0 & a-bi \\ 0 & 0 & 0 & u-vi & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Its characteristic polynomial is

$$p_M(\lambda) = \lambda^6 - 2(au - bv)\lambda^3 + (u^2 + v^2)(a^2 + b^2 + c^2 + d^2).$$

Since $w \neq 0$, at least one of a, b, c or d is nonzero, and so $a^2 + b^2 + c^2 + d^2 \neq 0$. Choose $u, v \neq 0$ in $\mathbb{R}$ such that $au - bv = 0$ (there are infinitely many choices here). Then $C = (u^2 + v^2)(a^2 + b^2 + c^2 + d^2) > 0$ and

$$p_M(\lambda) = \lambda^6 + (u^2 + v^2)(a^2 + b^2 + c^2 + d^2)$$

has distinct roots given by $\sqrt[6]{C}\zeta_6^k i$ for $1 \leqslant k \leqslant 6$, where ζ_6 is a primitive sixth root of unity. Therefore $\chi(M)$ is diagonalizable, completing the proof. $\square$

On the other hand, we can show that the tensor

$$T = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right)$$

has rank 4 by Lemma 2.2.8, which shows that the bound of 4 is the best possible for the $3 \times 2 \times 3$ case.

3.4 The $3 \times 3 \times 3$ case

The proof for the $3 \times 3 \times 3$ case is a lot more complicated. Thus, we first consider some special cases.

Proposition 3.4.1. *Let T be a $3\times3\times3$ quaternion tensor. If any horizontal (frontal or lateral) slice of T has rank no more than 1, then* $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 5$.

Proof. By [45, Theorem 4.2, 4.3], any $2 \times 3 \times 3$ ($3 \times 2 \times 3$ or $3 \times 3 \times 2$) quaternion tensor has rank no more than 4. If T has a horizontal (frontal or lateral) slice with rank no more than 1, then T can be written as the sum of a $2 \times 3 \times 3$ ($3 \times 2 \times 3$ or $3 \times 3 \times 2$) tensor and that horizontal (frontal or lateral) slice, which shows that $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 4 + 1 = 5$. □

Proposition 3.4.2. *Let $T = (A; B; C)$ be a $3 \times 3 \times 3$ quaternion tensor. If A, B and C are all singular, then* $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 5$.

Proof. Since C is singular, by performing rank-preserving row and column operations, we can assume that

$$T = \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} ; \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} ; \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

If either a_{33} or b_{33} is nonzero, then by adding one of the first two frontal slices to the other, we can assume that both a_{33} and b_{33} are nonzero. Since $a_{33}, b_{33} \neq 0$, we can denote

$$k_1 = a_{13}a_{33}^{-1}, k_2 = a_{23}a_{33}^{-1}, k_3 = b_{13}b_{33}^{-1}, k_4 = b_{23}b_{33}^{-1}.$$

Then we can write $T = T_1 + T_2 + T_3$ where:

$$T_1 = \left(\begin{bmatrix} k_1 a_{31} & k_1 a_{32} & k_1 a_{33} \\ k_2 a_{31} & k_2 a_{32} & k_2 a_{33} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} k_3 b_{31} & k_3 b_{32} & k_3 b_{33} \\ k_4 b_{31} & k_4 b_{32} & k_4 b_{33} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_3 = \left(\begin{bmatrix} a_{11} - k_1 a_{31} & a_{12} - k_1 a_{32} & 0 \\ a_{21} - k_2 a_{31} & a_{22} - k_2 a_{32} & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} b_{11} - k_3 b_{31} & b_{12} - k_3 b_{32} & 0 \\ b_{21} - k_4 b_{31} & b_{22} - k_4 b_{32} & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

By [45, Theorem 3.4], any $2 \times 3 \times 2$ quaternion tensor has rank no more than 3, so we have

$$\operatorname{rank}_{\mathbb{H}}(T) \leqslant \operatorname{rank}_{\mathbb{H}}(T_1) + \operatorname{rank}_{\mathbb{H}}(T_2) + \operatorname{rank}_{\mathbb{H}}(T_3) \leqslant 1 + 1 + 3 = 5.$$

Otherwise, we may assume that $a_{33} = b_{33} = 0$. We then have

$$T = \left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix} ; \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & 0 \end{bmatrix} ; \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

By Proposition 3.4.1, we may assume that both the third horizontal slice (with entries a_{31}, a_{32}, b_{31} and b_{32}) and the third lateral slice (with entries a_{13}, a_{23}, b_{13} and b_{23}) have rank greater than 1. By performing rank-preserving row and column operations (switch rows and columns if necessary), we may assume further that $a_{31} = b_{32} = 1$ and $a_{32} = b_{31} = 0$. Similarly, we can also assume $a_{13} = b_{23} = 1$ and $a_{23} = b_{13} = 0$.

Therefore, we have

$$T = \left(\begin{bmatrix} a_{11} & a_{12} & 1 \\ a_{21} & a_{22} & 0 \\ 1 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & b_{12} & 0 \\ b_{21} & b_{22} & 1 \\ 0 & 1 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

Using the 1's in the first two frontal slices, we can use rank-preserving row and column operations to make $a_{12} = a_{21} = b_{12} = b_{21} = 0$. We can therefore assume that

$$T = \left(\begin{bmatrix} a_{11} & 0 & 1 \\ 0 & a_{22} & 0 \\ 1 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 1 \\ 0 & 1 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

Since the frontal slices A and B are singular by the assumption, we must have $a_{22} = b_{11} = 0$. So, we have

$$T = \left(\begin{bmatrix} a_{11} & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_{22} & 1 \\ 0 & 1 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

If $c_{11} = c_{21} = 0$, then it is obvious that T can written as the sum of 5 column vectors and thus have rank no more than 5. Otherwise, there exist $0 \neq x, y \in \mathbb{R}$ such that $xc_{11} + yc_{21} \neq 0$. Define s_1, s_2 by

$$s_1 := xc_{11} + yc_{21},$$

$$s_2 := xc_{12} + yc_{22}.$$

We can check by inspection that

$$c_{11} = -x^{-1}yc_{21} + x^{-1}s_1,$$

$$c_{22} = -y^{-1}xc_{12} + y^{-1}s_2.$$

Since $s_1 = xc_{11} + yc_{21} \neq 0$ and $x, y \neq 0$, we can then write $T = T_1 + T_2 + T_3 + T_4 + T_5$ where:

$$T_1 = \left(\begin{bmatrix} a_{11}+1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_2 = \left(\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} x^{-1}s_1 & 0 & -x^{-1}s_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_3 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} -x^{-1}yc_{21} & c_{12} & x^{-1}s_1 \\ (y^{-1}x)x^{-1}yc_{21} & -(y^{-1}x)c_{12} & -(y^{-1}x)x^{-1}s_1 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_4 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 \cdot s_1^{-1}s_2 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & y^{-1}s_1(s_1^{-1}s_2) & y^{-1}s_1 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$T_5 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_{22} - s_1^{-1}s_2 & 0 \\ 0 & 1 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

Therefore, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \leqslant \textstyle\sum_{i=1}^{5} \operatorname{rank}_{\mathbb{H}} T_i) = 1 + 1 + 1 + 1 + 1 = 5.$$

□

Since the case in Proposition 3.4.2 has been handled, we are left with those $3 \times 3 \times 3$ quaternion tensors with at least one invertible frontal slice to deal with. By performing rank-preserving row and column operations and switching frontal slices if necessary, we may assume further that the tensors we are left with to consider have the identity matrix as their first frontal slices. Namely, they are of the form $T = (I_3; B; C)$. To tackle the remaining cases, we will need to introduce some auxiliary lemmas.

Lemma 3.4.3. *Let $A, B \in M_n(\mathbb{H})$ be $n \times n$ matrices with entries in $\mathbb{H}$. If A is invertible, then there exists $x_0 \in \mathbb{R}$ such that $x_0 A + B$ is also invertible.*

Proof. Since A is invertible, it suffices to show that $x_0 I + BA^{-1}$ is invertible. Consider the $2n \times 2n$ complex adjoint matrix $\chi(BA^{-1})$. Its determinant is a degree $2n$ complex polynomial, which only has finitely many roots in $\mathbb{C}$. Therefore, there exists $x_0 \in \mathbb{R}$ such that $det\,(x_0 I_{2n} + \chi(BA^{-1})) \neq 0$, which implies that $x_0 I_{2n} + \chi(BA^{-1})$ is invertible. Since $x_0 \in \mathbb{R}$ and χ is an $\mathbb{R}$-algebra homomorphism, we have

$$\chi(x_0 I + BA^{-1}) = \chi(x_0 I) + \chi(BA^{-1}) = x_0 I_{2n} + \chi(BA^{-1}).$$

Therefore, $\chi(x_0 I + BA^{-1})$ is invertible, which shows that $x_0 I + BA^{-1}$ is also invertible. □

Lemma 3.4.4. *Let $A \in M_n(\mathbb{H})$ be an $n \times n$ matrix with entries in $\mathbb{H}$. Then there exist a lower-triangular matrix $L \in M_n(\mathbb{H})$ with unit pivots, an upper-triangular matrix $U \in M_n(\mathbb{H})$ with unit pivots, an invertible diagonal matrix $D \in M_n(\mathbb{H})$ and a permutation matrix $P \in M_n(\mathbb{R})$ such that $A = LPDU$.*

Proof. This is a direct result of [16, Proposition 1.3]. □

Lemma 3.4.5. *Let $T = (I_n; B; C)$ be a $n \times 3 \times n$ quaternion tensor. Then it can be reduced to the form*

$$S = \left(I_n; B_L; PC_U P^{-1}\right),$$

where B_L is an $n \times n$ invertible lower triangular quaternion matrix, C_U is an $n \times n$ invertible upper triangular quaternion matrix, and P is an $n \times n$ permutation matrix.

Proof. By Lemma 3.4.3, we can make B invertible by replacing B with $x_0 I + B$ for some $x_0 \in \mathbb{R}$. Similarly, C can also be assume to be invertible. Then we consider the Jordan Canonical Forms (see [49, 71]) of B and C, and denote them by J_B and J_C respectively. Let P_B and P_C be invertible matrices such that

$$B = P_B^{-1} J_B^* P_B$$

and

$$C = P_C^{-1} J_C P_C.$$

By Lemma 3.4.4, there exist a lower-triangular matrix $L \in M_n(\mathbb{H})$ with unit pivots, an upper-triangular matrix $U \in M_n(\mathbb{H})$ with unit pivots, an invertible diagonal matrix $D \in M_n(\mathbb{H})$ and a permutation matrix $P \in M_n(\mathbb{R})$ such that

$$P_B P_C^{-1} = LPDU$$

Since $P_B P_C^{-1}$ is invertible, so are L and U. We then have

$$L^{-1} P_B P_C^{-1} = PDU.$$

and thus the tensor T can be reduced to

$$\begin{aligned}
S :&= (L^{-1}P_B)T(L^{-1}P_B)^{-1} \\
&= (L^{-1}P_B)\,(I_3;\ B;\ C)\,(L^{-1}P_B)^{-1} \\
&= (L^{-1}P_B)\left(I_3\ ;\ P_B^{-1}J_B^*P_B\ ;\ P_C^{-1}J_CP_C\right)(L^{-1}P_B)^{-1} \\
&= \left(I_3\ ;\ L^{-1}J_B^*L\ ;\ \left(L^{-1}P_BP_C^{-1}\right)J_C\left(L^{-1}P_BP_C^{-1}\right)^{-1}\right) \\
&= \left(I_3\ ;\ L^{-1}J_B^*L\ ;\ (PDU)\,J_C\,(PDU)^{-1}\right) \\
&= \left(I_3; L^{-1}J_B^*L; P\left(DUJ_CU^{-1}D^{-1}\right)P^{-1}\right).
\end{aligned}$$

Define $B_L = L^{-1}J_B^*L$ and $C_U = DUJ_CU^{-1}D^{-1}$. Since L and J_B^* are lower-triangular, so is B_L. Similarly, since U and J_C are upper-triangular and D is diagonal, we see that $DUJ_CU^{-1}D^{-1}$ is also upper-triangular. □

Now, we are ready to tackle the general cases.

Proposition 3.4.6. *Let $T = (I_3; B; C)$ be a $3 \times 3 \times 3$ quaternion tensor. Then $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 5$.*

Proof. By Lemma 3.4.5, T can be reduced to

$$S = \left(I_3; B_L; PC_UP^{-1}\right) = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}; P \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} P^{-1} \right),$$

where P is a 3×3 permutation matrix and $b_{11}, b_{22}, b_{33}, c_{11}, c_{22}, c_{33} \neq 0$. Then the rest of the proof boils down to dealing with all possible forms of PC_UP^{-1}. Because P is a 3×3 permutation matrix, it has 6 possible patterns. Namely, we have $P \in \{P_1, P_2, P_3, P_4, P_5, P_6\}$ where:

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, P_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$P_4 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, P_5 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, P_6 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

In the following, we will show that $\operatorname{rank}_{\mathbb{H}}(S) \leqslant 5$ in each case.

Case 1: $P = P_1$. We have

$$PC_U P^{-1} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix},$$

and thus

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \right).$$

Subcase 1-1: $b_{31}, c_{13} \neq 0$. If so, we can immediately decompose S into the sum of simple tensors $S = S_1 + S_2 + S_3 + S_4 + S_5$ where:

$$S_1 := \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_2 := \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_{22} - b_{21}b_{31}^{-1}b_{32} - c_{23}c_{13}^{-1}c_{12} & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_3 := \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b_{33} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c_{33} \end{bmatrix} \right),$$

$$S_4 := \left(\begin{bmatrix} 0 & 0 & 0 \\ (b_{21}b_{31}^{-1})b_{31} & (b_{21}b_{31}^{-1})b_{32} & 0 \\ b_{31} & b_{32} & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_5 := \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & c_{13}(c_{13}^{-1}c_{12}) & c_{13} \\ 0 & c_{23}(c_{13}^{-1}c_{12}) & c_{23} \\ 0 & 0 & 0 \end{bmatrix} \right).$$

Subcase 1-2: $b_{33} = c_{13} = 0$. We first choose an $x \in \mathbb{R}$ such that

$$(b_{11} + x)(c_{11} + x) \neq 0$$

and

$$\frac{\|b_{21}c_{12}\|}{\|(b_{11} + x)(c_{11} + x)\|} \neq 1,$$

which implies

$$1 + b_{21}(b_{11} + x)^{-1}(c_{11} + x)^{-1}c_{12} \neq 0.$$

Therefore, by adding x times the first frontal slice to the second and third, we may just assume that

$$1 + b_{21}b_{11}^{-1}c_{11}^{-1}c_{12} \neq 0.$$

We then perform rank-preserving row and column operations to make $b_{21} = c_{12} = 0$:

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & 0 \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \right)$$

$$\longrightarrow \left(\begin{bmatrix} 1 & 0 & 0 \\ -b_{21}b_{11}^{-1} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & 0 \\ * & * & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \right)$$

$$\longrightarrow \left(\begin{bmatrix} 1 & -c_{11}^{-1}c_{12} & 0 \\ -b_{21}b_{11}^{-1} & 1 + b_{21}b_{11}^{-1}c_{11}^{-1}c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} b_{11} & * & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} ; \begin{bmatrix} c_{11} & 0 & 0 \\ * & * & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \right).$$

Let us denote $a_{22} = 1 + b_{21}b_{11}^{-1}c_{11}^{-1}c_{12}$ and relabel the resulting tensor as

$$S^{(1)} := \left(\begin{bmatrix} 1 & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} b_{11} & b_{12} & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} ; \begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c'_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \right).$$

Since $a_{22} \neq 0$, we can then write $S^{(1)} = S_1 + S_2 + S_3 + S_4 + S_5$ where:

$$S_1 := \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} 0 & b_{12} & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & 0 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_2 := \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ c_{21} & c'_{22} & c_{23} \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_3 := \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b_{33} \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c_{33} \end{bmatrix} \right),$$

$$S_4 := \left(\begin{bmatrix} 1 - (a_{12}a_{22}^{-1})a_{21} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} c_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_5 := \left(\begin{bmatrix} (a_{12}a_{22}^{-1})a_{21} & (a_{12}a_{22}^{-1})a_{22} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).$$

Subcase 1-3: $b_{31} = 0$ and $c_{13} \neq 0$ or $b_{31} \neq 0$ and $c_{13} = 0$. If so, we may just assume that $b_{31} = 0$ and $c_{13} \neq 0$, since otherwise we can interchange the second and the third frontal slices and work with the transpose of matrices, or alternatively use the similarity transformation on S via

$$P_6 \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} P_6^{-1} = \begin{bmatrix} b_{33} & b_{32} & b_{31} \\ 0 & b_{22} & b_{21} \\ 0 & 0 & b_{11} \end{bmatrix}$$

and

$$P_6 \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} P_6^{-1} = \begin{bmatrix} c_{33} & 0 & 0 \\ c_{23} & c_{22} & 0 \\ c_{13} & c_{12} & c_{11} \end{bmatrix}$$

to reduce it to the case of $b_{33} = 0$ and $c_{13} \neq 0$. So, in any case, we can have

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \right).$$

By Proposition 3.4.1, we can assume that both the third horizontal slice (with entries $1, b_{32}, b_{33}$ and c_{33}) and the first lateral slice (with entries $1, b_{11}, b_{21}$ and c_{11}) have rank greater than 1, which implies $b_{21}, b_{32} \neq 0$. By performing the similarity transformation via

$$\begin{bmatrix} b_{21} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & b_{32}^{-1} \end{bmatrix} \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} b_{21} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & b_{32}^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} b_{21}b_{11}b_{21}^{-1} & 0 & 0 \\ 1 & b_{22} & 0 \\ 0 & 1 & b_{32}^{-1}b_{33}b_{32} \end{bmatrix},$$

we may assume further that $b_{21} = b_{32} = 1$. Therefore, we have

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ 1 & b_{22} & 0 \\ 0 & 1 & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \right).$$

We then decompose S into the sum of simple tensors according to the following situations:

1-3-1: If $c_{12} \neq 0$, define

$$s = c_{22} - b_{11}^{-1}c_{11}b_{11} - c_{23}c_{13}^{-1}c_{12}.$$

Since $b_{11}, b_{22}, c_{12}, c_{13} \neq 0$, we can then write $S = S_1 + S_2 + S_3 + S_4 + S_5$ where:

$$S_1 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b_{33} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c_{33} + b_{22}^{-1}sc_{12}^{-1}c_{13} \end{bmatrix} \right),$$

$$S_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & c_{12} & c_{12}(c_{12}^{-1}c_{13}) \\ 0 & c_{23}c_{13}^{-1}c_{12} & c_{23}c_{13}^{-1}c_{12}(c_{12}^{-1}c_{13}) \\ 0 & -b_{22}^{-1}s & -b_{22}^{-1}s(c_{12}^{-1}c_{13}) \end{bmatrix} \right),$$

$$S_3 = \left(\begin{bmatrix} 0 & 0 & 0 \\ -b_{11}^{-1} & (-b_{11}^{-1})(-b_{11}) & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ -b_{11}^{-1}c_{11} & -(b_{11}^{-1}c_{11})(-b_{11}) & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_4 = \left(\begin{bmatrix} 1 & 0 & 0 \\ b_{11}^{-1} \cdot 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{11}^{-1}b_{11} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & 0 & 0 \\ b_{11}^{-1}c_{11} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_5 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{22}^{-1}b_{22} & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & s & 0 \\ 0 & b_{22}^{-1}s & 0 \end{bmatrix} \right).$$

1-3-2: If $c_{23} \neq 0$, this reduces to **1-3-1** by transposing all the frontal slices across the anti-diagonal (use the conjugate transpose and the similar transformation induced by P_6).

