



**HAL**  
open science

# Non-negative Tensor Approximations

Yang Qi, Pierre Comon

► **To cite this version:**

| Yang Qi, Pierre Comon. Non-negative Tensor Approximations. 2014. hal-01015519v2

**HAL Id: hal-01015519**

**<https://hal.science/hal-01015519v2>**

Preprint submitted on 17 Jul 2014 (v2), last revised 12 Apr 2016 (v5)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# NON-NEGATIVE TENSOR APPROXIMATIONS

YANG QI, PIERRE COMON

ABSTRACT. Necessary conditions are derived for a rank- $r$  tensor to be a best rank- $r$  approximation of a given tensor. It is shown that a positive tensor with rank  $> 1$  has a unique rank one approximation, and that a non negative tensor generally has a unique low-rank nonnegative approximate. We discuss the notion of  $r$ -singular values and their corresponding  $r$ -singular vector tuples, which is closely related to best rank- $r$  approximations. We then show that a generic tensor has a finite number of  $r$ -singular vector tuples for some  $r$ .

## 1. INTRODUCTION

Nonnegative tensors are widely used in several fields, including *e.g.*, hyperspectral imaging, spectrography, chemometrics, statistics, data mining and machine learning among others; see [23, 20, 16, 17, 2, 13] and references therein. In particular, nonnegative tensors appear in the decomposition of a joint distribution of discrete variables when they are independent conditionally to another discrete latent variable [16, 24].

The decomposition of tensors into a sum of rank-one terms is addressed therein, and a key issue is that such a decomposition be unique. Our motivation and our contributions are described in this section after some preliminary definitions.

**1.1. Definitions and notation.** A tensor of order  $n$  is often assimilated to a  $n$ -way array of numbers, which sometimes hides intrinsic properties. In particular, this array of coordinates is meaningful only if bases of underlying vector spaces have been defined in the first place. But beyond this obvious fact, this assimilation had led to some confusion concerning the definition of the uniqueness concept. For this reason, a slightly more formal definition is useful.

**Definition 1.** Denote  $V_i$  linear spaces of finite dimension  $d_i$  constructed on the same field  $\mathbb{K}$ ,  $1 \leq i \leq n$ , and  $\times_{i=1}^n V_i$  the set of  $n$ -tuples of vectors. Then the tensor product  $V = \otimes_{i=1}^n V_i$  is the free linear space spanned by  $\times_{i=1}^n V_i$  quotiented by an equivalence class, which imposes the property below  $\forall \alpha_i \in \mathbb{K}$ :

$$(1.1) \quad (\alpha_1 u_1, \alpha_2 u_2, \dots, \alpha_n u_n) \equiv \left( \prod_{i=1}^n \alpha_i \right) (u_1, u_2, \dots, u_n)$$

More details on the definition of tensor spaces may be found in [15, 11], and some simple examples in [7]. In this paper,  $\mathbb{K}$  will be the complex or the real field. The latter equivalence class yields what is often referred to as a “scaling indeterminacy”

---

Y.Qi and P.Comon are with GIPSA-Lab, CNRS UMR5216, Grenoble Campus, BP.46, F-38402 St Martin d’Heres cedex, France.

The project is funded by the European Research Council under the European Community’s Seventh Framework Program FP7/2007-2013 Grant Agreement no. 320594.

in the engineering literature. A direct consequence appears in the definition of “decomposable” tensors, also sometimes called “pure” tensors:

**Definition 2.** A decomposable tensor is of the form  $\otimes_{i=1}^n u_i$ ,  $u_i \in V_i$ .

From this definition, it is indeed clear that a decomposable tensor is not represented uniquely by a  $n$ -tuple of vectors. For instance, a rank-1 matrix can be represented by a pair  $(u_1, u_2)$  as well as by any other pair  $(\alpha_1 u_1, \alpha_2 u_2)$ , provided  $\alpha_1 \alpha_2 = 1$ .

Now, it is always possible to write a tensor as a sum of a finite number of decomposable tensors. When the number of terms is minimal, this decomposition is often qualified “Canonical Polyadic” (CP), and the minimal number  $R$  of terms defines the *tensor rank*:

**Definition 3.**  $\forall T \in \otimes_{i=1}^n V_i$ ,  $\exists u_i^{(r)} \in V_i$ ,  $1 \leq r \leq R$ , such that

$$(1.2) \quad T = \sum_{r=1}^R D^{(r)} \text{ with } D^{(r)} = u_1^{(r)} \otimes \cdots \otimes u_n^{(r)}$$

The minimal value of  $R$  is denoted  $\text{rank}\{T\}$ .

