# 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, we consider $3 \times 3 \times 3$ arrays with real or complex entries, and provide a 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 the field $\mathbb{F}_2$ with two elements; 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 A few years later, Rocci [7] circulated a simplified proof of this result, based on Kruskal's unpublished hand-written notes. This result has been part of the 'folklore' of multilinear algebra for more than two decades, but complete details of the proof appear never to have been published.

In §§2–4, we consider $3 \times 3 \times 3$ arrays with real or complex entries, and provide a self-contained proof that the maximum rank is 5. (Our proof in the complex case also holds over any algebraically closed field of characteristic $\neq 2$.)

In §5 we consider the same 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 use computer algebra to 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:

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

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 $X_1, \dots, X_r$ from left to right:

$$X = \left[ \; X_1 \; \middle| \; \cdots \; \middle| \; X_r \; \right] = \left[ \begin{array}{ccc|c|ccc} x_{111} & \cdots & x_{1q1} & \cdots & x_{11r} & \cdots & x_{1qr} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \cdots \\ x_{p11} & \cdots & x_{pq1} & \cdots & x_{p1r} & \cdots & x_{pqr} \end{array} \right].$$

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

*Key words and phrases.* Kruskal's theorem, 3-dimensional arrays, tensor decompositions, canonical forms.

Given three column vectors,

$$\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \in \mathbb{F}^p, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_q \end{bmatrix} \in \mathbb{F}^q, \qquad \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_r \end{bmatrix} \in \mathbb{F}^r,$$

their outer product $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}$ is the $p \times q \times r$ array whose $ijk$ entry is $a_i b_j c_k$. A simple tensor (also called a decomposable tensor) is an outer product of nonzero vectors. A fundamental problem is to represent the array $X$ as a sum of simple tensors:

$$X = \sum_{i=1}^{n} \mathbf{a}^{(i)} \otimes \mathbf{b}^{(i)} \otimes \mathbf{c}^{(i)}.$$

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

$$\big( (\alpha, \beta, \gamma) \cdot X \big)_{ijk} = x_{\alpha(i)\beta(j)\gamma(k)}.$$

More generally, the rank does not change if we apply a change of basis in each direction. Given invertible matrices

$$A = (a_{i_1 i_2}) \in GL(p, \mathbb{F}), \qquad B = (b_{j_1 j_2}) \in GL(q, \mathbb{F}), \qquad C = (c_{k_1 k_2}) \in GL(r, \mathbb{F}),$$

we form another $p \times q \times r$ array by

$$\big( (A, B, C) \cdot X \big)_{i_1 j_1 k_1} = \sum_{i_2=1}^{p} \sum_{j_2=1}^{q} \sum_{k_2=1}^{r} a_{i_1 i_2} b_{j_1 j_2} c_{k_1 k_2} x_{i_2 j_2 k_2}.$$

The rank also 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

$$(\delta \cdot X)_{i_{\delta(1)} i_{\delta(2)} i_{\delta(3)}} = x_{i_1 i_2 i_3}.$$

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

In §§2–4, the base field $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. In §5 the base field is the field $\mathbb{F}_2$ with two elements.

## 2. ten Berge's theorem on $2 \times 2 \times 2$ arrays

The results in this section are taken from ten Berge [8], who considers only the case $\mathbb{F} = \mathbb{R}$. With very minor changes indicated below (related to the roots of the quadratic polynomial in the proof of Theorem 2.9), ten Berge's proof also applies to the case $\mathbb{F} = \mathbb{C}$. 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

$$X = \sum_{i=1}^{n} \mathbf{a}^{(i)} \otimes \mathbf{b}^{(i)} \otimes \mathbf{c}^{(i)}, \quad \text{where} \quad \mathbf{a}^{(i)}, \mathbf{b}^{(i)}, \mathbf{c}^{(i)} \in \mathbb{F}^2 \text{ for } 1 \leq i \leq n.$$

We express this decomposition in terms of three $2 \times n$ matrices $A, B, C$:

$$A = \begin{bmatrix} \mathbf{a}^{(1)} & \cdots & \mathbf{a}^{(n)} \end{bmatrix}, \quad B = \begin{bmatrix} \mathbf{b}^{(1)} & \cdots & \mathbf{b}^{(n)} \end{bmatrix}, \quad C = \begin{bmatrix} \mathbf{c}^{(1)} & \cdots & \mathbf{c}^{(n)} \end{bmatrix}.$$

**Lemma 2.1.** [8, p. 632] *The rank of a nonzero $2\times2\times2$ 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.*

*Proof.* The first frontal slice $X_1$ has the form

$$\begin{aligned} X_1 &= \sum_{i=1}^{n} \begin{bmatrix} a_1^{(i)} b_1^{(i)} c_1^{(i)} & a_1^{(i)} b_2^{(i)} c_1^{(i)} \\ a_2^{(i)} b_1^{(i)} c_1^{(i)} & a_2^{(i)} b_2^{(i)} c_1^{(i)} \end{bmatrix} = \sum_{i=1}^{n} \begin{bmatrix} A_{1i}\, c_1^{(i)} B_{i1}^t & A_{1i}\, c_1^{(i)} B_{i2}^t \\ A_{2i}\, c_1^{(i)} B_{i1}^t & A_{2i}\, c_1^{(i)} B_{i2}^t \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^n A_{1i}\, c_1^{(i)} B_{i1}^t & \sum_{i=1}^n A_{1i}\, c_1^{(i)} B_{i2}^t \\ \sum_{i=1}^n A_{2i}\, c_1^{(i)} B_{i1}^t & \sum_{i=1}^n A_{2i}\, c_1^{(i)} B_{i2}^t \end{bmatrix} = A\, C_1 B^t, \end{aligned}$$

where $C_1$ is the $n \times n$ diagonal matrix whose diagonal entries $c_1^{(1)}$, $c_1^{(2)}$, $\dots$, $c_1^{(n)}$ 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. $\square$

**Definition 2.2.** We call the $2\times2\times2$ array $X$ **superdiagonal** if it has one of these forms for $\alpha, \beta \in \mathbb{F} \setminus \{0\}$:

$$\left[\begin{array}{cc|cc} \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & \alpha & 0 & 0 \\ 0 & 0 & \beta & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & 0 & 0 & \beta \\ \alpha & 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cc|cc} 0 & 0 & \beta & 0 \\ 0 & \alpha & 0 & 0 \end{array}\right].$$

**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

$$\left[\begin{array}{cc|cc} \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \end{array}\right] = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \beta \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

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

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

We obtain

$$X_1 = ADB^t = \begin{bmatrix} a_1 d b_1 & a_1 d b_2 \\ a_2 d b_1 & a_2 d b_2 \end{bmatrix}, \qquad X_2 = AEB^t = \begin{bmatrix} a_1 e b_1 & a_1 e b_2 \\ a_2 e b_1 & a_2 e b_2 \end{bmatrix},$$

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 $\geq 2$. $\square$

**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 $X_1 = d(\mathbf{a} \otimes \mathbf{b})$ where $\mathbf{a}$ and $\mathbf{b}$ are nonzero vectors in $\mathbb{F}^2$; 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 $X_2 = \mathbf{a} \otimes \mathbf{b}$ for some nonzero vectors $\mathbf{a}, \mathbf{b} \in \mathbb{F}^2$. But then $X = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{e}_1$ where $\mathbf{e}_1 = [0, 1]^t$.

