ON KRUSKAL’S THEOREM THAT EVERY $3 \times 3 \times 3$ ARRAY HAS RANK AT MOST 5

MURRAY R. BREMNER AND JIAXIONG HU

Abstract. In the first part of this paper, we consider $3 \times 3 \times 3$ arrays with complex entries, and provide a complete self-contained proof of Kruskal’s theorem that the maximum rank is 5. In the second part, we provide a complete classification of the canonical forms of $3 \times 3 \times 3$ arrays over $\mathbb{F}_2$; in particular, we obtain explicit examples of such arrays with rank 6.

In 1989, Kruskal [6, page 10] stated without proof that every $3 \times 3 \times 3$ array with real entries has rank at most 5. A few years later, Rocci [7] circulated a simplified proof of this result, based on Kruskal’s unpublished hand-written notes. The details of this argument appear never to have been published. In sections 2–4, we consider $3 \times 3 \times 3$ arrays with complex entries, and provide a complete self-contained proof that the maximum rank is 5.

In section 5 we consider this problem over the field $\mathbb{F}_2$ with two elements. A remarkable fact, first noted by von zur Gathen [9], is that in this case there exist $3 \times 3 \times 3$ arrays of rank 6. We provide a complete classification of the canonical forms of $3 \times 3 \times 3$ arrays over $\mathbb{F}_2$; in particular, we obtain explicit examples of such arrays with rank 6.

We use without reference many basic results on multidimensional arrays which can be found in de Silva and Lim [3] and Kolda and Bader [5].

1. Preliminaries on 3-dimensional arrays

We consider a $p \times q \times r$ array $X$ with entries in an arbitrary field $\mathbb{F}$ of scalars:

$$X = [x_{ijk}], \quad x_{ijk} \in \mathbb{F}, \quad 1 \leq i \leq p, \quad 1 \leq j \leq q, \quad 1 \leq k \leq r.$$ 

By a slice of $X$ we mean any (2-dimensional) submatrix obtained by fixing one index. Fixing $i$ gives a horizontal slice, fixing $j$ gives a vertical slice, and fixing $k$ gives a frontal slice. The matrix form of $X$ is the $p \times qr$ matrix obtained by concatenating the frontal slices $X_1, \ldots, X_r$ from left to right:

$$X = \begin{bmatrix}
  X_1 \\
  \vdots \\
  X_r
\end{bmatrix} =
\begin{bmatrix}
  x_{111} & \cdots & x_{1q1} & \cdots & x_{11r} & \cdots & x_{1qr} \\
  \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\
  x_{p11} & \cdots & x_{pq1} & \cdots & x_{p1r} & \cdots & x_{pqr}
\end{bmatrix}.$$

2010 Mathematics Subject Classification. Primary 15A69. Secondary 15-04, 15A03, 15A21, 20G20, 20G40.

Key words and phrases. 3-dimensional arrays, tensor decomposition, canonical forms.
Given three column vectors,
\[
\begin{align*}
    a &= \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \in \mathbb{F}^p, \\
    b &= \begin{bmatrix} b_1 \\ \vdots \\ b_q \end{bmatrix} \in \mathbb{F}^q, \\
    c &= \begin{bmatrix} c_1 \\ \vdots \\ c_r \end{bmatrix} \in \mathbb{F}^r,
\end{align*}
\]
their outer product \( a \otimes b \otimes c \) is the \( p \times q \times r \) array whose \( ijk \) entry is \( a_ib_jc_k \). A simple tensor is an outer product of nonzero vectors. A fundamental problem is to represent the array \( X \) as a sum of simple tensors:
\[
X = \sum_{i=1}^{n} a^{(i)} \otimes b^{(i)} \otimes c^{(i)}.
\]
The rank of the array \( X \) is the smallest non-negative integer \( n \) for which this decomposition is possible. The rank is 0 if and only if every entry of the array is 0; the rank is 1 if and only if the array is a simple tensor.

The rank does not change if we permute the slices in each direction. Given permutations \( \alpha \in S_p, \beta \in S_q, \gamma \in S_r \), we form another \( p \times q \times r \) array by
\[
(\alpha, \beta, \gamma) \cdot X = x_{\alpha(i)\beta(j)\gamma(k)}.
\]
More generally, the rank does not change if we apply a change of basis in each direction. Given invertible matrices
\[
A = (a_{i_1i_2}) \in GL(p, \mathbb{F}), \quad B = (b_{j_1j_2}) \in GL(q, \mathbb{F}), \quad C = (c_{k_1k_2}) \in GL(r, \mathbb{F}),
\]
we form another \( p \times q \times r \) array by
\[
(A, B, C) \cdot X = a_{i_1i_2}b_{j_1j_2}c_{k_1k_2}x_{i_1j_1k_1}.
\]
The rank does not change if we permute the directions; however, this permutes the dimensions \( p, q, r \) and hence may give a different ordered triple \( (p, q, r) \). If we write the dimensions as \( p_1 \times p_2 \times p_3 \) with corresponding indices \( i_1, i_2, i_3 \) then applying a permutation \( \delta \in S_3 \) gives an array of size \( p_{\delta(1)} \times p_{\delta(2)} \times p_{\delta(3)} \) defined by
\[
(\delta \cdot X)_{i_{\delta(1)}j_{\delta(2)}k_{\delta(3)}} = x_{i_1j_1k_1}.
\]

In the rest of this paper, we often use these rank-preserving transformations without further comment.

2. Ten Berge’s theorem on \( 2 \times 2 \times 2 \) arrays

The results in this section are taken with minor changes from ten Berge [8]. However, for us the base field is \( \mathbb{C} \) whereas for ten Berge it is \( \mathbb{R} \). We recall these results in detail since they are essential to the analysis of \( 3 \times 3 \times 3 \) arrays. For \( 2 \times 2 \times 2 \) arrays, the rank decomposition takes the form
\[
X = \sum_{i=1}^{n} a^{(i)} \otimes b^{(i)} \otimes c^{(i)}, \quad \text{where} \quad a^{(i)}, b^{(i)}, c^{(i)} \in \mathbb{C}^2 \quad \text{for} \quad 1 \leq i \leq n.
\]
We express this decomposition in terms of three \( 2 \times n \) matrices \( A, B, C \):
\[
A = \begin{bmatrix} a^{(1)} & \cdots & a^{(n)} \end{bmatrix}, \quad B = \begin{bmatrix} b^{(1)} & \cdots & b^{(n)} \end{bmatrix}, \quad C = \begin{bmatrix} c^{(1)} & \cdots & c^{(n)} \end{bmatrix}.
\]

Lemma 2.1. [8, p. 632] The rank of a nonzero \( 2 \times 2 \times 2 \) array \( X \) is the least integer \( n \geq 1 \) such that the frontal slices \( X_1, X_2 \) have the form \( X_1 = ADB^t \), \( X_2 = AEB^t \) where \( A, B \) are \( 2 \times n \) matrices and \( D, E \) are \( n \times n \) diagonal matrices.
The first frontal slice $X_1$ has the form

$$X_1 = \sum_{i=1}^{n} \begin{bmatrix} a_1^{(i)} b_1^{(i)} c_1^{(i)} \\ a_2^{(i)} b_2^{(i)} c_2^{(i)} \end{bmatrix} = \sum_{i=1}^{n} \begin{bmatrix} A_{11} c_1^{(i)} B_{11} \\ A_{21} c_1^{(i)} B_{12} \end{bmatrix} = A C_1 B^t,$$

where $C_1$ is the $n \times n$ diagonal matrix whose diagonal entries $c_1^{(1)}, c_1^{(2)}, \ldots, c_1^{(n)}$ come from row 1 of $C$. Similarly, for the second frontal slice we have $X_2 = A C_2 B^t$, where $C_2$ is the $n \times n$ diagonal matrix whose diagonal entries come from the second row of $C$. Conversely, if the two frontal slices $X_1$ and $X_2$ can be written as $A C_1 B^t$ and $A C_2 B^t$ where $A$ and $B$ are $2 \times n$ matrices and $C_1$ and $C_2$ are $n \times n$ diagonal matrices, then $X$ has the given decomposition.