To summarise, in the CP decomposition, pure tensors  $D^{(r)}$  can be written in different ways in  $\times_n V_n$ , but have a unique representation in  $\otimes_n V_n$ . This is the basic difference between tensor and cartesian products between linear spaces [15, 11, 7].

We shall be mainly concerned by real nonnegative tensors. More precisely, denote  $V_i^+$  the subset of vectors with nonnegative entries in  $V_i$ , and  $V^+$  the subset of nonnegative tensors in  $V$ . The CP decomposition restricted to nonnegative decomposable tensors yields another definition of tensor rank:

**Definition 4.**  $\forall T \in V^+$ ,  $\exists u_i^{(r)} \in V_i^+$  such that  $T = \sum_{r=1}^R u_1^{(r)} \otimes \cdots \otimes u_n^{(r)}$ . The minimal value of  $R$  is called the nonnegative rank of  $T$ , and is denoted  $\text{rank}_+\{T\}$ .

An obvious property is that the nonnegative rank of a nonnegative tensor is always at least as large as its rank.

**1.2. Motivation.** An important issue is whether the CP decomposition defined in (1.2) is unique or not, in the sense that the set of decomposable tensors  $\{D^{(r)}, 1 \leq r \leq R\}$  is unique for a given  $T$  or for all  $T$  is some given class. Clearly, the above mentioned scaling indeterminacy is hence not part of the uniqueness problem. There already exist in the literature known sufficient conditions ensuring uniqueness of the CP decomposition. We have for instance [14, 18, 6]:

**Theorem 5** (Kruskal). *The CP decomposition of a tensor  $T$  is unique if*

$$\text{rank}\{T\} \leq \frac{1 + \sum_i (\kappa_i - 1)}{2}$$

where  $\kappa_i$  denote the so-called Kruskal’s rank of loading matrices, which correspond generically to dimensions  $d_i$  if  $\text{rank}\{T\} \geq d_i$ .

An another condition, easier to satisfy because the upper bound is larger, is sometimes preferred. However, contrary to Kruskal’s condition, it only guarantees CP uniqueness in an almost sure sense (*i.e.* for almost all tensors satisfying the constraint):

**Theorem 6** (Chiantini et al.). *The CP decomposition of a generic tensor  $T$  is unique if*

$$\text{rank}\{T\} < \left\lceil \frac{\prod_i d_i}{1 + \sum_i (d_i - 1)} \right\rceil$$

The authors of [6] also strengthen the above result by a prior compression of tensor  $T$ . The consequence is that dimensions  $d_i$  can be replaced in Prop. 6 by the multilinear ranks of  $T$ .

Yet, these results do not apply neither to nonnegative decompositions nor to low-rank approximations. The purpose of this paper is to provide some first results in these directions. Therefore, it will be necessary to distinguish between exact and approximate CP decompositions. Note that when a low-rank approximation is unique, its CP decomposition does not *a priori* need to be. This fact may raise problems in some applications.

**1.3. Contributions.** The next section addresses general low-rank approximations, whereas Section 3 is devoted to the case of rank-one approximations. In particular, Proposition 8 states the almost sure uniqueness of the best low-rank approximation of any nonnegative tensor in  $V^+$ , and Prop. 15 states the uniqueness of the best rank-one approximation of any positive tensor. The last section eventually points out the links existing between  $r$ -singular tuples and best rank- $r$  approximations. These results complement those obtained in [8, 9].

## 2. EXISTENCE AND UNIQUENESS OF APPROXIMATIONS

Given a nonnegative tensor  $T \in V^+$ , we consider the best rank- $r$  approximations of  $T$ , where  $r$  is less than the non-negative rank of  $T$ .

**Definition 7.** For a fixed positive integer  $r$ , let  $D_r = \{X \in V^+ | \text{rank}(X) \leq r\}$ , where  $\text{rank}(X)$  means the non-negative rank of  $X$ , and let  $d(T) = \inf_{X \in D_r} \|T - X\|$ , where  $\|\cdot\|$  is the  $l^2$ -norm.

According to [16],  $D_r$  is a closed set, thus for any  $T \notin D_r$ , there is some  $T^* \in D_r$  such that  $\|T - T^*\| = d(T)$ . The following result is based on [9, Corollary 18], and we give it a proof for completeness.

**Proposition 8.** Almost every  $T \in V^+$  with nonnegative rank  $> r$  has a unique best rank- $r$  approximation.

*Proof.* For any  $T, T' \in V_1 \otimes \cdots \otimes V_n$ ,  $|d(T) - d(T')| \leq \|T - T'\|$ , i.e.  $d$  is Lipschitz, thus differentiable a.e. in  $V_1 \otimes \cdots \otimes V_n$  by Rademacher's theorem.