1-3-3: If $c_{12} = c_{23} = 0$, then we choose an $x \in \mathbb{R}\backslash\{-1, 0, 1\}$ such that

$$s := x^2 + (b_{22} - b_{33})x + 1 \neq 0.$$

Since $s, x, x^2 - 1 \neq 0$, we can then write $S = S_1 + S_2 + S_3 + S_4 + S_5$ where:

$$S_1 = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} - c_{13} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} c_{13} & 0 & c_{13} \\ xs^{-1}(c_{33} - c_{22}) & 0 & xs^{-1}(c_{33} - c_{22}) \\ 0 & 0 & 0 \end{bmatrix} \right).$$

$$S_3 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{-1}{x^2-1} & \frac{-x}{x^2-1} \\ 0 & \frac{x}{x^2-1} & \frac{x^2}{x^2-1} \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1-b_{33}x}{x(x^2-1)} & \frac{1-b_{33}x}{x^2-1} \\ 0 & \frac{b_{33}x-1}{x^2-1} & \frac{(b_{33}x-1)x}{x^2-1} \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{-c_{33}}{x^2-1} & \frac{-c_{33}x}{x^2-1} \\ 0 & \frac{c_{33}x}{x^2-1} & \frac{c_{33}x^2}{x^2-1} \end{bmatrix} \right),$$

$$S_4 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{x^2}{x^2-1} & \frac{x}{x^2-1} \\ 0 & \frac{-x}{x^2-1} & \frac{-1}{x^2-1} \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{(b_{33}-x)x^2}{x^2-1} & \frac{(b_{33}-x)x}{x^2-1} \\ 0 & \frac{(x-b_{33})x}{x^2-1} & \frac{x-b_{33}}{x^2-1} \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{c_{33}x^2}{x^2-1} & \frac{c_{33}x}{x^2-1} \\ 0 & \frac{-c_{33}x}{x^2-1} & \frac{-c_{33}}{x^2-1} \end{bmatrix} \right),$$

$$S_5 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ 1 & \frac{s}{x} & 1 \\ 0 & 0 & 0 \end{bmatrix} ; \begin{bmatrix} 0 & 0 & 0 \\ xs^{-1}(c_{22}-c_{33}) & c_{22}-c_{33} & xs^{-1}(c_{22}-c_{33}) \\ 0 & 0 & 0 \end{bmatrix} \right).$$

Based on all the discussion above, if $P = P_1$, we have

$$\operatorname{rank}_{\mathbb{H}}(S) \leqslant \sum_{i=1}^{5} \operatorname{rank}_{\mathbb{H}}(S_i) = 1 + 1 + 1 + 1 + 1 = 5.$$

Case 2: $P = P_2$. We have

$$PC_UP^{-1} = \begin{bmatrix} c_{11} & c_{13} & c_{12} \\ 0 & c_{33} & 0 \\ 0 & c_{23} & c_{22} \end{bmatrix},$$

and thus

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} ; \begin{bmatrix} c_{11} & c_{13} & c_{12} \\ 0 & c_{33} & 0 \\ 0 & c_{23} & c_{22} \end{bmatrix} \right).$$

Subcase 2-1: $b_{31}, c_{13} \neq 0$. we can immediately decompose S into the sum of simple tensors $S = S_1 + S_2 + S_3 + S_4 + S_5$ where:

$$S_1 = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_{22} - b_{21}b_{31}^{-1}b_{32} & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_{33} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_3 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b_{33} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c_{22} - c_{23}c_{13}^{-1}c_{12} \end{bmatrix} \right),$$

$$S_4 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ b_{21} & b_{21}(b_{31}^{-1}b_{32}) & 0 \\ b_{31} & b_{31}(b_{31}^{-1}b_{32}) & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_5 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & c_{13} & c_{12} \\ 0 & 0 & 0 \\ 0 & (c_{23}c_{13}^{-1})c_{13} & (c_{23}c_{13}^{-1})c_{12} \end{bmatrix} \right).$$

Subcase 2-2: $b_{31}, c_{13} = 0$. If so, then

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & 0 & c_{12} \\ 0 & c_{33} & 0 \\ 0 & c_{23} & c_{22} \end{bmatrix} \right).$$

We can then write $S = S_1 + S_2 + S_3 + S_4 + S_5$ where:

$$S_1 = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & c_{12} \\ 0 & 0 & 0 \\ 0 & 0 & -c_{23} - c_{33} + c_{22} \end{bmatrix} \right),$$

$$S_3 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ b_{21} & -b_{33} + b_{23} + b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_4 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_{33} - b_{23} & 0 \\ 0 & -b_{33} + b_{23} & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_{33} & 0 \\ 0 & -c_{33} & 0 \end{bmatrix} \right),$$

$$S_5 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & b_{33} & b_{33} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & c_{23} + c_{33} & c_{23} + c_{33} \end{bmatrix} \right).$$

Subcase 2-3: $b_{31} = 0$ and $c_{13} \neq 0$. If so, then

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{13} & c_{12} \\ 0 & c_{33} & 0 \\ 0 & c_{23} & c_{22} \end{bmatrix} \right).$$

By Proposition 3.4.1, we can assume that first lateral slice (with entries $1, b_{11}, b_{21}$

and c_{11}) has rank greater than 1, which imples $b_{21} \neq 0$. Since $b_{11} \neq 0$, we can assume $b_{11} = b_{21}$ by performing the following similarity transformation:

$$\begin{bmatrix} b_{21}b_{11}^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} b_{21}b_{11}^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} b_{21}b_{11}b_{21}^{-1} & 0 & 0 \\ b_{21}b_{11}b_{21}^{-1} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix}.$$

Therefore, we have

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{11} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{11} & c_{13} & c_{12} \\ 0 & c_{33} & 0 \\ 0 & c_{23} & c_{22} \end{bmatrix} \right).$$

Define

$$s := c_{23} - b_{32}b_{22}^{-1}(c_{33} - c_{11}).$$

Since $c_{13}, b_{22} \neq 0$, we can then write $S = S_1 + S_2 + S_3 + S_4 + S_5$ where:

$$S_1 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b_{33} \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c_{22} - sc_{13}^{-1}c_{12} \end{bmatrix} \right),$$

$$S_2 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & c_{13} & c_{13}(c_{13}^{-1}c_{12}) \\ 0 & 0 & 0 \\ 0 & s & s(c_{13}^{-1}c_{12}) \end{bmatrix} \right),$$

$$S_3 = \left(\begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ -c_{11} & c_{11} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_4 = \left(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{11} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} c_{11} & 0 & 0 \\ c_{11} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right),$$

$$S_5 = \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & (b_{32}b_{22}^{-1})b_{22} & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_{33} - c_{11} & 0 \\ 0 & (b_{32}b_{22}^{-1})(c_{33} - c_{11}) & 0 \end{bmatrix} \right).$$

Subcase 2-4: $b_{31} \neq 0$ and $c_{13} = 0$ If so, we can perform the following rank-preserving operations:

$$\begin{aligned} S &= \left(I_3; B_L; P_2 C_U P_2^{-1}\right) \\ &\longrightarrow (P_2 P_6)^{-1} \left(I_3; B_L; P_2 C_U P_2^{-1}\right) (P_2 P_6) \\ &\longrightarrow \left(I_3; (P_2 P_6)^{-1} B_L (P_2 P_6); P_6 C_U P_6\right) \\ &\longrightarrow \left(I_3; P_6 C_U P_6^{-1}; (P_2 P_6)^{-1} B_L (P_2 P_6);\right) \end{aligned}$$

$$= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; P_6 \begin{bmatrix} c_{11} & c_{12} & 0 \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} P_6^{-1}; P_4^{-1} \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} P_4 \right)$$

$$= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} c_{33} & 0 & 0 \\ c_{23} & c_{22} & 0 \\ 0 & c_{12} & c_{11} \end{bmatrix}; \begin{bmatrix} b_{22} & 0 & b_{21} \\ b_{32} & b_{22} & b_{31} \\ 0 & 0 & b_{11} \end{bmatrix} \right).$$

By transposing all the frontal slices across the anti-diagonal (use the conjugate transpose and the similar transformation induced by P_6), this reduces to **Subcase 2-3**.

Based on all the discussion above, if $P = P_2$, we have

$$\operatorname{rank}_{\mathbb{H}}(S) \leqslant \sum_{i=1}^{5} \operatorname{rank}_{\mathbb{H}}(S_i) = 1 + 1 + 1 + 1 + 1 = 5.$$

Case 3: $P = P_3$. We have

$$PC_UP^{-1} = \begin{bmatrix} c_{22} & 0 & c_{23} \\ c_{12} & c_{11} & c_{13} \\ 0 & 0 & c_{33} \end{bmatrix},$$

and thus

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{22} & 0 & c_{23} \\ c_{12} & c_{11} & c_{13} \\ 0 & 0 & c_{33} \end{bmatrix} \right).$$

By transposing all the frontal slices across the anti-diagonal (use the conjugate transpose and the similar transformation induced by P_6), this reduces to **Case 2**.

Case 4: $P = P_4$. We have

$$PC_UP^{-1} = \begin{bmatrix} c_{22} & c_{23} & 0 \\ 0 & c_{33} & c_{13} \\ c_{12} & c_{13} & c_{11} \end{bmatrix},$$

and thus

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{22} & c_{23} & 0 \\ 0 & c_{33} & 0 \\ c_{12} & c_{13} & c_{11} \end{bmatrix} \right).$$

Since the third lateral slice of S has rank no more than 1, by Lemma 3.4.1, we have $\mathrm{rank}_{\mathbb{H}}(S) \leqslant 5$.

Case 5: $P = P_5$. We have

$$PC_UP^{-1} = \begin{bmatrix} c_{33} & 0 & 0 \\ c_{13} & c_{11} & c_{12} \\ c_{23} & 0 & c_{22} \end{bmatrix},$$

and thus

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{33} & 0 & 0 \\ c_{13} & c_{11} & c_{12} \\ c_{23} & 0 & c_{22} \end{bmatrix} \right).$$

Since the first horizontal slice of S has rank no more than 1, by Lemma 3.4.1, we have $\mathrm{rank}_{\mathbb{H}}(S) \leqslant 5$.

Case 6: $P = P_6$. We have

$$PC_UP^{-1} = \begin{bmatrix} c_{33} & 0 & 0 \\ c_{32} & c_{22} & 0 \\ c_{31} & c_{12} & c_{11} \end{bmatrix},$$

and thus

$$S = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}; \begin{bmatrix} c_{33} & 0 & 0 \\ c_{32} & c_{22} & 0 \\ c_{31} & c_{12} & c_{11} \end{bmatrix} \right).$$

Since the first horizontal slice of S has rank no more than 1, by Lemma 3.4.1, we have $\mathrm{rank}_{\mathbb{H}}(S) \leqslant 5$. □

By Proposition 3.4.2 and Proposition 3.4.6 we have $\mathrm{max.rank}_{\mathbb{H}}(3,3,3) \leqslant 5$. Next, we show that $\mathrm{max.rank}_{\mathbb{H}}(3,3,3) \geqslant 5$. First we use Theorem 3.3.3 to obtain a lower bound for $n \times n \times n$ quaternion tensors, and the we apply the this bound for $n = 3$.

Proposition 3.4.7. *For any* $n \in \mathbb{N}^+$, *we have* $\mathrm{max.rank}_{\mathbb{H}}(n,n,n) \geqslant 2n - 1$.

Proof. Let E_{ij} denote the $n \times n$ matrix that has 1 at the (i, j) entry and 0 elsewhere. Then

$$\mathcal{S} = \{I_n\} \cup \{E_{ij} \mid 1 \leqslant j < i \leqslant n\}$$

is a left $\mathbb{H}$-linearly independent set. Let T be the tensor uniquely determined by the horizontal slices: $R_1(T) = I_n$ and $R_2(T), ..., R_n(T)$ being any $n-1$ distinct elements in $\mathcal{S}$. Denote the left $\mathbb{H}$-module generated by $\{R_2(T), R_3(T), ..., R_{N_1}(T)\}$ by M. Then we have $\dim_{\mathbb{H}} R(T) = n$. By Theorem 3.3.3, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}}(R_1(T) + A) + \dim_{\mathbb{H}} R(T) - 1 = n + n - 1 = 2n - 1.$$

□

Applying Proposition 3.4.7 for n=3, we have $\text{max.rank}_{\mathbb{H}}(3, 3, 3) \geqslant 5$. Therefore, we now have

Theorem 3.4.8. $\text{max.rank}_{\mathbb{H}}(3, 3, 3) = 5$.

3.5 The $2 \times 4 \times 2$ case (a counterexample)

It is well known that tensor rank is dependent on the underlying field. For example, the maximal rank of any real $2 \times 2 \times 2 \times 2$ tensor is 5, while the maximal rank for a complex tensor of the same size is 4 (these bounds were shown earlier in [38] and [6] respectively). In this section, we give an example of a complex tensor that has a strictly smaller rank over the quaternions than the complex field.

Theorem 3.5.1. *Any complex $2 \times n \times 2$ tensor has rank no more than 3 over the quaternions.*

Proof. Let $T \in \mathbb{C}^{2 \times n \times 2}$. It suffices to show that there exist $\vec{a}_i, \vec{c}_i \in \mathbb{H}^2$, $i = 1, 2, 3$, such that the standard $\mathbb{R}$-basis of $M_2(\mathbb{C})$ is contained in

$$V := \operatorname{span}_{\mathbb{R}} \{\vec{a}_i \otimes b \otimes \vec{c}_i \mid b \in \mathbb{H}, i = 1, 2, 3\}.$$

We can check that for any $x_{11}, x_{12}, x_{21}, x_{22}, y_{11}, y_{12}, y_{21}, y_{22} \in \mathbb{R}$, we have

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ j & \dfrac{-j-\sqrt{3}k}{2} & \dfrac{-j+\sqrt{3}k}{2} \end{bmatrix} \begin{bmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{bmatrix} \begin{bmatrix} 1 & j \\ 1 & \dfrac{-j+\sqrt{3}k}{2} \\ 1 & \dfrac{-j-\sqrt{3}k}{2} \end{bmatrix},$$

where

$$\begin{aligned} b_1 &= \frac{x_{11} - x_{22} - (x_{12} + x_{21})j}{3}, \\ b_2 &= \frac{2x_{11} + x_{22} - \sqrt{3}x_{22}i + (x_{12} + x_{21})j - \sqrt{3}(x_{12} - x_{21})k}{6}, \\ b_3 &= \frac{2x_{11} + x_{22} + \sqrt{3}x_{22}i + (x_{12} + x_{21})j + \sqrt{3}(x_{12} - x_{21})k}{6}. \end{aligned}$$

and

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} i = \begin{bmatrix} 1 & 1 & 1 \\ j & \dfrac{-j-\sqrt{3}k}{2} & \dfrac{-j+\sqrt{3}k}{2} \end{bmatrix} \begin{bmatrix} b_4 & 0 & 0 \\ 0 & b_5 & 0 \\ 0 & 0 & b_6 \end{bmatrix} \begin{bmatrix} 1 & j \\ 1 & \dfrac{-j+\sqrt{3}k}{2} \\ 1 & \dfrac{-j-\sqrt{3}k}{2} \end{bmatrix},$$

where

$$\begin{aligned} b_4 &= \frac{(y_{11} + y_{22})i - (y_{12} - y_{21})k}{3}, \\ b_5 &= \frac{-\sqrt{3}y_{22} + (2y_{11} - y_{22})i + \sqrt{3}(y_{12} + y_{21})j + (y_{12} - y_{21})k}{6}, \\ b_6 &= \frac{\sqrt{3}y_{22} + (2y_{11} - y_{22})i - \sqrt{3}(y_{12} + y_{21})j + (y_{12} - y_{21})k}{6}. \end{aligned}$$

Set

$$\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ j & \dfrac{-j-\sqrt{3}k}{2} & \dfrac{-j+\sqrt{3}k}{2} \end{bmatrix},$$

and

$$\begin{bmatrix} \vec{c_1}^T \\ \vec{c_2}^T \\ \vec{c_3}^T \end{bmatrix} = \begin{bmatrix} 1 & j \\ 1 & \dfrac{-j+\sqrt{3}k}{2} \\ 1 & \dfrac{-j-\sqrt{3}k}{2} \end{bmatrix}.$$

Then, any $T \in \mathbb{H}^{2\times n\times 2}$ can be written as $T = \sum_{i=1}^{3} \vec{a}_i \otimes \vec{b}_i \otimes \vec{c}_i$ for some $\vec{b}_i \in \mathbb{H}^n$, $i = 1, 2, 3$, which shows that $\mathrm{rank}_{\mathbb{H}}(T) \leqslant 3$. □

Corollary 3.5.2. *If* ${}_{\mathbb{R}}\mathbb{W} \leqslant {}_{\mathbb{R}}\mathbb{H}$ *is a* 2*-dimensional* $\mathbb{R}$*-subspace, then for any* $T \in \mathbb{H}^{2\times n\times 2}$ *with entries in* $\mathbb{W}$*, we have* $\mathrm{rank}_{\mathbb{H}}(T) \leqslant 3$*.*

Proof. Denote an $\mathbb{R}$-basis of $\mathbb{W}$ by $\{x_1, x_2\} \subset \mathbb{H}$. Set $u = x_1^{-1}x_2$. Then $\dfrac{u - \Re(u)}{||u - \Re(u)||}$ is a unit quaternion and thus equivalent to i, i.e., there exists $p \in \mathbb{H}$ such that

$$i = p^{-1}\frac{u - \Re(u)}{||u - \Re(u)||}p.$$

Then for any $x = r_1x_1 + r_2x_2 \in \mathbb{W}$, where $r_1, r_2 \in \mathbb{R}$, we have

$$\begin{aligned} p^{-1}x_1^{-1}xp &= p^{-1}x_1^{-1}(r_1x_1 + r_2x_2)p \\ &= r_1 + r_2p^{-1}x_1^{-1}x_2p \\ &= r_1 + r_2p^{-1}up \\ &= r_1 + r_2p^{-1}(\Re(u) + u - \Re(u))p \\ &= r_1 + r_2\Re(u) + r_2p^{-1}(u - \Re(u))p \\ &= r_1 + r_2\Re(u) + r_2||u - \Re(u)||i \in \mathbb{C}. \end{aligned}$$

Let f be a $\mathbb{R}$-module isomorphism induced by

$$f : \mathbb{H}^{2\times n\times 2} \longrightarrow \mathbb{H}^{2\times n\times 2},$$

$$T \mapsto p^{-1}x_1^{-1}Tp.$$

By the previous discussion, for any $T \in \mathbb{H}^{2\times n\times 2}$ with all entries in $\mathbb{W}$, we have $f(T) \in \mathbb{C}^{2\times n\times 2}$. Hence, by Theorem 3.5.1 we have

$$f(T) = \sum_{i=1}^{3} \vec{a}_i \otimes \vec{b}_i \otimes \vec{c}_i$$

for some $\vec{a}_i, \vec{c}_i \in \mathbb{C}^2$ and $\vec{b}_i \in \mathbb{C}^n$, where $i = 1, 2, 3$. Thus, we have

$$T = f^{-1}(f(T)) = f^{-1}\left(\sum_{i=1}^{3} \vec{a}_i \otimes \vec{b}_i \otimes \vec{c}_i\right) = \sum_{i=1}^{3} f^{-1}(\vec{a}_i \otimes \vec{b}_i \otimes \vec{c}_i)$$

implies that $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 3$. □

Example 3.5.3. The following tensor

$$T = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

has rank 4 over the complex field while it only has rank 3 over the quaternions.

