THE TENSOR RANK PROBLEM OVER THE QUATERNIONS

YG LIANG, SERGIO DA SILVA, YANG ZHANG

Abstract. We provide a nontrivial bound on the rank of any tensor $T$ over the quaternions $\mathbb{H}$ in the $n_1 \times n_2 \times n_3$ cases where $2 \leq n_i \leq 3$. We describe a decomposition of $T$ into 3 simple tensors in the $2 \times 2 \times 2$ case. We also show that the upper bound is the best possible for some of the cases, and we provide various partial results involving tensor decompositions over $\mathbb{C}$ and $\mathbb{H}$.

Determining the rank of any given tensor can generally be a difficult question to answer. While there are known results for specific cases over the complex and real numbers, the problem remains a largely open question for tensors over the real quaternion algebra $\mathbb{H} = \{a_0 + a_1i + a_2j + a_3k | i^2 = j^2 = k^2 = ijk = -1; a_0, a_1, a_2, a_3 \in \mathbb{R}\}$.

In this article, we will consider tensors of the form $\mathbb{H}^{n_1} \otimes_\mathbb{H} \mathbb{H}^{n_2} \otimes_\mathbb{H} \mathbb{H}^{n_3}$, which is the set of multiway arrays in $\mathbb{H}^{n_1 \times n_2 \times n_3}$ that has an $M_{n_1}(\mathbb{H}) \times M_{n_2}(\mathbb{R}) \times M_{n_3}(\mathbb{H})$ action defined by $(a_1, a_2, a_3) \cdot (h_1, h_2, h_3) = (a_1 h_1, a_2 h_2, h_3 a_3)$. In other words, we are considering $(M_{n_1}(\mathbb{H}), M_{n_3}(\mathbb{H}))$-bimodules and have an additional real action on frontal slices. The tensor rank problem asks how to minimally decompose $T \in \mathbb{H}^{n_1} \otimes_\mathbb{H} \mathbb{H}^{n_2} \otimes_\mathbb{H} \mathbb{H}^{n_3}$ into a sum of simple tensors. This can involve determining the rank of $T$, but also includes questions about the existence and uniqueness of some minimal decomposition. We address some of these questions for the cases where $n_i \leq 3$.

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 rank. Similarly, tensor networks play a large role in algebraic statistics, and the separation of variables using tensor decomposition appears in numerical analysis (see Part I in [5]). We also see tensors used in questions from data mining and machine learning (see [6]). Not all problems are commutative in nature however, so understanding tensors in the noncommutative case is also important.

Key words and phrases. Tensor rank, Quaternions, Tensor decomposition.

Research supported in part by a PIMS postdoctoral fellowship, Canada NSERC and UM Interdisciplinary/New Directions Research Collaboration Initiation Grants.
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 [9], 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 [4] and [1] respectively). In some recent work involving $\mathbb{H}$, a simulatneous diagonalization result in [3] produces solutions to a specific generalized Sylvester quaternion matrix equation, while expanded work in [11] provides solutions for a two-sided coupled Sylvester-type equation in a similar setting. These results however assume that the rank of the tensors being used are known. It is therefore essential to find ways of determining the rank of tensors over $\mathbb{H}$ if one hopes to utilize these results. Our goal is to provide explicit criteria which can easily be checked via a computer program, as opposed to criteria which theoretically determines the rank of a given tensor.

Our results begin by analyzing the $2 \times 2 \times 2$ case. It is known that over the complex and real numbers, the maximum rank of any tensor is 3 (see [8]). In the quaternion case, we show that maximum rank is also 3, and we provide a convenient decomposition into simple tensors. We will then provide a bound on the rank for tensors over $\mathbb{H}$ in the remaining $n_1 \times n_2 \times n_3$ cases, where $2 \leq n_i \leq 3$. Finally, we provide results involving the decompositions for some these tensors, as well as examples of tensors which achieve the maximal bound. The results on tensor rank bounds can be summarized with the following theorem.

**Theorem 0.1.** Let $T \in \mathbb{H}^{n_1} \otimes \mathbb{H}^{n_2} \otimes \mathbb{H}^{n_3}$ where $2 \leq n_i \leq 3$. Then in the $2 \times 2 \times 2$ case, there exists an explicit decomposition of $T$ into 3 simple tensors, so $\operatorname{rank}(T) \leq 3$. Furthermore, $\operatorname{rank}(T) \leq 3$ if $n_i = 3$ for exactly one $i$, $\operatorname{rank}(T) \leq 4$ if $n_i = 2$ for exactly one $i$, and $\operatorname{rank}(T) \leq 6$ in the $3 \times 3 \times 3$ case.

1. Preliminaries

A multiway array $T = (T_{i_1 i_2 \ldots i_K})$ where $1 \leq i_1 \leq N_1, \ldots, 1 \leq i_K \leq N_K$ is called a $K$-way tensor of size $(N_1, N_2, \ldots, N_K)$. 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

$$T = \begin{pmatrix}
    a_{111} & a_{121} & a_{131} \\
    a_{211} & a_{221} & a_{231} \\
    a_{311} & a_{321} & a_{331}
\end{pmatrix}, \quad \begin{pmatrix}
    a_{112} & a_{122} & a_{132} \\
    a_{212} & a_{222} & a_{232} \\
    a_{312} & a_{322} & a_{332}
\end{pmatrix}$$

or

$$T = \begin{pmatrix}
    \tilde{a}_{11} & \tilde{a}_{12} & \tilde{a}_{13} \\
    \tilde{a}_{21} & \tilde{a}_{22} & \tilde{a}_{23} \\
    \tilde{a}_{31} & \tilde{a}_{32} & \tilde{a}_{33}
\end{pmatrix},$$
where \( \vec{a}_{ij} = (a_{ij1}, a_{ij2}) \), \( i, j = 1, 2, 3 \). Here the first lateral slice is the \( 3 \times 2 \) matrix defined by the vectors \( \vec{a}_{i1} \), and the first horizontal slice is the \( 2 \times 3 \) matrix defined by the \( \vec{a}_{1j} \).