For a general  $T \in V^+$ , there is an open neighbourhood  $B(T, \delta)$  of  $T$  contained in  $V^+$ , so  $d$  is differentiable a.e. in  $V^+$ . Assume that  $d$  is differentiable at  $T \in V^+$ , for any  $U \in V_1 \otimes \cdots \otimes V_n$ , let  $\partial d_T^2(U)$  be the differential of  $d^2$  at  $T$  along the direction  $U$ , let  $\|T - T^*\| = d(T)$ . Then,

$$d^2(T + tU) = d^2(T) + t\partial d_T^2(U) + O(t^2)$$

$$d^2(T + tU) \leq \|T + tU - T^*\|^2 = d^2(T) + 2t\langle U, T - T^* \rangle + t^2 \|U\|^2.$$

Therefore, for any  $t$ ,  $t\partial d_T^2(U) \leq 2t\langle U, T - T^* \rangle$ , then

$$\partial d_T^2(U) = 2\langle U, T - T^* \rangle.$$

Assume  $T'$  is another best rank- $r$  approximation of  $T$ , then

$$2\langle U, T - T^* \rangle = \partial d_T^2(U) = 2\langle U, T - T' \rangle,$$

which implies  $\langle T' - T^*, U \rangle = 0$  for any  $U$ , i.e.  $T' = T^*$ .  $\square$

**Proposition 9.** The nonnegative tensors which have nonnegative rank  $> r$  and do not have a unique best rank- $r$  approximation form a semi-algebraic set which does not contain an open set and is contained in some hypersurface.

*Proof.*  $D_r$  is the image of the polynomial map

$$\begin{aligned} \phi_r: (V_1^{\geq 0} \times \cdots \times V_n^{\geq 0})^r &\rightarrow V^+ \\ (u_{1,1}, \dots, u_{n,1}, \dots, u_{1,r}, \dots, u_{n,r}) &\mapsto \sum_{j=1}^r u_{1,j} \otimes \cdots \otimes u_{n,j} \end{aligned}$$

hence  $D_r$  is semi-algebraic by the Tarski-Seidenberg theorem [19], then the proposition follows by [10, Theorem 3.4].  $\square$

Now we study the necessary condition for  $\sum_{i=1}^r T_i$  to be a best rank- $r$  approximation.

**Lemma 10.** *If  $\text{rank}(T) > r$ , and  $\sum_{i=1}^k T_i$  is a best rank- $r$  approximation, where  $T_i = u_{1,i} \otimes \cdots \otimes u_{n,i}$  and  $\|\sum_{i=1}^r T_i\| = 1$ , and assume  $d(u_{i,j}) = \min \|T - \alpha \sum_{i=1}^r T_i\|$  is smooth at each  $u_{i,j}$ , then*

$$(2.1) \quad \langle T, \otimes_{k \neq i} u_{k,j} \rangle = \lambda \langle \sum_{l=1}^r T_l, \otimes_{k \neq i} u_{k,j} \rangle,$$

with  $\lambda = \langle T, \sum_{l=1}^r T_l \rangle$ , where  $\langle, \rangle$  denotes the contraction.

*Proof.* Let  $L$  denote the line in  $V_1 \otimes \cdots \otimes V_n$  spanned by  $\sum_{i=1}^r T_i$ , and  $L^\perp$  denote the orthogonal complement of  $L$ . Denote the projection of  $T$  to  $L$  by  $\text{Proj}_L(T)$ , then

$$\|T\|^2 = \|\text{Proj}_L(T)\|^2 + \|\text{Proj}_{L^\perp}(T)\|^2,$$

$$\left\| T - \alpha \sum_{i=1}^r T_i \right\|^2 = \|T - \text{Proj}_L(T)\|^2 = \|\text{Proj}_{L^\perp}(T)\|^2 = \|T\|^2 - \|\text{Proj}_L(T)\|^2,$$

so to compute  $\min \|T - \alpha \sum_{i=1}^r T_i\|$  is equivalent to compute  $\max \text{Proj}_L(T)$  which is  $\max \langle T, \sum_{i=1}^r T_i \rangle$ .