Case 2: No slice is zero. Since $X_1$ is nonzero and singular, we have $X_1 = \mathbf{a} \otimes \mathbf{b}$ where $\mathbf{a} = [a_1, a_2]^t$ and $\mathbf{b} = [b_1, b_2]^t$ 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, $\mathbf{b} = [1, \lambda]^t$ for some $\lambda \in \mathbb{F}$; thus $X_1 = [\mathbf{a}|\lambda\mathbf{a}]$.

Subcase 2(a): $\lambda = 0$. Since the first vertical slice is singular, $X = [\mathbf{a}, \mathbf{0}|\mu\mathbf{a}, \mathbf{d}]$ for some $\mu \in \mathbb{F}$ and some $\mathbf{d}$; we have $\mathbf{d} \neq \mathbf{0}$ since the second vertical slice is nonzero.

If $\mu \neq 0$ then since $X_2$ is singular, there is $\nu \in \mathbb{F}\backslash\{0\}$ such that $X = [\mathbf{a}, \mathbf{0}|\mu\mathbf{a}, \nu\mathbf{a}]$. 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 $\nu = 0$; but then $X = [\mathbf{a}, \mathbf{0}|\mu\mathbf{a}, \mathbf{0}]$, so the second vertical slice is zero, giving a contradiction.

If $\mu = 0$ then $X = [\mathbf{a}, \mathbf{0}|\mathbf{0}, \mathbf{d}]$. 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 $X = [\mathbf{a}, \lambda\mathbf{a}|\mu\mathbf{a}, \mathbf{d}]$. But $\mathbf{a} \neq \mathbf{0}$ and the second vertical slice is singular, so $\mathbf{d} = \nu\mathbf{a}$ for some $\nu \in \mathbb{F}$, giving $X = [\mathbf{a}, \lambda\mathbf{a}|\mu\mathbf{a}, \nu\mathbf{a}]$. Since either $a_1 \neq 0$ or $a_2 \neq 0$ (or both), singularity of the horizontal slices implies that $\nu = \lambda\mu$. Then $X = [\mathbf{a}, \lambda\mathbf{a}|\mu\mathbf{a}, \lambda\mu\mathbf{a}] = [a_1, a_2]^t \otimes [1, \lambda]^t \otimes [1, \mu]^t$ 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.** [8, 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

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

Consider the following matrices:

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

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

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 $x_{ijk}$:

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

**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 $\Delta(X) > 0$ (for $\mathbb{F} = \mathbb{R}$) or $\Delta(X) \neq 0$ (for $\mathbb{F} = \mathbb{C}$) then $X$ has rank 2;*
(b) *if $\Delta(X) = 0$ then $X$ has rank 3.*

*Proof.* First, assume that $X_2 = \lambda X_1$ for some $\lambda \in \mathbb{F}$. Since $X_1$ is non-singular, it has rank 2, and hence $X_1 = \mathbf{a}^{(1)} \otimes \mathbf{b}^{(1)} + \mathbf{a}^{(2)} \otimes \mathbf{b}^{(2)}$. Writing $\mathbf{c} = [\,1, \lambda\,]^t$ then we see that $X$ has rank 2: $X = X_1 \otimes \mathbf{c} = \mathbf{a}^{(1)} \otimes \mathbf{b}^{(1)} \otimes \mathbf{c} + \mathbf{a}^{(2)} \otimes \mathbf{b}^{(2)} \otimes \mathbf{c}$.

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

$$A = \left[\ \mathbf{a}^{(1)} \quad \mathbf{a}^{(2)}\ \right], \quad B = \left[\ \mathbf{b}^{(1)} \quad \mathbf{b}^{(2)}\ \right], \quad D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}, \quad E = \begin{bmatrix} e_1 & 0 \\ 0 & e_2 \end{bmatrix}.$$

But $X_1 = ADB^t$, $X_2 = AEB^t$ gives $X_1 = d_1\mathbf{a}_1\mathbf{b}_1^t + d_2\mathbf{a}_2\mathbf{b}_2^t$, $X_2 = e_1\mathbf{a}_1\mathbf{b}_1^t + e_2\mathbf{a}_2\mathbf{b}_2^t$. 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 $d_1e_2 - d_2e_1 \neq 0$, and so $X_2 - d_1^{-1}e_1X_1$, $X_2 - d_2^{-1}e_2X_1$ are distinct. We calculate

$$\begin{aligned} & X_2 - d_1^{-1}e_1X_1 = e_1\mathbf{a}_1\mathbf{b}_1^t + e_2\mathbf{a}_2\mathbf{b}_2^t - d_1^{-1}e_1\big(d_1\mathbf{a}_1\mathbf{b}_1^t + d_2\mathbf{a}_2\mathbf{b}_2^t\big) \\ & = e_1\mathbf{a}_1\mathbf{b}_1^t + e_2\mathbf{a}_2\mathbf{b}_2^t - e_1\mathbf{a}_1\mathbf{b}_1^t - d_1^{-1}d_2e_1\mathbf{a}_2\mathbf{b}_2^t = d_1^{-1}(d_1e_2 - d_2e_1)\mathbf{a}_2\mathbf{b}_2^t, \\ & X_2 - d_2^{-1}e_2X_1 = e_1\mathbf{a}_1\mathbf{b}_1^t + e_2\mathbf{a}_2\mathbf{b}_2^t - d_2^{-1}e_2\big(d_1\mathbf{a}_1\mathbf{b}_1^t + d_2\mathbf{a}_2\mathbf{b}_2^t\big) \\ & = e_1\mathbf{a}_1\mathbf{b}_1^t + e_2\mathbf{a}_2\mathbf{b}_2^t - d_1d_2^{-1}e_2\mathbf{a}_1\mathbf{b}_1^t - e_2\mathbf{a}_2\mathbf{b}_2^t = -d_2^{-1}(d_1e_2 - d_2e_1)\mathbf{a}_1\mathbf{b}_1^t. \end{aligned}$$

It follows that these two matrices are singular. Hence the quadratic polynomial $\det(X_2 - \lambda X_1)$ has two distinct roots in $\mathbb{F}$, namely $\lambda = d_1^{-1}e_1$ and $\lambda = d_2^{-1}e_2$. But this determinant is

$$\begin{aligned} & (x_{111}x_{221} - x_{121}x_{211})\lambda^2 - (x_{111}x_{222} + x_{112}x_{221} - x_{121}x_{212} - x_{122}x_{211})\lambda \\ & \quad + (x_{112}x_{222} - x_{122}x_{212}), \end{aligned}$$

and the discriminant of this quadratic polynomial is $\Delta(X)$. Thus if $X$ has rank 2 then either $\Delta(X) > 0$ when $\mathbb{F} = \mathbb{R}$, or $\Delta(X) \neq 0$ when $\mathbb{F} = \mathbb{C}$. (This is the only place where we need to distinguish $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$.)