Proof. Since all 4 frontal slices of T are $\mathbb{C}$-linearly independent, by [39, Corollary 1], we have $\operatorname{rank}_{\mathbb{C}}(T) \geqslant 1 + 4 - 1 = 4$. However, by Theorem 3.5.1 we have that $\operatorname{rank}_{\mathbb{H}}(T) \leqslant 3$. Therefore, T has different ranks over the complex field $\mathbb{C}$ and the real quaternion algebra $\mathbb{H}$. □

Chapter 4

Higher Quaternionic Tensors

In this chapter, we discuss some general properties of quaternionic tensors with higher orders.

4.1 Definitions and Notations

Let's first fix some notations:

1. For $k, p \in \mathbb{N}^+$ with $k \leqslant p$ and $N_1, N_2, ..., N_k \in \mathbb{N}^+$, denote

   $$[[\mathcal{I}_k \leqslant p]] = \left\{(i_1, i_2, ..., i_k) \in \mathbb{N}^k \mid 1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant p\right\},$$

   and

   $$[[N_1, N_2, .., N_k]] := \left\{(i_1, i_2, ..., i_k) \in \mathbb{N}^k \mid 1 \leqslant i_j \leqslant N_j,\ j = 1, 2, ..., k\right\}.$$

2. For any matrix $A \in M_{N_1 \times N_2}(\mathbb{H})$, denote $A^{\leqslant m}$ to be the sub-matrix consisting the first m rows of A and ${}^{m<}A$ to be the sub-matrix consisting the last $N_1 - m$ rows of A, i.e.,

   $$A^{\leqslant m} := \begin{bmatrix} I_m & O_{m \times (N_1 - m)} \end{bmatrix} A, \quad {}^{m<}A := \begin{bmatrix} O_{(N_1 - m) \times m} & I_{N_1 - m} \end{bmatrix} A.$$

Similarly, we denote $A_{\leqslant m}$ to be the sub-matrix consisting the first m columns of A and ${}_{m<}A$ to be the sub-matrix consisting the last $N_2 - m$ columns of A, i.e.,

$$A_{\leqslant m} := A \begin{bmatrix} I_m \\ O_{(N_2-m)\times m} \end{bmatrix}, \quad {}_{m<}A := A \begin{bmatrix} O_{m\times(N_2-m)} \\ I_{N_2-m} \end{bmatrix}.$$

For integers $i_1, \ldots, i_r$ and $j_1, \ldots, j_s$ with $1 \leqslant i_1 < \cdots < i_r \leqslant N_1$ and $1 \leqslant j_1 < \cdots < j_s \leqslant N_2$, we denote the $r \times s$ matrix consisting of i_1-th, i_2-th $, \ldots, i_r$-th rows and j_1-th, j_2-th $, \ldots, j_s$-th columns of A by $A^{=\{i_1,\ldots,i_r\}}_{=\{j_1,\ldots,j_s\}}$.

3. For any $n \in \mathbb{N}^+$, denote $\{\vec{e}_{n,i} \mid i = 1, 2, ..., n\}$ to be the set of standard basis of $\mathbb{H}^{n\times 1}$.

Before our formal discussion, let me introduce the definition of tensor matrices for sake of convenience. A similar definition for block tensors can be found in [48].

Definition 4.1.1. Let $M_{m\times n}(\mathbb{H}^{N_1\times\cdots\times N_p})$, simply $M_n(\mathbb{H}^{N_1\times\cdots\times N_p})$ when $m = n$, denote the collection of all $m \times n$ matrices with whose entries are $N_1 \times \cdots \times N_p$ quaternionic tensors. An $m \times n$ matrix A with $N_1 \times \cdots \times N_p$ quaternion tensor entries is denoted by $A = [A_{ij}]_{\otimes}$ or

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}_{\otimes},$$

where $A_{ij} \in \mathbb{H}^{N_1\times\cdots\times N_p}$ for $i = 1, 2, ..., m$ and $j = 1, 2, ..., n$. A is called a **tensor matrix**. The transpose of $A = [A_{ij}]_{\otimes}$ is defined as $A^T = [A_{ji}]_{\otimes}$, The multiplication of two matrices with tensor entries is defined similarly to the usual multiplication of matrices. That is, if A is an $m \times n$ matrix with $N_1 \times \cdots \times N_p$ tensor entries and B is an $n \times k$ matrix with $M_1 \times \cdots \times M_q$ tensor entries, then the matrix product $C =$

$[C_{ij}]_{\otimes} = A \star B$ is defined to be the $m \times k$ matrix with $N_1 \times \cdots \times N_p \times M_1 \times \cdots \times M_q$ tensor entries

$$C_{ij} = \sum_{l=1}^{n} A_{il} \otimes B_{lj},$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, l$.

Because the outer product $\otimes$ is associative and distributive over addition, we can see that for $A \in M_{m\times n}(\mathbb{H}^{N_1\times\cdots\times N_p})$, $B \in M_{n\times k}(\mathbb{H}^{M_1\times\cdots\times M_q})$, and $C \in M_{k\times l}(\mathbb{H}^{L_1\times\cdots\times L_s})$, we have

$$(A \star B) \star C = A \star (B \star C).$$

The left and right scalar multiplications are defined in a similar way as in the quaternion matrix case. That is, for $A = [A_{ij}]_{\otimes} \in M_{m\times n}(\mathbb{H}^{N_1\times\cdots\times N_p})$ and $t \in \mathbb{H}^{M_1\times\cdots\times M_q}$, we have

$$t \cdot A := [t \otimes A_{ij}]_{\otimes} \quad \text{and} \quad A \cdot t := [A_{ij} \otimes t]_{\otimes}.$$

Similarly, the defined scalar product is associative. Also, it is compatible with the matrix product, i.e., for $A \in M_{m\times n}(\mathbb{H}^{N_1\times\cdots\times N_p})$, $B_{n\times k}(\mathbb{H}^{M_1\times\cdots\times M_q})$, $t \in (\mathbb{H}^{L_1\times\cdots\times L_s})$ and $s \in (\mathbb{H}^{K_1\times\cdots\times K_r})$, we have

1. $(t \cdot (s \cdot A)) = ((t \cdot s) \cdot A)$;
2. $(A \cdot (s \cdot t)) = ((A \cdot s) \cdot t)$;
3. $(t \cdot (A \star B)) = ((t \cdot A) \star B)$;
4. $(A \star (B \cdot t)) = ((A \star B) \cdot t)$.

The addition $A + B$ is defined entry-wise if A and B have the same sizes and their entries of tensors are of the same order, i.e., for $A, B \in M_{m\times n}\left(\mathbb{H}^{N_1\times\cdots\times N_p}\right)$ with $A = [A_{ij}]_{\otimes}$ and $B = [B_{ij}]_{\otimes}$, we have

$$A + B := [A_{ij}]_{\otimes} + [B_{ij}]_{\otimes} = [A_{ij} + B_{ij}]_{\otimes}.$$

Since the outer operator $\otimes$ is distributive over addition, it is easy to see that we have

1. $(A + B) \star C = A \star C + B \star C$;

2. $C \star (A + B) = C \star A + C \star B$;

3. $(A + B) \cdot t = A \cdot t + B \cdot t$;

4. $t \cdot (A + B) = t \cdot A + t \cdot B$.

Since quaternion matrices can be viewed as matrices with entries of 1-tensors and the multiplication product of two quaternions can be viewed as the tensor product of two 1-tensors, we can see that the multiplication of quaternion matrices is a just special case of the multiplication of matrices of tensors. That is, when $N_1 = ... = N_p = 1$, we have

$$M_{m\times n}\left(\mathbb{H}^{N_1\times\cdots\times N_p}\right) = M_{m\times n}\left(\mathbb{H}^{1\times\cdots\times 1}\right) = M_{m\times n}\left(\mathbb{H}\right),$$

and for any $A \in M_{m\times n}(\mathbb{H})$, $B \in M_{n\times k}(\mathbb{H})$, we have

$$A \star B = AB.$$

Note: It is important to distinguish the notations of tensor matrix and partitioned matrix. For example, let $\vec{a_1}, \vec{a_2}, \vec{a_3}, \vec{a_4} \in \mathbb{H}^{3\times 1}$. The notation

$$A = \begin{bmatrix} \vec{a_1} & \vec{a_2} & \vec{a_3} & \vec{a_4} \end{bmatrix}$$

denotes a 3×4 matrix with entries in $\mathbb{H}$, i,e,. $A \in M_{3\times 4}(\mathbb{H})$, while the notation

$$A' = \begin{bmatrix} \vec{a_1} & \vec{a_2} & \vec{a_3} & \vec{a_4} \end{bmatrix}_{\otimes}$$

denotes a 1×4 tensor matrix with entries of 3-dimensional vectors, i.e., $A' \in M_{1\times 4}(\mathbb{H}^{3\times 1})$. Therefore, $A \neq A'$. Moreover, for a tensor $T \in \mathbb{H}^{N_1\times\cdots\times N_p}$, T can not be viewed as a 1×1 tensor matrix unless T is a quaternionic scalar , i.e., $T \neq [T]_{\otimes}$

unless $N_1 = \cdots = N_p = 1$. For example, let $\vec{v} \in \mathbb{H}^{4\times 1}$ be a 4-dimensional vector, we have $\vec{v} \neq [\vec{v}]_{\otimes}$, because $\vec{v}$ is a 4×1 matrix with entries in $\mathbb{H}$ while $[\vec{v}]_{\otimes}$ is a 1×1 tensor matrix with a entry of a 4-dimensional vector.

Definition 4.1.2. For any $m, i \in \mathbb{N}^+$ with $m \leqslant p$ and $i \leqslant N_m$, define

$$\begin{aligned} \mathrm{Proj}(m, i, \cdot) : \mathbb{H}^{N_1\times\cdots\times N_p} &\longrightarrow \mathbb{H}^{N_1\times\cdots\times N_{m-1}\times N_{m+1}\times\cdots\times N_p}, \\ \vec{v}_1 \otimes \cdots \otimes \vec{v}_p &\mapsto (\vec{v}_1 \otimes \cdots \otimes v_{m,i} \otimes \cdots \otimes \vec{v}_p)\,, \end{aligned}$$

where $v_{m,i} \in \mathbb{H}$ is the i-th coordinate of the vector $\vec{v}_m \in \mathbb{H}^{N_m}$. For any $T \in \mathbb{H}^{N_1\times\cdots\times N_p}$, the image $\mathrm{Proj}(m, i, T) \in \mathbb{H}^{N_1\times\cdots\times N_{m-1}\times N_{m+1}\times\cdots\times N_p}$ is called the i-th m-slice of the tensor T. For a given m,

$$\mathrm{Proj}_{\mathbb{R}}(m, \cdot, T) := \mathrm{span}_{\mathbb{R}}\{\mathrm{Proj}(m, i, T) \mid i = 1, 2, ..., N_m\}$$

is the $\mathbb{R}$-space generated by the collection of all m-slices, called the **space of m-slices**. The dimension of $\mathrm{Proj}(m, \cdot, T)$ (as $\mathbb{R}$-space) is denoted by $\dim_m(T)$. When $m = 1$, $\mathrm{Proj}(1, i, T)$ is denoted by $R_i(T)$ for $i = 1, 2, ..., N_1$, called the i-th row of T.

$$R(T) := \mathrm{span}_{\mathbb{H}}\{R_i(T) \mid i = 1, 2, ..., N_1\}$$

is the left $\mathbb{H}$-module generated by all rows of T, called the **row space** of T. The dimension of $R(T)$ (as left $\mathbb{H}$-module) is denoted by $\dim_{\mathbb{H}} R(T)$. Similarly, when $m = p$, $\mathrm{Proj}(p, i, T)$ is denoted by $C_i(T)$ for $i = 1, 2, ..., N_p$, called the i-th column of T.

$$C(T) := \mathrm{span}_{\mathbb{H}}\{C_i(T) \mid i = 1, 2, ..., N_p\}$$

is the right $\mathbb{H}$-module generated by all columns of T, called the **column space** of T. The dimension of $C(T)$ (as right $\mathbb{H}$-module) is denoted by $\dim_{\mathbb{H}} C(T)$.

We can see $\mathrm{Proj}(m, i, \cdot)$ defines an $(\mathbb{H}, \mathbb{H})$-bimodule homomorphism. It fixes the m-th component and maps a tensor to another tensor with a lower order. The next definition introduces more complex projections by composing different simple projections.

Definition 4.1.3. Let $p \in \mathbb{N}^{+}$. For any $\tau = (i_1, ..., i_k) \in [[\mathcal{I}_k \leqslant p]]$ and $\sigma = (j_1, ..., j_k) \in [[N_{i_1}, N_{i_2}, ..., N_{i_k}]]$, define

$$\mathrm{Proj}(\tau, \sigma, \cdot) = \mathrm{Proj}(i_1, j_1, \cdot) \circ \mathrm{Proj}(i_2, j_2, \cdot) \circ \cdots \circ \mathrm{Proj}(i_k, j_k, \cdot),$$

where $\circ$ is the composition of functions, for example,

$$\left(\mathrm{Proj}(i_1, j_1, \cdot) \circ \mathrm{Proj}(i_2, j_2, \cdot)\right)(T) := \mathrm{Proj}\left(i_1, j_1, \mathrm{Proj}(i_2, j_2, T)\right).$$

The following lemma shows the relationship between multilinear maps and projection maps.

Lemma 4.1.4. *Let $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, $A_i \in M_{M_i \times N_i}(\mathbb{R})$ for $i = 2, ..., p-1$, and $A_j \in M_{M_j \times N_j}(\mathbb{H})$ for $j = 1, p$. Then, we have when $m < p$*

$$\begin{bmatrix} \mathrm{Proj}\left(m, 1, T \times_m A_m\right) \\ \mathrm{Proj}\left(m, 2, T \times_m A_m\right) \\ \vdots \\ \mathrm{Proj}\left(m, N_m, T \times_m A_m\right) \end{bmatrix}_{\otimes} = A_m \star \begin{bmatrix} \mathrm{Proj}(m, 1, T) \\ \mathrm{Proj}(m, 2, T) \\ \vdots \\ \mathrm{Proj}(m, N_m, T) \end{bmatrix}_{\otimes}; \tag{4.1.1}$$

and when $m = p$

$$\begin{bmatrix} \text{Proj}\,(m, 1, T \times_m A_m) \\ \text{Proj}\,(m, 2, T \times_m A_m) \\ \vdots \\ \text{Proj}\,(m, N_m, T \times_m A_m) \end{bmatrix}_{\otimes}^{T} = \begin{bmatrix} \text{Proj}(m, 1, T) \\ \text{Proj}(m, 2, T) \\ \vdots \\ \text{Proj}(m, N_m, T) \end{bmatrix}_{\otimes}^{T} \star A_p^T. \tag{4.1.2}$$

Proof. Denote $\mathcal{P} = \begin{bmatrix} \text{Proj}(m, 1, T) \\ \text{Proj}(m, 2, T) \\ \vdots \\ \text{Proj}(m, N_m, T) \end{bmatrix}_{\otimes}$. When $m < p$, we have

$$\text{Proj}\,(m, j, T \times_m A_m) = T \times_m A_m \times_m (\vec{e}_{N_m,j})^T$$

$$= \sum_{i=1}^{r} \vec{v_1}^{(i)} \otimes \cdots \otimes (\vec{e}_{N_m,j})^T A_m \vec{v_m}^{(i)} \otimes \cdots \otimes \vec{v_p}^{(i)}$$

$$= \sum_{i=1}^{r} \left(\vec{v_1}^{(i)} \otimes \cdots \otimes \vec{v}_{m-1} \right) \cdot (\vec{e}_{N_m,j})^T A_m \vec{v_m}^{(i)} \cdot \left(\vec{v}_{m+1} \otimes \cdots \otimes \vec{v_p}^{(i)} \right)$$

$$= \sum_{i=1}^{r} (\vec{e}_{N_m,j})^T A_m \star \left((\vec{v}_1 \otimes \cdots \otimes \vec{v}_{m-1}) \cdot \vec{v_m}^{(i)} \cdot (\vec{v}_{m+11} \otimes \cdots \otimes \vec{v}_p) \right)$$

$$= (\vec{e}_{N_m,j})^T A_m \star \sum_{i=1}^{r} (\vec{v}_1 \otimes \cdots \otimes \vec{v}_{m-1}) \cdot \vec{v_m}^{(i)} \cdot (\vec{v}_{m+11} \otimes \cdots \otimes \vec{v}_p)$$

$$= (\vec{e}_{N_m,j})^T A_m \star \mathcal{P},$$

where $j = 1, 2, ..., N_m$. This implies Equation (4.1.1). When $m = p$, we have

$$\text{Proj}\,(p, j, T \times_p A_p) = T \times_p A_p \times_p (\vec{e}_{N_p,j})^T$$

$$= T \times_p \left((\vec{e}_{N_p,j})^T A_p \right)$$

$$= \sum_{i=1}^{r} \vec{v_1}^{(i)} \otimes \cdots \otimes \vec{v}_{p-1}^{(i)} \otimes \left((\vec{v_p}^{(i)})^T \left((\vec{e}_{N_p,j})^T A_p \right)^T \right)^T$$

$$= \sum_{i=1}^{r} \vec{v_1}^{(i)} \otimes \cdots \otimes \vec{v}_{p-1}^{(i)} \otimes \left((\vec{v_p}^{(i)})^T A_p^T \vec{e}_{N_p,j} \right)^T$$

$$= \sum_{i=1}^{r} \left(\vec{v_1}^{(i)} \otimes \cdots \otimes \vec{v}_{p-1}^{(i)} \right) \cdot \left((\vec{v_p}^{(i)})^T A_p^T \vec{e}_{N_p,j} \right)$$

$$= \sum_{i=1}^{r} \left(\left(\vec{v_1}^{(i)} \otimes \cdots \otimes \vec{v}_{p-1}^{(i)} \right) \cdot (\vec{v_p}^{(i)})^T \right) \star \left(A_p^T \vec{e}_{N_p,j} \right)$$

$$= \left(\sum_{i=1}^{r} \left(\vec{v_1}^{(i)} \otimes \cdots \otimes \vec{v}_{p-1}^{(i)} \right) \cdot (\vec{v_p}^{(i)})^T \right) \star \left(A_p^T \vec{e}_{N_p,j} \right)$$

$$= \mathcal{P}^T \star \left(A_p^T \vec{e}_{N_p,j} \right),$$

where $j = 1, 2, ..., N_m$. This implies Equation (4.1.2). □

Lemma 4.1.5. *Let $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, $A_i \in M_{M_i \times N_i}(\mathbb{R})$ for $i = 2, ..., p-1$, and $A_j \in M_{M_j \times N_j}(\mathbb{H})$ for $j = 1, p$. For any $m \leqslant p$ and $j \leqslant N_m$,*

if $A_m = I_{N_m}$, we have

$$\text{Proj}\left(m, j, \left(\bigotimes_{i=1}^{p} A_i \right) T \right) = \left(\bigotimes_{i \neq m}^{p} A_i \right) \text{Proj}\,(m, j, T).$$