Consider the Lagrangian:

$$(2.2) \quad \phi = \langle T, \sum_{i=1}^r T_i \rangle - \lambda \left( \left\| \sum_{i=1}^r T_i \right\| - 1 \right),$$

then  $\frac{\partial \phi}{\partial u_{i,j}} = 0$  gives

$$(2.3) \quad \langle T, \otimes_{k \neq i} u_{k,j} \rangle = \lambda \langle \sum_{l=1}^r T_l, \otimes_{k \neq i} u_{k,j} \rangle$$

with  $\lambda = \langle T, \sum_{l=1}^r T_l \rangle$  for all  $1 \leq i \leq n$ ,  $1 \leq j \leq r$ .  $\square$

## 3. RANK ONE APPROXIMATION

Lemma 10 motivates us to propose the following definition:

**Definition 11.** For  $T \in V_1 \otimes \cdots \otimes V_n$ ,  $(\lambda, u_1 \otimes \cdots \otimes u_n)$  is called a nonnegative singular pair of  $T$  if  $\lambda \geq 0$ , and for all  $i \in \{1, \dots, n\}$ ,  $0 \neq u_i \geq 0$ , and

$$(3.1) \quad \langle T, \otimes_{j \neq i} u_j \rangle = \lambda u_i.$$

**Lemma 12.** A nonnegative tensor  $T$  has a nonnegative singular pair.

*Proof.* Let  $u_i = (u_{i,1}, \dots, u_{i,d_i})$  be the coordinate of  $u_i$ . Let  $D = \{(u_1, \dots, u_n) \mid u_{i,j} \geq 0, \sum_{i,j} u_{i,j} = 1\}$ , then  $D$  is a compact convex set. Define

$$\phi: D \rightarrow D$$

$$(u_1, \dots, u_n) \mapsto \left( \frac{\langle T, u_2 \otimes \cdots \otimes u_n \rangle}{\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l}, \dots, \frac{\langle T, u_1 \otimes \cdots \otimes u_{n-1} \rangle}{\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l} \right)$$

If  $\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l = 0$ , then  $\langle T, \otimes_{j \neq i} u_j \rangle = 0$  for all  $i$ , i.e.  $\lambda = 0$ .

If  $\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l > 0$ , by Brouwer's Fixed Point Theorem, there is some  $u_1 \otimes \cdots \otimes u_n$  such that  $\langle T, \otimes_{j \neq i} u_j \rangle = \lambda u_i$ , where  $\lambda = \sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l$ .  $\square$

**Lemma 13.** If  $T$  is positive,  $T$  has a nonnegative pair  $(\lambda, u_1 \otimes \cdots \otimes u_n)$  with  $\lambda > 0$ . If  $u_1 \otimes \cdots \otimes u_n$  has unit length, then  $u_1 \otimes \cdots \otimes u_n$  is unique and every  $u_i > 0$ .

*Proof.* Let  $I_i = \{j \mid u_{i,j} \neq 0\}$ , and  $\alpha = \min\{u_{i,j} \mid 1 \leq i \leq n, j \in I_i\}$ . For any  $i$  and  $k$ ,  $\lambda u_{i,k} = \langle T, \otimes_{j \neq i} u_j \rangle_k \geq \alpha^{n-1} \sum_{l_j \in I_j} T_{l_1 \dots l_{i-1} k l_{i+1} \dots l_n} > 0$ .

Assume  $T$  had two positive singular vector tuples  $u_1 \otimes \cdots \otimes u_n$  and  $v_1 \otimes \cdots \otimes v_n$  corresponding to  $\lambda$ ,

$$\langle T, u_1 \otimes \cdots \otimes u_{i-1} \otimes \widehat{u}_i \otimes u_{i+1} \otimes \cdots \otimes u_n \rangle = \lambda u_i,$$

$$\langle T, v_1 \otimes \cdots \otimes v_{i-1} \otimes \widehat{v}_i \otimes v_{i+1} \otimes \cdots \otimes v_n \rangle = \lambda v_i.$$

Let  $\alpha_i = \max\{\alpha \in \mathbb{R}_{\geq 0} \mid u_i - \alpha v_i \geq 0\}$  and  $\beta_i = \max\{\beta \in \mathbb{R}_{\geq 0} \mid v_i - \beta u_i \geq 0\}$ . Since  $\|u_i\| = \|v_i\| = 1$  and  $u_i, v_i > 0$ , then  $0 < \alpha_i, \beta_i \leq 1$ . Therefore

$$\lambda u_i = \langle T, \otimes_{j \neq i} u_j \rangle \geq \langle T, \otimes_{j \neq i} \alpha_j v_j \rangle = \lambda \prod_{j \neq i} \alpha_j \cdot v_i,$$

$$\lambda v_i = \langle T, \otimes_{j \neq i} v_j \rangle \geq \langle T, \otimes_{j \neq i} \beta_j u_j \rangle = \lambda \prod_{j \neq i} \beta_j \cdot u_i.$$

By the maximality of  $\alpha_i$ ,  $\frac{\prod_{j \neq i} \alpha_j}{\alpha_i} \leq 1$  for each  $i$ , thus  $\alpha_i = 1$ , and similarly,  $\beta_i = 1$ .  $\square$