Conversely, suppose that $\Delta(X) \neq 0$. Then $\det(X_2 - \lambda X_1)$ has two distinct roots, say $\lambda_1$, $\lambda_2$ in $\mathbb{F}$. We have two nonzero singular matrices $X_2 - \lambda_1X_1$, $X_2 - \lambda_2X_1$.

These matrices both have rank 1, and so we can write

$$(\lambda_1-\lambda_2)^{-1}(X_2-\lambda_2X_1)=\mathbf{u}_1\mathbf{v}_1^t, \qquad -(\lambda_1-\lambda_2)^{-1}(X_2-\lambda_1X_1)=\mathbf{u}_2\mathbf{v}_2^t.$$

Then we have $X_1=\mathbf{u}_1\mathbf{v}_1^t+\mathbf{u}_2\mathbf{v}_2^t$ and $X_2=\lambda_1\mathbf{u}_1\mathbf{v}_1^t+\lambda_2\mathbf{u}_2\mathbf{v}_2^t$, which imply that $X=\mathbf{u}_1\otimes\mathbf{v}_1\otimes[1,\lambda_1]^t+\mathbf{u}_2\otimes\mathbf{v}_2\otimes[1,\lambda_2]^t$. Thus if $\Delta(X)>0$ (for $\mathbb{F}=\mathbb{R}$) or $\Delta(X)\neq 0$ (for $\mathbb{F}=\mathbb{C}$) then $X$ has rank 2. □

**Example 2.10.** Consider these arrays, where $X$ is the limit as $a\to 0$ of $Y(a)$:

$$X=\left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right], \qquad Y(a)=\left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & a^2 & 0 \end{array}\right].$$

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 $[\,I\,|\,X_2]$ 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 $[\,I\,|\,X_2]$ 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 $[\,I\,|\,X_2]$ 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 $\mathbb{F}=\mathbb{R}$ or $\mathbb{F}=\mathbb{C}$ have frontal slices $A$ and $B$:

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

Ja'Ja' [4, Corollary 3.4.1] has shown that the rank of a $p\times p\times 2$ array is at most $\lfloor 3p/2\rfloor$. 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

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

and hence

$$[A|B]=\mathbf{a}^{(1)}\otimes\mathbf{b}^{(1)}\otimes\mathbf{e}_1+\mathbf{a}^{(2)}\otimes\mathbf{b}^{(2)}\otimes\mathbf{e}_1+\mathbf{a}^{(3)}\otimes\mathbf{b}^{(3)}\otimes\mathbf{e}_2+\mathbf{a}^{(4)}\otimes\mathbf{b}^{(4)}\otimes\mathbf{e}_2.$$

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 $\geq 2$.

*We first consider the case* $\mathbb{F}=\mathbb{R}$. There exists an invertible matrix $E$ such that $R=E^{-1}CE$ is the rational canonical form of $C$. 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 $[I|R]$, where the second frontal slice $R$ still has rank $\geq 2$. It remains to show that any such array has rank $\leq 4$. The rational canonical form of $C$ is a block diagonal matrix with blocks $C_1, \ldots, C_k$; each $C_i$ is the companion matrix of a polynomial $f(t)^m$ where $f(t)$ is a monic irreducible divisor of the characteristic polynomial of $C$ and $m$ is a positive integer. Over $\mathbb{R}$, the possible characteristic polynomials of a $3 \times 3$ matrix and the corresponding rational canonical forms are as follows:

$$(1) \qquad (x^2 - ax - b)(x - c), \quad a^2 + 4b < 0, \qquad \begin{bmatrix} 0 & b & 0 \\ 1 & a & 0 \\ 0 & 0 & c \end{bmatrix},$$

$$(2) \qquad (x - a)(x - b)(x - c), \quad a, b, c \text{ distinct}, \qquad \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix},$$

$$(3a) \qquad (x - a)^2(x - b), \quad a \neq b, \qquad \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix},$$

$$(3b) \qquad (x - a)^2(x - b), \quad a \neq b, \qquad \begin{bmatrix} 0 & -a^2 & 0 \\ 1 & 2a & 0 \\ 0 & 0 & b \end{bmatrix},$$

$$(4a) \qquad (x - a)^3 \qquad \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix},$$

$$(4b) \qquad (x - a)^3 \qquad \begin{bmatrix} 0 & -a^2 & 0 \\ 1 & 2a & 0 \\ 0 & 0 & a \end{bmatrix},$$

$$(4c) \qquad (x - a)^3 \qquad \begin{bmatrix} 0 & 0 & a^3 \\ 1 & 0 & -3a^2 \\ 0 & 1 & 3a \end{bmatrix},$$

Clearly (3a) and (4a) are special cases of (2); and (3b) and (4b) are special cases of (1). It remains to consider (1), (2) and (4c).

In case (1), we have

$$\begin{aligned} [I|R] &= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & b & 0 \\ 0 & 1 & 0 & 1 & a & 0 \\ 0 & 0 & 1 & 0 & 0 & c \end{array}\right] \\ &= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & b & 0 \\ 0 & 1 & 0 & 1 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & c \end{array}\right]. \end{aligned}$$

The second array has rank 1, and decomposing the first array is equivalent to decomposing a $2 \times 2 \times 2$ array, which has rank $\leq 3$ by Theorem 2.9.