Proof. Denote $T = \sum_{i=1}^{r} \vec{v_1}^{(i)} \otimes \cdots \otimes \vec{v_p}^{(i)}$, where $\vec{v_j}^{(i)} = (v_{j,1}^{(i)}, ..., v_{j,N_j}^{(i)})^T \in \mathbb{H}^{N_j}$, $j = 1, 2, ..., p$. If $A_m = I_{N_m}$, we have

$$\text{Proj}\left(m, j, \left(\bigotimes_{i=1}^{p} A_i \right) T \right)$$

$$= \text{Proj}\left(m, j, \sum_{i=1}^{r} A_1 \vec{v_1}^{(i)} \otimes \cdots \otimes I_{N_m} \vec{v_m}^{(i)} \otimes \cdots \otimes \left((\vec{v_p}^{(i)})^T A_p^T \right)^T \right)$$

$$= \sum_{i=1}^{r} A_1 \vec{v_1}^{(i)} \otimes \cdots \otimes (\vec{e}_{N_m,j})^T I_n \vec{v_m}^{(i)} \otimes \cdots \otimes \left((\vec{v_p}^{(i)})^T A_p^T \right)^T$$

$$= \left(\bigotimes_{i \neq m}^{p} A_i \right) \left(\sum_{i=1}^{r} \vec{v_1}^{(i)} \otimes \cdots \otimes (\vec{e}_{N_m,j})^T I_n \vec{v_m}^{(i)} \otimes \cdots \otimes \vec{v_p}^{(i)} \right)$$

$$= \left(\bigotimes_{i \neq m}^{p} A_i \right) \text{Proj}(m, j, T).$$

□

Corollary 4.1.6. *Let $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, $A_i \in M_{M_i \times N_i}(\mathbb{R})$ for $i = 2, ..., p-1$, and $A_j \in M_{M_j \times N_j}(\mathbb{H})$ for $j = 1, p$. For any $m \leqslant p$ and $j \leqslant N_m$ (not necessarily $A_m = I_{N_m}$), we have*

$$\begin{bmatrix} \text{Proj}\Big(m, 1, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \text{Proj}\Big(m, 2, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \vdots \\ \text{Proj}\Big(m, N_m, (\bigotimes_{i=1}^{p} A_i)T\Big) \end{bmatrix}_{\otimes} = A_m \star \begin{bmatrix} (\bigotimes_{i \neq m}^{p} A_i)\text{Proj}(m, 1, T) \\ (\bigotimes_{i \neq m}^{p} A_i)\text{Proj}(m, 2, T) \\ \vdots \\ (\bigotimes_{i \neq m}^{p} A_i)\text{Proj}(m, N_m, T) \end{bmatrix}_{\otimes}.$$

when $m = p$, we have

$$\begin{bmatrix} \text{Proj}\Big(m, 1, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \text{Proj}\Big(m, 2, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \vdots \\ \text{Proj}\Big(m, N_m, (\bigotimes_{i=1}^{p} A_i)T\Big) \end{bmatrix}_{\otimes}^{T} = \begin{bmatrix} (\bigotimes_{i \neq m}^{p} A_i)\text{Proj}(m, 1, T) \\ (\bigotimes_{i \neq m}^{p} A_i)\text{Proj}(m, 2, T) \\ \vdots \\ (\bigotimes_{i \neq m}^{p} A_i)\text{Proj}(m, N_m, T) \end{bmatrix}_{\otimes}^{T} \star A_p^T.$$

Proof. For $j = 1, 2, ..., N_m$, we have by Lemma 4.1.5 that

$$\begin{aligned} \text{Proj}\left(m, j, \left(\bigotimes_{i=1}^{p} A_i\right) T\right) &= \text{Proj}\left(m, j, \left(\bigotimes_{i=1}^{m-1} A_i \otimes I_{N_m} \bigotimes_{i=m+1}^{p} A_i\right) (T \times_m A_m)\right) \\ &= \left(\bigotimes_{i \neq m}^{p} A_i\right) \text{Proj}\left(m, j, (T \times_m A_m)\right), \end{aligned}$$

which implies (when $m < p$)

$$\begin{bmatrix} \mathrm{Proj}\Big(m, 1, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \mathrm{Proj}\Big(m, 2, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \vdots \\ \mathrm{Proj}\Big(m, N_m, (\bigotimes_{i=1}^{p} A_i)T\Big) \end{bmatrix}_{\otimes} = \begin{bmatrix} (\bigotimes_{i\neq m}^{p} A_i)\mathrm{Proj}\,(m, 1, (T \times_m A_m)) \\ (\bigotimes_{i\neq m}^{p} A_i)\mathrm{Proj}\,(m, 2, (T \times_m A_m)) \\ \vdots \\ (\bigotimes_{i\neq m}^{p} A_i)\mathrm{Proj}\,(m, N_m, (T \times_m A_m)) \end{bmatrix}_{\otimes}$$

$$= A_m \star \begin{bmatrix} (\bigotimes_{i\neq m}^{p} A_i)\mathrm{Proj}(m, 1, T) \\ (\bigotimes_{i\neq m}^{p} A_i)\mathrm{Proj}(m, 2, T) \\ \vdots \\ (\bigotimes_{i\neq m}^{p} A_i)\mathrm{Proj}(m, N_m, T) \end{bmatrix}_{\otimes}$$

and (when $m = p$)

$$\begin{bmatrix} \mathrm{Proj}\Big(p, 1, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \mathrm{Proj}\Big(p, 2, (\bigotimes_{i=1}^{p} A_i)T\Big) \\ \vdots \\ \mathrm{Proj}\Big(p, N_m, (\bigotimes_{i=1}^{p} A_i)T\Big) \end{bmatrix}_{\otimes}^{T} = \begin{bmatrix} (\bigotimes_{i=1}^{p-1} A_i)\mathrm{Proj}\,(m, 1, (T \times_p A_p)) \\ (\bigotimes_{i=1}^{p-1} A_i)\mathrm{Proj}\,(m, 2, (T \times_p A_p)) \\ \vdots \\ (\bigotimes_{i=1}^{p-1} A_i)\mathrm{Proj}\,(m, N_m, (T \times_p A_p)) \end{bmatrix}_{\otimes}^{T}$$

$$= \begin{bmatrix} (\bigotimes_{i=1}^{p-1} A_i)\mathrm{Proj}(m, 1, T) \\ (\bigotimes_{i=1}^{p-1} A_i)\mathrm{Proj}(m, 2, T) \\ \vdots \\ (\bigotimes_{i=1}^{p-1} A_i)\mathrm{Proj}(m, N_m, T) \end{bmatrix}_{\otimes}^{T} \star A_p^T.$$

□

Corollary 4.1.7. *Let* $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, $\tau = (i_1, ..., i_k) \in [[\mathcal{I}_k \leqslant p]]$, $\sigma \in [[N_{i_1}, N_{i_2}, ..., N_{i_k}]]$, $A_i \in M_{M_i \times N_i}(\mathbb{R})$ *for* $i = 2, ..., p-1$, *and* $A_j \in M_{M_j \times N_j}(\mathbb{H})$ *for* $j = 1, p$. *If* $A_m = I_{N_m}$

for $m = i_1, i_2, ..., i_k$, *then we have*

$$\text{Proj}\left(\tau, \sigma, \left(\bigotimes_{m=1}^{p} A_m\right) T\right) = \left(\bigotimes_{m \neq i_1,...,i_k}^{p} A_m\right) \text{Proj}\left(\tau, \sigma, T\right).$$

Proof. Use induction on k with Lemma 4.1.5. □

Remark 4.1.8. Denote ordered n-tuples

$$\mathcal{R}(T) = (R_1(T), R_2(T) \cdots, R_{N_1}(T))$$

and

$$\mathcal{C}(T) = \left(C_1(T), C_2(T), \cdots, C_{N_p}(T)\right).$$

An immediately consequence is that T is uniquely determined by $\mathcal{R}(T)$ or $\mathcal{C}(T)$. To illustrate, any tensor $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$ can be written as

$$T = \sum_{i=1}^{N_1} \vec{e}_{N_1,i} \otimes R_i(T) = \sum_{i=1}^{N_p} C_i(T) \otimes \vec{e}_{N_p,i}.$$

Similarly, for any other given $m = 2, 3, .., p-1$, we can denote

$$\text{Proj}(m, i, T) = \sum_j A_{ij} \otimes B_{ij}$$

for $i = 1, 2, ..., N_m$, where $A_{ij} \in \mathbb{H}^{N_1 \times \cdots \times N_{m-1}}$ and $B_{ij} \in \mathbb{H}^{N_{m+1} \times \cdots \times N_p}$. Then we can write

$$T = \sum_{i=1}^{N_m} \left(\sum_j A_{ij} \otimes \vec{e}_{N_m,i} \otimes B_{ij}\right).$$

This shows that T is also uniquely determined by

$$\{\text{Proj}(m, i, T) \mid i = 1, 2, ..., N_m\}.$$

Thus, for any given $\tau = (i_1, ..., i_k) \in [[\mathcal{I}_k \leqslant p]]$, we can use induction on k to show that T is uniquely determined by

$$\{\operatorname{Proj}(\tau, \sigma, T) \mid \sigma \in [[N_{i_1}, ..., N_{i_k}]]\}.$$

Therefore, we have the following lemma.

Lemma 4.1.9. *Let $T_1, T_2 \in \mathbb{H}^{N_1 \times \cdots \times N_p}$. Then the following are equivalent:*

1. $T_1 = T_2$;
2. $\mathcal{R}(T_1) = \mathcal{R}(T_2)$;
3. $\mathcal{C}(T_1) = \mathcal{C}(T_2)$;
4. *There exists $m \in \{1, 2, ..., p\}$ such that $\operatorname{Proj}(m, i, T_1) = \operatorname{Proj}(m, i, T_2)$ for $i = 1, 2, ..., N_m$;*
5. *There exists $\tau = (i_1, ..., i_k) \in [[\mathcal{I}_k \leqslant p]]$ such that $\operatorname{Proj}(\tau, \sigma, T_1) = \operatorname{Proj}(\tau, \sigma, T_2)$ for all $\sigma \in [[N_{i_1}, ..., N_{i_k}]]$.*

4.2 General Properties

We start with giving some results for general quaternion tensors.

Theorem 4.2.1. *Let $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$. Then the following are equivalent:*

1. $\operatorname{rank}_{\mathbb{H}}(T) \leqslant r$;
2. *$R(T)$ is contained in a left $\mathbb{H}$-module generated by r simple tensors;*
3. *$C(T)$ is contained in a right $\mathbb{H}$-module generated by r simple tensors;*

4. *For some* $m \in \{1, 2, ..., p\}$, *there exist* r *simple tensors* $A_1, ..., A_r \in \mathbb{H}^{N_1 \times \cdots \times N_{m-1}}$ *and* r *simple tensors* $C_1, ..., C_r \in \mathbb{H}^{N_{m+1} \times \cdots \times N_p}$ *such that* $\mathrm{Proj}(m, \cdot, T)$ *is contained in the* $\mathbb{R}$*-module generated by*

$$\{A_i \otimes b \otimes C_i \mid b \in \mathbb{H}, i = 1, 2, ..., r\}.$$

Proof. (1) $\Longrightarrow$ (2) and (3): Suppose $\mathrm{rank}_{\mathbb{H}}(T) \leqslant r$. Then T can be written as

$$T = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_r \end{bmatrix}_{\otimes} \star \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_r \end{bmatrix}_{\otimes} \star \begin{bmatrix} \vec{c}_1 \\ \vec{c}_2 \\ \vdots \\ \vec{c}_r \end{bmatrix}_{\otimes},$$

where $\vec{a}_j \in \mathbb{H}^{N_1 \times 1}$, $\vec{c}_j \in \mathbb{H}^{N_p \times 1}$ and $B_j \in \mathbb{H}^{N_2 \times \cdots \times N_{p-1}}$, for $j = 1, 2, ..., r$, are all simple tensors. Denote $\vec{a}_j = \begin{bmatrix} a_{j1} & a_{j2} & ... & a_{j,N_1} \end{bmatrix}^T$, $\vec{c}_j = \begin{bmatrix} c_{j1} & c_{j2} & ... & c_{j,N_p} \end{bmatrix}^T$ for $j = 1, 2, ..., r$, and let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_1 1} & a_{N_1 2} & \cdots & a_{N_1,r} \end{bmatrix}, \quad C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1r} \\ c_{21} & c_{22} & \cdots & c_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N_p 1} & c_{N_p 2} & \cdots & c_{N_p,r} \end{bmatrix}.$$

so that we can write

$$\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_r \end{bmatrix}_{\otimes} = \begin{bmatrix} \vec{e}_{N_1,1} & \vec{e}_{N_1,2} & \dots & \vec{e}_{N_1,N_1} \end{bmatrix}_{\otimes} \star A$$

and

$$\begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \dots & \vec{c}_r \end{bmatrix}_{\otimes} = \begin{bmatrix} \vec{e}_{N_p,1} & \vec{e}_{N_p,2} & \dots & \vec{e}_{N_p,N_p} \end{bmatrix}_{\otimes} \star C$$

where $\{\vec{e}_{N_1,i}\}_{i=1}^{N_1}$ and $\{\vec{e}_{N_p,i}\}_{i=1}^{N_p}$ are the standard basis for $\mathbb{H}^{N_1}$ and $\mathbb{H}^{N_p}$. Therefore, we can write

$$T = \begin{bmatrix} \vec{e}_{N_1,1} & \vec{e}_{N_1,2} & \dots & \vec{e}_{N_1,N_1} \end{bmatrix}_{\otimes} \star A \star \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_r \end{bmatrix}_{\otimes} \star \begin{bmatrix} \vec{c}_1 \\ \vec{c}_2 \\ \vdots \\ \vec{c}_r \end{bmatrix}_{\otimes} \qquad \text{(i)}$$

$$= \begin{bmatrix} \vec{e}_{N_1,1} & \vec{e}_{N_1,2} & \dots & \vec{e}_{N_1,N_1} \end{bmatrix}_{\otimes} \star A \star \begin{bmatrix} B_1 \otimes \vec{c}_1 \\ B_2 \otimes \vec{c}_2 \\ \vdots \\ B_r \otimes \vec{c}_r \end{bmatrix}_{\otimes}$$

and

$$T = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_r \end{bmatrix}_{\otimes} \star \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_r \end{bmatrix}_{\otimes} \star C^T \star \begin{bmatrix} \vec{e}_{N_p,1} \\ \vec{e}_{N_p,2} \\ \vdots \\ \vec{e}_{N_p,N_p} \end{bmatrix}_{\otimes} \qquad \text{(ii)}$$

$$= \begin{bmatrix} \vec{a}_1 \otimes B_1 & \vec{a}_2 \otimes B_2 & \dots & \vec{a}_r \otimes B_r \end{bmatrix}_{\otimes} \star C^T \star \begin{bmatrix} \vec{e}_{N_p,1} \\ \vec{e}_{N_p,2} \\ \vdots \\ \vec{e}_{N_p,N_p} \end{bmatrix}_{\otimes} .$$

By definition, Equation (i) shows

$$\begin{bmatrix} R_1(T) \\ R_2(T) \\ \vdots \\ R_{N_1}(T) \end{bmatrix}_{\otimes} = A \star \begin{bmatrix} B_1 \otimes \vec{c}_1 \\ B_2 \otimes \vec{c}_2 \\ \vdots \\ B_r \otimes \vec{c}_r \end{bmatrix}_{\otimes}$$

which implies that for $i = 1, 2, ..., N_1$, $R_i(T)$ is an element in the left $\mathbb{H}$-module generated by

$$\{B_1 \otimes \vec{c}_1, B_2 \otimes \vec{c}_2, ..., B_r \otimes \vec{c}_r\}$$

consisting of simple tensors. Similarly, Equation (ii) shows

$$\begin{bmatrix} C_1(T) & C_2(T) & \dots & C_{N_p}(T) \end{bmatrix}_{\otimes} = \begin{bmatrix} \vec{a}_1 \otimes B_1 & \vec{a}_2 \otimes B_2 & \dots & \vec{a}_r \otimes B_r \end{bmatrix}_{\otimes} \star C^T,$$

which implies that for $j = 1, 2, ..., N_p$, $C_i(T)$ is an element in the right $\mathbb{H}$-module generated by

$$\{\vec{a}_1 \otimes B_1, \vec{a}_2 \otimes B_2, ..., \vec{a}_r \otimes B_r\}$$

consisting of simple tensors.

(2) $\Longrightarrow$ (4) and (3) $\Longrightarrow$ (4) are both obvious, because $\{a_i \otimes b \otimes c_i \mid b \in \mathbb{H}, i = 1, 2, ..., r\}$ becomes a left $\mathbb{H}$-module when $m = 1$ and becomes a right $\mathbb{H}$-module when $m = p$.

(4) $\Longrightarrow$ (1): Suppose there are simple tensors simple tensors $A_1, ..., A_r \in \mathbb{H}^{N_1 \times \cdots \times N_{m-1}}$ and $C_1, ..., C_r \in \mathbb{H}^{N_{m+1} \times \cdots \times N_p}$ such that $\mathrm{Proj}(m, \cdot, T)$ is contained in the $\mathbb{R}$-module generated by $\{A_j \otimes b \otimes C_j \mid b \in \mathbb{H}, j = 1, 2, ..., r\}$. Then for $i = 1, 2, ..., N_m$ we can write

$$\mathrm{Proj}(m, i, T) = \sum_{j=1}^{r} A_j \otimes b_{ij} \otimes C_j$$

for some $b_{ij} \in \mathbb{H}$, i.e.,

$$\mathrm{Proj}(m, i, T) = \begin{bmatrix} A_1 & A_2 & \dots & A_r \end{bmatrix}_{\otimes} \star \begin{bmatrix} b_{i1} & & & \\ & b_{i2} & & \\ & & \ddots & \\ & & & b_{ir} \end{bmatrix} \star \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_r \end{bmatrix}_{\otimes}.$$

This implies

$$T = \begin{bmatrix} A_1 & A_2 & \dots & A_r \end{bmatrix}_{\otimes} \star \begin{bmatrix} \sum_{i=1}^{N_m} b_{i1}\vec{e}_{N_m,i} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sum_{i=1}^{N_m} b_{ir}\vec{e}_{N_m,i} \end{bmatrix}_{\otimes} \star \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_r \end{bmatrix}_{\otimes}.$$

Therefore, we have $\text{rank}_{\mathbb{H}}(T) \leqslant r$. □

Corollary 4.2.2. *For any $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, we have*

1. $\text{rank}_{\mathbb{H}}(T) \geqslant \dim_{\mathbb{H}} R(T)$; $\text{rank}_{\mathbb{H}}(T) = \dim R(T)$ *if and only if there exists a basis of $R(T)$ consisting of r simple tensors;*

2. $\text{rank}_{\mathbb{H}}(T) \geqslant \dim_{\mathbb{H}} C(T)$; $\text{rank}_{\mathbb{H}}(T) = \dim C(T)$ *if and only if there exists a basis of $C(T)$ consisting of r simple tensors.*

In 1977, Kruskal ([39]) generalized a matrix theorem of Frobenius. We now show that this can also be generalized to quaternionic tensors of an arbitrary order.