**Remark 14.** Lemma 11 and Lemma 13 are an analogue of the Perron-Frobenius theorem [4, 8].

**Theorem 15.** A positive tensor  $T$  with rank  $> 1$  has a unique best rank one nonnegative approximation.

*Proof.* Since the smooth function

$$\begin{aligned} \varphi: \mathbb{S}^{d_1-1} \times \dots \times \mathbb{S}^{d_n-1} &\rightarrow \mathbb{R}_{\geq 0} \\ (u_1, \dots, u_n) &\mapsto \langle T, u_1 \otimes \dots \otimes u_n \rangle \end{aligned}$$

reaches its maximal value at some  $(u_1, \dots, u_n) \geq 0$ , where  $\mathbb{S}^{d_i-1}$  is unit sphere in  $V_i$ , then the critical points of the Lagrangian

$$\langle T, u_1 \otimes \dots \otimes u_n \rangle - \sum_{i=1}^n \lambda_i (\|u_i\| - 1)$$

give us

$$(3.2) \quad \langle T, u_1 \otimes \dots \otimes u_{i-1} \otimes \widehat{u}_i \otimes u_{i+1} \otimes \dots \otimes u_n \rangle = \lambda_i u_i$$

and  $\langle T, u_1 \otimes \dots \otimes u_n \rangle = \lambda_i$  gives us  $\lambda_1 = \dots = \lambda_n$ , denoted by  $\lambda$ .

Since  $\lambda$  is maximal,  $\lambda > 0$  and  $u_1 \otimes \dots \otimes u_n$  is unique by 13. Hence this unique critical point  $u_1 \otimes \dots \otimes u_n$  yields the best rank one approximation.  $\square$

After knowing the existence and uniqueness of a best rank- $r$  approximation, we want to find out this approximation explicitly. A first idea is to compute these approximations inductively: for  $k < r$ , let  $\sum_{i=1}^k T_i$  be the best rank- $k$  approximation of  $T$ , then we hope to obtain the best rank- $r$  approximation by computing the rank- $(r-k)$  approximation of  $T - \sum_{i=1}^k T_i$ . It has been already shown in [21] that this ‘‘deflation procedure’’ does not work for real or complex tensors of order strictly larger than 2. The following proposition shows that this doesn’t work either for nonnegative tensors.

**Proposition 16.** A best rank-2 approximation of a general  $T \in V^+$  can not be obtained by a sequence of rank-1 approximations.

*Proof.* We prove by contradiction. Assume

$$\|T - \alpha \cdot u_1 \otimes \dots \otimes u_n\| = \min_{X \in D_1} \|T - X\|,$$

$$\|T - \beta \cdot (\lambda u_1 \otimes \dots \otimes \lambda u_n + \lambda v_1 \otimes \dots \otimes \lambda v_n)\| = \min_{X \in D_2} \|T - X\|.$$

where  $\|u_i\| = \|\lambda u_1 \otimes \dots \otimes \lambda u_n + \lambda v_1 \otimes \dots \otimes \lambda v_n\| = 1$ ,  $\|v_1\| = \dots = \|v_n\| = \nu$ , and  $\alpha = \beta \lambda^n$ , then by Lemma 10 we have

$$\langle T, \otimes_{k \neq i} u_k \rangle = \alpha u_i,$$

$$\langle T, \otimes_{k \neq i} \lambda u_k \rangle = \beta \langle \lambda u_1 \otimes \dots \otimes \lambda u_n + \lambda v_1 \otimes \dots \otimes \lambda v_n, \otimes_{k \neq i} \lambda u_k \rangle,$$

$$\langle T, \otimes_{k \neq i} \lambda v_k \rangle = \beta \langle \lambda u_1 \otimes \dots \otimes \lambda u_n + \lambda v_1 \otimes \dots \otimes \lambda v_n, \otimes_{k \neq i} \lambda v_k \rangle,$$

which implies  $\prod_{k \neq i} \langle u_k, v_k \rangle = 0$  and  $\langle T, \otimes_{k \neq i} v_k \rangle = \alpha \nu^{2n-2} v_i$ .