**Note:** We should mention that this notation is different from the one used by some authors cited in this article (for example, in [10], \( N_1 \) indicates the number of frontal slices).

**Definition 1.1.** A nonzero \( K \)-tensor \( T = (T_{i_1i_2...i_K}) \) is called a **simple tensor** if there exist vectors
\[
\vec{a}_1 = (a_{11}, a_{12}, \ldots, a_{1N_1}), \\
\vec{a}_2 = (a_{21}, a_{22}, \ldots, a_{2N_2}), \\
\vdots \\
\vec{a}_K = (a_{K1}, a_{K2}, \ldots, a_{KN_K}),
\]
such that \( T = (T_{i_1i_2...i_K}) = (a_{1i_1} a_{2i_2} \ldots a_{Ki_K}) \). We will also denote this by
\[
T = \vec{a}_1 \otimes \vec{a}_2 \otimes \cdots \otimes \vec{a}_K.
\]

**Example 1.2.** For the \( 2 \times 3 \times 2 \) tensor
\[
T = \begin{pmatrix} 2 & 3 & -2 & -3 \\ 8 & 12 & 4 & 6 \end{pmatrix},
\]
there exist vectors
\[
\vec{a} = (a_1, a_2) = (1, 4), \quad \vec{b} = (b_1, b_2, b_3) = (1, -1, 2) \quad \text{and} \quad \vec{c} = (c_1, c_2) = (2, 3)
\]
such that
\[
T = \begin{pmatrix}
\begin{bmatrix} 1 & 1 & 2 \\ 4 & 1 & 2 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 3 \\ 4 & 1 & 3 \end{bmatrix} \\
\begin{bmatrix} 1 & (−1) & 2 \\ 4 & (−1) & 2 \end{bmatrix} & \begin{bmatrix} 1 & (−1) & 3 \\ 4 & (−1) & 3 \end{bmatrix} \\
\begin{bmatrix} 1 & 2 & 2 \\ 4 & 2 & 2 \end{bmatrix} & \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 3 \end{bmatrix}
\end{pmatrix}
= \begin{pmatrix}
\begin{bmatrix} a_1 b_1 c_1 & a_1 b_1 c_2 \\ a_2 b_1 c_1 & a_2 b_1 c_2 \end{bmatrix} \\
\begin{bmatrix} a_1 b_2 c_1 & a_1 b_2 c_2 \\ a_2 b_2 c_1 & a_2 b_2 c_2 \end{bmatrix} \\
\begin{bmatrix} a_1 b_3 c_1 & a_1 b_3 c_2 \\ a_2 b_3 c_1 & a_2 b_3 c_2 \end{bmatrix}
\end{pmatrix}
= \vec{a} \otimes \vec{b} \otimes \vec{c}.
\]

Therefore, \( T \) is a simple tensor. \( \square \)

**Definition 1.3.** Let \( T \) be a nonzero \( K \)-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 \( K \)-tensors. We will say that \( T \) has rank \( n \) and denote this by \( \text{rank}(T) = n \).

An immediate consequence of this definition is that \( \text{rank}(T+S) \leq \text{rank}(T) + \text{rank}(S) \). 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 subsequent sections 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 [2] for this type of tensor decomposition in the quaternion case).

In many of the results in this article, 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 operation (see [12, Section 1.3.3] for an exposition on the difficulties of defining tensors over the quaternions). For example, in applying the two column operations shown below, we have increased the rank of the matrix:

\[
\begin{bmatrix}
i & i \\
i + j & i + j
\end{bmatrix}_{c_2 \rightarrow c_2 j} \frac{c_2 \rightarrow c_2 j}{c_1 \rightarrow c_1} \begin{bmatrix}
-k & k \\
-1 - k & -1 + k
\end{bmatrix}.
\]

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 \( \text{rank}(T) \).

**Lemma 1.4.** Let \( T \) be an \( N_1 \times \ldots \times N_p \) 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 \( \text{rank}((A_1 \otimes \ldots \otimes A_p)(T)) \leq \text{rank}(T) \). Furthermore, if \( A_1 \otimes \ldots \otimes A_p \) is invertible, then \( \text{rank}((A_1 \otimes \ldots \otimes A_p)(T)) = \text{rank}(T) \).

**Proof.** Suppose \( \text{rank}(T) = n \), 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

\[
(A_1 \otimes \ldots \otimes A_p)(T) = \sum_{i=1}^{n} A_1 v_{i_1} \otimes \ldots \otimes A_p v_{i_p},
\]

which implies \( \text{rank}((A_1 \otimes \ldots \otimes A_p)(T)) \leq n \). \( \Box \)

A similar proof works if we replace \( \mathbb{H} \) with \( \mathbb{C} \). Allowing the \( A_i \) to have entries in \( \mathbb{H} \) would no longer define a multilinear map. However, using row operations with left \( \mathbb{H} \)
multiplication (or column operations with right \( \mathbb{H} \) multiplication) coming from nonsingular matrices does preserve rank.

Let us fix 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 \( A = A_1 + A_2j \), 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

\[
\chi_A = \begin{bmatrix}
A_1 & A_2 \\
-A_2 & A_1
\end{bmatrix}.
\]

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 ([7]).