**Definition 2.2.** We call the $2 \times 2 \times 2$ array $X$ superdiagonal if it has one of these forms for $\alpha, \beta \in \mathbb{C} \setminus \{0\}$:

$\begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix}$, $\begin{bmatrix} 0 & \alpha & 0 \\ 0 & 0 & \beta \\ \alpha & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & \beta \\ 0 & 0 & 0 \\ \alpha & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & \beta \\ 0 & 0 & 0 \\ 0 & \alpha & 0 \end{bmatrix}$.

**Lemma 2.3.** [8 p. 632] A superdiagonal array has rank 2.

**Proof.** By applying permutations of the slices, we may assume that $X$ has the first form. It is then clear that the array has rank $\leq 2$ since

$$\begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

To find the general form of an array of rank 1 according to Lemma 2.1, we set

$$A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad D = \begin{bmatrix} d \end{bmatrix}, \quad E = \begin{bmatrix} e \end{bmatrix}.$$

We obtain

$$X_1 = A D B^t = \begin{bmatrix} a_1 d b_1 \\ a_2 d b_1 \end{bmatrix}, \quad X_2 = A E B^t = \begin{bmatrix} a_1 e b_1 \\ a_2 e b_2 \end{bmatrix},$$

or more simply $X_1 = d(AB^t)$ and $X_2 = e(AB^t)$. Thus $X_1$ and $X_2$ are scalar multiples of the same matrix of rank 1. This does not hold for a superdiagonal array, which therefore has rank $\geq 2$.

**Lemma 2.4.** [8 p. 632] Let $X$ be a nonzero $2 \times 2 \times 2$ array which is not superdiagonal. Then $X$ has rank 1 if and only if all six of its slices are singular.

**Proof.** ($\Rightarrow$) We show that if $X$ has a non-singular slice, then its rank is $\geq 2$. By permuting the directions, we may assume that a frontal slice is non-singular. By permuting the frontal slices, we may assume that $X_1$ is non-singular. If the rank of $X$ is 1 then as in the proof of Lemma 2.3 we have $X_1 = d(a \otimes b)$ where $a$ and $b$ are nonzero vectors in $\mathbb{C}^2$; but this matrix is clearly singular.

($\Leftarrow$) We show that if all six slices are singular then $X$ has rank 1.

Case 1: Some slice is zero; by permuting the directions and slices we may assume that $X_1 = 0$. Since $X_2$ is nonzero and singular we have $X_2 = a \otimes b$ for some nonzero vectors $a, b \in \mathbb{C}^2$. But then $X = a \otimes b \otimes e_1$ where $e_1 = [0, 1]^t$.

Case 2: No slice is zero. Since $X_1$ is nonzero and singular, we have $X_1 = a \otimes b$ where $a = [a_1, a_2]^t$ and $b = [b_1, b_2]^t$ are nonzero vectors. By transposing the
vertical slices of \( X \) if necessary, we may assume that the first column of \( X_1 \) is nonzero. Equivalently, \( b = [1, \lambda]^t \) for some \( \lambda \in \mathbb{C} \); thus \( X_1 = [a|\lambda a] \).

Subcase 2(a): \( \lambda = 0 \). Since the first vertical slice is singular, \( X = [a, 0|\mu a, d] \) for some \( \mu \in \mathbb{C} \) and some \( d \); we have \( d \neq 0 \) since the second vertical slice is nonzero. If \( \mu \neq 0 \) then since \( X_2 \) is singular, there is \( \nu \in \mathbb{C} \setminus \{0\} \) such that \( X = [a, 0|\mu a, \nu a] \). In this case, since the horizontal slices are nonzero, we have \( a_1 \neq 0, a_2 \neq 0 \). Since the horizontal slices are singular, it follows that \( \mu = 0, \nu = 0 \). But then the second frontal slice is zero, giving a contradiction. If \( \mu = 0 \) then \( X = [a, 0|0, d] \). In this case, since the horizontal slices are nonzero and singular, \( X \) must be a superdiagonal array, again giving a contradiction.

Subcase 2(b): \( \lambda \neq 0 \). We have \( X = [a, \lambda a|\mu a, d] \). But \( a \neq 0 \) and the second vertical slice is singular, so \( d = \nu a \) for some \( \nu \in \mathbb{C} \), giving \( X = [a, \lambda a|\mu a, \nu a] \). Since either \( a_1 \neq 0 \) or \( a_2 \neq 0 \) (or both), singularity of the horizontal slices implies that \( \nu = \lambda \mu \). Then \( X = [a, \lambda a|\lambda a, \lambda a] = [a_1, a_2]^t \otimes [1, \lambda]^t \otimes [1, \mu]^t \) has rank 1. \( \square \)

**Remark 2.5.** We now have a partial algorithm for computing the rank of \( X \). If \( X \) is the zero array then \( X \) has rank 0. If \( X \) is a superdiagonal array then \( X \) has rank 2. If \( X \) is nonzero and not superdiagonal, and all of its slices are singular, then \( X \) has rank 1. It remains to consider an array \( X \) with a non-singular slice; by permuting the directions and the slices, we may assume that \( X_1 \) is non-singular.

**Lemma 2.6.** \cite{p.632-633} The rank of a \( 2 \times 2 \times 2 \) array \( X \) is at most 3.

**Proof.** It remains to prove that if the first frontal slice \( X_1 \) is non-singular, then the rank is at most 3. We construct an explicit decomposition with \( n \leq 3 \). We write

\[
X_1 = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \quad Y_2 = X_2 X_1^{-1} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}.
\]

Consider the following matrices:

\[
A = \begin{bmatrix} 1 & 0 & y_{12} \\ 0 & 1 & y_{21} \end{bmatrix}, \quad B = \begin{bmatrix} x_{11} & x_{21} & x_{11} + x_{21} \\ x_{12} & x_{22} & x_{12} + x_{22} \end{bmatrix} = X_1^t \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},
\]

\[
D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} y_{11} - y_{12} & 0 & 0 \\ 0 & y_{22} - y_{21} & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

We then verify by direct calculation that

\[
ADB^t = \begin{bmatrix} 1 & 0 & y_{12} \\ 0 & 1 & y_{21} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} X_1 = X_1,
\]

\[
AEB^t = \begin{bmatrix} 1 & 0 & y_{12} \\ 0 & 1 & y_{21} \end{bmatrix} \begin{bmatrix} y_{11} - y_{12} & 0 & 0 \\ 0 & y_{22} - y_{21} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} X_1 = Y_2 X_1 = X_2.
\]

We now apply Lemma 2.1 to complete the proof. \( \square \)