Theorem 4.2.3. *Let $T \in \mathbb{H}^{N_1 \times \cdots \times N_p}$, $A_i \in M_{M_i \times N_i}(\mathbb{R})$ for $i = 2, ..., p-1$, and $A_j \in M_{M_j \times N_j}(\mathbb{H})$ for $j = 1, p$. Then we have*

$$\text{rank}_{\mathbb{H}}(T) \geqslant \text{rank}_{\mathbb{H}} \left(\left(\bigotimes_{i=2}^{p} A_i \right) T \right) + \dim_{\mathbb{H}} R(A_1 T) - \dim_{\mathbb{H}} R \left(\left(\bigotimes_{i=1}^{p} A_i \right) T \right)$$

and

$$\text{rank}_{\mathbb{H}}(T) \geqslant \text{rank}_{\mathbb{H}} \left(\left(\bigotimes_{i=1}^{p-1} A_i \right) T \right) + \dim_{\mathbb{H}} C(T A_p^T) - \dim_{\mathbb{H}} C \left(\left(\bigotimes_{i=1}^{p} A_i \right) T \right).$$

Proof. Let $\text{rank}_{\mathbb{H}}(T) = r$. By Theorem 4.2.1, there exists a set of r simple tensors

$\{S_1, S_2, ..., S_r\}$ such that

$$\begin{bmatrix} R_1(T) \\ R_2(T) \\ \vdots \\ R_{N_1}(T) \end{bmatrix}_{\otimes} = A \star \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_r \end{bmatrix}_{\otimes}$$

for some $A \in M_{N_1 \times r}(\mathbb{H})$. Without loss of generality, we may also assume that $S_1, ..., S_r$ are linear independent. Let ${}_{\mathbb{H}}M$ be the left $\mathbb{H}$-module generated by $S_1, ..., S_r$. Then

$$\dim_{\mathbb{H}} M = r = \operatorname{rank}(T).$$

Noticing by Corollary 4.1.6 that

$$\begin{bmatrix} R_1\left(A_1 T\right) \\ R_2\left(A_1 T\right) \\ \vdots \\ R_{M_1}\left(A_1 T\right) \end{bmatrix}_{\otimes} = A_1 \star \begin{bmatrix} R_1(T) \\ R_2(T) \\ \vdots \\ R_{N_1}(T) \end{bmatrix}_{\otimes},$$

we have

$${}_{\mathbb{H}}R(A_1 T) \leqslant {}_{\mathbb{H}}R(T) \leqslant {}_{\mathbb{H}}M \leqslant {}_{\mathbb{H}}\mathbb{H}^{N_2 \times \cdots \times N_{p-1} \times N_p}.$$

Let $\sigma := \bigotimes_{i=2}^{p} A_i$ denote the left $\mathbb{H}$-module homomorphism induced by

$$\begin{aligned} \sigma : \mathbb{H}^{N_2 \times \cdots \times N_{p-1} \times N_p} &\longrightarrow \mathbb{H}^{M_2 \times \cdots \times M_{p-1} \times M_p}, \\ \vec{v}_2 \otimes \cdots \otimes \vec{v}_p &\mapsto A_2 \vec{v}_2 \otimes \cdots \otimes A_{p-1} \vec{v}_{p-1} \otimes (\vec{v}_p{}^T A_p^T)^T. \end{aligned}$$

Noticing by Corollary 4.1.7 that

$$\left(\bigotimes_{i=2}^{p} A_i\right) R(A_1 T) = R\left(\left(\bigotimes_{i=1}^{p} A_i\right) T\right)$$

and

$$\left(\bigotimes_{i=2}^{p} A_i\right) R(T) = R\left(\left(\bigotimes_{i=2}^{p} A_i\right) T\right),$$

we have the following commutative diagram:

$$\begin{array}{ccccccc}
R(A_1T) & \hookrightarrow & R(T) & \hookrightarrow & M & \hookrightarrow & \mathbb{H}^{N_2\times\cdots\times N_{p-1}\times N_p} \\
\downarrow \sigma|_{R(A_1T)} & & \downarrow \sigma|_{R(T)} & & \downarrow \sigma|_M & & \downarrow \sigma \\
R\left(\left(\bigotimes_{i=1}^{p} A_i\right) T\right) & \hookrightarrow & R\left(\left(\bigotimes_{i=2}^{p} A_i\right) T\right) & \hookrightarrow & \sigma(M) & \hookrightarrow & \mathbb{H}^{N_2\times\cdots\times N_{p-1}\times M_2}
\end{array}.$$

Since $R\left(\left(\bigotimes_{i=2}^{p} A_i\right) T\right)$ is contained in $\sigma(M)$ which is generated by a set of simple tensors $\{\sigma(S_1), \sigma(S_2), ..., \sigma(S_r)\}$, by Theorem 4.2.1, we have

$$\operatorname{rank}_{\mathbb{H}}\left(\left(\bigotimes_{i=2}^{p} A_i\right) T\right) \leqslant \dim_{\mathbb{H}} \sigma(M).$$

Moreover,

$$\dim_{\mathbb{H}} \ker(\sigma|_{R(A_1T)}) \leqslant \dim_{\mathbb{H}} \ker(\sigma|_M)$$

implies

$$\dim_{\mathbb{H}} R\left(A_1T\right) - \dim_{\mathbb{H}} R\left(\left(\bigotimes_{i=1}^{p} A_i\right) T\right) \leqslant \dim_{\mathbb{H}} M - \dim_{\mathbb{H}} \sigma(M),$$

which suggests

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \dim_{\mathbb{H}} R\left(A_1T\right) + \operatorname{rank}_{\mathbb{H}}\left(\left(\bigotimes_{i=2}^{p} A_i\right) T\right) - \dim_{\mathbb{H}} R\left(\left(\bigotimes_{i=1}^{p} A_i\right) T\right).$$

Proving in a similar fashion, we can also show

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \operatorname{rank}_{\mathbb{H}}\left(\left(\bigotimes_{i=1}^{p-1} A_i\right) T\right) + \dim_{\mathbb{H}} C(TA_p^T) - \dim_{\mathbb{H}} C\left(\left(\bigotimes_{i=1}^{p} A_i\right) T\right).$$

□

By letting $A_i = I_{N_i}$ for $i = 2, 3, ..., p-1$ in Theorem 4.2.3, we have the following corollary, which is the classical inequality by Frobenius about matrix ranks ([46, Theorem 5.66])

$$\operatorname{rank}(X) \geqslant \operatorname{rank}(UX) + \operatorname{rank}(XV) - \operatorname{rank}(UXV),$$

as well as the one by Kruskal ([39, Theorem 1]) in 1977 for complex 3-tensors.

Corollary 4.2.4. *Let $T \in \mathbb{H}^{N_1\times\cdots\times N_p}$, $U \in M_{M_1\times N_1}(\mathbb{H})$, $V \in M_{N_p\times M_2}(\mathbb{H})$. Then we have*

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \dim_{\mathbb{H}} R(UT) + \operatorname{rank}_{\mathbb{H}}(TV) - \dim_{\mathbb{H}} R(UTV)$$

and

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \operatorname{rank}_{\mathbb{H}}(UT) + \dim_{\mathbb{H}} C(TV) - \dim_{\mathbb{H}} C(UTV).$$

Before moving on to the next theorem, let's introduce some notations. For any tensor $T \in \mathbb{H}^{N_1\times N_2\times\cdots\times N_p}$, denote $T^{\leqslant m}$ to be the sub-tensor consisting the first m rows of T and ${}^{m<}T$ to be the sub-tensor consisting the last $N_1 - m$ rows of T, i.e.,

$$T^{\leqslant m} := \begin{bmatrix} I_m & O_{m\times(N_1-m)} \end{bmatrix} T, \qquad {}^{m<}T := \begin{bmatrix} O_{(N_1-m)\times m} & I_{N_1-m} \end{bmatrix} T.$$

Similarly, we have

$$T_{\leqslant m} := T \begin{bmatrix} I_m \\ O_{(N_p-m)\times m} \end{bmatrix}, \qquad {}_{m<}T := T \begin{bmatrix} O_{m\times(N_p-m)} \\ I_{N_p-m} \end{bmatrix}.$$

The next theorem is extremely useful for determining lower bounds of maximal rank of quaternionic tensors.

Theorem 4.2.5. *Let $T \in \mathbb{H}^{N_1\times N_2\times\cdots\times N_p}$. Then we have*

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{B\in\mathbb{H}^{m\times(N_1-m)}} \operatorname{rank}_{\mathbb{H}}\left(T^{\leqslant m} + B({}^{m<}T)\right) + \dim_{\mathbb{H}} R(T) - m.$$

Proof. Denote $\operatorname{rank}_{\mathbb{H}}(T) = r$. By Theorem 4.2.1, there exist $A \in \mathbb{H}^{N_1 \times r}$ and simple tensors $s_1, s_2, .., s_r \in \mathbb{H}^{N_2 \times \cdots \times N_p}$ such that

$$\begin{bmatrix} R_1(T) \\ R_2(T) \\ \vdots \\ R_{N_1}(T) \end{bmatrix}_{\otimes} = A \star \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_r \end{bmatrix}_{\otimes} .$$

Without loss of generality, we can assume $\dim_{\mathbb{H}} R(T) = N_1$. Then by Corollary 4.2.2, we have $\operatorname{rank}_{\mathbb{H}}(A) \geqslant \dim_{\mathbb{H}} R(T) = N_1$. Let Q_1 be a product of some elementary matrices $P_{ij}(k)$ and Q_2 be a product of some elementary matrices P_k and P_{ij} such that

$$Q_1 A Q_2 = \begin{bmatrix} I_{N_1} & B_{N_1 \times (r-N_1)} \end{bmatrix}.$$

Denote

$$M = \begin{bmatrix} I_{N_1} & O_{N_1 \times (r-N_1)} \end{bmatrix} Q_1,$$

and

$$\begin{bmatrix} S'_1 \\ S'_2 \\ \vdots \\ S'_r \end{bmatrix}_{\otimes} = Q_2^{-1} \star \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_r \end{bmatrix}_{\otimes} .$$

Then we have

$$\begin{bmatrix} R_1(MT) \\ R_2(MT) \\ \vdots \\ R_{N_1}(MT) \end{bmatrix}_{\otimes} = M \star \begin{bmatrix} R_1(T) \\ R_2(T) \\ \vdots \\ R_{N_1}(T) \end{bmatrix}_{\otimes}$$

$$= (MQ_1^{-1})(Q_1AQ_2)Q_2^{-1} \star \begin{bmatrix} S'_1 \\ S'_2 \\ \vdots \\ S'_r \end{bmatrix}_{\otimes}$$

$$= \begin{bmatrix} I_m & O_{m\times(N_1-m)} \end{bmatrix} \begin{bmatrix} I_{N_1} & B_{N_1\times(r-N_1)} \end{bmatrix} \star \begin{bmatrix} S'_1 \\ S'_2 \\ \vdots \\ S'_r \end{bmatrix}_{\otimes}$$

$$= \begin{bmatrix} I_m & O_{m\times N_1-m} & B'_{m\times(r-N_1)} \end{bmatrix} \star \begin{bmatrix} S'_1 \\ S'_2 \\ \vdots \\ S'_r \end{bmatrix}_{\otimes},$$

which implies

$$\text{rank}_{\mathbb{H}}(MT) \leqslant r - (N_1 - m).$$

On the other hand, denote

$$Q_1 = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix},$$

where $Q_{11} \in \mathbb{H}^{m\times m}$, $Q_{12} \in \mathbb{H}^{m\times(N_1-m)}$, $Q_{21} \in \mathbb{H}^{(N_1-m)\times m}$, and $Q_{22} \in \mathbb{H}^{(N_1-m)\times(N_1-m)}$.

By the construction of Q_1, we have that Q_{11} is invertible. Hence, we have

$$\begin{aligned} Q_{11}^{-1}MT &= Q_{11}^{-1}\begin{bmatrix} I_m & O_{m\times(N_1-m)} \end{bmatrix}\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \\ &= \begin{bmatrix} Q_{11}^{-1} & O_{m\times(N_1-m)} \end{bmatrix}\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} T \\ &= \begin{bmatrix} I_m & Q_{11}^{-1}Q_{12} \end{bmatrix} T \\ &= T^{\leqslant m} + Q_{11}^{-1}Q_{12}\left({}^{m<}T\right), \end{aligned}$$

which implies

$$\begin{aligned} \operatorname{rank}_{\mathbb{H}}(MT) &= \operatorname{rank}_{\mathbb{H}}(Q_{11}^{-1}MT) \\ &= \operatorname{rank}_{\mathbb{H}}\left(T^{\leqslant m} + Q_{11}^{-1}Q_{12}\left({}^{m<}T\right)\right) \\ &\geqslant \min_{B\in\mathbb{H}^{m\times(N_1\times m)}} \operatorname{rank}_{\mathbb{H}}\left(T^{\leqslant m} + B({}^{m<}T)\right). \end{aligned}$$

Therefore, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{B\in\mathbb{H}^{m\times(N_1-m)}} \operatorname{rank}_{\mathbb{H}}\left(T^{\leqslant m} + B({}^{m<}T)\right) + \dim_{\mathbb{H}} R(T) - m.$$

□

4.3 Normal Forms and maximal rank of $m \times 2 \times n$ quaternionic tensors

The maximal rank of $m \times 2 \times n$ complex tensors was obtained in 2009 by Sumi et al. ([58, Theorem 4.3]). They improved this result given by JaJa ([32]) by determining the rank of such tensors in comparison with their Kronecker canonical forms over the complex and real number fields. In this section, We give the maximal rank of

$m \times 2 \times n$ tensors over the quaternions as well as their canonical forms by using Kronecker canonical forms over the quaternions.

Definition 4.3.1 ([49], Page 153)**.** Let $\mathbb{H}(t)^{m\times n}$ be the set of all $m \times n$ matrices with entries in $\mathbb{H}$, which will be called matrix polynomials with the standard operations of addition, right and left multiplication by quaternions, and matrix multiplication: If $A(t) \in \mathbb{H}(t)^{m\times n}$ and $B(t) \in \mathbb{H}(t)^{n\times p}$, then $A(t)B(t) \in \mathbb{H}(t)^{m\times p}$. Two matrix polynomials of degree at most one $A_1 + tB_1$ and $A_2 + tB_2$, where $A_1, B_1, A_2, B_2 \in \mathbb{H}^{m\times n}$, are called **matrix pencils**.

Definition 4.3.2 ([49], Page 161)**.** The matrix pencils $A_j + tB_j, j = 1, 2$, are called strictly equivalent if

$$A_1 = PA_2Q, \quad B_1 = PB_2Q$$

for some invertible matrices $P \in \mathbb{H}^{m\times m}$ and $Q \in \mathbb{H}^{n\times n}$.

Theorem 4.3.3 ([49], Page 161)**.** *Every pencil $A+tB \in \mathbb{H}(t)^{m\times n}$ is strictly equivalent to a matrix pencil with the block diagonal form:*

$$\begin{aligned} 0_{u\times v} &\oplus L_{\varepsilon_1\times(\varepsilon_1+1)} \oplus \cdots \oplus L_{\varepsilon_p\times(\varepsilon_p+1)} \oplus L^T_{\eta_1\times(\eta_1+1)} \oplus L^T_{\eta_q\times(\eta_q+1)} \\ &\oplus (I_{k_1} + tJ_{k_1}(0)) \oplus \cdots \oplus (I_{k_r} + tJ_{k_r}(0)) \\ &\oplus (tI_{\ell_1} + J_{\ell_1}(\alpha_1)) \oplus \cdots \oplus (tI_{\ell_s} + J_{\ell_s}(\alpha_s)), \end{aligned}$$

where

$$L_{\varepsilon\times(\varepsilon+1)}(t) := \begin{bmatrix} 0_{\varepsilon\times 1} & I_\varepsilon \end{bmatrix} + t \begin{bmatrix} I_\varepsilon & 0_{\varepsilon\times 1} \end{bmatrix} \in \mathbb{H}^{\varepsilon\times(\varepsilon+1)},$$

$\varepsilon_1 \leqslant \cdots \leqslant \varepsilon_p$; $\eta_1 \leqslant \cdots \leqslant \eta_q; k_1 \leqslant \cdots \leqslant k_r$, are positive integers, and $\alpha_1, \ldots, \alpha_s \in \mathbb{H}$, and I_k denotes the $k \times k$ identity matrix and $J_k(\alpha_i)$ denotes the $k \times k$ Jordan block with eigenvalue α_i, $i = 1, 2, ..., s$.

Moreover, the integers u, v, and $\varepsilon_i, \eta_j, k_w$ are uniquely determined by the pair

A, B, and the part

$$(tI_{\ell_1} + J_{\ell_1}(\alpha_1)) \oplus \cdots \oplus (tI_{\ell_s} + J_{\ell_s}(\alpha_s))$$

is uniquely determined by A and B up to an arbitrary permutation of the diagonal blocks and up to replacing α_j with any quaternion similar to α_j in each $J_{\ell_j}(\alpha_j)$.

Applying Definition 4.3.2 and Theorem 4.3.3, we obtain the canonical forms of $m \times 2 \times n$ quaternion tensors in the following theorem.

Theorem 4.3.4. *Let T be a $m \times 2 \times n$ quaternionic tensor. Then there exist matrices $P \in M_m(\mathbb{H})$ and $Q \in M_n(\mathbb{H})$ such that T can be transformed into a block diagonal form:*

$$T' = PTQ = \operatorname{Diag}((S_1; T_1), \ldots, (S_r; T_r)),$$

where each $(S_j; T_j)$ is one of the following:

1. *zero tensor $(O_{k\times l}; O_{k\times l}) \in \mathbb{H}^{k\times 2\times l}, k, l \geqslant 0$, $(k, l) \neq (0, 0)$;*
2. *$(J_k(a); I_k) \in \mathbb{H}^{k\times 2\times l}$, $a \in \mathbb{C}$, $k \geqslant 1$;*
3. *$(I_k; J_k(0)) \in \mathbb{H}^{k\times 2\times l}, k \geqslant 1$;*
4. *$([O_{k\times 1}, I_k]; [I_k, O_{k\times 1}]) \in \mathbb{H}^{k\times 2\times (k+1)}, k \geqslant 1$;*
5. $\left(\begin{bmatrix} O_{1\times k} \\ I_k \end{bmatrix}; \begin{bmatrix} I_k \\ O_{1\times k} \end{bmatrix} \right) \in \mathbb{H}^{(k+1)\times 2\times k}, k \geqslant 1,$

where I_k denotes the $k \times k$ identity matrix and $J_k(a)$ denotes the $k \times k$ Jordan block with eigenvalue a.

Proof. Since all the Jordan Blocks in the Kronecker canonical given in Theorem 4.3.3 can take standard eigenvalues (complex numbers with positive real parts), we see that for any given $m \times 2 \times n$ quaternionic tensor $T_1 = (A_1; B_1)$, there exist invertible matrices $P \in M_m(\mathbb{H})$ and $Q \in M_n(\mathbb{H})$ such that

$$A_2 = PA_1Q \in M_{m\times n}(\mathbb{C})$$

and

$$B_2 = PB_1Q \in M_{m\times n}(\mathbb{C}).$$

Therefore, T_1 can be transformed into a $m \times 2 \times n$ complex tensor:

$$PT_1Q = P(A_1; B_1)Q = (PA_1Q; PB_1Q) = (A_2; B_2) \in \mathbb{C}^{m\times 2\times n}.$$

According to the Kronecker canonical forms given in Theorem 4.3.3, we have obtained the corresponding canonical forms above for $m \times 2 \times n$ quaternionic tensors. □

Since each of the $m\times 2\times n$ quaternionic tensors has a canonical form as a complex tensor, we also obtain an upper bound of the maximal rank of $m\times 2\times n$ quaternionic tensors.

Corollary 4.3.5. *Let $m, n \in \mathbb{N}^+$. Then we have*

$$\text{max.rank}_{\mathbb{H}}(m, 2, n) \leqslant \min\left\{n + \left\lfloor \frac{m}{2} \right\rfloor, m + \left\lfloor \frac{n}{2} \right\rfloor, 2m, 2n\right\}.$$

Proof. Let T be any given $m \times 2 \times n$ quaternionic tensor. By Theorem 4.3.4, there exist invertible matrices $P \in M_m(\mathbb{H})$ and $Q \in M_n(\mathbb{H})$ such that

$$T' = PTQ$$

is a $m \times 2 \times n$ complex tensor. By [60, Theorem 5.5], any $m \times 2 \times n$ complex tensor can be written as the sum of no more than

$$s := \min\left\{n + \left\lfloor \frac{m}{2} \right\rfloor, m + \left\lfloor \frac{n}{2} \right\rfloor, 2m, 2n\right\}$$

complex tensors. Hence, we can write

$$T' = \sum_{i=1}^{s} T_i,$$

which implies

$$T = P^{-1}T'Q^{-1} = P^{-1}\left(\sum_{i=1}^{s} T_i\right)Q^{-1} = \sum_{i=1}^{s} P^{-1}T_iQ^{-1}.$$

Therefore, we have $\operatorname{rank}_{\mathbb{H}}(T) \leqslant s = \min\left\{n + \left\lfloor \frac{m}{2} \right\rfloor, m + \left\lfloor \frac{n}{2} \right\rfloor, 2m, 2n\right\}$, which implies the desired inequality. $\square$

Next, we show that the upper bound we obtain in Corollary 4.3.5 is the best possible for quaternion tensors.