**Lemma 1.5.** An \( n \times n \) matrix \( A \) with entries in \( \mathbb{H} \) is diagonalizable if and only if \( \chi_A \) is diagonalizable.

We conclude this section with two auxiliary lemmas which are important in the sections that follow.

**Lemma 1.6.** Let \( T = (A_1; A_2; \ldots ; A_p) \) be an \( m \times p \times n \) tensor. Then \( \text{rank}(T) \leq r \) 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 [10, Proposition 2.1]. First suppose that \( \text{rank}(T) \leq r \), 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 \( 1 \leq i \leq r \), so that \( A_k = \sum_{i=1}^{r} a_i^T b_{ik} c_{iT} \). We will define \( P, Q \) and \( D_k \) by

\[
P = [\vec{a}_1, \vec{a}_2, \ldots , \vec{a}_r], \quad Q = \begin{bmatrix}
\vec{c}_1^T \\
\vec{c}_2^T \\
\vdots \\
\vec{c}_r^T
\end{bmatrix}, \quad D_k = \text{diag}(b_{1k}, b_{2k}, \ldots , b_{rk}).
\]

Then

\[
PD_kQ = [\vec{a}_1, \vec{a}_2, \ldots , \vec{a}_r] \begin{bmatrix}
b_{1k} & b_{2k} & \cdots & b_{rk}
\end{bmatrix} \begin{bmatrix}
c_{1k}^T \\
c_{2k}^T \\
\vdots \\
c_{rk}^T
\end{bmatrix} = \sum_{i=1}^{r} a_i b_{ik} c_{iT} = A_k,
\]

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:

\[
A_k = PD_kQ = \sum_{i=1}^{r} \vec{a}_i b_{ik} c_i^T \implies T = \sum_{i=1}^{r} \vec{a}_i \otimes \vec{b}_i \otimes \vec{c}_i.
\]

Hence \( \text{rank}(T) \leq r \), completing the proof. \( \square \)

**Lemma 1.7.** Let \( T = (A_1; A_2; \ldots; A_p) \) be an \( n \times p \times n \) tensor, where \( A_1 \) is nonsingular. Then \( \text{rank}(T) = n \) if and only if \( \{ A_j A_1^{-1} \mid j = 2, 3, \ldots, p \} \) can be simultaneously diagonalized.

**Proof.** The following proof is similar to that found in [10, Proposition 2.5]. First suppose that \( \text{rank}(T) = n \). Then by Lemma 1.6, 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 \( \text{rank}(T') = n \), and since multiplication by invertible matrices is rank preserving, \( \text{rank}(T) = n \) as required. \( \square \)

2. **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. A priori, we could guarantee a trivial bound of 4, so it would be useful to start with a motivating argument on why an attempt to reduce this bound to 3 is justified. Considering the tensor rank problem from the algebro-geometric perspective, it is natural to try and compute the **generic rank**, which is the minimum \( r \) such that the set of all tensors of rank at most \( r \) is a Zariski dense set in \( \mathbb{H}^2 \otimes_{\mathbb{H}} \mathbb{H}^2 \otimes_{\mathbb{H}} \mathbb{H}^2 \).

In the complex case, one can show that the generic rank is 2 using an argument from [5, Proposition 12.4.3.2]. We can view \( T \in A \otimes B \otimes C = \mathbb{C}^2 \otimes_{\mathbb{C}} \mathbb{C}^2 \otimes_{\mathbb{C}} \mathbb{C}^2 \) as a map from \( A^* \to B \otimes C \). Then \( \text{rank}(T) \) is the number of rank one matrices needed to span \( T(A^*) \subset B \otimes C \) as a vector space (see [5, Theorem 3.1.1.1]). We can projectivize to get \( \mathbb{P}T(A^*) \subset \mathbb{P}(B \otimes C) \) and consider the Segre embedding
\[ \sigma : \mathbb{C}P^1 \times \mathbb{C}P^1 \longrightarrow \mathbb{C}P^3 \]

\[ [a, b] \times [c, d] \mapsto [ac, ad, bc, bd] = [w, x, y, z]. \]

The space of simple tensors is isomorphic to \( \sigma(\mathbb{C}P^1 \times \mathbb{C}P^1) \), which is a projective subvariety of \( \mathbb{C}P^3 = \mathbb{P}(B \otimes C) \). It is not difficult to show that \( wz = xy \) generates the ideal of the image of \( \sigma \). Notice that \( \mathbb{P}T(A^*) \) is generally a degree 1 hypersurface in \( \mathbb{P}(B \otimes C) \) since \( \mathbb{P}T(A) \) is a linear subspace. It will intersect \( \sigma(\mathbb{C}P^1 \times \mathbb{C}P^1) \) at \( \deg(\sigma(\mathbb{C}P^1 \times \mathbb{C}P^1)) = 2 \) many points by Bézout’s theorem. Since these points correspond to simple tensors in the preimage, we can write \( T \) as the sum of 3 simple tensors. This shows that the generic rank in the complex case is 2.

If we try and apply the same argument in the quaternion case, we immediately run into problems. We could assume that \( w, x, y \) and \( z \) are not commutative so that the equation that they satisfy is \( wy^{-1} = xz^{-1} \) which can be simplified to \( wy^{-1}z = x \). If we restrict to the Zariski open set where \( y \) is invertible, we are left with a degree 3 polynomial. This would suggest that the generic rank in the noncommutative case is 3. However, making this formal proves to be more difficult than the complex case.