**Remark 2.7.** Lemmas 2.4 and 2.6 imply that if \( X \) has a non-singular slice then its rank is either 2 or 3. It remains to distinguish these two cases. As before, up to permuting the directions and the slices, we may assume that \( X_1 \) is non-singular.
Definition 2.8. For a $2 \times 2 \times 2$ array $X$, Cayley’s hyperdeterminant is the following homogeneous polynomial of degree 4 in the entries $x_{ijk}$:
\[
\Delta(X) = x_{111}^2 x_{222} + x_{112}^2 x_{221} + x_{121}^2 x_{212} + x_{122}^2 x_{211} \\
- 2 (x_{111} x_{112} x_{221} x_{222} + x_{111} x_{121} x_{212} x_{222} + x_{111} x_{122} x_{211} x_{222} \\
+ x_{112} x_{121} x_{212} x_{221} + x_{112} x_{122} x_{211} x_{222} + x_{121} x_{122} x_{211} x_{212}) \\
+ 4 (x_{111} x_{112} x_{221} x_{212} + x_{112} x_{121} x_{211} x_{222}).
\]

Theorem 2.9. [8 p. 633-634] Let $X$ be a $2 \times 2 \times 2$ array whose first frontal slice $X_1$ is non-singular. If $X_2$ is a scalar multiple of $X_1$, then $X$ has rank 2. If $X_2$ is not a scalar multiple of $X_1$, then (a) if $\Delta(X) \neq 0$ then $X$ has rank 2, and (b) if $\Delta(X) = 0$ then $X$ has rank 3.

Proof. First, assume that $X_2 = \lambda X_1$ for some $\lambda \in \mathbb{C}$. Since $X_1$ is non-singular, it has rank 2, and hence $X_1 = a^{(1)} \otimes b^{(1)} + a^{(2)} \otimes b^{(2)}$. Writing $c = [1, \lambda]^t$ then we see that $X$ has rank 2: $X = X_1 \otimes c = a^{(1)} \otimes b^{(1)} \otimes c + a^{(2)} \otimes b^{(2)} \otimes c$. Second, assume that $X_2$ is not a scalar multiple of $X_1$. We will find a necessary condition for $X$ to have rank 2. We apply Lemma 2.1 with $n = 2$ and write
\[
A = [ \begin{array}{cc} a^{(1)} & a^{(2)} \end{array} ], \quad B = [ \begin{array}{cc} b^{(1)} & b^{(2)} \end{array} ], \quad D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}, \quad E = \begin{bmatrix} e_1 & 0 \\ 0 & e_2 \end{bmatrix}.
\]
But $X_1 = ADB^t$, $X_2 = AEB^t$ gives $X_1 = d_1 a_1 b_1^t + d_2 a_2 b_2^t$, $X_2 = e_1 a_1 b_1^t + e_2 a_2 b_2^t$. Since $X_1$ is non-singular, it has rank 2, and so $d_1 \neq 0$, $d_2 \neq 0$. Since $X_2$ is not a scalar multiple of $X_1$, it follows that $E$ is not a scalar multiple of $D$. Hence $d_1 e_2 - d_2 e_1 \neq 0$, and so $X_2 - d_1^{-1}e_1 X_1$, $X_2 - d_2^{-1}e_2 X_1$ are distinct. We calculate
\[
\begin{align*}
X_2 - d_1^{-1}e_1 X_1 &= e_1 a_1 b_1^t + e_2 a_2 b_2^t - d_1^{-1}e_1 (d_1 a_1 b_1^t + d_2 a_2 b_2^t) \\
&= e_1 a_1 b_1^t + e_2 a_2 b_2^t - e_1 a_1 b_1^t + d_1^{-1}e_2 a_2 b_2^t = d_1^{-1}e_2 (d_2 e_2 - d_2 e_1) a_2 b_2^t, \\
X_2 - d_2^{-1}e_2 X_1 &= e_1 a_1 b_1^t + e_2 a_2 b_2^t - d_2^{-1}e_2 (d_1 a_1 b_1^t + d_2 a_2 b_2^t) \\
&= e_1 a_1 b_1^t + e_2 a_2 b_2^t - d_1 d_2^{-1}e_2 a_1 b_1^t - e_2 a_2 b_2^t = -d_1^{-1}d_2 (d_1 e_2 - d_2 e_1) a_1 b_1^t.
\end{align*}
\]
It follows that these two matrices are singular. Hence the quadratic polynomial $\det(X_2 - \lambda X_1)$ has two distinct roots, but this determinant is
\[
(x_{111} x_{221} - x_{112} x_{222})^2 - (x_{111} x_{222} + x_{112} x_{221} - x_{121} x_{212} - x_{122} x_{211}) \lambda
+ (x_{112} x_{222} - x_{122} x_{212}),
\]
and the discriminant is $\Delta(X)$. Thus if $X$ has rank 2 then $\Delta(X) \neq 0$.

Conversely, suppose that $\Delta(X) \neq 0$. Then $\det(X_2 - \lambda X_1)$ has two distinct roots, say $\lambda_1$, $\lambda_2$. We have two nonzero singular matrices $X_2 - \lambda_1 X_1$, $X_2 - \lambda_2 X_1$. These matrices both have rank 1, and so we can write
\[
(\lambda_1 - \lambda_2)^{-1}(X_2 - \lambda_2 X_1) = u_1 v_1^t, \quad -(\lambda_1 - \lambda_2)^{-1}(X_2 - \lambda_1 X_1) = u_2 v_2^t.
\]
Then we have $X_1 = u_1 v_1^t + u_2 v_2^t$ and $X_2 = \lambda_1 u_1 v_1^t + \lambda_2 u_2 v_2^t$, which imply that $X = u_1 \otimes v_1 \otimes [1, \lambda_1]^t + u_2 \otimes v_2 \otimes [1, \lambda_2]^t$. Thus if $\Delta(X) \neq 0$ then $X$ has rank 2. □

Example 2.10. Consider these arrays, where $X$ is the limit as $a \to 0$ of $Y(a)$:
\[
X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad Y(a) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.
\]
Clearly $X_1$ is non-singular, and $X_2$ is not a scalar multiple of $X_1$. But $\Delta(X) = 0$, and so by Theorem 2.9 the rank of $X$ is 3. For $Y(a)$, the first frontal slice is non-singular and the second frontal slice is not a scalar multiple of the first, but $\Delta(Y(a)) = 4a^2$ which is nonzero for $a \neq 0$. Hence if $a \neq 0$ then $Y(a)$ has rank 2. Thus $X$ is the limit of arrays of rank 2, and so the border rank of $X$ is 2.

It follows from Ja’Ja’ [4, Lemma 3.1] that an array $[I \mid X_2]$ has rank 2 if and only if $X_2$ is similar to a diagonal matrix. The same paper [4, Theorem 3.2] implies that if $X_2$ is the companion matrix of a quadratic polynomial $f(t)$ then $[I \mid X_2]$ has rank 2 if and only if $f(t)$ has two distinct roots; otherwise, it has rank 3. In our example, $X_2$ is the companion matrix of $f(t) = t^2$, so $[I \mid X_2]$ has rank 3. This example is the case $n = 2$ of the pair of bilinear forms in the proof of [4, Theorem 3.5]. A result of von zur Gathen [9, Theorem 4] implies that the maximal bilinear complexity of two $2 \times 2$ matrices over any field is at least 3.