Let  $\tilde{v}_i = \frac{v_i}{\nu}$ , then  $\langle T, \otimes_{k \neq i} \tilde{v}_k \rangle = \alpha \nu^n \tilde{v}_i$ , i.e.  $(\alpha, u_1 \otimes \dots \otimes u_n)$  and  $(\alpha \nu^n, \tilde{v}_1 \otimes \dots \otimes \tilde{v}_n)$  are orthogonal normalized singular pairs. By Lemma 13, for a general  $T$ ,  $u_i > 0$  and  $v_j > 0$ , which contradicts that  $\prod \langle u_i, v_i \rangle = 0$ .  $\square$

4.  $r$ -SINGULAR VECTOR TUPLES

In [9], Friedland and Ottaviani introduce the definition of singular vector tuples, and compute the number of them for a generic tensor. Due to the relationship with rank- $r$  approximations (Lemma 10), we would like to consider the vector tuples  $(u_{1,1}, \dots, u_{n,1}, \dots, u_{1,r}, \dots, u_{n,r}) \in (V_1 \times \dots \times V_n)^r$  satisfying

$$(4.1) \quad \begin{cases} \langle T, \otimes_{k \neq i} u_{k,j} \rangle = \lambda \langle \sum_{j=1}^r u_{1,j} \otimes \dots \otimes u_{n,j}, \otimes_{k \neq i} u_{k,j} \rangle \\ \langle \sum_{j=1}^r u_{1,j} \otimes \dots \otimes u_{n,j}, \sum_{j=1}^r u_{1,j} \otimes \dots \otimes u_{n,j} \rangle = 1 \end{cases}$$

for some  $\lambda$ , where  $\langle, \rangle$  denotes the contraction.

**Definition 17.** There is a unique polynomial  $\phi(\lambda)$  such that (4.1) has a nontrivial solution over  $\mathbb{C}$  if and only if  $\phi = 0$ . In fact  $\phi(\lambda)$  is the resultant of (4.1), and is called the  $r$ -characteristic polynomial of  $T$ . The zeros  $\lambda$  of  $\phi$  are called normalized  $r$ -singular values, and the vector tuples  $(u_{1,1}, \dots, u_{n,1}, \dots, u_{1,r}, \dots, u_{n,r})$  satisfying (4.1) are called the normalized  $r$ -singular vector tuples corresponding to  $\lambda$ .

**Remark 18.** When  $r = 1$ , an  $r$ -singular vector tuple is a singular vector tuple. When  $n = 2$ , if we do not require  $r$ -singular vector tuples to be normalized, and assume  $\lambda = 1$ , let  $T = \sum_i \sqrt{\sigma_i} u_i \otimes \sqrt{\sigma_i} v_i$  be a singular value decomposition, then  $r$  singular vectors give an  $r$ -singular vector tuples  $(\sqrt{\sigma_1} u_1, \sqrt{\sigma_1} v_1, \dots, \sqrt{\sigma_r} u_r, \sqrt{\sigma_r} v_r)$  satisfying  $\langle T, \sqrt{\sigma_i} u_i \rangle = \sigma_i(\sqrt{\sigma_i} v_i) = \langle \sum_j \sqrt{\sigma_j} u_j \otimes \sqrt{\sigma_j} v_j, \sqrt{\sigma_i} u_i \rangle$  and  $\langle T, \sqrt{\sigma_i} v_i \rangle = \sigma_i(\sqrt{\sigma_i} u_i) = \langle \sum_j \sqrt{\sigma_j} u_j \otimes \sqrt{\sigma_j} v_j, \sqrt{\sigma_i} v_i \rangle$ .

It is shown in [3] the set of normalized eigenvalues of a tensor is either finite or the complement of a finite set. Similarly, we have

**Proposition 19.** For any tensor  $T$ , the set of normalized  $r$ -singular values  $\theta(T)$  is either a finite set or the complement of a finite set.

*Proof.* Consider the projection  $\pi : \mathbb{C} \times (V_1 \times \dots \times V_n)^r \rightarrow \mathbb{C}$ , then  $\theta(T)$  is the image of the variety defined by (4.1) under  $\pi$ . By Chevalley's theorem [12],  $\theta(T)$  is a constructible subset of  $\mathbb{C}$ , thus either a finite set or the complement of a finite set in  $\mathbb{C}$ .  $\square$

The number of normalized singular values and singular vector tuples of a tensor  $T$  is related to the number of decompositions of  $T$ : assume the distance function  $d(T, \sigma_r(\mathbb{P}V_1 \times \dots \times \mathbb{P}V_n))$  reaches the minimal value at some point  $T_0 \in \sigma_r(\mathbb{P}V_1 \times \dots \times \mathbb{P}V_n)$  and is smooth at  $T_0$ , then any decomposition of  $T_0$  gives a  $r$ -singular vector tuple of  $T$ . If  $T_0$  has an infinite number of decompositions, then  $T$  has an infinite number of  $r$ -singular vector tuples. We need the following definition.

**Definition 20** ([1, 5, 22]). Let  $X \subset \mathbb{P}^N$  be a reduced, irreducible projective variety of dimension  $n$ .  $X$  is called  $r$ -defective if  $\dim \sigma_r(X) < \min\{N, rn + r - 1\}$ .