Proposition 4.3.6. *Let $m, n \in \mathbb{N}^+$ with $m \leqslant n$. Then we have*

1. $\operatorname{max.rank}_{\mathbb{H}}(m, 2, n) \geqslant m + \left\lfloor \frac{n}{2} \right\rfloor$, *if $m \leqslant n \leqslant 2m$;*

2. $\operatorname{max.rank}_{\mathbb{H}}(m, 2, n) \geqslant 2m$, *if $2m < n$.*

Proof. 1. Since $n \leqslant 2m$, we have $\left\lfloor \frac{n}{2} \right\rfloor \leqslant m$. Denote $k = \left\lfloor \frac{n}{2} \right\rfloor$ and let

$$T := \left(\begin{bmatrix} I_m & O_{m\times(n-k)} \end{bmatrix}; \begin{bmatrix} O_{k\times(n-k)} & I_k \\ O_{(m-k)\times(n-k)} & O_{(m-k)\times k} \end{bmatrix} \right).$$

Then when n is odd, we have $n = 2k + 1$, and we can write $T = (A_1; A_2)$, where

$$A_1 = \begin{bmatrix} I_{k+1} & O_{k\times k} \\ O_{(m-k-1)\times(k+1)} & (I_{(2k-m+1)})^{\leqslant m-k-1} \end{bmatrix}$$

and

$$A_2 = \begin{bmatrix} O_{(k+1)\times(k+1)} & (I_{k+1})^{\leqslant k} \\ O_{(m-k-1)\times(k+1)} & O_{(m-k-1)\times(2k-m+1)} \end{bmatrix},$$

which shows that for any $B \in M_{(m-k-1)\times(2k+1)}(\mathbb{H})$ all $2k + 1$ columns of $T^{\leqslant k+1} + B\left({}^{k+1<}T\right)$ are linearly independent (in the right-$\mathbb{H}$ module). So, we have

$$\min_{B\in\mathbb{H}^{(m-k-1)\times(2k+1)}} \dim_{\mathbb{H}} C\left(T^{\leqslant k+1} + B\left({}^{k+1<}T\right)\right) = 2k + 1.$$

Noticing that $\dim_{\mathbb{H}} R(T) = m$, by Theorem 4.2.5, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{B\in\mathbb{H}^{(m-k-1)\times(2k+1)}} \operatorname{rank}_{\mathbb{H}}\left(T^{\leqslant k+1} + B\left({}^{k+1<}T\right)\right) + \dim_{\mathbb{H}} R(T) - (k+1)$$

$$\geqslant \min_{B\in\mathbb{H}^{(m-k-1)\times(2k+1)}} \dim_{\mathbb{H}} C\left(T^{\leqslant k+1} + B({}^{k+1<}T)\right) + \dim_{\mathbb{H}} R(T) - (k+1)$$

$$\geqslant 2k + 1 + m - (k+1)$$

$$= m + k = m + \left\lfloor \frac{n}{2} \right\rfloor.$$

When n is even, we have $n = 2k$, and we can write

$$T = \left(\begin{bmatrix} I_k & O_{k\times k} \\ O_{(m-k)\times k} & (I_k)^{\leqslant m-k} \end{bmatrix}; \begin{bmatrix} O_{k\times k} & I_k \\ O_{(m-k)\times k} & O_{(m-k)\times k} \end{bmatrix} \right)$$

which shows that for any $B \in M_{(m-k)\times 2k}(\mathbb{H})$ all $2k$ columns of $T^{\leqslant k} + B\left({}^{k<}T\right)$ are linearly independent (in the right-$\mathbb{H}$ module). So, we have

$$\min_{B\in\mathbb{H}^{(m-k)\times 2k}} \dim_{\mathbb{H}} C\left(T^{\leqslant k} + B\left({}^{k<}T\right)\right) = 2k.$$

Noticing that $\dim_{\mathbb{H}} R(T) = m$, by Theorem 4.2.5, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{B\in\mathbb{H}^{(m-k)\times 2k}} \operatorname{rank}_{\mathbb{H}}\left(T^{\leqslant k} + B({}^{k<}T)\right) + \dim_{\mathbb{H}} R(T) - k$$

$$\geqslant \min_{B\in\mathbb{H}^{(m-k)\times 2k}} \dim_{\mathbb{H}} C\left(T^{\leqslant k} + B({}^{k<}T)\right) + \dim_{\mathbb{H}} R(T) - k$$

$$\geqslant 2k + m - k$$

$$= m + k = m + \left\lfloor \frac{n}{2} \right\rfloor.$$

Therefore, we have $\operatorname{rank}_{\mathbb{H}}(T) = m + \left\lfloor \frac{n}{2} \right\rfloor$ in any case, which shows

$$\text{max.rank}_{\mathbb{H}}(m, 2, n) \geqslant m + \left\lfloor \frac{n}{2} \right\rfloor.$$

2. Since $2m < n$, we have $m \leqslant \left\lfloor \frac{n}{2} \right\rfloor$. Let

$$T := \left(\begin{bmatrix} I_m & O_{m\times m} & O_{m\times(n-2m)} \end{bmatrix}; \begin{bmatrix} O_{m\times m} & I_m & O_{m\times(n-2m)} \end{bmatrix} \right).$$

Since all $2m$ columns of $T_{\leqslant 2m}$ are linearly independent (in the right-$\mathbb{H}$ module), we have

$$\operatorname{rank}_{\mathbb{H}}(T) \geqslant \operatorname{rank}_{\mathbb{H}}(T_{\leqslant 2m}) \geqslant \dim_{\mathbb{H}} C(T_{\leqslant 2m}) = 2m,$$

which shows

$$\text{max.rank}_{\mathbb{H}}(m, 2, n) \geqslant 2m.$$

□

By Proposition 4.3.6, Proposition 2.2.12 and Corollary 4.3.5, we have

Theorem 4.3.7. *Let $m, n \in \mathbb{N}^+$. Then*

$$\text{max.rank}_{\mathbb{H}}(m, 2, n) = \min\left\{ n + \left\lfloor \frac{m}{2} \right\rfloor, m + \left\lfloor \frac{n}{2} \right\rfloor, 2m, 2n \right\}.$$

4.4 Maximal Rank of Higher Quaternionic Tensors

For real and complex tensors with higher orders, some upper bounds were obtained in the past few decades. We now show that some of the results can be extended to the quaternion case. We start with introducing the following lemmas.

Lemma 4.4.1. *Let $A, B, X, Y \in M_n(\mathbb{H})$ be $n \times n$ matrices with entries in $\mathbb{H}$. If X is nonsingular and the eigenvalues of $X^{-1}Y$ are in different conjugacy classes of quaternions, then there exists $r \in \mathbb{R}$ such that*

1. *$A + rX$ is nonsingular;*

2. *The eigenvalues of* $(A+rX)^{-1}(B+rY)$ *are also in different conjugacy classes of quaternions;*

3. $(A+rX)^{-1}(B+rY)$ *is diagonalizable.*

Proof. Let $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be the standard eigenvalues (principle eigenvalues) of $X^{-1}Y$ that are in distinct conjugacy classes. By [71, Corollary 5.1], the eigenvalues of $\chi(X^{-1}Y)$ appear in conjugate pairs (real roots appear twice). So we can denote

$$\begin{aligned} f(\lambda) &= \det\left(\lambda I_{2n} - \chi(X^{-1}Y)\right) \\ &= \prod_{i=1}^{n}(\lambda-\lambda_i)(\lambda-\lambda_i^*) \\ &= \prod_{i=1}^{n}(\lambda-\lambda_i)^2 \\ &= \lambda^{2n} + \sum_{k=0}^{2n-1} a_k\lambda^k. \end{aligned}$$

Let $\epsilon_0 > 0$ be sufficiently small such that the open balls $B(\lambda_i, \epsilon_0)$, $i = 1, 2, ..., n$, containing the roots λ_i's are disjoint. By [64, Theorem 1], there exists $\delta > 0$ such that any degree $2n$ polynomial,

$$\lambda^{2n} + \sum_{k=0}^{2n-1} b_k\lambda^k,$$

with $|a_k - b_k| < \delta$, $k = 0, ..., n-1$, has exactly the same numbers of roots (counting multiplicities) in $B(\lambda_i, \epsilon_0)$, $i = 1, ..., n$. For sufficiently small ϵ, $\epsilon A + X$ is nonsingular and the coefficients of of the polynomial

$$q_\epsilon(\lambda) = \det\left(\lambda I_{2n} - \chi\left((\epsilon A + X)^{-1}(\epsilon B + Y)\right)\right) = z^{2n} + \sum_{k=0}^{2n-1} b_k\lambda^k$$

satisfy $|a_k - b_k| < \delta$. Therefore, $q_\epsilon(\lambda)$ has exactly the same numbers of roots (counting multiplicities) in $B(\lambda_i, \epsilon_0)$, $i = 1, ..., n$. Noticing

$$\begin{aligned} q(\lambda) &= \det\left(\lambda I_{2n} - \chi\left((\epsilon A + X)^{-1}(\epsilon B + Y)\right)\right) \\ &= \det\left(\lambda I_{2n} - \chi\left((\epsilon A + X)^{-1}\epsilon(B + \frac{1}{\epsilon}Y)\right)\right) \\ &= \det\left(\lambda I_{2n} - \chi\left((A + \frac{1}{\epsilon}X)^{-1}(B + \frac{1}{\epsilon}Y)\right)\right) \end{aligned}$$

we see that those roots are also eigenvalues of $\chi\left((A + \frac{1}{\epsilon}X)^{-1}(B + \frac{1}{\epsilon}Y)\right)$. For those eigenvalues that are distinct, their associated eigenvectors are linearly independent. For those real eigenvalue(s) that appear(s) twice, by [71, Corollary 6.3], each of them has two linearly independent eigenvectors. Therefore, $\chi\left((A + \frac{1}{\epsilon}X)^{-1}(B + \frac{1}{\epsilon}Y)\right)$ is diagonalizable, and thus $(A + \frac{1}{\epsilon}X)^{-1}(B + \frac{1}{\epsilon}Y)$ is also diagonalizable. Letting $r = \frac{1}{\epsilon}$, we see that $(A + rX)^{-1}(B + rY)$ satisfies (1),(2) and (3). □

Lemma 4.4.2. *Let $m, n \in \mathbb{N}^+$ with $m \leqslant n$, and $A, B \in M_m(\mathbb{H})$. If A is invertible and $A^{-1}B$ is diagonalizable, then we have*

$$\operatorname{rank}_{\mathbb{H}}\left(\left(\begin{bmatrix} A & X \end{bmatrix}; \begin{bmatrix} B & Y \end{bmatrix}\right)\right) \leqslant n$$

for any $X, Y \in M_{m\times(n-m)}(\mathbb{H})$.

Proof. Let $P \in M_m(\mathbb{H})$ be an invertible matrix such that $D := P^{-1}(A^{-1}B)P$ is a diagonal matrix. Then we have

$$
\begin{aligned}
&\operatorname{rank}_{\mathbb{H}}\left(\begin{bmatrix} A & X \end{bmatrix}; \begin{bmatrix} B & Y \end{bmatrix}\right) \\
=&\operatorname{rank}_{\mathbb{H}}\left(P^{-1}A^{-1}\left(\begin{bmatrix} A & X \end{bmatrix}; \begin{bmatrix} B & Y \end{bmatrix}\right)\begin{bmatrix} P & -A^{-1}X \\ O & I_{n-m} \end{bmatrix}\right) \\
=&\operatorname{rank}_{\mathbb{H}}\left(\left(\begin{bmatrix} I_m & O_{m\times(n-m)} \end{bmatrix}; \begin{bmatrix} D & P^{-1}A^{-1}(Y-BA^{-1}X) \end{bmatrix}\right)\right) \\
\leqslant&\operatorname{rank}_{\mathbb{H}}(I_m; D) + \operatorname{rank}_{\mathbb{H}}\left(O_{m\times(n-m)}; P^{-1}A^{-1}(Y-BA^{-1}X)\right) \\
\leqslant& m + (n-m) = n.
\end{aligned}
$$

□

In 2010, Sumi, Miyazaki and Sakata gave some upper bounds for $m \times p \times n$ complex tensors (see [59, Theorem 1]). We now show that these bounds also work for quaternion tensors.

Theorem 4.4.3. *Let $m, n, p \in N^+$ with $m \leqslant n$. Then we have*

$$\text{max.rank}_{\mathbb{H}}(m, p, n) \leqslant \left\lfloor \frac{(p-1)n}{2} \right\rfloor + m$$

if p is even, and

$$\text{max.rank}_{\mathbb{H}}(m, p, n) \leqslant \left\lfloor \frac{(p-2)n}{2} \right\rfloor + 2m$$

if p is odd.

Proof. For any $T \in \mathbb{H}^{m\times p\times n}$, denote $T = (A_1; A_2; \ldots; A_p)$, where $A_i \in M_{m\times n}(\mathbb{H})$ for $i = 1, 2, ..., p$. Denote $k = \lfloor n/2 \rfloor$. By Theorem 4.3.4, there exist invertible matrices $P_1 \in M_m(\mathbb{H})$ and $Q_1 \in M_n(\mathbb{H})$ such that

$$(A'_1; A'_2) := P_1(A_1; A_2)Q_1 \in \mathbb{C}^{m\times 2\times n}.$$

Then by [58, Corollary 4.9], there exists a tensor

$$C := (C_1, C_2) \in \mathbb{C}^{m\times 2\times n}$$

with $\operatorname{rank}_{\mathbb{H}}(C) \leqslant k$ such that $A_1' - C_1$ and $A_2' - C_2$ are simultaneously equivalent to diagonal matrices, i.e., there exists $P_2 \in M_m(\mathbb{H})$ and $Q_2 \in M_n(\mathbb{H})$ such that

$$P_2(A_1' - C_1; A_2' - C_2)Q_2 = (\begin{bmatrix} D_1 & O_{m\times(n-m)} \end{bmatrix}; \begin{bmatrix} D_2 & O_{m\times(n-m)} \end{bmatrix}),$$

where $D_1, D_2 \in M_m(\mathbb{C})$ are diagonal matrices. This implies

$$(P_2A_1'Q_2 - \begin{bmatrix} D_1 & O_{m\times(n-m)} \end{bmatrix}; P_2A_2'Q_2 - \begin{bmatrix} D_2 & O_{m\times(n-m)} \end{bmatrix}) = P_2CQ_2$$

Let

$$S := P_2P_1TQ_1Q_2 = (B_1; B_2; ...; B_p).$$

If p is even, denote $p = 2q$ for some $q \in \mathbb{N}^+$. Then for each $i \in \{2, ..., q\}$ by Lemma 4.4.1 there exist diagonal matrices $D_{2i-1}, D_{2i} \in M_m(\mathbb{H})$ such that

$$((B_{2i-1})_{\leqslant m} - D_{2i-1})^{-1}((B_{2i})_{\leqslant m} - D_{2i})$$

is diagonalizable, and thus by Lemma 4.4.2 we have

$$\operatorname{rank}_{\mathbb{H}}\left(B_{2i-1} - \begin{bmatrix} D_{2i-1} & O_{m\times(n-m)} \end{bmatrix}; B_{2i} - \begin{bmatrix} D_{2i} & O_{m\times(n-m)} \end{bmatrix}\right) \leqslant n,$$

for $i = 2, ..., q$. Since we have

$$\operatorname{rank}_{\mathbb{H}}\left(B_1 - \begin{bmatrix} D_1 & O_{m\times(n-m)} \end{bmatrix}; B_2 - \begin{bmatrix} D_1 & O_{m\times(n-m)} \end{bmatrix}\right) = \operatorname{rank}_{\mathbb{H}}(P_2CQ_2) \leqslant k,$$

we can now bound the rank of the tensor T in the following

$$\begin{aligned}
\operatorname{rank}_{\mathbb{H}}(T) &= \operatorname{rank}_{\mathbb{H}}(S) \\
&= \operatorname{rank}_{\mathbb{H}}\left(B_1; B_2; B_3; ...; B_p\right) \\
&= \operatorname{rank}_{\mathbb{H}}\left(B_1 - \begin{bmatrix} D_1 & O_{m\times(n-m)} \end{bmatrix}; ...; B_p - \begin{bmatrix} D_p & O_{m\times(n-m)} \end{bmatrix}\right) \\
&\quad + \operatorname{rank}_{\mathbb{H}}\left((D_1; ...; D_p)\right) \\
&\leqslant \sum_{i=1}^{q} \operatorname{rank}_{\mathbb{H}}\left(B_{2i-1} - \begin{bmatrix} D_{2i-1} & O_{m\times(n-m)} \end{bmatrix}; B_{2i} - \begin{bmatrix} D_{2i} & O_{m\times(n-m)} \end{bmatrix}\right) \\
&\quad + \operatorname{rank}_{\mathbb{H}}\left((D_1; ...; D_p)\right) \\
&\leqslant k + (q-1)n + m \\
&= \left\lfloor \frac{(p-1)n}{2} \right\rfloor + m.
\end{aligned}$$

If p is odd, we can use the same approach for the first $p-1$ frontal slices of T. Then we have

$$\begin{aligned}
\operatorname{rank}_{\mathbb{H}}(T) &= \operatorname{rank}_{\mathbb{H}}(A_1; ...; A_p) \\
&\leqslant \operatorname{rank}_{\mathbb{H}}(A_1; ...; A_{p-1}) + \operatorname{rank}_{\mathbb{H}}(A_p) \\
&\leqslant \left\lfloor \frac{((p-1)-1)n}{2} \right\rfloor + m + m \\
&= \left\lfloor \frac{(p-2)n}{2} \right\rfloor + 2m.
\end{aligned}$$

□

In 1979, Atkinson and Stephens ([3, Theorem 2]) showed that

$$\text{max.rank}_{\mathbb{C}}(m, n, mn-k) = mn - k^2 + \text{max.rank}_{\mathbb{C}}(k, k, k^2 - k),$$

for complex 3-tensors. We now show that this result can be extended to the case of complex and quaternionic tensors with higher orders.