3. Some Lemmas on $3 \times 3 \times 2$ and $3 \times 3 \times 3$ Arrays

Let the $3 \times 3 \times 2$ array over $\mathbb{C}$ have frontal slices $A$ and $B$:

$$[A|B] = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} & b_{11} & b_{12} & b_{13} \\
  a_{21} & a_{22} & a_{23} & b_{21} & b_{22} & b_{23} \\
  a_{31} & a_{32} & a_{33} & b_{31} & b_{32} & b_{33}
\end{bmatrix}$$

Ja’Ja’ [4 Corollary 3.4.1] has shown that the rank of a $p \times p \times 2$ array is at most $[3p/2]$. We give an elementary proof of this result in the case $p = 3$.

**Lemma 3.1.** The rank of a $3 \times 3 \times 2$ array is at most 4.

**Proof.** The maximum rank of a $3 \times 3$ matrix is 3. If both $A$ and $B$ have rank $\leq 2$, then it is straightforward to express $[A|B]$ as a sum of $\leq 4$ simple tensors. We have

$$A = a^{(1)} \otimes b^{(1)} + a^{(2)} \otimes b^{(2)}, \quad B = a^{(3)} \otimes b^{(3)} + a^{(4)} \otimes b^{(4)},$$

and hence

$$[A|B] = a^{(1)} \otimes b^{(1)} \otimes e_1 + a^{(2)} \otimes b^{(2)} \otimes e_1 + a^{(3)} \otimes b^{(3)} \otimes e_2 + a^{(4)} \otimes b^{(4)} \otimes e_2.$$  

We now assume that both $A$ and $B$ have rank $\geq 2$, and that either $A$ or $B$ has rank 3. Interchanging $A$ and $B$ if necessary, we assume that $A$ has rank 3, so that $A$ is invertible. Left multiplication of $A$ and $B$ by $A^{-1}$ (that is, applying a change of basis in the first direction) gives the array $[I|C]$ where the second frontal slice $C = A^{-1}B$ still has rank $\geq 2$. There exists an invertible matrix $E$ such that $J = E^{-1}CE$ is the Jordan canonical form of $C$. (Here we use the assumption that the base field is algebraically closed.) Clearly $E^{-1}IE = I$, so we act on $[I|C]$ by $E^{-1}$ along the first direction and by $E$ along the second direction, to obtain $[I|J]$, where the second frontal slice $J$ still has rank $\geq 2$. It remains to show that any such array has rank $\leq 4$. There are three cases for the Jordan canonical form of a $3 \times 3$ matrix $J$. Case 1: Three $1 \times 1$ Jordan blocks; $J$ is a diagonal matrix:

$$[I|J] = \begin{bmatrix}
  1 & 0 & 0 & d_1 & 0 & 0 \\
  0 & 1 & 0 & 0 & d_2 & 0 \\
  0 & 0 & 1 & 0 & 0 & d_3
\end{bmatrix}$$

Then clearly the rank is $\leq 3$:

$$[I|J] = e_1 \otimes e_1 \otimes \begin{bmatrix} d_1 \end{bmatrix} + e_2 \otimes e_2 \otimes \begin{bmatrix} d_2 \end{bmatrix} + e_3 \otimes e_3 \otimes \begin{bmatrix} d_3 \end{bmatrix}.$$
Case 2: One $2 \times 2$ block and one $1 \times 1$ block:

$$[I|J] = \begin{bmatrix} 1 & 0 & 0 & d_1 & 1 & 0 \\ 0 & 1 & 0 & 0 & d_1 & 0 \\ 0 & 0 & 1 & 0 & 0 & d_2 \end{bmatrix}$$

We have $[I|J] = [I|D] + [O|F]$ where $D$ is a diagonal matrix and $F$ is the matrix unit $E_{12}$. By the previous case, $[I|D]$ has rank $\leq 3$, and clearly $[O|F]$ has rank 1.

Case 3: One $3 \times 3$ block:

$$[I|J] = \begin{bmatrix} 1 & 0 & 0 & d_1 & 1 & 0 \\ 0 & 1 & 0 & 0 & d_1 & 1 \\ 0 & 0 & 1 & 0 & 0 & d_1 \end{bmatrix}$$

We add $-d_1$ times the first frontal slice to the second frontal slice; that is, we change basis along the third direction by the matrix

$$\begin{bmatrix} 1 & 0 \\ -d_1 & 1 \end{bmatrix}$$

We obtain this array:

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

It remains to prove that this array has rank $\leq 4$. We have the following explicit representation as a sum of four simple tensors:

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

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

This completes the proof. \qed

Let the $3 \times 3 \times 3$ array $T$ over $\mathbb{C}$ have frontal slices $A$, $B$ and $C$:

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

**Lemma 3.2. Kruskal’s One-Edge Lemma.** If the array $T$ has parallel slices $D$ and $E$ for which there exists a nonzero vector $x$ such that $Dx = Ex = 0$ or $D^t x = E^t x = 0$, then $\text{rank}(T) \leq 5$. 
Proof. Permuting the directions if necessary, we may assume that \( D \) and \( E \) are frontal slices. Permuting the frontal slices if necessary, we may assume that \( D \) and \( E \) are the first and second frontal slices \( A \) and \( B \). Suppose that \( Ax = Bx = 0 \) where \( x \neq 0 \). Let \( X \) be a \( 3 \times 3 \) non-singular matrix which has \( x \) as its first column. (Extend the set \( \{ x \} \) to a basis \( \{ x, y, z \} \) of \( \mathbb{C}^3 \) and let \( X = [x|y|z] \).

Acting on \( T \) along the second direction gives \([AX]BX[CX] \), but \( Ax = Bx = 0 \), so

\[
[AX]BX[CX] = \begin{bmatrix}
0 & a_1' & 0 & b_1' & 0 & c_1'' & 0 & 0 & 0 \\
0 & a_2' & 0 & b_2' & 0 & c_2'' & 0 & 0 & 0 \\
0 & a_3' & 0 & b_3' & 0 & c_3'' & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

The first term is a simple tensor,

\[
\begin{bmatrix}
c_{11} \\
c_{21} \\
c_{31}
\end{bmatrix} \otimes \begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix} \otimes \begin{bmatrix}
0 \\
1
\end{bmatrix},
\]

and so it remains to prove that the second term has rank \( \leq 4 \). To write the second term as a sum of simple tensors it suffices to decompose this \( 3 \times 2 \times 3 \) array:

\[
\begin{bmatrix}
a_1'' & b_1' & c_1'' \\
a_2'' & b_2' & c_2'' \\
a_3'' & b_3' & c_3'' \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

Transposing the second and third directions, we may consider the \( 3 \times 3 \times 2 \) array,

\[
\begin{bmatrix}
a_1'' & b_1' & c_1'' \\
a_2'' & b_2' & c_2'' \\
a_3'' & b_3' & c_3'' \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

The claim now follows from Lemma 3.1.

If \( A'x = B'x = 0 \), then we transpose the matrices \( A, B \) and \( C \) and use the analogous reasoning; this can also be expressed in terms of a transposition of the first two directions in the array \( T = [A|B|C] \).

\[\square\]

Lemma 3.3. Kruskal’s Two-Edge Lemma. If the array \( T \) has frontal slices \( D \) and \( E \) for which there exist nonzero vectors \( x \) and \( y \) such that \( Dx = y'D = 0 \) and \( y'E \neq 0 \), then \( \text{rank}(T) \leq 5 \).