In case (2), $R$ is a diagonal matrix, and hence the rank is $\leq 3$:

$$[I|R] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & a & 0 & 0 \\ 0 & 1 & 0 & 0 & b & 0 \\ 0 & 0 & 1 & 0 & 0 & c \end{array}\right]$$

$$= \mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \begin{bmatrix} 1 \\ a \end{bmatrix} + \mathbf{e}_2 \otimes \mathbf{e}_2 \otimes \begin{bmatrix} 1 \\ b \end{bmatrix} + \mathbf{e}_3 \otimes \mathbf{e}_3 \otimes \begin{bmatrix} 1 \\ c \end{bmatrix}.$$

In case (4c), we adapt the argument of Ja'ja' [4, Theorem 3.2, pages 453-454]. We recall the matrix $C$, and define the matrix $D$:

$$C = \begin{bmatrix} 0 & 0 & a^3 \\ 1 & 0 & -3a^2 \\ 0 & 1 & 3a \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & a^3 \\ 0 & 0 & -3a^2-1 \\ 0 & 0 & 3a \end{bmatrix}, \quad C - D = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

The matrix $C - D$ is the companion matrix of the polynomial $t^3 - t$, which has 3 distinct real roots 1, 0 and $-1$. Hence there exists an invertible matrix $P$ such that

$$P^{-1}(C-D)P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} = E_{11} - E_{33}, \qquad P = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

From this we obtain

$$C = D + (C-D) = D + P(E_{11} - E_{33})P^{-1} = D + PE_{11}P^{-1} - PE_{33}P^{-1},$$

and this gives

$$[\, I \,|\, C \,] = [\, O \,|\, D \,] + P[\, E_{11} \,|\, E_{11} \,]P^{-1} + P[\, E_{22} \,|\, O \,]P^{-1} + P[\, E_{33} \,|\, -E_{33} \,]P^{-1}.$$

Clearly each of these 4 terms is an array of rank $\le 1$, showing that the rank of $[I|C]$ is at most 4. We obtain the following explicit decomposition in case (4c) into a sum of simple tensors:

$$[\, I \,|\, C \,] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & a^3 \\ 0 & 1 & 0 & 1 & 0 & -3a^2 \\ 0 & 0 & 1 & 0 & 1 & 3a \end{array}\right]$$

$$= \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & a^3 \\ 0 & 0 & 0 & 0 & 0 & -3a^2-1 \\ 0 & 0 & 0 & 0 & 0 & 3a \end{array}\right] + \frac{1}{2} \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right]$$

$$+ \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \end{array}\right] + \frac{1}{2} \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{array}\right].$$

This completes the proof in the real case.

*We next consider the case* $\mathbb{F} = \mathbb{C}$. There exists an invertible matrix $E$ such that $J = E^{-1}CE$ is the Jordan canonical form of $C$. 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 $\ge 2$. It remains to show that any such array has rank $\le 4$. 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:

$$[I|J] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & d_1 & 0 & 0 \\ 0 & 1 & 0 & 0 & d_2 & 0 \\ 0 & 0 & 1 & 0 & 0 & d_3 \end{array}\right]$$

We apply the same argument as in case (2) when $\mathbb{F} = \mathbb{R}$.

Case 2: One $2 \times 2$ block and one $1 \times 1$ block:

$$[I|J] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & d_1 & 1 & 0 \\ 0 & 1 & 0 & 0 & d_1 & 0 \\ 0 & 0 & 1 & 0 & 0 & d_2 \end{array}\right]$$

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

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

$$[I|J] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & d_1 & 1 & 0 \\ 0 & 1 & 0 & 0 & d_1 & 1 \\ 0 & 0 & 1 & 0 & 0 & d_1 \end{array}\right]$$

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:

$$[\, I \mid E_{12}+E_{23} \,] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

It remains to prove that this array has rank $\leq 4$. We have the following explicit representation of $[\, I \mid E_{12}+E_{23} \,]$ as a sum of four simple tensors:

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -1 & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$
$$+ \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$$
$$= \begin{bmatrix} 1 \\ \frac{1}{2} \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ -\frac{1}{2} \\ 1 \end{bmatrix} \otimes \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$
$$+ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

This completes the proof in the complex case. □

**Remark 3.2.** The end of the proof of Lemma 3.1 in the complex case is the only place in the proof of Kruskal's theorem for $\mathbb{F} = \mathbb{C}$ where we need to assume that the characteristic of $\mathbb{F}$ is not 2. Therefore our proof is also valid over any algebraically closed field of characteristic 0 or $p > 2$.

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

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

**Lemma 3.3.** ***Kruskal's One-Edge Lemma.*** *If the array $T$ has parallel slices $D$ and $E$ for which there exists a nonzero vector $\mathbf{x}$ such that $D\mathbf{x} = E\mathbf{x} = 0$ or $D^t\mathbf{x} = E^t\mathbf{x} = 0$, then* $\mathrm{rank}(T) \leq 5$.

*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 $A\mathbf{x} = B\mathbf{x} = 0$ where $\mathbf{x} \neq 0$. Let $X$ be a $3\times3$ non-singular matrix which has $\mathbf{x}$ as its first column. (Extend the set $\{\mathbf{x}\}$ to a basis $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ of $\mathbb{F}^3$ and let $X = [\mathbf{x}|\mathbf{y}|\mathbf{z}]$.) Acting on $T = [A|B|C]$ by $X$ along the second direction gives $[AX|BX|CX]$, but $A\mathbf{x} = B\mathbf{x} = 0$, so

$$
[AX|BX|CX] = \left[\begin{array}{ccc|ccc|ccc} 0 & a'_{12} & a'_{13} & 0 & b'_{12} & b'_{13} & c'_{11} & c'_{12} & c'_{13} \\ 0 & a'_{22} & a'_{23} & 0 & b'_{22} & b'_{23} & c'_{21} & c'_{22} & c'_{23} \\ 0 & a'_{32} & a'_{33} & 0 & b'_{32} & b'_{33} & c'_{31} & c'_{32} & c'_{33} \end{array}\right]
$$

$$
= \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & c'_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & c'_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & c'_{31} & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|ccc|ccc} 0 & a'_{12} & a'_{13} & 0 & b'_{12} & b'_{13} & 0 & c'_{12} & c'_{13} \\ 0 & a'_{22} & a'_{23} & 0 & b'_{22} & b'_{23} & 0 & c'_{22} & c'_{23} \\ 0 & a'_{32} & a'_{33} & 0 & b'_{32} & b'_{33} & 0 & c'_{32} & c'_{33} \end{array}\right]
$$

The first term is a simple tensor,

$$
\begin{bmatrix} c'_{11} \\ c'_{21} \\ c'_{31} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},
$$

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:

$$
\left[\begin{array}{cc|cc|cc} a'_{12} & a'_{13} & b'_{12} & b'_{13} & c'_{12} & c'_{13} \\ a'_{22} & a'_{23} & b'_{22} & b'_{23} & c'_{22} & c'_{23} \\ a'_{32} & a'_{33} & b'_{32} & b'_{33} & c'_{32} & c'_{33} \end{array}\right]
$$

Transposing the second and third directions, we may consider this $3 \times 3 \times 2$ array:

$$
\left[\begin{array}{ccc|ccc} a'_{12} & b'_{12} & c'_{12} & a'_{13} & b'_{13} & c'_{13} \\ a'_{22} & b'_{22} & c'_{22} & a'_{23} & b'_{23} & c'_{23} \\ a'_{32} & b'_{32} & c'_{32} & a'_{33} & b'_{33} & c'_{33} \end{array}\right]
$$

The claim now follows from Lemma 3.1.