In our case, a general  $T \in V_1 \otimes \dots \otimes V_n$  has a finite number of decompositions if and only if  $\mathbb{P}V_1 \times \dots \times \mathbb{P}V_n$  is not  $r$ -defective. We will focus on such  $r$ 's. By the result of [9], we see 0 is not a normalized  $r$ -singular value for a generic tensor. In fact, we have

**Theorem 21.** *A generic tensor in  $V_1 \otimes \cdots \otimes V_n$  has a finite number of simple normalized  $r$ -singular vector tuples corresponding to nonzero  $r$ -singular values for any  $r$  such that  $\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n$  is not  $r$ -defective.*

## 5. PROOF OF THEOREM 21

For simplicity we assume  $\lambda = 1$ , and do not require the corresponding  $r$ -singular vector tuples to be normalized, i.e. we use the following equivalent definition:

**Definition 22.** A vector tuple  $(u_{1,1}, \dots, u_{n,1}, \dots, u_{1,r}, \dots, u_{n,r}) \in (V_1 \times \cdots \times V_n)^r$  with  $u_{i,j} \neq 0$  for all  $i, j$  is called a  $r$ -singular vector tuple of  $T$  if

$$(5.1) \quad \langle T, \otimes_{k \neq i} u_{k,j} \rangle = \left\langle \sum_{j=1}^r u_{1,j} \otimes \cdots \otimes u_{n,j}, \otimes_{k \neq i} u_{k,j} \right\rangle$$

for all  $i, j$ .

We use the vector bundle method introduced by Friedland and Ottaviani in [9] to prove the theorem. Let  $M$  be a nonsingular complex variety, and  $E \xrightarrow{\pi} M$  be a holomorphic vector bundle on  $M$  with  $\dim M = \text{rank} E$ . Let  $S \subset H^0(M, E)$  be a finite dimensional subspace, and  $M \times S \xrightarrow{\nu} E$  be the evaluation map  $(p, s) \mapsto s(p)$ . We say  $S$  generates  $E$  if  $\{s(p) | s \in S\} = \pi^{-1}(p)$  for all  $p \in M$ .

**Lemma 23.** *If there exists an open subset  $U \subset M$  such that  $S_U$  generates  $H^0(U, E)$ ,  $T_\nu : T_p M \times T_s S \rightarrow T_{s(p)} E$  is surjective on  $U \times S_U$ , and for a generic  $\sigma \in S$ , the zero locus of  $\sigma$ ,  $Z_\sigma$ , is noetherian and contained in  $U$ , then  $Z_\sigma$  consists of a finite number of simple points.*

*Proof.* Let  $\tau$  be the zero section of  $E$ , and  $Z_U := \{(p, s_U) \in U \times S_U | s_U(p) = 0\} \subset \nu^{-1}(\tau)$ . Since  $\nu$  is dominant and  $T_\nu$  is surjective on  $U \times S_U$ ,  $Z_U$  has dimension  $\dim M + \dim S - \text{rank} E$ . Let  $p : \nu^{-1}(\tau) \rightarrow S$  be the projection, and  $p_U : Z_U \rightarrow S_U$  the restriction on  $U$ . Since for a generic  $\sigma \in S$ ,  $Z_\sigma$  is isomorphic to  $p_U^{-1}(\sigma)$ . By the generic smoothness theorem [12],  $Z_\sigma$  is a smooth 0-dimensional subvariety of  $M$ . Since  $Z_\sigma$  is noetherian, then  $Z_\sigma$  is of a finite number of simple points.  $\square$

For convenience, we generalize Lemma 23 to a vector bundle over a singular variety. Let  $E$  be a vector bundle over a variety  $M$  with fiber isomorphic to  $V$ , and  $U$  an open subset of  $M$ . Assume  $S \subset H^0(M, E)$  is a finite dimensional subvariety and generates  $E|_U$ , and let  $M \times S \xrightarrow{\nu} V$  be the evaluation map  $(p, s) \mapsto s(p)$ .

**Lemma 24.** *Assume for any  $v \in V$ ,  $\nu^{-1}(v) \subset U \times S_U$  has the same dimension, and for a generic  $\sigma \in S$ , the zero locus of  $\sigma$ ,  $Z_\sigma$ , is noetherian. In  $U$ ,*

- (1) *if  $\text{rank} E = \dim M$ ,  $Z_\sigma|_U$  is of a finite number of simple points.*
- (2) *if  $\text{rank} E > \dim M$ ,  $Z_\sigma|_U$  is empty.*

*Proof.* Since  $\nu$  is dominant and each fibre of  $\nu$  has same dimension, then  $Z_U$  has the relative dimension  $\dim M + \dim S - \text{rank} E$ .  $Z_\sigma|_U$  is isomorphic to  $p_U^{-1}(\sigma)$ , and is noetherian, so when  $\text{rank} E = \dim M$ , by the generic smoothness theorem,  $Z_\sigma|_U$  is a smooth 0-dimensional variety, thus consists of a finite number of simple points. When  $\text{rank} E > \dim M$ ,  $Z_\sigma|_U$  is empty.  $\square$

**Remark 25.** This lemma is an analogue of the "Bertini-type" theorem [9].

We follow the idea of [9, Theorem 6] to propose the following lemma as a refinement of Lemma 24.