Proof. As before, we may assume that \( D \) and \( E \) are the first and second frontal slices \( A \) and \( B \). Let \( x \) and \( y \) satisfy the conditions of the lemma. We choose vectors \( u_2, u_3, v_2, v_3 \) such that \( U = [x|u_2|u_3] \) and \( V = [y|v_2|v_3] \) are nonsingular. Then

\[
V^tAU = \begin{bmatrix}
0 & 0 & 0 \\
0 & ** & * \\
0 & * & *
\end{bmatrix}, \quad V^tBU = \begin{bmatrix}
\alpha & * & * \\
* & * & * \\
* & * & *
\end{bmatrix}, \quad \text{and} \quad V^tCU = \begin{bmatrix}
\beta & * & * \\
* & * & * \\
* & * & *
\end{bmatrix},
\]

where \( \alpha = y^tBx \neq 0 \), but \( \beta \) can be 0, and * denotes unspecified elements (which are not necessarily equal). If \( \beta = 0 \) then we add \( B \) to \( C \) to make \( \beta = \alpha \neq 0 \). Equivalently, we change basis along the third direction in \( T \) by the matrix

\[
\begin{bmatrix}
1 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]
Since \( \alpha \neq 0 \), we can construct a matrix \( X \) of rank 1 which has the same first row and first column as \( V^t BU \); explicitly,

\[
V^t BU = \begin{bmatrix}
\alpha & \alpha' & \alpha''
\gamma & * & *
\delta & * & *
\end{bmatrix}, \quad X = \begin{bmatrix}
\frac{\alpha}{\gamma} & \frac{\alpha'}{\gamma} & \frac{\alpha''}{\gamma}
\frac{\alpha}{\delta} & \frac{\alpha'}{\delta} & \frac{\alpha''}{\delta}
\end{bmatrix}
\]

Similarly, we can construct a matrix \( Y \) of rank 1 which has the same first row and first column as \( V^t CU \). Then the two arrays \([0]X[0]\) and \([0]Y[0]\) also have rank 1 as \(3 \times 3 \times 3\) arrays; that is, they are simple tensors. We now see that

\[
V^t[A|B|C]U - [0]X[0] - [0]Y[0] = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
0 & * & 0 & * & 0 & * & 0 & 0 & *
0 & * & 0 & * & 0 & * & 0 & 0 & *
\end{bmatrix}
\]

It remains to decompose a \(2 \times 2 \times 2\) array, and this requires at most three simple tensors according to the results of Section 2.

\[ \square \]

4. PROOF OF KRUSKAL’S THEOREM ON \(3 \times 3 \times 3\) ARRAYS

**Theorem 4.1.** Every \(3 \times 3 \times 3\) array \( T = [A|B|C] \) over \( \mathbb{C} \) has rank \( \leq 5 \).

**Proof.** If \( C = 0 \) then the problem reduces to considering a \(3 \times 3 \times 2\) array, which has rank \( \leq 4 \) by Lemma 3.1. We assume from now on that \( C \neq 0 \). Consider the \(3 \times 3\) matrix \( A - \lambda C \); its determinant is a nonconstant polynomial in \( \lambda \), which has a root over \( \mathbb{C} \). (Here again we use the assumption that the base field is algebraically closed.) Thus by subtracting a multiple of \( C \) from \( A \), we may ensure that \( A \) is singular, and so rank(\( A \)) \( \leq 2 \). Equivalently, we change basis along the third direction in \( T \) by the matrix

\[
\begin{bmatrix}
1 & 0 & -\lambda \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

The same considerations apply to \( B \). We assume from now on that the first and second frontal slices of \( T \) both have rank \( \leq 2 \).

Suppose that some frontal slice has rank \( \leq 1 \); up to permuting these slices, we may assume that rank(\( A \)) \( \leq 1 \). If rank(\( A \)) = 0, then\( A \) is the 0 matrix, and we have a \(3 \times 3 \times 2\) array, which has rank \( \leq 4 \) by Lemma 3.1. If rank(\( A \)) = 1, then the array \([A|0|0]\) has rank 1; subtracting this simple tensor from \( T \) leaves a \(3 \times 3 \times 2\) array which has rank \( \leq 4 \), and so \( T \) has rank \( \leq 5 \).

We may now assume that \( A \) and \( B \) have rank 2, and that \( C \) has rank 2 or 3.

**Case 1:** \( A \), \( B \) and \( C \) all have rank 2. It follows that there exist nonzero vectors \( x_1, x_2, x_3 \) and \( y_1, y_2, y_3 \) (basis vectors for the right and left nullspaces) such that

\[
A x_1 = B x_2 = C x_3 = 0, \quad y_1^t A = y_2^t B = y_3^t C = 0.
\]

Then for the \(3 \times 3\) matrices \( X = [x_1|x_2|x_3] \) and \( Y = [y_1|y_2|y_3] \) we have

\[
(1) \quad Y^t AX = \begin{bmatrix}
0 & 0 & 0 \\
0 & * & * \\
0 & * & *
\end{bmatrix}, \quad Y^t BX = \begin{bmatrix}
* & 0 & * \\
0 & 0 & 0 \\
* & 0 & *
\end{bmatrix}, \quad Y^t CX = \begin{bmatrix}
* & * & 0 \\
* & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}.
\]

If the conditions of Lemma 3.3 are satisfied for any two frontal slices, then the proof is complete. Otherwise, it follows that

\[
(2) \quad y_i^t A x_i = y_i^t B x_i = y_i^t C x_i = 0, \quad \text{for all } i = 1, 2, 3.
\]
Consider these three subcases:

Subcase 1.1: Two columns of $X$ are linearly dependent (that is, one column is a scalar multiple of another). Then Lemma 3.2 completes the proof.

Subcase 1.2: The matrix $X$ has rank 2, but no two columns are linearly dependent. Then $x_1$ and $x_2$ are linearly independent, and so $x_3 = \beta x_1 + \gamma x_2$ for some $\beta, \gamma \in \mathbb{C} \setminus \{0\}$. We choose vectors $u, v_2, v_3$ such that the matrices $U = [x_1|x_2|u]$ and $V = [y_1|v_2|v_3]$ are invertible. Then for some $\delta \in \mathbb{C}$ we have

$$V^t A U = \begin{bmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}, \quad V^t B U = \begin{bmatrix} 0 & 0 & * \\ * & 0 & * \\ * & * & * \end{bmatrix}, \quad V^t C U = \begin{bmatrix} 0 & \delta & * \\ * & * & * \\ * & * & * \end{bmatrix}.$$

(The $(1, 1)$ entries of $V^t B U$ and $V^t C U$ are zero; otherwise Lemma 3.3 would apply.) But $C x_3 = 0$ implies $\beta C x_1 + \gamma C x_2 = 0$, and so the first two columns of $V^t C U$ are linearly dependent, implying $\delta = 0$. Hence the first three rows of $V^t A U$, $V^t B U$ and $V^t C U$ are linearly dependent. We subtract the simple tensor in which the first horizontal slice is the same as that of $V^t [A|B|C] U$ and the second and third horizontal slices are zero. There remains an array in which the first horizontal slice is zero, and the second and third horizontal slices are the same as those of $V^t [A|B|C] U$. But the rank of this $2 \times 3 \times 3$ array is at most 4 by Lemma 3.1.