If $A^t\mathbf{x} = B^t\mathbf{x} = 0$, 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 3.4.** ***Kruskal's Two-Edge Lemma.*** *If the array $T$ has frontal slices $D$ and $E$ for which there exist nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ such that $D\mathbf{x} = \mathbf{y}^tD = 0$ and $\mathbf{y}^tE\mathbf{x} \neq 0$, then* $\mathrm{rank}(T) \leq 5$.

*Proof.* As before, we may assume that $D$ and $E$ are the first and second frontal slices $A$ and $B$. Let $\mathbf{x}$ and $\mathbf{y}$ satisfy the conditions of the lemma. We choose vectors $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{v}_2$, $\mathbf{v}_3$ such that $U = [\,\mathbf{x}|\mathbf{u}_2|\mathbf{u}_3\,]$ and $V = [\,\mathbf{y}|\mathbf{v}_2|\mathbf{v}_3\,]$ 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 $\alpha = \mathbf{y}^t B \mathbf{x} \neq 0$, 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

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

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,

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

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 $[0|0|Y]$ also have rank 1 as $3 \times 3 \times 3$ arrays; that is, they are simple tensors. We now see that

$$V^t[A|B|C]U - [0|X|0] - [0|0|Y] = \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & 0 & * & * & 0 & * & * \\ 0 & * & * & 0 & * & * & 0 & * & * \end{array}\right]$$

It remains to decompose the $2 \times 2 \times 2$ array corresponding to the symbols $*$, and this requires at most three simple tensors by Lemma 2.6. $\square$

## 4. Proof of Kruskal's theorem on $3 \times 3 \times 3$ arrays

As before, $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

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

*Proof.* If any of the frontal slices is zero, 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 $A$, $B$ and $C$ are all nonzero.

If exactly two or three of the frontal slices are singular, then by permuting the frontal slices we may assume that $A$ and $B$ are singular.

The alternative is that exactly zero or one of the frontal slices are singular, which means that at least two of the frontal slices are non-singular. In this case, by permuting the frontal slices we may assume that $C$ is non-singular. Consider the $3 \times 3$ matrix $A - \lambda C$; its determinant is a cubic polynomial in $\lambda$, since the coefficient of $\lambda^3$ is $\det(C) \neq 0$. If $\mathbb{F} = \mathbb{R}$ then this polynomial has a root in $\mathbb{R}$ because its degree is odd. If $\mathbb{F} = \mathbb{C}$ then this polynomial has a root in $\mathbb{C}$ by algebraic closure. Thus by subtracting a multiple of $C$ from $A$, we may ensure that $A$ is singular, and so $\operatorname{rank}(A) \leq 2$. Saying the same thing a different way, we are changing 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 $B$.

*Assumption 1:* We may now assume that the first and second frontal slices of $T$ are both singular, and hence both have rank $\leq 2$. (See Remark 4.2 below.)

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

*Assumption 2:* We may now assume that every frontal slice has rank $\geq 2$.

Combining Assumptions 1 and 2, we may assume from now on that $A$ and $B$ have rank 2 and $C$ has rank 2 or 3. This gives two main cases for the rest of the proof, depending on the rank of $C$.

Case 1: $A$, $B$ and $C$ all have rank 2. It follows that there exist nonzero vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ (basis vectors for the right and left nullspaces) such that

$$A\mathbf{x}_1 = B\mathbf{x}_2 = C\mathbf{x}_3 = 0, \qquad \mathbf{y}_1^t A = \mathbf{y}_2^t B = \mathbf{y}_3^t C = 0.$$

Then for the $3 \times 3$ matrices $X = [\mathbf{x}_1|\mathbf{x}_2|\mathbf{x}_3]$ and $Y = [\mathbf{y}_1|\mathbf{y}_2|\mathbf{y}_3]$ we have

$$(1) \qquad Y^tAX = \begin{bmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}, \quad Y^tBX = \begin{bmatrix} * & 0 & * \\ 0 & 0 & 0 \\ * & 0 & * \end{bmatrix}, \quad Y^tCX = \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

If the conditions of Lemma 3.4 are satisfied for any two frontal slices, then the proof is complete. Otherwise, it follows that

$$(2) \qquad \mathbf{y}_i^t A\mathbf{x}_i = \mathbf{y}_i^t B\mathbf{x}_i = \mathbf{y}_i^t C\mathbf{x}_i = 0, \text{ for all } i = 1, 2, 3.$$

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 3.3 completes the proof.

*Subcase 1.2:* The matrix $X$ has rank 2, but no two columns are linearly dependent. Then $\mathbf{x}_1$ and $\mathbf{x}_2$ are linearly independent, and so $\mathbf{x}_3 = \beta\mathbf{x}_1 + \gamma\mathbf{x}_2$ for some $\beta, \gamma \in \mathbb{F} \setminus \{0\}$. We choose vectors $\mathbf{u}, \mathbf{v}_2, \mathbf{v}_3$ such that the matrices $U = [\mathbf{x}_1|\mathbf{x}_2|\mathbf{u}]$ and $V = [\mathbf{y}_1|\mathbf{v}_2|\mathbf{v}_3]$ are invertible. Then for some $\delta \in \mathbb{C}$ we have

$$V^tAU = \begin{bmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}, \qquad V^tBU = \begin{bmatrix} 0 & 0 & * \\ * & 0 & * \\ * & 0 & * \end{bmatrix}, \qquad V^tCU = \begin{bmatrix} 0 & \delta & * \\ * & * & * \\ * & * & * \end{bmatrix}.$$

(The $(1,1)$ entries of $V^tBU$ and $V^tCU$ are zero; otherwise Lemma 3.4 would apply.) But $C\mathbf{x}_3 = 0$ implies $\beta C\mathbf{x}_1 + \gamma C\mathbf{x}_2 = 0$, and so the first two columns of $V^tCU$ are linearly dependent, implying $\delta = 0$. Hence the first three rows of $V^tAU$, $V^tBU$ and $V^tCU$ 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:

$$Y^tAX = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \beta \\ 0 & \alpha & 0 \end{bmatrix}, \quad Y^tBX = \begin{bmatrix} 0 & 0 & \delta \\ 0 & 0 & 0 \\ \gamma & 0 & 0 \end{bmatrix}, \quad Y^tCX = \begin{bmatrix} 0 & \zeta & 0 \\ \epsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