The Segre map in general is defined on vector spaces by \( \mathbb{P}(V) \times \mathbb{P}(W) \longrightarrow \mathbb{P}(V \otimes W) \). We could extend this idea to quaternionic vector spaces so that \( V \times W = \mathbb{H}P^m \times \mathbb{H}P^n \), except that \( V \otimes W \) would no longer be quaternionic vector space (see [12, Section 1.3.3]), so \( \mathbb{P}(V \otimes H W) \) is ill-defined. However, we could use the tensor product over the complex numbers, and replace \( \mathbb{P}(V \otimes H W) \) with the Grassmannian \( \text{Gr}_2(V \otimes_C W) \). All of this taken together suggests that the generic rank over \( \mathbb{H} \) is 3.

**Conjecture 2.1.** The generic rank of a tensor in \( \mathbb{H}^2 \otimes_H \mathbb{H}^2 \otimes_H \mathbb{H}^2 \) is 3.

With this result in hand, it makes sense to try and prove that any \( 2 \times 2 \times 2 \) quaternion tensor has rank at most 3. The authors first proved this bound by writing out a system of polynomial equations that a tensor having rank 3 must satisfy. Using some basic results from algebraic geometry, we were able to prove the existence of a solution. This allowed us to arrive at the explicit and more convenient decomposition found below.

**Proposition 2.2.** 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:
One could ask whether this tensor decomposition is unique up to rescaling or permutation indeterminacy (see [6] 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 article. 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 2.3.** Let $T$ be a $2 \times 2 \times 2$ quaternion tensor. Then $\text{rank}(T) \leq 3$.

**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 1.4 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 $\text{rank}(T) \leq 3$ follows from Proposition 2.2 since the rank of each $T_i$ is no more than 1. \hfill \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 1.7. Therefore, the bound on the rank for $2 \times 2 \times 2$ quaternion tensors is the best possible.

### 3. 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 the authors are aware, an explicit decomposition like the one below has not been made readily available.

**Theorem 3.1.** Let $T$ be a complex $2 \times 3 \times 2$ tensor

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

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

\[
\begin{align*}
S_1 &= A_2B_3C_4 - A_2B_4C_3 - A_3B_2C_4 + A_3B_4C_2 + A_4B_2C_3 - A_4B_3C_2, \\
S_2 &= A_1B_2C_4 - A_1B_4C_2 - A_2B_1C_4 + A_2B_4C_1 + A_4B_1C_2 - A_4B_2C_1, \\
T_1 &= A_1B_3C_4 - A_1B_4C_3 - A_2B_3C_4 + A_2B_4C_1 + A_4B_3C_1 - A_4B_4C_1, \\
T_2 &= A_1B_2C_3 - A_1B_3C_2 - A_2B_1C_3 + A_2B_4C_1 + A_3B_1C_2 - A_3B_2C_1.
\end{align*}
\]

We can check by inspection that

\[
\begin{align*}
A_2T_1 - A_1S_1 + A_4T_2 &= A_3S_2, \\
B_2T_1 - B_1S_1 + B_4T_2 &= B_3S_2, \\
C_2T_1 - C_1S_1 + C_4T_2 &= C_3S_2.
\end{align*}
\]

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:

\[
\begin{align*}
T_1 &= \begin{bmatrix} A_2T_1S_1^{-1} & A_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B_2T_1S_1^{-1} & B_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} C_2T_1S_1^{-1} & C_2 \\ 0 & 0 \end{bmatrix}, \\
T_2 &= \begin{bmatrix} A_4T_2S_2^{-1} & A_4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B_4T_2S_2^{-1} & B_4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} C_4T_2S_2^{-1} & C_4 \\ 0 & 0 \end{bmatrix}, \\
T_3 &= \begin{bmatrix} (A_1S_1 - A_2T_1)S_1^{-1} & 0 \\ -(A_1S_1 - A_2T_1)S_2^{-1} & 0 \end{bmatrix} \begin{bmatrix} (B_1S_1 - B_2T_1)S_1^{-1} & 0 \\ -(B_1S_1 - B_2T_1)S_2^{-1} & 0 \end{bmatrix} \begin{bmatrix} (C_1S_1 - C_2T_1)S_1^{-1} & 0 \\ -(C_1S_1 - C_2T_1)S_2^{-1} & 0 \end{bmatrix}.
\end{align*}
\]

\[\square\]
If we try and generalize Theorem 2.3 or Theorem 3.1 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 2.3 for example, we only needed real row and column operations. As mentioned in Section 1, 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.** 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 to reduce $T$ to a form that is easier to decompose into simple tensors.

$$T \rightarrow \left( \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 1 + i \end{bmatrix}; \begin{bmatrix} 0 & -1 + j & 0 \\ 0 & -i + k & 1 + j \end{bmatrix} \right) \quad \text{using } C_2 \rightarrow C_2j$$

$$\rightarrow \left( \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \end{bmatrix}; \begin{bmatrix} 0 & -1 + j & 1 + i - j + k \\ 0 & -i + k & i - k \end{bmatrix} \right) \quad \text{using } C_3 \rightarrow C_3 - C_2(1 + i) + C_1(j + k)$$

$$\rightarrow \left( \begin{bmatrix} 1 & 3k & 0 \\ 0 & 3 & 0 \end{bmatrix}; \begin{bmatrix} 0 & -1 - j + 2k & 1 + i - j + k \\ 0 & 2 - i + k & i - k \end{bmatrix} \right) \quad \text{using } C_2 \rightarrow 3C_2 - C_3(-2 + i - k)$$

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

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

Therefore, $\text{rank}(T) \leq 3$. \hfill \Box

**Theorem 3.3.** Let $T$ be a $2 \times 2 \times 3$ or a $3 \times 2 \times 2$ quaternion tensor. Then $\text{rank}(T) \leq 3$.