Subcase 1.3: The matrix $X$ has rank 3. If $Y$ has rank $\leq 2$, then we replace each frontal slice $A$, $B$, $C$ by its transpose (equivalently, we interchange the first two directions of $T$), and then we may apply one of the previous subcases. So we assume that $Y$ has rank 3. Combining (1) and (2) gives three matrices of rank 2:

$$Y^t A X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \beta \\ 0 & \alpha & 0 \end{bmatrix}, \quad Y^t B X = \begin{bmatrix} 0 & 0 & \delta \\ 0 & 0 & 0 \\ \gamma & 0 & 0 \end{bmatrix}, \quad Y^t C X = \begin{bmatrix} 0 & \zeta & 0 \\ \epsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

If we transpose the first two columns of $X$ (that is, interchange the first two vertical slices of $T$) then we obtain

$$Y^t A X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \beta \\ \alpha & 0 & 0 \end{bmatrix}, \quad Y^t B X = \begin{bmatrix} 0 & 0 & \delta \\ 0 & 0 & 0 \\ 0 & \gamma & 0 \end{bmatrix}, \quad Y^t C X = \begin{bmatrix} \zeta & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We subtract from $Y^t [A|B|C] X$ the following sum of three simple tensors:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha & -\beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \zeta & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We obtain the following array:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 & \gamma & 0 \\ 0 & \beta & \beta & 0 & 0 & 0 & 0 \end{bmatrix}$$
The three frontal slices are linear combinations of these two matrices of rank 1:
\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 1 \\
0 & 1 & 1
\end{bmatrix}, \quad \begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}.
\]

Therefore (3) is the sum of two simple tensors:
\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & \beta & \beta & \gamma & \gamma & 0 \\
0 & \beta & \beta & 0 & \gamma & \gamma
\end{bmatrix} + \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\beta & 0 & 0 & -\gamma & 0 \\
0 & 0 & 0 & 0 & 0 & \epsilon
\end{bmatrix}
\]

We now have a decomposition of the original $3 \times 3 \times 3$ array $T$ into a sum of at most 5 simple tensors.

**Case 2:** $A$ and $B$ have rank 2 but $C$ has rank 3.

**Subcase 2.1:** There exist $\alpha, \beta \in \mathbb{C}$ such that $\alpha A + \beta B + C$ has rank 2. This corresponds to changing basis in $T$ along the third direction by the matrix
\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\alpha & \beta & 1
\end{bmatrix}
\]

Then we are back in Case 1. Such scalars may not exist; a simple example is
\[
A = B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

**Subcase 2.2:** The matrix $\alpha A + \beta B + C$ has rank 3 for all $\alpha, \beta \in \mathbb{C}$. There exist nonzero vectors $x_1, x_2, x_3$ and $y_1, y_2, y_3$ such that
\[
Ax_1 = Bx_2 = 0, \quad Cx_3 = Ax_2, \quad y_1^t A = y_2^t B = 0, \quad y_3^t C = y_1^t A.
\]

If we can apply Lemma 3.3 then we are done. So we may assume that Lemma 3.3 does not apply, and hence we must have
\[
y_1^t Ax_1 = y_1^t Bx_1 = y_1^t Cx_1 = 0, \quad y_2^t Ax_2 = y_2^t Bx_2 = y_2^t Cx_2 = 0.
\]

We write $X = [x_1|x_2|x_3]$ and $Y = [y_1|y_2|y_3]$.

**Subcase 2.2.1:** We have linear dependence of $x_1$ and $x_2$, or of $y_1$ and $y_2$, or both. Then the result follows from Lemma 3.2.

**Subcase 2.2.2:** We have linear independence of $x_1$ and $x_2$, and of $y_1$ and $y_2$. Then both matrices $X$ and $Y$ have rank $\geq 2$.

Assume $X$ has rank 2. Then $x_3 = \gamma x_1 + \delta x_2$ for some $\gamma, \delta \in \mathbb{C}$. There exist nonzero vectors $u, v$ such that $U = [x_1|x_2|u], \quad V = [y_1|y_2|v]$ both have rank 3. Then using the previous equations we have
\[
V^t AU = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \zeta \\ 0 & * & * \end{bmatrix}, \quad V^t BU = \begin{bmatrix} 0 & 0 & \zeta \\ 0 & 0 & 0 \\ \epsilon & 0 & * \end{bmatrix}, \quad V^t CU = \begin{bmatrix} 0 & \theta & * \\ \eta & 0 & * \\ * & * & * \end{bmatrix}.
\]

Since $B$ has rank 2, it follows that $\epsilon \neq 0$ and $\zeta \neq 0$. We have
\[
V^t Ax_2 = V^t Cx_3 = V^t C(\gamma x_1 + \delta x_2) = \gamma V^t Cx_1 + \delta V^t Cx_2,
\]

and so the second column of $V^t AU$ is a linear combination of the first two columns of $V^t CU$. Since $x_3 \neq 0$, it follows that $\gamma$ and $\delta$ are not both 0, and so $\eta = 0$ or $\theta = 0$ (or both). In either case, adding a multiple of $B$ to $C$ (that is, changing basis along the third direction), and applying the same change of basis matrices $V^t$ and
along the first and second directions, gives an array in which the third frontal slice has rank 2. This contradicts the assumption that \( \alpha A + \beta B + C \) has rank 3.

Assume \( X \) has rank 3. If \( Y \) has rank 2, then we interchange the first and second directions of the array, which amounts to applying the usual matrix transpose to the frontal slices \( A, B, C \). Equivalently, we interchange \( X \) and \( Y \), which reduces to the previous paragraph. So we may assume that \( Y \) also has rank 3.

Using the previous equations, together with
\[
y_t^3 A x_3 = y_t^3 C x_3 = y_t^3 A x_2,
\]
we obtain
\[
Y^t AX = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \gamma \\ 0 & \gamma & \delta \end{bmatrix}.
\]

Using the previous equations, we obtain
\[
Y^t BX = \begin{bmatrix} 0 & 0 & \zeta \\ 0 & 0 & 0 \\ \epsilon & 0 & \eta \end{bmatrix}.
\]

Using the previous equations, together with
\[
y_1^t C x_1 = 0, \quad y_1^t C x_3 = y_1^t A x_2 = 0,
\]
\[
y_2^t C x_2 = 0, \quad y_2^t C x_3 = y_2^t A x_2 = 0,
\]
\[
y_3^t C x_1 = y_3^t A x_1 = 0, \quad y_3^t C x_2 = y_3^t A x_2 = 0,
\]
we obtain
\[
Y^t CX = \begin{bmatrix} 0 & \lambda & 0 \\ \kappa & 0 & 0 \\ 0 & 0 & \gamma \end{bmatrix}.
\]

If \( \delta \neq 0 \) (respectively \( \eta \neq 0 \)) then we add a multiple of \( A \) (respectively \( B \)) to \( C \) to eliminate \( \gamma \) and obtain an array for which \( Y^t CX \) has rank 2; but this contradicts the assumption that \( \alpha A + \beta B + C \) has rank 3. So we may assume that \( \delta = \eta = 0 \):
\[
Y^t AX = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \gamma \\ 0 & \gamma & 0 \end{bmatrix}, \quad Y^t BX = \begin{bmatrix} 0 & 0 & \zeta \\ 0 & 0 & 0 \\ \epsilon & 0 & 0 \end{bmatrix}, \quad Y^t CX = \begin{bmatrix} 0 & \lambda & 0 \\ \kappa & 0 & 0 \\ 0 & 0 & \gamma \end{bmatrix}.
\]

Interchanging the first and second vertical slices of \( T = [A|B|C] \), and applying the same transformations, amounts to interchanging the first and second columns in each of the above matrices. We now have this array:
\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & \gamma \\
\gamma & 0 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 0 & \zeta/\lambda \\
0 & 0 & 0 \\
0 & \epsilon/\gamma & 0
\end{bmatrix}.
\]
But \( \lambda, \kappa, \gamma \) are all nonzero by our assumption that the third frontal slice has rank 3.