If we interchange 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:

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha & -\beta & -\beta & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] +$$

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & \delta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\gamma & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\gamma & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & \zeta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

We obtain the following array:

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 & \gamma & 0 & \epsilon & 0 \\ 0 & \beta & \beta & 0 & \gamma & \gamma & 0 & 0 & 0 \end{array}\right] \tag{3}$$

The three frontal slices are linear combinations of these two matrices of rank 1:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore the array (3) is the sum of two simple tensors:

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta & \beta & 0 & \gamma & \gamma & 0 & 0 & 0 \\ 0 & \beta & \beta & 0 & \gamma & \gamma & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\beta & 0 & 0 & -\gamma & 0 & 0 & \epsilon & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

From this we obtain 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 (so $C$ is invertible). If there exist $\alpha, \beta \in \mathbb{F}$ such that $\alpha A + \beta B + C$ has rank $\leq 1$, then we are back in the cases considered before Assumption 2.

*Subcase 2.1:* There exist $\alpha, \beta \in \mathbb{F}$ 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

$$A = B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

*Subcase 2.2:* The matrix $\alpha A + \beta B + C$ has rank 3 for all $\alpha, \beta \in \mathbb{F}$. There exist nonzero vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ such that

$$A\mathbf{x}_1 = B\mathbf{x}_2 = 0, \qquad C\mathbf{x}_3 = A\mathbf{x}_2, \qquad \mathbf{y}_1^t A = \mathbf{y}_2^t B = 0, \qquad \mathbf{y}_3^t C = \mathbf{y}_2^t A.$$

If we can apply Lemma 3.4, then we are done. So we may assume that Lemma 3.4 does not apply, and hence we must have

$$\mathbf{y}_1^t A\mathbf{x}_1 = \mathbf{y}_1^t B\mathbf{x}_1 = \mathbf{y}_1^t C\mathbf{x}_1 = 0, \qquad \mathbf{y}_2^t A\mathbf{x}_2 = \mathbf{y}_2^t B\mathbf{x}_2 = \mathbf{y}_2^t C\mathbf{x}_2 = 0.$$

We write $X = [\mathbf{x}_1|\mathbf{x}_2|\mathbf{x}_3]$ and $Y = [\mathbf{y}_1|\mathbf{y}_2|\mathbf{y}_3]$.

*Subsubcase 2.2.1:* At least one of the pairs $\{\mathbf{x}_1, \mathbf{x}_2\}$, $\{\mathbf{y}_1, \mathbf{y}_2\}$ is linearly dependent. Then the result follows from Lemma 3.3.

*Subsubcase 2.2.2:* Both pairs $\{\mathbf{x}_1, \mathbf{x}_2\}$, $\{\mathbf{y}_1, \mathbf{y}_2\}$ are linearly independent. Then both matrices $X$ and $Y$ have rank $\geq 2$. The rest of the proof deals with this subsubcase.

Assume $X$ has rank 2. Then $\mathbf{x}_3 = \gamma\mathbf{x}_1 + \delta\mathbf{x}_2$ for some $\gamma, \delta \in \mathbb{F}$. There exist nonzero vectors $\mathbf{u}$, $\mathbf{v}$ such that $U = [\mathbf{x}_1|\mathbf{x}_2|\mathbf{u}]$ and $V = [\mathbf{y}_1|\mathbf{y}_2|\mathbf{v}]$ both have rank 3. Then using the previous equations we have

$$V^tAU = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & * \\ 0 & * & * \end{bmatrix}, \qquad V^tBU = \begin{bmatrix} 0 & 0 & \zeta \\ 0 & 0 & 0 \\ \epsilon & 0 & * \end{bmatrix}, \qquad V^tCU = \begin{bmatrix} 0 & \theta & * \\ \eta & 0 & * \\ * & * & * \end{bmatrix}.$$

Since $B$ has rank 2, it follows that $\epsilon \neq 0$ and $\zeta \neq 0$. We have

$$V^tA\mathbf{x}_2 = V^tC\mathbf{x}_3 = V^tC(\gamma\mathbf{x}_1 + \delta\mathbf{x}_2) = \gamma V^tC\mathbf{x}_1 + \delta V^tC\mathbf{x}_2,$$

and so the second column of $V^tAU$ is a linear combination of the first two columns of $V^tCU$. Since $\mathbf{x}_3 \neq 0$ (assumed at the start of subcase 2.2), and $\mathbf{x}_3 = \gamma\mathbf{x}_1 + \delta\mathbf{x}_2$ (assumed at the start of this paragraph), it follows that $\gamma, \delta$ are not both 0, and so at least one of $\eta$, $\theta$ is zero. 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 $U$ along the first and second directions, gives an array in which the third frontal slice has rank 2. But this new third frontal slice has the form $\beta B + C$ for some $\beta \in \mathbb{F}$, and this contradicts the assumption (at the start of subcase 2.2) that $\alpha A + \beta B + C$ has rank 3 for all $\alpha, \beta \in \mathbb{F}$.

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.

Assuming that $X$ and $Y$ both have rank 3 (and hence are invertible), and using the previous equations, together with

$$\mathbf{y}_2^t A\mathbf{x}_3 = \mathbf{y}_3^t C\mathbf{x}_3 = \mathbf{y}_3^t A\mathbf{x}_2,$$

we obtain

$$Y^tAX = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \gamma \\ 0 & \gamma & \delta \end{bmatrix}, \qquad Y^tBX = \begin{bmatrix} 0 & 0 & \zeta \\ 0 & 0 & 0 \\ \epsilon & 0 & \eta \end{bmatrix}.$$

Using the previous equations, together with

$$\begin{aligned} &\mathbf{y}_1^t C\mathbf{x}_1 = 0, && \mathbf{y}_1^t C\mathbf{x}_3 = \mathbf{y}_1^t A\mathbf{x}_2 = 0, \\ &\mathbf{y}_2^t C\mathbf{x}_2 = 0, && \mathbf{y}_2^t C\mathbf{x}_3 = \mathbf{y}_2^t A\mathbf{x}_2 = 0, \\ &\mathbf{y}_3^t C\mathbf{x}_1 = \mathbf{y}_2^t A\mathbf{x}_1 = 0, && \mathbf{y}_3^t C\mathbf{x}_2 = \mathbf{y}_2^t A\mathbf{x}_2 = 0, \end{aligned}$$

we obtain

$$Y^t C X = \begin{bmatrix} 0 & \lambda & 0 \\ \kappa & 0 & 0 \\ 0 & 0 & \gamma \end{bmatrix}.$$

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 $Y^t C X$ has rank 2; but this contradicts the assumption that $\alpha A + \beta B + C$ has rank 3. So we may assume that $\delta = \eta = 0$:

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

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:

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & \zeta & \lambda & 0 & 0 \\ 0 & 0 & \gamma & 0 & 0 & 0 & 0 & \kappa & 0 \\ \gamma & 0 & 0 & 0 & \epsilon & 0 & 0 & 0 & \gamma \end{array}\right].$$

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:

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & \zeta/\lambda & 1 & 0 & 0 \\ 0 & 0 & \gamma/\kappa & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & \epsilon/\gamma & 0 & 0 & 0 & 1 \end{array}\right]. \tag{4}$$

We now need to consider the cases $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$ separately.