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

$$A = [\vec{a} \quad \vec{b} \quad \vec{c}], \quad B = [\vec{a} \quad \vec{e} \quad \vec{f}],$$

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

$$\text{rank}(T) \leq \text{rank}((A; 0)) + \text{rank}((0; B)) \leq 1 + 2 = 3.$$
Therefore, let us assume that \( \text{rank}(A) = \text{rank}(B) = 2 \). Without loss of generality, we can assume that \( \text{rank}(\vec{a}, \vec{b}) = 2 \). Then, by a column operation, we have

\[
T = ([\vec{a} \vec{b} \vec{c}] ; [\vec{d} \vec{e} \vec{f}]) \longrightarrow ([\vec{a} \vec{b} 0] ; [\vec{d} \vec{e} \vec{g}]).
\]

If \( \vec{g} = 0 \), then \( \text{rank}(T) \leq 3 \) by Theorem 2.3. Let us assume then that \( \vec{g} \neq 0 \), and without loss of generality, that \( \text{rank}(\vec{e}, \vec{g}) = 2 \) (since \( B \) has full rank). By further column operations, we have

\[
([\vec{a} \vec{b} 0] ; [\vec{d} \vec{e} \vec{g}]) \longrightarrow ([\vec{h} \vec{b} 0] ; [0 \vec{e} \vec{g}]).
\]

Since \( \text{rank}(\vec{h}, \vec{b}) = 2 \), we can write

\[
\vec{e} = \vec{hc}_1 + \vec{bc}_2, \quad \vec{g} = \vec{hc}_3 + \vec{bc}_4.
\]

If \( c_3 \neq 0 \), then

\[
([\vec{h} \vec{b} 0] ; [0 \vec{e} \vec{g}]) = ([\vec{h} \vec{b} 0] ; [0 \vec{hc}_1 + \vec{bc}_2 \vec{hc}_3 + \vec{bc}_4])
\]

\[
\longrightarrow ([\vec{h} \vec{b} 0] ; [0 \vec{bc}_5 \vec{hc}_3 + \vec{bc}_4]),
\]

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

\[
([\vec{h} \vec{b} 0] ; [0 \vec{e} \vec{g}]) = ([\vec{h} \vec{b} 0] ; [0 \vec{hc}_1 + \vec{bc}_2 \vec{bc}_4])
\]

\[
\longrightarrow ([\vec{h} \vec{hc}_6 + \vec{bc}_7 0] ; [0 (\vec{hc}_6 + \vec{bc}_7)c_8 \vec{bc}_4]),
\]

using the column operation \( C_2 \rightarrow C_2 + C_1 k_1 + C_3 k_3 \) 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 in a similar fashion by simply applying a rotation of the tensor and working with \( A^T \) and \( B^T \) instead.

\[ \square \]

**Theorem 3.4.** Let \( T \) be a \( 2 \times 3 \times 2 \) quaternion tensor. Then \( \text{rank}(T) \leq 3 \).

**Proof.** Let

\[
T = (A; B; C) = \left([a_{11} a_{12} ; b_{11} b_{12}] ; [c_{11} c_{12}]ight).
\]

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

\[
T = \left([1 0 ; b'_{11} b'_{21}] ; [c'_{11} c'_{12}]ight).
\]

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 a row operation we can assume that \( b'_{22}c'_{22} \neq 0 \), and we have the following decomposition into simple tensors \( T = T_1 + T_2 + T_3 \):
In both cases we see that \( \text{rank}(T) \leq 3 \).

Therefore, we may assume that \( A, B \) and \( C \) are nonsingular. We can also assume that every vector in the lateral direction is nonzero (otherwise the argument above bounds the rank by 3). By performing row and column operations, we can assume that

\[
T_1 = \begin{pmatrix} 0 & 0 \\ b'_{12}(b'_{22})^{-1}b'_{21} & b'_{12} \\ 0 & 0 \end{pmatrix},
T_2 = \begin{pmatrix} 0 & 0 \\ c'_{12}(c'_{22})^{-1}c'_{21} & c'_{12} \\ 0 & 0 \end{pmatrix},
T_3 = \begin{pmatrix} 0 & 0 \\ c'_{11}(c'_{22})^{-1}c'_{21} & c'_{11} \end{pmatrix}.
\]

If \( b_{11} = b_{22} = c_{11} = c_{22} = 0 \), then we have the following decomposition into simple tensors \( T = T_1 + T_2 + T_3 \), where

\[
T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}.
\]

If \( b_{11} = b_{22} = c_{11} = c_{22} = 0 \), then we have the following decomposition into simple tensors \( T = T_1 + T_2 + T_3 \), where

\[
T_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},
T_2 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix},
T_3 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}.
\]

so that \( \text{rank}(T) \leq 3 \). Therefore let us assume that one of these coordinates is nonzero, which without loss of generality we can choose \( b_{11} \neq 0 \). The \( b_{21} = 0 \) is trivial, so let's assume that \( b_{21} \) is nonzero. Then by [13, Theorem 3.1], there exists a nonzero quaternion \( x \) such that

\[-xb_{12}x - xb_{11} + b_{22}x + b_{21} = 0.\]

Therefore, we have

\[
\begin{pmatrix} 1 & 0 \\ -x & 1 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} = \begin{pmatrix} b_{11} + b_{12}x & b_{12} \\ 0 & -xb_{12} + b_{22} \end{pmatrix}.
\]

Since \( b_{11} + b_{12}x \neq 0 \), we may again use [13, Theorem 3.1] to show that there exists a quaternion \( y \) such that
\[
\begin{bmatrix}
1 & -y \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
b_{11} + b_{12}x & b_{12} \\
0 & -xb_{12} + b_{22}
\end{bmatrix}
\begin{bmatrix}
1 & y \\
0 & 1
\end{bmatrix}
= \begin{bmatrix}
b_{11} + b_{12}x & 0 \\
0 & -xb_{12} + b_{22}
\end{bmatrix}.
\]