We scale the first, second and third horizontal slices by \( 1/\lambda, 1/\kappa, 1/\gamma \) respectively:
\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & \zeta/\lambda \\
1 & 0 & 0
\end{bmatrix}.
\]
From this array we subtract the following array of rank 2, where the bars denote complex conjugates:

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \gamma/\kappa & 0 & 0 & -\epsilon/\gamma & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & -\zeta/\lambda & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\times
\begin{bmatrix}
\gamma/\kappa & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\epsilon/\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta/\lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

and obtain

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta/\lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\times
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\epsilon/\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

The second frontal slice (which we still denote by \(B\)) is now Hermitian, and so its Jordan canonical form \(J = E^{-1}BE\) is a diagonal matrix. Changing basis along the first and second directions by \(E^{-1}\) and \(E\) respectively, we obtain

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mu & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\nu & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\xi & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\end{bmatrix}
\]

This array clearly has rank 3, and the proof is complete. \(\Box\)

5. Arrays over the field with two elements

In this section we use computer algebra to classify the canonical forms of \(3 \times 3 \times 3\) arrays \(X = [x_{ijk}]\) over the field \(\mathbb{F}_2\) with two elements. We use the term tensor for such an array to avoid confusion with the data structures called arrays in Maple. The flattening of \(X\) is the row vector \(\text{flat}(X) = [x_{111}, \ldots, x_{ijk}, \ldots, x_{333}]\), where the entries are in lexic order by subscripts. Conversely, the unflattening of such a row vector is the corresponding tensor. We encode \(X\) as the non-negative integer whose representation in base 2 is \(\text{flat}(X)\). Conversely, the decoding of an integer in the range \(0, \ldots, 2^{27} - 1\) is the corresponding tensor. The lexic order on flattenings coincides with the natural order on integers. The minimal element of a set of tensors is defined in terms of this total order. We identify \(X\) with an element of \(\mathbb{F}_2^3 \otimes \mathbb{F}_2^3 \otimes \mathbb{F}_2^3\).

The direct product of general linear groups \(\text{GL}_3(\mathbb{F}_2) \times \text{GL}_3(\mathbb{F}_2) \times \text{GL}_3(\mathbb{F}_2)\) acts on \(\mathbb{F}_2^3 \otimes \mathbb{F}_2^3 \otimes \mathbb{F}_2^3\), and the canonical form of a tensor is the minimal element in its orbit under this group action. The finite group \(\text{GL}_3(\mathbb{F}_2)\) has order 168, and is generated by two elements: the cyclic permutation \(e_1 \mapsto e_2, e_2 \mapsto e_3, e_3 \mapsto e_1\), and the row operation \(e_1 \mapsto e_1 + e_2, e_2 \mapsto e_2, e_3 \mapsto e_3\). The group \(\text{GL}_3(\mathbb{F}_2) \times \text{GL}_3(\mathbb{F}_2) \times \text{GL}_3(\mathbb{F}_2)\) has order 4741632 and is generated by 6 elements.

For a tensor \(X\) over \(\mathbb{F}_2\), we use the spinning algorithm to compute its orbit. In the following pseudocode, \(\mathcal{O}\) is the current value of the orbit, \(\mathcal{L}\) contains the new elements computed during the previous iteration, and \(\mathcal{N}\) contains the new elements computed during the current iteration:

1. \(\mathcal{O} \leftarrow \emptyset; \mathcal{L} \leftarrow \{X\}\)
2. while \(\mathcal{L} \neq \emptyset\) do:
   a. \(\mathcal{O} \leftarrow \mathcal{O} \cup \mathcal{L}\)
   b. \(\mathcal{N} \leftarrow \emptyset;\) for \(Y \in \mathcal{L}\) do for \(M \in \mathcal{G}\) do: \(\mathcal{N} \leftarrow \mathcal{N} \cup \{M \cdot Y\}\)
   c. \(\mathcal{L} \leftarrow \mathcal{N} \setminus \mathcal{O}\)
MURRAY R. BREMNER AND JIAXIONG HU

We first create a large Maple array, called orbitarray, with $2^{27} - 1$ entries. The indices of orbitarray correspond to nonzero tensors: for an index $i$ we first decode $i$ by writing it as a binary numeral of 27 bits (adding leading 0s if necessary), and then unflatten this binary numeral to obtain the corresponding tensor. To start, every entry of orbitarray is set to 0. We then perform the following iteration:

(1) $\omega \leftarrow 0$, $i \leftarrow 0$
(2) While $i < 2^{27} - 1$ do:
   (a) $i \leftarrow i + 1$
   (b) if orbitarray[$i$] = 0 then
       (i) $\omega \leftarrow \omega + 1$
       (ii) findorbit[$i$]

Procedure findorbit takes the index $i$, decodes and unflattens it to the corresponding tensor $X$, uses the spinning algorithm to generate the orbit $O(X)$, and sets the corresponding entries of orbitarray to the orbit index $\omega$. Upon termination, $\omega$ equals the total number of orbits for the group action, and orbitarray represents the function which assigns to each tensor the index number of its orbit.

The natural order of the index numbers of the orbits agrees with the lex order on the minimal elements in the orbits (the canonical forms of the tensors).

The next step is to compute the ranks of the orbits. We create another Maple array, called linkarray, of the same size as orbitarray. We use the data from orbitarray to set entry $i$ of linkarray (representing the tensor $X$) equal to the index $j$ of the next tensor in lex order in the orbit containing $X$. We then create another Maple array of the same size, called rankarray, and initialize every entry to 0. We generate all simple tensors (tensor products of nonzero vectors) and set the corresponding entries of rankarray to 1. Each index $i$ for which rankarray[$i$] = 0 represents the encoding of a tensor of rank 1. Let $E$ denote the minimal tensor of rank 1: its flattening is $[0, \ldots, 0, 1]$. We then perform the following iteration:

(1) oldrank $\leftarrow 0$, finished $\leftarrow$ false
(2) While not finished do:
   (a) oldrank $\leftarrow$ oldrank + 1, finished $\leftarrow$ true
   (b) For each index $i$ for which rankarray[$i$] = oldrank, do:
       (i) Let $X$ be the unflattening of the decoding of $i$.
       (ii) Set $Y \leftarrow X + E$: this amounts to changing the rightmost bit of the flattening of $X$ from 0 to 1 or from 1 to 0.
       (iii) Let $j$ be the encoding of the flattening of $Y$. Thus $j = i + 1$ if $i$ is even, and $j = i - 1$ if $i$ is odd.
       (iv) If rankarray[$j$] = 0, then $Y$ has rank oldrank + 1. In this case:
           • Use linkarray to store oldrank + 1 in every entry of rankarray corresponding to the tensors in the orbit of $Y$.
           • finished $\leftarrow$ false

The iteration terminates when every entry of rankarray contains a positive integer, which is the rank of the corresponding (nonzero) tensor.