If $\mathbb{F} = \mathbb{R}$ then from the array (4) we subtract the following array of rank 2:

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma/\kappa & 0 & 0 & -\epsilon/\gamma & 0 & 0 & 0 \\ 1 & 0 & 0 & -\zeta/\lambda & 0 & 0 & 0 & 0 & 0 \end{array}\right] =$$

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} \gamma/\kappa \\ -\epsilon/\gamma \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -\zeta/\lambda \\ 0 \end{bmatrix}.$$

We obtain

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & \zeta/\lambda & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \epsilon/\gamma & 0 & 1 & 0 \\ 0 & 0 & 0 & \zeta/\lambda & \epsilon/\gamma & 0 & 0 & 0 & 1 \end{array}\right].$$

The second frontal slice (which we still denote by $B$) is now symmetric, hence diagonalizable over $\mathbb{R}$, and so there exists an invertible matrix $E$ such that $E^{-1} B E$ is a diagonal matrix. Changing basis along the first and second directions by $E^{-1}$ and $E$ respectively, we obtain

$$\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & \mu & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \nu & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \xi & 0 & 0 & 1 \end{array}\right].$$

This array clearly has rank 3, and the proof over $\mathbb{R}$ is complete.

If $\mathbb{F} = \mathbb{C}$ then from the array (4) we subtract the following array of rank 2, where the bars denote complex conjugates:

$$
\left[\begin{array}{ccc|ccc|ccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \gamma/\kappa & 0 & 0 & -\overline{\epsilon/\gamma} & 0 & 0 & 0 \\
1 & 0 & 0 & -\overline{\zeta/\lambda} & 0 & 0 & 0 & 0 & 0
\end{array}\right] =
$$

$$
\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} \gamma/\kappa \\ -\overline{\epsilon/\gamma} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -\overline{\zeta/\lambda} \\ 0 \end{bmatrix}.
$$

We obtain

$$
\left[\begin{array}{ccc|ccc|ccc}
0 & 0 & 0 & 0 & 0 & \zeta/\lambda & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \overline{\epsilon/\gamma} & 0 & 1 & 0 \\
0 & 0 & 0 & \overline{\zeta/\lambda} & \epsilon/\gamma & 0 & 0 & 0 & 1
\end{array}\right].
$$

The second frontal slice (which we still denote by $B$) is now Hermitian, hence diagonalizable over $\mathbb{C}$, and so its Jordan canonical form $J = E^{-1}BE$ is a diagonal matrix. The rest of the proof is the same as in the case $\mathbb{F} = \mathbb{R}$. $\square$

**Remark 4.2.** Rocci [7] states at the start of the proof of his main theorem that "by subtracting multiples of $C$ from $A$ and $B$, we can ensure that $A$ and $B$ are both singular." However, this is not true, even in the complex case, as the following example shows, for which $\det(A - \lambda C) = -1$ for all $\lambda$:

$$
A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
$$

## 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 $\mathrm{flat}(X) = [x_{111}, \dots, x_{ijk}, \dots, x_{333}]$, where the entries are in lex 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 $\mathrm{flat}(X)$. Conversely, the decoding of an integer in the range $0, \dots, 2^{27}-1$ is the corresponding tensor. The lex 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 $GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2)$ 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 $GL_3(\mathbb{F}_2)$ 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 $GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2) \times GL_3(\mathbb{F}_2)$ 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) $\mathcal{L} \leftarrow \mathcal{N} \setminus \mathcal{O}$

(3) return $\mathcal{O}$

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 $\mathcal{O}(X)$, 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`$[i] = 1$ represents the encoding of a tensor of rank 1. Let $E$ denote the minimal tensor of rank 1: its flattening is $[0, \dots, 0, 1]$. 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

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

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 $\mathcal{O}$ with index number $i$, we apply the elements of $S_3$ to obtain tensors $X_1 = X, \dots, X_6$. We then use the Maple arrays, which we have already computed, to find the index numbers $i_1 = i, \dots, i_6$ of the small orbits containing these tensors. We conclude that the union $\mathcal{O}_{i_1} \cup \cdots \cup \mathcal{O}_{i_6}$ 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 $\mathcal{O}_{i_1}, \dots, \mathcal{O}_{i_6}$. There are 115 (nonzero) small orbits and 55 (nonzero) large orbits:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| rank | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| # small | 1 | 1 | 4 | 18 | 44 | 45 | 3 |
| # large | 1 | 1 | 2 | 8 | 18 | 23 | 3 |
| # tensors | 1 | 343 | 43218 | 2372286 | 47506872 | 83670048 | 624960 |
| percent | 0.0000 | 0.0003 | 0.0322 | 1.7675 | 35.3954 | 62.3390 | 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. The authors thank J. M. F. ten Berge for sending us the unpublished notes by Rocci [7], and the two anonymous referees for numerous suggestions that substantially improved the paper.

## References

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

[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 Journal on Matrix Analysis and Applications* 30 (2008) 1084–1127.

[4] J. Ja'Ja': Optimal evaluation of pairs of bilinear forms. *SIAM Journal on Computing* 8 (1979) 443–462.