If we apply the same operations as above to the tensor \( T \), we get the resulting tensor

\[
S = \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ; \begin{bmatrix} b_{11} + b_{12}x & 0 \\ 0 & -xb_{12} + b_{22} \end{bmatrix} ; \begin{bmatrix} c'_{11} & c'_{12} \\ c'_{21} & c'_{22} \end{bmatrix} \right).
\]

The result follows using the decomposition \( S = S_1 + S_2 + S_3 \) with \( \text{rank}(S_i) \leq 1 \):

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

\[
S_2 = \left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} ; \begin{bmatrix} b_{11} + b_{12}x & 0 \\ 0 & 0 \end{bmatrix} ; \begin{bmatrix} c'_{11} - 1 & 0 \\ 0 & 0 \end{bmatrix} \right),
\]

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

\[\square\]

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 tensors that have rank 3.

### 4. 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 subsequently 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 4.1.** 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 \( x_0A + B \) is singular.

**Proof.** By [13, 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 \( x_0I + BA^{-1} \) is singular. Therefore, \( x_0A + B \) is also singular. \( \square \)

**Theorem 4.2.** Let \( T \) be a \( 2 \times 3 \times 3 \) or \( 3 \times 3 \times 2 \) quaternion tensor. Then \( \text{rank}(T) \leq 4 \).

**Proof.** Let

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

\[
T = T_1 + T_2,
\]

where
Consider the matrices

\[
T_1 = \begin{pmatrix}
[a_{11} & a_{12} & a_{13}] \\
0 & 0 & 0 \\
\end{pmatrix};
\begin{pmatrix}
a_{21} & a_{22} & a_{23} \\
0 & 0 & 0 \\
\end{pmatrix};
\begin{pmatrix}
a_{31} & a_{32} & a_{33} \\
0 & 0 & 0 \\
\end{pmatrix},
\]

\[
T_2 = \begin{pmatrix}
0 & 0 & 0 \\
b_{11} & b_{12} & b_{13} \\
\end{pmatrix};
\begin{pmatrix}
0 & 0 & 0 \\
b_{21} & b_{22} & b_{23} \\
\end{pmatrix};
\begin{pmatrix}
b_{31} & b_{32} & b_{33} \\
0 & 0 & 0 \\
\end{pmatrix}.
\]

Notice that \(\text{rank}(T_1) = \text{rank}(A)\) and \(\text{rank}(T_2) = \text{rank}(B)\). If both \(A\) and \(B\) have rank at most 2, then

\[
\text{rank}(T) \leq \text{rank}(T_1) + \text{rank}(T_2) \leq 2 + 2 = 4.
\]

Therefore, without loss of generality, let us assume that \(\text{rank}(A) = 3\). By Lemma 4.1, there exists \(x_0 \in \mathbb{H}\) such that

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

is singular. Therefore \(\text{rank}(C) \leq 2\) and we can assume that

\[
C = \begin{pmatrix}
c_{11} & c_{12} & c_{13} \\
c_{21} & c_{22} & c_{23} \\
c_{31} & c_{32} & c_{33} \\
\end{pmatrix} \rightarrow \begin{pmatrix}
c_{11} & c_{12} & 0 \\
c_{21} & c_{22} & 0 \\
c_{31} & c_{32} & 0 \\
\end{pmatrix}.
\]

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

\[
T = \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
b_{11} & b_{12} & b_{13} \\
\end{pmatrix};
\begin{pmatrix}
a_{21} & a_{22} & a_{23} \\
b_{21} & b_{22} & b_{23} \\
\end{pmatrix};
\begin{pmatrix}
a_{31} & a_{32} & a_{33} \\
b_{31} & b_{32} & b_{33} \\
\end{pmatrix},
\]

\[
\rightarrow \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
c_{11} & c_{12} & c_{13} \\
\end{pmatrix};
\begin{pmatrix}
a_{21} & a_{22} & a_{23} \\
c_{21} & c_{22} & c_{23} \\
\end{pmatrix};
\begin{pmatrix}
a_{31} & a_{32} & a_{33} \\
c_{31} & c_{32} & c_{33} \\
\end{pmatrix},
\]

\[
\rightarrow \begin{pmatrix}
a_{11} & a_{12} & a_{13}' \\
c_{11} & c_{12} & 0 \\
\end{pmatrix};
\begin{pmatrix}
a_{21} & a_{22} & a_{23}' \\
c_{21} & c_{22} & 0 \\
\end{pmatrix};
\begin{pmatrix}
a_{31} & a_{32} & a_{33}' \\
c_{31} & c_{32} & 0 \\
\end{pmatrix}.
\]