To reduce the number of orbits, we consider the larger group

$$G = (GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2)) \rtimes S_3,$$

where the symmetric group $S_3$ permutes the three directions. We first compute the small orbits obtained by the action of $GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2)$ and then
apply the permutations to determine which small orbits combine to make a single large orbit. Given the canonical form $X$ of a small orbit $O$ with index number $i$, we apply the elements of $S_3$ to obtain tensors $X_1 = X, \ldots, X_6$. We then use the Maple arrays, which we have already computed, to find the index numbers $i_1 = i, \ldots, i_6$ of the small orbits containing these tensors. We conclude that the union $O_{i_1} \cup \cdots \cup O_{i_6}$ is a large orbit for the action of $G$. The canonical form for this large orbit is the smallest (in lex order) of the canonical forms of $O_{i_1}, \ldots, O_{i_6}$. There are 115 (nonzero) small orbits and 55 (nonzero) large orbits:

| rank | # small | # tensors | percent |
|------|---------|-----------|---------|
| 0    | 1       | 1         | 0.0000  |
| 1    | 1       | 343       | 0.0003  |
| 2    | 1       | 43218     | 0.0322  |
| 3    | 1       | 2372286   | 1.7675  |
| 4    | 8       | 47506872  | 35.3954 |
| 5    | 18      | 83670048  | 62.3390 |
| 6    | 23      | 624960    | 0.4656  |

For the large orbit sizes and canonical forms, see Table 1. This computation took just under 282 minutes with Maple 16 on a Lenovo ThinkCentre M91p Tower 7052A8U i7-2600 CPU (Quad Core 3.40/3.80GHz) using Windows 7 Professional 64-bit with 16 gigabytes of RAM.

For similar results for other tensor formats over $\mathbb{F}_2$, see [1, 2].

6. ACKNOWLEDGEMENTS

The first author was partially supported by a Discovery Grant from NSERC. We thank J. M. F. ten Berge for sending us the unpublished notes by Rocci [7].

REFERENCES

[1] M. R. Bremner, J. Hu: Canonical forms of small tensors over $\mathbb{F}_2$. arXiv:1206.5179v1 [math.CO]

[2] M. R. Bremner, S. G. Stavrou: Canonical forms of $2 \times 2 \times 2$ and $2 \times 2 \times 2 \times 2$ tensors over $\mathbb{F}_2$ and $\mathbb{F}_3$. Linear and Multilinear Algebra (to appear).

[3] V. de Silva, L.-H. Lim: Tensor rank and the ill-posedness of the best low-rank approximation problem. SIAM J. Matrix Anal. Appl. 30 (2008) 1084–1127.

[4] J. Ji’Ja’i: Optimal evaluation of pairs of bilinear forms. SIAM J. Comput. 8 (1979) 443–462.

[5] T. G. Kolda, B. W. Bader: Tensor decompositions and applications. SIAM Rev. 51 (2009) 455–500.

[6] J. B. Kruskal: Rank, decomposition, and uniqueness for 3-way and $N$-way arrays. In: R. Coppi, S. Bolasco (editors), Multiway Data Analysis, 7–18. North-Holland, Amsterdam, 1989.

[7] R. Rocci: A $3 \times 3 \times 3$ array has maximum rank 5. (Microsoft Word document, no date.)

[8] J. M. F. ten Berge: Kruskal’s polynomial for $2 \times 2 \times 2$ arrays, and a generalization to $2 \times n \times n$ arrays. Psychometrika 56 (1991) 631–636.

[9] J. von zur Gathen: Maximal bilinear complexity and codes. Linear Algebra Appl. 144 (1991) 49–61.
\[
\begin{array}{cccc}
\# & \text{rank} & \text{size} & \text{canonical form} \\
1 & 1 & 343 & 1 \ldots 1 \\
2 & 2 & 6174 & 1 \ldots 1 \\
3 & 2 & 37044 & 1 \ldots 1 \\
4 & 3 & 3528 & 1 \ldots 1 \\
5 & 3 & 4116 & 1 \ldots 1 \\
6 & 3 & 18522 & 1 \ldots 1 \\
7 & 3 & 148176 & 1 \ldots 1 \\
8 & 3 & 222264 & 1 \ldots 1 \\
9 & 3 & 592704 & 1 \ldots 1 \\
10 & 3 & 592704 & 1 \ldots 1 \\
11 & 3 & 790272 & 1 \ldots 1 \\
12 & 4 & 148176 & 1 \ldots 1 \\
13 & 4 & 199568 & 1 \ldots 1 \\
14 & 4 & 222264 & 1 \ldots 1 \\
15 & 4 & 263424 & 1 \ldots 1 \\
16 & 4 & 44528 & 1 \ldots 1 \\
17 & 4 & 592704 & 1 \ldots 1 \\
18 & 4 & 1185408 & 1 \ldots 1 \\
19 & 4 & 1778112 & 1 \ldots 1 \\
20 & 4 & 1778112 & 1 \ldots 1 \\
21 & 4 & 2370816 & 1 \ldots 1 \\
22 & 4 & 2370816 & 1 \ldots 1 \\
23 & 4 & 592704 & 1 \ldots 1 \\
24 & 4 & 1185408 & 1 \ldots 1 \\
25 & 4 & 1778112 & 1 \ldots 1 \\
26 & 4 & 1778112 & 1 \ldots 1 \\
27 & 4 & 1778112 & 1 \ldots 1 \\
28 & 4 & 2370816 & 1 \ldots 1 \\
29 & 4 & 2370816 & 1 \ldots 1 \\
30 & 4 & 1778112 & 1 \ldots 1 \\
31 & 5 & 28224 & 1 \ldots 1 \\
32 & 5 & 148176 & 1 \ldots 1 \\
33 & 5 & 148176 & 1 \ldots 1 \\
34 & 5 & 169344 & 1 \ldots 1 \\
35 & 5 & 592704 & 1 \ldots 1 \\
36 & 5 & 1185408 & 1 \ldots 1 \\
37 & 5 & 1580544 & 1 \ldots 1 \\
38 & 5 & 1580544 & 1 \ldots 1 \\
39 & 5 & 1778112 & 1 \ldots 1 \\
40 & 5 & 1778112 & 1 \ldots 1 \\
41 & 5 & 2370816 & 1 \ldots 1 \\
42 & 5 & 2370816 & 1 \ldots 1 \\
43 & 5 & 2370816 & 1 \ldots 1 \\
44 & 5 & 3556224 & 1 \ldots 1 \\
45 & 5 & 4741632 & 1 \ldots 1 \\
46 & 5 & 4741632 & 1 \ldots 1 \\
47 & 5 & 4741632 & 1 \ldots 1 \\
48 & 5 & 4741632 & 1 \ldots 1 \\
49 & 5 & 4741632 & 1 \ldots 1 \\
50 & 5 & 7112448 & 1 \ldots 1 \\
51 & 5 & 7112448 & 1 \ldots 1 \\
52 & 5 & 7112448 & 1 \ldots 1 \\
53 & 5 & 7112448 & 1 \ldots 1 \\
54 & 5 & 7112448 & 1 \ldots 1 \\
55 & 5 & 7112448 & 1 \ldots 1 \\
\end{array}
\]

Table 1. Large orbits of $3 \times 3 \times 3$ tensors over $F_2$