[5] T. G. Kolda, B. W. Bader: Tensor decompositions and applications. *SIAM Review* 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 and Its Applications* 144 (1991) 49–61.

| # | rank | size | canonical form |
|---|---|---|---|
| 1 | 1 | 343 | . . . . . . . . . . . . . . . . . . . . . . . . . . 1 |
| 2 | 2 | 6174 | . . . . . . . . . . . . . . . . . . . . . . . 1 . 1 . |
| 3 | 2 | 37044 | . . . . . . . . . . . . . . . . . 1 . . . . 1 . . . . |
| 4 | 3 | 3528 | . . . . . . . . . . . . . . . . . . . . 1 . 1 . 1 . . |
| 5 | 3 | 4116 | . . . . . . . . . . . . . . 1 . 1 . . . . . 1 . . 1 1 |
| 6 | 3 | 18522 | . . . . . . . . . . . . . . . . . 1 . . . . . 1 . 1 . |
| 7 | 3 | 148176 | . . . . . . . . . . . . . . 1 . 1 . . . . . 1 . 1 . . |
| 8 | 3 | 222264 | . . . . . . . . . . . . . . . . . 1 . . . . 1 . 1 . . |
| 9 | 3 | 592704 | . . . . . . . . . . . . . . . . . 1 . 1 . 1 . . . . . |
| 10 | 3 | 592704 | . . . . . . . . . . . . . . 1 . 1 . 1 . . . . . . 1 . |
| 11 | 3 | 790272 | . . . . . . . . 1 . . . . 1 . . . . 1 . . . . . . . . |
| 12 | 4 | 148176 | . . . . . . . . . . . . . . 1 . 1 . . . 1 . . . 1 . . |
| 13 | 4 | 197568 | . . . . . 1 . 1 . . . 1 . . . 1 . . . 1 . 1 1 . . 1 . |
| 14 | 4 | 222264 | . . . . . . . . . . . . . . . . . 1 . . 1 . 1 . 1 . . |
| 15 | 4 | 263424 | . . . . . . . . 1 . 1 . 1 . . . . . 1 . . 1 1 . . . . |
| 16 | 4 | 444528 | . . . . . . . . . . . . . . 1 . 1 . . . 1 . 1 . 1 . . |
| 17 | 4 | 592704 | . . . . . . . . . . . . . . 1 . 1 . 1 . . . 1 . . 1 1 |
| 18 | 4 | 1185408 | . . . . . . . . 1 . . . . . 1 . 1 . 1 . . . . . . . . |
| 19 | 4 | 1778112 | . . . . . . . . . . . . . . 1 . 1 . . . 1 1 . . . . . |
| 20 | 4 | 1778112 | . . . . . . . . 1 . . . . . . . 1 . . . 1 1 . . . . . |
| 21 | 4 | 1778112 | . . . . . . . . 1 . . . . 1 . . . . . 1 1 1 . . 1 . . |
| 22 | 4 | 2370816 | . . . . . . . . 1 . 1 . 1 . . . . . 1 . . 1 1 . . 1 . |
| 23 | 4 | 2370816 | . . . . . . . . 1 . 1 . 1 . . . . . 1 . . 1 1 1 . 1 . |
| 24 | 4 | 3556224 | . . . . . . . . 1 . . . . . 1 . 1 . 1 . . . . . . 1 . |
| 25 | 4 | 4741632 | . . . . . . . . 1 . . . . 1 . . . . 1 . . . . . . 1 . |
| 26 | 4 | 4741632 | . . . . . . . . 1 . . . . 1 . 1 . . 1 . . . . 1 . . . |
| 27 | 4 | 7112448 | . . . . . . . . 1 . . . . . 1 . 1 . 1 . . . 1 . . . . |
| 28 | 4 | 7112448 | . . . . . . . . 1 . . . . 1 . . . . 1 . . . . 1 . 1 . |
| 29 | 4 | 7112448 | . . . . . . . . 1 . . . . 1 . 1 . . . 1 1 1 . . . . . |
| 30 | 5 | 28224 | . . . . . 1 . 1 . . . 1 . . . 1 . . . 1 . 1 . . . . . |
| 31 | 5 | 148176 | . . . . . . . . 1 . . . . . . . 1 . . . 1 . 1 . 1 . . |
| 32 | 5 | 148176 | . . . . . . . . 1 . . . . . 1 . 1 . . . 1 . 1 . 1 . . |
| 33 | 5 | 169344 | . . . . . . . . . . . 1 . 1 . 1 . . . 1 . 1 . . . 1 1 |
| 34 | 5 | 592704 | . . . . . 1 . 1 . . . 1 . . . 1 . . . 1 . 1 1 . . 1 1 |
| 35 | 5 | 1185408 | . . . . . 1 . 1 . . . 1 . . . 1 . . . 1 . 1 . . . 1 . |
| 36 | 5 | 1580544 | . . . . . 1 . 1 . . . 1 . . . 1 . . 1 . . 1 1 . . . . |
| 37 | 5 | 1580544 | . . . . . 1 . 1 . . . 1 . . . 1 . . 1 . . 1 1 . . . 1 |
| 38 | 5 | 1778112 | . . . . . . . . 1 . . . . . 1 . 1 . . . 1 1 . . . . . |
| 39 | 5 | 1778112 | . . . . . . . . 1 . . . . 1 . 1 . . . . 1 1 . . 1 1 . |
| 40 | 5 | 2370816 | . . . . . 1 . 1 . . . 1 . . . 1 . . . 1 . 1 1 . 1 . . |
| 41 | 5 | 2370816 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . 1 . . 1 . . . . 1 |
| 42 | 5 | 2370816 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . 1 . . 1 . . 1 . 1 |
| 43 | 5 | 3556224 | . . . . . . . . 1 . . . . . 1 . 1 . . 1 . 1 . . . . . |
| 44 | 5 | 4741632 | . . . . . . . . 1 . 1 . 1 . . . . . 1 . . 1 1 1 1 . . |
| 45 | 5 | 4741632 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . . 1 . 1 . . 1 . . |
| 46 | 5 | 4741632 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . 1 . . . . . . 1 . |
| 47 | 5 | 4741632 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . 1 . . . . . . 1 1 |
| 48 | 5 | 4741632 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . 1 . . 1 . . . 1 . |
| 49 | 5 | 4741632 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . 1 . . 1 . . 1 1 . |
| 50 | 5 | 7112448 | . . . . . . . . 1 . . 1 . 1 . 1 . . 1 . . . . 1 . 1 . |
| 51 | 5 | 14224896 | . . . . . . . . 1 . . . . 1 . 1 . . 1 . . . . 1 . 1 . |
| 52 | 5 | 14224896 | . . . . . . . . 1 . . 1 . 1 . 1 . . . 1 . 1 . . . . . |
| 53 | 6 | 32256 | . . 1 . 1 . 1 . . . 1 . 1 . . . 1 1 1 . . . 1 1 1 1 . |
| 54 | 6 | 197568 | . . . . . 1 . 1 . . . 1 . . . 1 . . . 1 . 1 . . . . 1 |
| 55 | 6 | 395136 | . . . . . 1 . 1 . . . 1 . 1 . 1 . . . 1 . 1 . . . 1 1 |

TABLE 1. Large orbits of $3 \times 3 \times 3$ tensors over $\mathbb{F}_2$

Department of Mathematics and Statistics, University of Saskatchewan, Canada
*E-mail address*: bremner@math.usask.ca

Department of Mathematics, Simon Fraser University, Canada
*E-mail address*: hujiaxiong@gmail.com