If \(a_{13}' = a_{23}' = a_{33}' = 0\), then we immediately have \(\text{rank}(T) \leq 4\). Otherwise, we may that assume \(a_{13}' \neq 0\). By column operations, the tensor can be reduced to

\[
\begin{pmatrix}
0 & 0 & a_{13}' \\
c_{11} & c_{12} & 0 \\
\end{pmatrix};
\begin{pmatrix}
a_{21}' & a_{22}' & a_{23}' \\
c_{21} & c_{22} & 0 \\
\end{pmatrix};
\begin{pmatrix}
a_{31}' & a_{32}' & a_{33}' \\
c_{31} & c_{32} & 0 \\
\end{pmatrix}.
\]
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 \( \text{rank}(T) \leq 4 \) by Theorem 3.3. Therefore, by switching columns if necessary, we can assume that \( c_{12} \neq 0 \). After performing further column operations we can reduce \( T \) to

\[
\begin{pmatrix}
0 & 0 & a'_{13} \\
0 & c_{12} & 0 \\
0 & c_{12} & 0
\end{pmatrix} ;
\begin{pmatrix}
0 & 0 & 0 \\
0 & c_{22} & -c'_{21}(a''_{21})^{-1}a''_{22} & 0 \\
0 & c_{32} & -c'_{31}(a''_{31})^{-1}a''_{32} & 0
\end{pmatrix} ;
\begin{pmatrix}
a''_{31} & a''_{32} & a''_{33} \\
c''_{31} & c''_{32} & 0
\end{pmatrix}.
\]

If \( a''_{31} = a''_{21} = 0 \), then \( T \) is the sum of 4 simple tensors defined from the 4 nonzero \( 1 \times 3 \) vectors in the lateral direction, and we immediately have \( \text{rank}(T) \leq 4 \). 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 \( \text{rank}(T) \leq 4 \):

\[
T_1 = \begin{pmatrix}
0 & 0 & 0 \\
0 & c_{12} & 0 \\
0 & c_{12} & 0
\end{pmatrix} ;
\begin{pmatrix}
0 & 0 & 0 \\
0 & c_{22} - c'_{21}(a''_{21})^{-1}a''_{22} & 0 \\
0 & c_{32} - c'_{31}(a''_{31})^{-1}a''_{32} & 0
\end{pmatrix} ;
\begin{pmatrix}
a''_{31} & a''_{32} & a''_{33} \\
c''_{31} & c''_{32} & 0
\end{pmatrix},
\]

\[
T_2 = \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix} ;
\begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix} ;
\begin{pmatrix}
a''_{21} & a''_{21}(a''_{31})^{-1}a''_{32} & 0 \\
c''_{21} & c''_{21}(a''_{31})^{-1}a''_{32} & 0
\end{pmatrix},
\]

\[
T_3 = \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix} ;
\begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix} ;
\begin{pmatrix}
a''_{21} & a''_{21}(a''_{31})^{-1}a''_{32} & 0 \\
c''_{21} & c''_{21}(a''_{31})^{-1}a''_{32} & 0
\end{pmatrix},
\]

\[
T_4 = \begin{pmatrix}
0 & 0 & a'_{13} \\
0 & 0 & a'_{23} \\
0 & 0 & a'_{33}
\end{pmatrix} ;
\begin{pmatrix}
0 & 0 & a'_{13} \\
0 & 0 & a'_{23} \\
0 & 0 & a'_{33}
\end{pmatrix} ;
\begin{pmatrix}
0 & 0 & a'_{13} \\
0 & 0 & a'_{23} \\
0 & 0 & a'_{33}
\end{pmatrix}.
\]

Finally, we can repeat the above arguments for the \( 3 \times 3 \times 2 \) case using the transpose of the matrices involved.

It is not clear whether the bound of 4 is the best possible for the \( 2 \times 3 \times 3 \) and \( 3 \times 3 \times 2 \) cases. On the other hand, we can show that the tensor

\[
T = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

has rank 4 by Lemma 1.7 and Theorem 4.3 below. 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 4.3.** Let \( T \) be a \( 3 \times 2 \times 3 \) quaternion tensor. Then \( \text{rank}(T) \leq 4 \).

**Proof.** Let

\[
T = (A; B) = \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{pmatrix} ;
\begin{pmatrix}
b_{11} & b_{12} & b_{13} \\
b_{21} & b_{22} & b_{23} \\
b_{31} & b_{32} & b_{33}
\end{pmatrix}.
\]
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

$$
\begin{pmatrix}
1 & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{pmatrix}
; 
\begin{pmatrix}
b_{11} & b_{12} & b_{13} \\
b_{21} & b_{22} & b_{23} \\
b_{31} & b_{32} & b_{33}
\end{pmatrix},
$$

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 3.3, \(\text{rank}(T) \leq 3 + 1 \leq 4\). For the same reason, if any horizontal or lateral slice of the tensor does not have maximal rank, then \(\text{rank}(T) \leq 4\). For example, the first lateral slice of $T$ is the matrix

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

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 \(\text{rank}(T) \leq 4\). 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:

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

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

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

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

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

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

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

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

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

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 $\text{rank}(T) \leq 4$. 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 1.5, the resulting matrix \( M \) is diagonalizable if and only if its complex adjoint \( \chi_M \) is. The adjoint of \( M \) is the complex matrix

\[
\chi_M = \begin{bmatrix}
0 & 0 & a + bi & 0 & 0 & c + di \\
u + vi & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -c + di & 0 & 0 & a - bi \\
0 & 0 & 0 & u - vi & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}.
\]

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

\[
C = (u^2 + v^2)(a^2 + b^2 + c^2 + d^2) > 0 \quad \text{and}
\]

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

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

5. The 3 × 3 × 3 Case

For real or complex 3 × 3 × 3 tensors, it is known that the maximal rank is 5 (see [10, Section 3.4]). However, a general decomposition of such tensors into 5 simple tensors is not provided. Although there are several subcases considered in [10, Section 3.4], we will only provide a decomposition for the main subcase (namely [10, Equation 3.4.2]).