Theorem 4.4.4. *If* $k \leqslant n$, $m = \prod_{i=1}^{p} N_i$ *then*

$$\text{max.rank}_{\mathbb{H}}(mn-k, N_1, ..., N_p, n) = m(n-k) + \text{max.rank}_{\mathbb{H}}(mk-k, N_1, ..., N_p, k).$$

Proof. First, we show that

$$\text{max.rank}_{\mathbb{H}}(mn-k, N_1, ..., N_p, n) \leqslant m(n-k) + \text{max.rank}_{\mathbb{H}}(mk-k, N_1, ..., N_P, k).$$

Let $T \in \mathbb{H}^{(mn-k)\times N_1 \times \cdots \times N_p \times n}$ be any tensor. By Theorem 4.2.1, it suffices to show that $R(T)$ is contained in a left $\mathbb{H}$-module generated by $m(n-k)+\text{max.rank}_{\mathbb{H}}(mk-k, N_1, ..., N_P, k)$ simple tensors. Without loss of generality, we may assume $\dim_{\mathbb{H}} R(T) = mn-k$. Let $\tau = (1, 2, ..., p)$ and denote $I = [[N_1, ..., N_p]]$. For each $\sigma \in I$, define left $\mathbb{H}$-modules

$${}_{\mathbb{H}}Y_\sigma = \text{span}_{\mathbb{H}}\left\{S \in \mathbb{H}^{N_1\times\cdots\times N_p\times n} \mid \text{Proj}(\tau, \sigma', S) = \vec{0},\ \text{if } \sigma' \in I \text{ with } \sigma' \neq \sigma\right\}.$$

Then we have $\forall \sigma \in I$, Y_σ only contains simple tensors, and $\dim\left({}_{\mathbb{H}}Y_\sigma\right) = n$. Since

$$\begin{aligned}\dim\left({}_{\mathbb{H}}R(T) \cap {}_{\mathbb{H}}Y_\sigma\right) &= \dim({}_{\mathbb{H}}R(T)) + \dim({}_{\mathbb{H}}Y_\sigma) - \dim\left({}_{\mathbb{H}}R(T) + {}_{\mathbb{H}}Y_\sigma\right)\\ &\geqslant (mn-k) + n - mn\\ &= n-k,\end{aligned}$$

we can find a set of $n-k$ left $\mathbb{H}$-linearly independent simple tensors in each $R(T)\cap Y_\sigma$, which we denote by

$$\mathcal{B}_\sigma = \{T_{1,\sigma}, ..., T_{n-k,\sigma}\},\ \ \forall \sigma \in I.$$

Then $\bigcup_{\sigma\in I} \mathcal{B}_\sigma$ can be extended to a basis of $R(T)$ by adding $(mn-k) - m(n-k) = (m-1)k$ linearly independent tensors, i.e., there exist $A_1, ..., A_{(m-1)k} \in \mathbb{H}^{N_1\times\cdots\times N_p\times n}$

such that

$$_{\mathbb{H}}R(T) = \text{span}_{\mathbb{H}}\left(\bigcup_{\sigma \in I} \mathcal{B}_\sigma \cup \{A_1, .., A_{(m-1)k}\}\right).$$

Define left $\mathbb{H}$-submodules of $_{\mathbb{H}}\mathbb{H}^{1\times n}$ by

$$_{\mathbb{H}}V_\sigma = \text{span}_{\mathbb{H}}\left\{\text{Proj}(\tau, \sigma, T_{i,\sigma}) \mid i = 1, ..., n-k, \right\}, \forall \sigma \in I.$$

Since for each $\sigma \in I$, $B_{1,\sigma}, ..., B_{n-k,\sigma}$ are left $\mathbb{H}$-linearly independent, and

$$\text{Proj}(\tau, \sigma', T_{i,\sigma}) = \vec{0},\ \forall \sigma' \in I \text{ with } \sigma' \neq \sigma,\ \forall i \in \{1, 2, ..., n-k\},$$

we can see that $\text{Proj}(\tau, \sigma, T_{1,\sigma}), ..., \text{Proj}(\tau, \sigma, T_{n-k,\sigma})$ are also left $\mathbb{H}$-linearly independent, which implies

$$\dim {_{\mathbb{H}}V_\sigma} = n-k,\ \forall \sigma \in I.$$

Thus, there exists a k-dimensional left $\mathbb{H}$-submodule $_{\mathbb{H}}U$ of $_{\mathbb{H}}\mathbb{H}^{1\times n}$ such that

$$_{\mathbb{H}}U + {_{\mathbb{H}}V_\sigma} = {_{\mathbb{H}}\mathbb{H}^{1\times n}},\ \forall \sigma \in I.$$

Therefore, for any $\sigma \in I$ and $j \in \{1, 2, ..., (m-1)k\}$, we can write

$$\text{Proj}(\tau, \sigma, A_j) = \vec{u}_{\sigma,j} + \sum_i a_{ij,\sigma}\text{Proj}(\tau, \sigma, T_{i,\sigma}),$$

for some $\vec{u}_{\sigma,j} \in U$ and $a_{ij,\sigma} \in \mathbb{H}$. For each $j = 1, 2, ..., (m-1)k$, let

$$A'_j = A_j - \sum_{i,\sigma} a_{ij,\sigma} T_{i,\sigma}.$$

Then we have

$$\text{Proj}(\tau, \sigma, A'_j) \in U,\ \forall \sigma \in I\ ,\ \forall j \in \{1, 2, ..., (m-1)k\},$$

Since $U \leqslant \mathbb{H}^{1\times n}$ and $\dim({_{\mathbb{H}}U}) = k$ Then there exists $n-k$ linearly independent vectors $\vec{w}_1, ..., \vec{w}_{n-k}$ such that

$$\vec{u}^T \cdot \vec{w}_i = 0,\ \forall \vec{u}^T \in U,\ \forall i \in \{1, 2, ..., n-k\},$$

Let $Q \in \mathbb{H}^{n\times n}$ be a non-singular matrix whose last $n-k$ columns are $\vec{w}_1, ..., \vec{w}_{n-k}$. Then, by Corollary 4.1.7 we have

$$\text{Proj}(\tau, \sigma, A'_j Q) = \text{Proj}(\tau, \sigma, A'_j)Q = \begin{bmatrix} *_{1\times k} & O_{1\times(n-k)} \end{bmatrix},$$

for any $\sigma \in I$ and $j \in \{1, 2, ..., (m-1)k\}$. This implies

$$C_i\left(A'_j Q\right) = \mathcal{O},\ \forall j \in \{1, ..., (m-1)k\},\ \forall i \in \{k+1, ..., n\}$$

which means $A'_j Q$ can be viewed as $\left(A'_j Q\right)_{C \leqslant k}$, an $N_1 \times \cdots \times N_p \times k$ tensor, for $j = 1, 2, ..., (m-1)k$. Therefore, we have

$$\begin{aligned} \text{rank}_{\mathbb{H}}\left(\sum_{j=1}^{(m-1)k} \vec{e}_j \otimes A'_j\right) &= \text{rank}_{\mathbb{H}}\left(\left(\sum_{j=1}^{(m-1)k} \vec{e}_j \otimes A'_j\right) Q\right) \\ &= \text{rank}_{\mathbb{H}}\left(\sum_{j=1}^{(m-1)k} \vec{e}_j \otimes A'_j Q\right) \\ &= \text{rank}_{\mathbb{H}}\left(\sum_{j=1}^{(m-1)k} \vec{e}_j \otimes \left(A'_j Q\right)_{C \leqslant k}\right) \\ &\leqslant \text{max.rank}_{\mathbb{H}}\left((m-1)k, N_1, ..., N_p, k\right), \end{aligned}$$

which implies that $A'_1, ..., A'_{(m-1)k}$ are contained in a left $\mathbb{H}$-module generated by $\text{max.rank}_{\mathbb{H}}\left((m-1)k, N_1, ..., N_p, k\right)$ simple tensors. Since we have, by the replacement theorem, that

$${}_{\mathbb{H}}R(T) = \text{span}_{\mathbb{H}}\left(\bigcup_{\sigma \in I} \mathcal{B}_\sigma \cup \{A'_1, .., A'_{(m-1)k}\}\right).$$

and the fact that each $\mathcal{B}_\sigma$ is a set of $n-k$ simple tensors. We see that $R(T)$ can be generated by $m(n-k) + \text{max.rank}_{\mathbb{H}}\left((m-1)k, N_1, ..., N_p, k\right)$ simple tensors.

Therefore, we have

$$\operatorname{rank}_{\mathbb{H}}(T) \leqslant m(n-k) + \max.\operatorname{rank}_{\mathbb{H}}\left((m-1)k, N_1, ..., N_p, k\right).$$

Next, we show that

$$\max.\operatorname{rank}_{\mathbb{H}}(mn-k, N_1, ..., N_p, n) \geqslant m(n-k) + \max.\operatorname{rank}_{\mathbb{H}}(mk-k, N_1, ..., N_P, k).$$

Let $T, S \in \mathbb{H}^{mn-k, N_1\times\cdots\times N_p, n}$ be a tensor such that both $T^{\leqslant mk-k}_{\leqslant k} \in \mathbb{H}^{mk-k, N_1\times\cdots\times N_p, k}$ and ${}^{mk-k<}_{k<}S \in \mathbb{H}^{mk-k, N_1\times\cdots\times N_p, k}$ attain the maximal rank with ${}^{mk-k<}_{k<}T = \mathbf{0}$ and $S^{\leqslant mk-k}_{\leqslant k} = \mathbf{0}$. Therefore, we have

$$\operatorname{rank}_{\mathbb{H}}(T^{\leqslant mk-k}_{\leqslant k}) = \max.\operatorname{rank}_{\mathbb{H}}((m-1)k, N_1, ..., N_p, k)$$

and

$$\operatorname{rank}_{\mathbb{H}}\left({}^{mk-k<}_{k<}S\right) = \max.\operatorname{rank}_{\mathbb{H}}((m(n-k), N_1, ..., N_p, n-k),$$

which implies $\dim_{\mathbb{H}} R(T) = (m-1)k$ and $\dim_{\mathbb{H}} R(S) = m(n-k)$. Denote $T' = T+S$. Noticing $\dim_{\mathbb{H}} R(T') = \dim_{\mathbb{H}} R(T) + \dim_{\mathbb{H}} R(S)$, by Theorem 4.2.5, we have

$$\begin{aligned}
\operatorname{rank}_{\mathbb{H}}(T') &\geqslant \min_{B\in\mathbb{H}^{m\times(N_1-m)}} \operatorname{rank}_{\mathbb{H}}\left((T')^{\leqslant mk-k} + B\left({}^{mk-k<}(T')\right)\right) \\
&\qquad + \dim_{\mathbb{H}} R(T') - (mk-k) \\
&\geqslant \min_{B\in\mathbb{H}^{m\times(N_1-m)}} \operatorname{rank}_{\mathbb{H}}\left(\left((T')^{\leqslant mk-k} + B\left({}^{mk-k<}(T')\right)\right)^{\leqslant mk-k}\right) \\
&\qquad + \dim_{\mathbb{H}} R(T') - (mk-k) \\
&= \operatorname{rank}_{\mathbb{H}}\left((T')^{\leqslant mk-k}\right)^{\leqslant mk-k} + \dim_{\mathbb{H}} R(T) + \dim_{\mathbb{H}} R(S) - (mk-k) \\
&= \operatorname{rank}_{\mathbb{H}}\left((T')^{\leqslant mk-k}\right) + (m-1)k + m(n-k) - (mk-k) \\
&= \max.\operatorname{rank}_{\mathbb{H}}((m-1)k, N_1, ..., N_p, k) + m(n-k).
\end{aligned}$$

□

As a result of Theorem 4.4.4, we have the following corollary for the maximal rank of quaternionic tensors with order 3.

Corollary 4.4.5. *If $k \leqslant n$, then we have*

$$\text{max.rank}_{\mathbb{H}}(mn-k,m,n) = m(n-k) + \text{max.rank}_{\mathbb{H}}(mk-k,m,k).$$

Meanwhile, we have also obtained a generalized version of [3, Theorem 2].

Corollary 4.4.6. *If $k \leqslant m_1 \leqslant m_2 \leqslant ... \leqslant m_p$, then we have*

$$\text{max.rank}_{\mathbb{C}}(\prod_{i=1}^{p} m_i - k, m_1, m_2, ..., m_p) = \prod_{i=1}^{p} m_i - k^p + \text{max.rank}_{\mathbb{C}}(k^p - k, k, ..., k).$$

Proof. Noticing that replacing $\mathbb{H}$ with $\mathbb{C}$ doesn't affect the proof of Theorem 4.4.4, we can repeat using Theorem 4.4.4 p times to prove the desired equality above. □

Example 4.4.7. By letting $k = 1, 2$ in Corollary 4.4.6, we have

$$\text{max.rank}_{\mathbb{C}}(\prod_{i=1}^{p} m_i - 1, m_1, m_2, ..., m_p) = \prod_{i=1}^{p} m_i - 1$$

as well as

$$\text{max.rank}_{\mathbb{C}}(\prod_{i=1}^{p} m_i - 2, m_1, m_2, ..., m_p) = \prod_{i=1}^{p} m_i - 2^p + \text{max.rank}_{\mathbb{C}}(2^p - 2, 2, ..., 2).$$

In 1978, Howell proved the following statements over a Principle Ideal Domain (see [30, Theorems 7 and 8] and [60, Proposition 5.4 and Theorem 5.12]). We now show that the result can also be extended to the quaternion case.

Theorem 4.4.8. *Let $m, n, N_1, ..., N_p \in \mathbb{N}^+$. Then we have*

$$\text{max.rank}_{\mathbb{H}}(m, N_1, N_2, ..., N_p, n) \leqslant \prod_{i=1}^{p} N_i + \text{max.rank}_{\mathbb{H}}\left(m-1, N_1, N_2, ., , , .N_p, n-1\right)$$

Proof. Let $\sigma_0 = (2, 3, 4, ..., p-1) \in [[\mathcal{I}_{p-1} \leqslant p+1]]$. For any $T \in \mathbb{H}^{n \times N_1 \times \cdots \times N_p \times m}$ and $\tau \in [[N_1, ..., N_p]]$, we have $\text{Proj}(\sigma_0, \tau, T) \in H^{m \times n}$, which can be viewd as an $m \times n$ quaternioinic matrix. Without loss of generality, we can assume its (m, n) entry is non-zero for all $\tau \in [[N_1, ..., N_p]]$, i.e.,

$$\text{Proj}(\sigma_0, \tau, T)_{mn} \neq 0, \quad \forall \tau \in [[N_1, ..., N_p]].$$

For each $\tau \in [[N_1, ..., N_p]]$, define

$$W_\tau = \text{Proj}(\sigma_0, \tau, T)_{mn}^{-1} \cdot R_m\left(\text{Proj}(\sigma_0, \tau, T)\right) \otimes C_n\left(\text{Proj}(\sigma_0, \tau, T)\right),$$

and let S be the tensor uniquely determined by

$$\{W_\tau \mid \tau \in [[N_1, ..., N_p]]\}.$$

Then we have

$$^{m-1<}(T-S) = 0, \quad {}_{n-1<}(T-S) = 0,$$

which implies

$$\text{rank}_{\mathbb{H}}(T-S) \leqslant \text{max.rank}_{\mathbb{H}}\left(m-1, N_1, N_2, ., , , .N_p, n-1\right).$$

Noticing

$$\text{rank}_{\mathbb{H}}(S) = \prod_{i=1}^{p} N_i,$$

we have

$$\begin{aligned}\operatorname{rank}_{\mathbb{H}}(T) &\leqslant \operatorname{rank}_{\mathbb{H}}(T-S)+\operatorname{rank}_{\mathbb{H}}(S)\\ &\leqslant \prod_{i=1}^{p} N_i + \text{max.rank}_{\mathbb{H}}\left(m-1, N_1, N_2, ., , , .N_p, n-1\right).\end{aligned}$$

□

Now, we can bound the rank of $n \times n \times n$ quaternionic tensors.

Corollary 4.4.9. $\text{max.rank}_{\mathbb{H}}(n,n,n) \leqslant \lceil 3n^2/4 \rceil$.

Proof. If we apply Theorem 4.4.8 $\lceil n/2 \rceil$ times for $p = 1$, then we have

$$\begin{aligned}\text{max.rank}_{\mathbb{H}}(n,n,n) &\leqslant n\lceil n/2\rceil + \text{max.rank}_{\mathbb{H}}(\lfloor n/2\rfloor, n, \lfloor n/2\rfloor)\\ &\leqslant n\lceil n/2\rceil + \lfloor n/2\rfloor^2 = \left\lceil 3n^2/4\right\rceil.\end{aligned}$$

□

In 1979, Atkinson and Stephens ([3]) stated without proof that any $n \times 3 \times n$ complex tensors have rank no more than $2n-1$. In 2010, Sumi et al. ([59, Theorem 5]) proved Atkinson and Stephens' assertion and extended their result over the real number field when n is odd . We now give a similar uppper bound for $n \times 3 \times n$ quaternion tensors.

Theorem 4.4.10. $\text{max.rank}_{\mathbb{H}}(n,3,n) \leqslant 2n$.

Proof. Let $T = (A; B; C) \in \mathbb{H}^{n\times 3\times n}$. Without loss of generality, we can assume

$$\operatorname{rank}_{\mathbb{H}}(A) \geqslant \operatorname{rank}_{\mathbb{H}}(B) \geqslant \operatorname{rank}_{\mathbb{H}}(C).$$

There exist invertible matrices $P_1, P_2 \in M_n(\mathbb{H})$ such that

$$P_1 A P_2 = I_r,$$

where $r = \text{rank}_{\mathbb{H}}(A) \leqslant n$. Then we have

$$T_1 := P_1 T P_2 = (P_1 A P_2; P_1 B P_2; P_2 C P_2) = (I_r; P_1 B P_2; P_1 C P_2).$$

Let $X = I_n$ and $Y = \text{diag}(y_1, ..., y_n)$ where $y_1, ..., y_n \in \mathbb{R}$ are distinct. By Lemma 4.4.1, there exists $r \in \mathbb{R}$ such that $(P_1 B P_2 + rX)$ is nonsingular and

$$(P_1 B P_2 + rX)^{-1}(P_1 C P_2 + rY)$$

is diagonalizable, i.e., there exist $Q \in M_n(\mathbb{H})$ such that

$$Q^{-1}(P_1 B P_2 + X)^{-1}(P_1 C P_2 + Y)Q = \text{diag}(\lambda_1, ..., \lambda_n).$$

Denote $S_1 = (I_r, X, Y)$, then we have

$$\begin{aligned} T_2 :=& Q^{-1}(P_1 B P_2 + rX)^{-1}(T_1 - S_1)Q \\ =& Q^{-1}(P_1 B P_2 + rX)^{-1}\Big(O_{n\times n};\ (P_1 B P_2 + rX);\ (P_1 C P_2 + rY)\Big)Q \\ =& Q^{-1}\Big(O_{n\times n};\ I_n;\ (P_1 B P_2 + rX)^{-1}(P_1 C P_2 + rY)\Big)Q \\ =& \Big(O_{n\times n};\ I_n;\ Q^{-1}(P_1 B P_2 + rX)^{-1}(P_1 C P_2 + rY)Q\Big) \\ =& \Big(O_{n\times n};\ I_n;\ \text{diag}(\lambda_1, ..., \lambda_n)\Big). \end{aligned}$$

Therefore, we have $\text{rank}_{\mathbb{H}}(S_1) \leqslant n$ and $\text{rank}_{\mathbb{H}}(T_2) = n$, which implies

$$\begin{aligned}
\text{rank}_{\mathbb{H}}(T) &= \text{rank}_{\mathbb{H}}(T_1) \\
&\leqslant \text{rank}_{\mathbb{H}}(S_1) + \text{rank}_{\mathbb{H}}(T_1 - S_1) \\
&\leqslant \text{rank}_{\mathbb{H}}(S_1) + \text{rank}_{\mathbb{H}}(T_2) \\
&\leqslant n + n = 2n.
\end{aligned}$$

□

In [60], it is shown for $2 \times 2 \times 2 \times 2$ tensors that the maximal rank is 5 over the real field $\mathbb{R}$ while the maximal rank is only 4 over the complex field $\mathbb{C}$. Now, we show that the maximal rank of $2 \times 2 \times 2 \times 2$ quaternionic tensors is bounded by 4 and 5.

Theorem 4.4.11. $4 \leqslant \text{max.rank}_{\mathbb{H}}(2, 2, 2, 2) \leqslant 5$.

Proof. For any $A \in \mathbb{H}^{2\times 2\times 2\times 2}$, denote

$$T = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|cc} a_{11} & a_{12} & b_{11} & a_{12} \\ a_{21} & a_{22} & b_{21} & a_{22} \\ \hline c_{11} & c_{12} & d_{11} & d_{12} \\ c_{21} & c_{22} & d_{21} & d_{22} \end{array}\right]$$

and the action of $(\text{GL}(2, \mathbb{H}) \times (\text{GL}(2, \mathbb{R}) \times (\text{GL}(2, \mathbb{R}) \times (\text{GL}(2, \mathbb{H}))$ be denoted by

$$(A_1, A_2, A_3, A_4) \cdot T = A_2 \star \left[\begin{array}{c|c} A_1 A A_4 & A_1 B A_4 \\ \hline A_1 C A_4 & A_1 D A_4 \end{array}\right] \star A_3$$

If any of the sub-tensors

$$\begin{bmatrix} A & B \end{bmatrix}, \begin{bmatrix} C & D \end{bmatrix}, \begin{bmatrix} A \\ C \end{bmatrix}, \begin{bmatrix} B \\ D \end{bmatrix},$$

has rank 2, then we have $\mathrm{rank}_{\mathbb{H}}(T) \leqslant 2+3=5$. Otherwise, we can assume without generality that $\mathrm{rank}_{\mathbb{H}}([A\ B]) = 3$, $\mathrm{rank}_{\mathbb{H}}(A) = 2$ and B is not diagonalizable. Let $P \in M_2(\mathbb{H})$ be an invertible matrix such that

$$P^{-1}(A^{-1}B)P = J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

where J is the Jordan canonical form of $A^{-1}B$. Then we have

$$\begin{aligned} T_1 :&= (P^{-1}A^{-1}, I_2, I_2, P) \cdot T \\ &= I_2 \star \left[\begin{array}{c|c} (P^{-1}A^{-1})AP & (P^{-1}A^{-1})BP \\ \hline (P^{-1}C^{-1})AP & (P^{-1}A^{-1})DP \end{array}\right] \star I_2 \\ &= \left[\begin{array}{c|c} I_2 & J \\ \hline P^{-1}C^{-1}AP & P^{-1}A^{-1}DP \end{array}\right] \\ &= \left[\begin{array}{cc|cc} 1 & 0 & \lambda & 1 \\ 0 & 1 & 0 & \lambda \\ \hline c'_{11} & c'_{12} & d'_{11} & d'_{12} \\ c'_{21} & c'_{22} & d'_{21} & d'_{22} \end{array}\right]. \end{aligned}$$

and $\mathrm{rank}_{\mathbb{H}}(S) = \mathrm{rank}_{\mathbb{H}}(T)$. Choose $\lambda_0 \in \mathbb{H}$ with $Re(\lambda_0) \neq Re(\lambda)$ and denote

$$B' = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda_0 \end{bmatrix}, C' = \begin{bmatrix} c'_{11} & c'_{12} \\ c'_{21} & c'_{22} \end{bmatrix}, D' = \begin{bmatrix} d'_{11} & d'_{12} \\ d'_{21} & d'_{22} \end{bmatrix},$$

and

$$S_1 = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \lambda-\lambda_0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Then we have

$$T_1 - S_1 = \left[\begin{array}{c|c} I_2 & B' \\ \hline C' & D' \end{array}\right].$$

By Lemma 4.4.1, there exists $r \in \mathbb{R}$ such that $(C'+rI_2)^{-1}(D'+rB')$ is diagonalizable. Let $Q \in M_2(\mathbb{H})$ such that

$$Q^{-1}(C'+rI_2)^{-1}(D'+rB')Q = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Therefore, we have

$$S := (I_2, I_2 + rE_{21}, I_2, I_2) \cdot (T_1 - S_1) = \left[\begin{array}{c|c} I_2 & B' \\ \hline C' + rI_2 & D' + rB' \end{array}\right].$$

Since

$$\operatorname{rank}_{\mathbb{H}}(S^{\leqslant 1}) = \operatorname{rank}_{\mathbb{H}}([I_2\ B']) = \operatorname{rank}_{\mathbb{H}}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & \lambda_0 \end{bmatrix}\right) = 2$$

and

$$
\begin{aligned}
\operatorname{rank}_{\mathbb{H}}({}^{1<}S) &= \operatorname{rank}_{\mathbb{H}}([C' + rI_2 \;\; D' + rB']) \\
&= \operatorname{rank}_{\mathbb{H}}([I_2 \;\; (C' + rI_2)^{-1}D' + rB']) \\
&= \operatorname{rank}_{\mathbb{H}}([Q^{-1}I_2Q \;\; Q^{-1}(C' + rI_2)^{-1}D' + rB'Q]) \\
&= \operatorname{rank}_{\mathbb{H}}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \; \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\right) = 2,
\end{aligned}
$$

we have

$$
\begin{aligned}
\operatorname{rank}_{\mathbb{H}}(T) &\leqslant \operatorname{rank}_{\mathbb{H}}(S_1) + \operatorname{rank}_{\mathbb{H}}(T_1 - S_1) \\
&= \operatorname{rank}_{\mathbb{H}}(S_1) + \operatorname{rank}_{\mathbb{H}}(S) \\
&\leqslant \operatorname{rank}_{\mathbb{H}}(S_1) + \operatorname{rank}_{\mathbb{H}}(S^{\leqslant 1}) + \operatorname{rank}_{\mathbb{H}}({}^{1<}S) \\
&= 1 + 2 + 2 = 5.
\end{aligned}
$$

On the other hand, consider the $2 \times 2 \times 2 \times 2$ tensor T determined by the horizontal slices:

$$
R_1(T) = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right), \; R_2(T) = \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).
$$

Since $R_1(T)$ and $R_2(T)$ are linear independent, we have $\dim_{\mathbb{H}} R(T) = 2$. For any scalar $q \in \mathbb{H}$, we know by Lemma 3.3.3 that the quaternionic tensor

$$
R_1(T) + qR_2(T) = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} q & 1 \\ 0 & q \end{bmatrix}\right)
$$

has rank 3. Denote the left $\mathbb{H}$-module generated by $\{R_2(T)\}$ by M. Again, by Lemma 3.3.3, we have

$$
\operatorname{rank}_{\mathbb{H}}(T) \geqslant \min_{A \in M} \operatorname{rank}_{\mathbb{H}} (R_1(T) + A) + \dim_{\mathbb{H}} R(T) - 1 = 3 + 2 - 1 = 4.
$$

Therefore, we have $4 \leqslant \text{max.rank}_{\mathbb{H}}(2, 2, 2, 2) \leqslant 5$. $\square$

Chapter 5

Future Work

Firstly, quaternionic 3-tensors with 3 frontal slices, namely $m \times 3 \times n$ tensors over the quaternions, would be a good direction for future work. In [58], the canonical forms and maximal rank of $m \times 2 \times n$ complex tensors were thoroughly studied. Corresponding results for the $m \times 2 \times n$ quaternion tensors were also obtained in my recent research (Theorem 4.3.7). However, the maximal rank and canonical forms of $m \times 3 \times n$ tensors still remain unknown. Sumi et al. ([59]) showed that the maximal rank of $n \times 3 \times n$ complex tensor is no more than $2n - 1$. In this thesis, it is shown that the maximal rank of $n \times 3 \times n$ quaternionic tensors is no more than $2n$ (Theorem 4.4.10), a slightly weaker bound. Since we have

$$\text{max.rank}_{\mathbb{H}}(m, 3, n) = \text{max.rank}_{\mathbb{C}}(m, 3, n)$$

for any $2 \leqslant m, n \leqslant 3$, and

$$\text{max.rank}_{\mathbb{H}}(m, 2, n) = \text{max.rank}_{\mathbb{C}}(m, 2, n)$$

for any $m, n \in \mathbb{N}^+$, it is natural to conjecture

$$\text{max.rank}_{\mathbb{H}}(m, p, n) = \text{max.rank}_{\mathbb{C}}(m, p, n)$$

for any $m, n, p \in \mathbb{N}^+$. Therefore, the first step would be to study $m \times 3 \times n$ quaternionic tensors.

Moreover, it would be interesting to study how the underlying base can change the tensor rank. We know that the maximal rank of any real $2 \times 2 \times 2 \times 2$ tensor is 5, while the maximal rank for a complex tensor of the same size is 4 (these bounds were shown earlier in [38] and [6] respectively). In this thesis, there is an example of a complex tensor that has a strictly smaller rank over the quaternions than the complex field (example 3.5.3). This example shows that a complex tensor could have smaller decomposition into quaternion tensors. Therefore, considering decomposition of complex tensors into quaternion tensors is another research direction.

Last but not least, apart from quaternionic 3-tensors, the tensor rank and tensor decomposition problems for higher quaternion tensors are also worth considering. For example, it is known that $\text{max.rank}_{\mathbb{R}}(2, 2, 2, 2) = 5$ and $\text{max.rank}_{\mathbb{C}}(2, 2, 2, 2) = 4$. In the quaternion case, however, the maximal rank of $2 \times 2 \times 2 \times 2$ tensors still remains unknown. Although determining the rank of higher quaternion tensors is considerably difficult, it is still a good idea to consider some upper bounds of the maximal rank of those tensors.

Bibliography

[1] C. J. Appellof and E. R. Davidson. Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents. *Analytical Chemistry*, 53:2053–2056, 1981.

[2] M. Atkinson and S. Lloyd. Bounds on the ranks of some 3-tensors. *Linear Algebra and its Applications*, 31:19–31, 1980.

[3] M. Atkinson and N. Stephens. On the maximal multiplicative complexity of a family of bilinear forms. *Linear Algebra and its Applications*, 27:1–8, 1979.

[4] M. R. Bremner and J. Hu. On kruskal's theorem that every $3\times 3\times 3$ array has rank at most 5. *Linear Algebra and its Applications*, 439(2):401–421, 2013.

[5] R. Bro. PARAFAC. Tutorial and applications. *Chemometrics and Intelligent Laboratory Systems*, 38:149–171, 10 1997.

[6] J.-L. Brylinski. Algebraic measures of entanglement. *Mathematics of Quantum Computation*, pages 3–24, 01 2002.

[7] J. D. Carroll and J.-J. Chang. Analysis of individual differences in multidimensional scaling via an n-way generalization of "eckart-young" decomposition. *Psychometrika*, 35(3):283–319, 1970.

[8] R. B. Cattell. "Parallel proportional profiles" and other principles for determining the choice of factors by rotation. *Psychometrika*, 9(4):267–283, 1944.

[9] R. B. Cattell. The three basic factor-analytic research designs—their interrelations and derivatives. *Psychological Bulletin*, 49(5):499, 1952.

[10] P. Comon. Independent component analysis, a new concept? *Signal processing*, 36(3):287–314, 1994.

[11] P. Comon. Tensor decompositions. *Mathematics in Signal Processing V*, pages 1–24, 2002.

[12] P. Comon, X. Luciani, and A. L. De Almeida. Tensor decompositions, alternating least squares and other tales. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 23(7-8):393–405, 2009.

[13] P. Comon and B. Mourrain. Decomposition of quantics in sums of powers of linear forms. *Signal Processing*, 53(2-3):93–107, 1996.

[14] P. Comon, J. M. Ten Berge, L. De Lathauwer, and J. Castaing. Generic and typical ranks of multi-way arrays. *Linear Algebra and its Applications*, 430(11-12):2997–3007, 2009.

[15] J. H. Conway and D. A. Smith. On quaternions and octonions. *On Quaternions and Octonions*, 1 2003.

[16] M. Copper. On the gauss decomposition of a matrix. *Linear Algebra and its Applications*, 177:175–189, 12 1992.

[17] L. De Lathauwer. A survey of tensor methods. In *2009 IEEE International Symposium on Circuits and Systems*, pages 2773–2776. IEEE, 2009.

[18] L. De Lathauwer and J. Castaing. Tensor-based techniques for the blind separation of ds–cdma signals. *Signal Processing*, 87(2):322–336, 2007.

[19] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

[20] L. De Lathauwer, B. De Moor, and J. Vandewalle. On the best rank-1 and rank-$(r_1, r_2, ..., r_n)$ approximation of higher-order tensors. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1324–1342, 2000.

[21] 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*, 30(3):1084–1127, 2008.

[22] A. Doostan, G. Iaccarino, and N. Etemadi. A least-squares approximation of high-dimensional uncertain systems. *Annual Research Briefs, Center for Turbulence Research, Stanford University*, pages 121–132, 2007.

[23] D. Y. Grigoryev. *Some new bounds on tensor rank.* Leningrad Department, Steklov Math. Inst., USSR Acad. of Science, 1978.

[24] J. Håstad. Tensor rank is np-complete. In *International Colloquium on Automata, Languages, and Programming*, pages 451–460. Springer, 1989.

[25] Z.-H. He, C. Navasca, and Q.-W. Wang. Tensor decompositions and tensor equations over quaternion algebra. *arXiv:1710.07552*, 2017.

[26] Z.-H. He, Q.-W. Wang, and Y. Zhang. A simultaneous decomposition for seven matrices with applications. *Journal of Computational and Applied Mathematics*, 349:93–113, 2019.

[27] R. Henrion. Body diagonalization of core matrices in three-way principal components analysis: Theoretical bounds and simulation. *Journal of Chemometrics*, 7:477–494, 11 1993.

[28] F. L. Hitchcock. The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematics and Physics*, 6(1-4):164–189, 1927.

[29] F. L. Hitchcock. Multiple invariants and generalized rank of a p-way matrix or tensor. *Journal of Mathematics and Physics*, 7:39–79, 4 1928.

[30] T. D. Howell. Global properties of tensor rank. *Linear Algebra and its Applications*, 22:9–23, 1978.

[31] L. Huang. Quaternion equations and quaternion polynomial matrices. *Master's Thesis at the University of Manitoba*, 2013.

[32] J. JáJá. Optimal evaluation of pairs of bilinear forms. *SIAM Journal on Computing*, 8(3):443–462, 1979.

[33] Y.-D. Kim and S. Choi. Nonnegative tucker decomposition. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007.

[34] D. E. Knuth. Seminumerical algorithms. *The Art of Computer Programming*, 2, 1997.

[35] T. G. Kolda. Orthogonal tensor decompositions. *SIAM Journal on Matrix Analysis and Applications*, 23(1):243–255, 2001.

[36] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[37] T. G. Kolda, B. W. Bader, and J. P. Kenny. Higher-order web link analysis using multilinear algebra. In *Fifth IEEE International Conference on Data Mining (ICDM'05)*, pages 8–pp. IEEE, 2005.

[38] X. Kong and Y.-L. Jiang. A note on the ranks of $2\times 2\times 2$ and $2\times 2\times 2\times 2$ tensors. *Linear and Multilinear Algebra*, 61(10):1348–1362, 2013.

[39] J. B. Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications*, 18(2):95–138, 1977.

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

[41] J. B. Kuipers. *Quaternions and Rotation Sequences: a Primer with Applications to Orbits, Aerospace, and Virtual Reality*. Princeton university press, 1999.

[42] J. M. Landsberg. The border rank of the multiplication of 2×2 matrices is seven. *Journal of the American Mathematical Society*, 19(2):447–459, 2006.

[43] J. M. Landsberg. *Tensors: Geometry and Applications*, volume 381. 2012.

[44] N. Le Bihan, S. Miron, and J. I. Mars. Music algorithm for vector-sensors array using biquaternions. *IEEE Transactions on Signal Processing*, 55(9):4523–4533, 2007.

[45] Y. Liang, S. Da Silva, and Y. Zhang. The tensor rank problem over the quaternions. *Linear Algebra and its Applications*, 620:37–60, 2021.

[46] L. Mirsky. *An Introduction to Linear Algebra*. Courier Corporation, 2012.

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

[48] S. Ragnarsson and C. F. Van Loan. Block tensor unfoldings. *SIAM Journal on Matrix Analysis and Applications*, 33:149–169, 1 2012.

[49] L. Rodman. *Topics in Quaternion Linear Algebra*. Princeton Series in Applied Mathematics, 2014.

[50] S. J. Sangwine and D. Alfsmann. Determination of the biquaternion divisors of zero, including the idempotents and nilpotents. *Advances in Applied Clifford Algebras 2010 20:2*, 20:401–410, 1 2010.

[51] S. J. Sangwine, T. A. Ell, and N. L. Bihan. Fundamental representations and algebraic properties of biquaternions or complexified quaternions. *Advances in Applied Clifford Algebras 2010 21:3*, 21:607–636, 11 2010.

[52] B. Savas and L. Eldén. Handwritten digit classification using higher order singular value decomposition. *Pattern Recognition*, 40(3):993–1003, 2007.

[53] H. D. Schutte and J. Wenzel. Hypercomplex numbers in digital signal processing. *Proceedings - IEEE International Symposium on Circuits and Systems*, 2:1557–1560, 1990.

[54] A. Shashua and T. Hazan. Non-negative tensor factorization with applications to statistics and computer vision. In *Proceedings of the 22nd International Conference on Machine Learning*, pages 792–799, 2005.

[55] A. K. Smilde, P. Geladi, and R. Bro. *Multi-way Analysis: Applications in the Chemical Sciences*. John Wiley & Sons, 2005.

[56] A. K. Smilde, Y. Wang, and B. R. Kowalski. Theory of medium-rank second-order calibration with restricted-tucker models. *Journal of Chemometrics*, 8:21–36, 1 1994.

[57] V. Strassen. Gaussian elimination is not optimal. *Numerische Mathematik*, 13(4):354–356, 1969.

[58] T. Sumi, M. Miyazaki, and T. Sakata. Rank of 3-tensors with 2 slices and kronecker canonical forms. *Linear Algebra and its Applications*, 431(10):1858–1868, 2009.

[59] T. Sumi, M. Miyazaki, and T. Sakata. About the maximal rank of 3-tensors over the real and the complex number field. *Annals of the Institute of Statistical Mathematics*, 62(4):807–822, 2010.

[60] T. Sumi, M. Miyazaki, and T. Sakata. *Algebraic and Computational Aspects of Real Tensor Ranks.* SpringerBriefs in Statistics. JSS Research Series in Statistics. Springer, [Tokyo], 2016.

[61] T. Sumi, T. Sakata, and M. Miyazaki. Rank of tensors with size $2 \times 2 \times \cdots \times 2$. *arXiv:1306.0708*, 2013.

[62] J. Sun, S. Papadimitriou, and S. Y. Philip. Window-based tensor analysis on high-dimensional and multi-aspect streams. In *Sixth International Conference on Data Mining (ICDM'06)*, pages 1076–1080. IEEE, 2006.

[63] L. R. Tucker. Some mathematical notes on three-mode factor analysis. *Psychometrika 1966 31:3*, 31:279–311, 9 1966.

[64] D. J. Uherka and A. M. Sergott. On the continuous dependence of the roots of a polynomial on its coefficients. *Source: The American Mathematical Monthly*, 84:368–370, 1977.

[65] M. A. O. Vasilescu and D. Terzopoulos. Multilinear image analysis for facial recognition. In *Object Recognition Supported by User Interaction for Service Robots*, volume 2, pages 511–514. IEEE, 2002.

[66] M. A. O. Vasilescu and D. Terzopoulos. Multilinear subspace analysis of image ensembles. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 2, pages II–93. IEEE, 2003.

[67] J. Voight. *Quaternion Algebras*, volume 288. Springer Nature, 2021.

[68] Q.-W. Wang, Z.-H. He, and Y. Zhang. Constrained two-sided coupled sylvester-type quaternion matrix equations. *Automatica*, 101:207–213, 2019.

[69] J. P. Ward. *Quaternions and Cayley Numbers*. Springer Netherlands, 1997.

[70] D. Widdows. Quaternion algebraic geometry. *Thesis at University of Oxford*, 2003.

[71] F. Zhang. Quaternions and matrices of quaternions. *Linear Algebra and Its Applications*, 251:21–57, 1997.