**Theorem 5.1.** Let \( T \) be a complex 3 × 3 × 3 tensor. Suppose that through a sequence of rank-preserving row and column operations, \( T \) can be reduced to the form

\[
T = \left( \begin{bmatrix} A_{11} & 0 & 1 \\ 0 & A_{22} & 0 \\ 1 & 0 & 0 \end{bmatrix} \right) \left( \begin{bmatrix} B_{11} & 0 & 0 \\ 0 & B_{22} & 1 \\ 0 & 1 & 0 \end{bmatrix} \right) \left( \begin{bmatrix} C_{11} & C_{12} & 0 \\ 0 & C_{21} & C_{22} \\ 0 & 0 & 0 \end{bmatrix} \right).
\]

where \( A_{11}, B_{22}, C_{22} \neq 0 \) and \( R = A_{11}C_{12} + B_{11}C_{22} \neq 0 \). Then \( T \) has a decomposition as the sum of the 5 simple tensors defined below.

**Proof.** Let \( S = C_{11}C_{22} - C_{12}C_{21} \). Let us define \( T_1, \ldots, T_5 \) by
Proof. Let $T = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \\ c_{11} & c_{12} & c_{13} \end{pmatrix}$, $T_1 = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ b_{21} & b_{22} & b_{23} \\ c_{21} & c_{22} & c_{23} \end{pmatrix}$, $T_2 = \begin{pmatrix} a_{31} & a_{32} & a_{33} \\ b_{31} & b_{32} & b_{33} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$, and write $T$ as the sum $T_1 + T_2 + T_3$, where

\[
T_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} a_{31} & a_{32} & a_{33} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]
Consider matrices

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

Notice that \(\text{rank}(T_1) = \text{rank}(A)\), \(\text{rank}(T_2) = \text{rank}(B)\), and \(\text{rank}(T_3) = \text{rank}(C)\), so if \(A\), \(B\) and \(C\) all have rank at most 2, then

\[
\text{rank}(T) \leq \text{rank}(T_1) + \text{rank}(T_2) + \text{rank}(T_3) \leq 2 + 2 + 2 = 6.
\]

Therefore, without loss of generality, we can assume that \(\text{rank}(A) = 3\). By Lemma 4.1, there exists \(x_0 \in \mathbb{H}\) such that

\[
D = x_0 A + C = \begin{bmatrix}
d_{11} & d_{12} & d_{13} \\
d_{21} & d_{22} & d_{23} \\
d_{31} & d_{32} & d_{33}
\end{bmatrix}
\]

is singular. Using rank-preserving row operations, we can reduce \(T\) to

\[
T \rightarrow \begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
b_{11} & b_{12} & b_{13} \\
d_{11} & d_{12} & d_{13}
\end{bmatrix}, \quad \begin{bmatrix}
a_{21} & a_{22} & a_{23} \\
b_{21} & b_{22} & b_{23} \\
d_{21} & d_{22} & d_{23}
\end{bmatrix}, \quad \begin{bmatrix}
a_{31} & a_{32} & a_{33} \\
b_{31} & b_{32} & b_{33} \\
d_{31} & d_{32} & d_{33}
\end{bmatrix}.
\]

Notice that the first 2 horizontal slices \(T_1 + T_2\) can viewed as a \(2 \times 3 \times 3\) tensor, which has rank at most 4 by Theorem 4.2, and the last slice

\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
d_{11} & d_{12} & d_{13}
\end{bmatrix}, \quad \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
d_{21} & d_{22} & d_{23}
\end{bmatrix}, \quad \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
d_{31} & d_{32} & d_{33}
\end{bmatrix}
\]

has the same rank as the matrix \(D\). Therefore, we have

\[
\text{rank}(T) \leq \text{rank}(T_1 + T_2) + \text{rank}(D) \leq 4 + 2 = 6.
\]

\square
THE TENSOR RANK PROBLEM OVER THE QUATERNIONS

References

[1] J. Brylinski, Algebraic measures of entanglement, Mathematics of Quantum Computation (200201), 3–24.
[2] Z. He, C. Navasca, and Q. Wang, Tensor decompositions and tensor equations over quaternion algebra, arxiv:1710.07552 (2017).
[3] Z. He, Q. Wang, and Y. Zhang, A simultaneous decomposition for seven matrices with applications, J. of Computational and Applied Math. 349 (2018), 93–113.
[4] X. Kong and Y. 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 (2013), 1348–1362.
[5] J. Landsberg, Tensors: Geometry and applications, 2012.
[6] S. Rabanser, O. Shchur, and S. Günnemann, Introduction to tensor decompositions and their applications in machine learning, arxiv:1711.10781 (2017).
[7] Leiba Rodman, Topics in quaternion linear algebra, Princeton Series in Applied Mathematics, 2014.
[8] T. Sumi, M. Miyazaki, and T. Sakata, Rank of 3-tensors with 2 slices and Kronecker canonical forms, Linear Algebra and its Applications 431 (2009), 1858–1868.
[9] , Rank of tensors with size $2 \times \ldots \times 2$, arXiv:1306.0708 (2013).
[10] , Algebraic and computational aspects of real tensor ranks, SpringerBriefs in Statistics. JSS Research Series in Statistics. Springer, [Tokyo], 2016.
[11] Q. Wang, Z. He, and Y. Zhang, Constrained two-sided coupled sylvester-type quaternion matrix equations, Automatica 101 (201903), 207–213.
[12] Dominic Widdows, Quaternion algebraic geometry, Thesis at University of Oxford (2003).
[13] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra and its Applications 251 (1997), 21–57.

YG Liang, University of Manitoba, Winnipeg MB
E-mail address: liangy1@myumanitoba.ca

Sergio Da Silva, University of Manitoba, Winnipeg MB
E-mail address: smd322@cornell.edu, sergio.dasilva@umanitoba.ca

Yang Zhang, University of Manitoba, Winnipeg MB
E-mail address: Yang.Zhang@umanitoba.ca