**Definition 26.** Let  $\pi : E \rightarrow M$  be a vector bundle over a variety  $M$  with  $\text{rank} E \geq \dim M$ , and  $S \subset H^0(M, E)$  be a finite dimensional variety. We say  $S$  almost generates  $E$  if there are a finite number of proper subvarieties  $Y_1, \dots, Y_k$  of  $M$ , over each  $Y_i$  there is a subbundle of  $E$ ,  $\pi_i : F_i \rightarrow Y_i$ , and there are subvarieties  $S_i \subset H^0(Y_i, F_i)$  of  $S$  such that:

- (1) On the open subset  $U = M \setminus (\cup_i Y_i)$ ,  $S$  generates  $E$ , and each fibre of the evaluation map  $\nu : S \times U \rightarrow E|_U$  has the same dimension.
- (2) If  $Y_i \subset Y_j$ ,  $F_i$  is a subbundle of  $F_j$ .
- (3) Let  $P_i$  be the set of all  $j$  such that  $Y_j$  is a proper subvariety of  $Y_i$ , on  $U_i = Y_i \setminus (\cup_{k \in P_i} Y_k)$ ,  $S_i$  generates  $F_i|_{U_i}$ , and each fibre of the evaluation map  $\nu : S_i \times U_i \rightarrow F_i|_{U_i}$  has the same dimension.
- (4)  $\text{rank} F_i \geq \dim Y_i$ .

**Lemma 27.** Let  $\pi : E \rightarrow M$  be a vector bundle over a variety  $M$  with  $\text{rank} E = \dim M$ ,  $S \subset H^0(M, E)$  be a finite dimensional variety, and  $\sigma \in S$  be a general section in  $S$ . Assume  $S$  almost generates  $E$ , and the zero locus of  $\sigma$ ,  $Z_\sigma$ , is noetherian, then  $Z_\sigma$  consists of a finite number of simple points.

*Proof.* Let  $F_i \rightarrow Y_i$  be subbundles of  $E$  over subvarieties of  $M$  satisfying Definition 26. By Lemma 24,  $Z_\sigma|_{U_i}$  is empty or of a finite number of simple points, and  $Z_\sigma|_U$  consists of a finite number of simple points. Hence  $Z_\sigma$  is of a finite number of simple points.  $\square$

**Remark 28.** Since we only use the dimension counting and the generic smoothness theorem in Lemma 24, we can generalize the lemma to the following setting with the same proof:

Let  $\pi : E \rightarrow M$  be a morphism between varieties, such that each fibre  $\pi^{-1}(p)$  is isomorphic to an affine space  $\mathbb{A}^n$ ,  $n$  is called the rank of  $E$ . Let  $H^0(M, E)$  be the set of morphisms  $s : M \rightarrow E$  such that  $\pi \circ s = \text{id}$ . We say  $S \subset H^0(M, E)$  generates  $E$  if  $\{s(p) | s \in S\} = \pi^{-1}(p)$  for all  $p \in M$ .

Similarly, Definition 26 and Lemma 27 can be generalized to this setting, which means we do not need to consider the transition morphisms.

For each  $1 \leq i \leq n$ , let  $V_i$  be a complex vector space with dimension  $d_i$ , and for each  $1 \leq j \leq r$ ,  $V_{i,j}$  be a complex vector space isomorphic to  $V_i$ . Let  $X_j$  denote  $\text{Seg}(\mathbb{P}V_{1,j} \times \dots \times \mathbb{P}V_{n,j})$ , and  $\alpha_{i,j} : X_j \rightarrow \mathbb{P}V_{i,j}$  be the natural projection. Let  $X = X_1 \times \dots \times X_r$ , and  $\beta_i : X \rightarrow X_i$  be the projection. Let  $T_{i,j}$  be the tautological

line bundle over  $\mathbb{P}V_{i,j}$ , and  $M = \bigoplus_{j=1}^r \beta_j^* \left( \bigotimes_{i=1}^n \alpha_{i,j}^*(T_{i,j}) \right)$  be a rank- $r$  vector bundle

$M \xrightarrow{\gamma} X$ . Let  $F_{i,j}$  be the trivial bundle over  $M$  with fibre  $V_{i,j}$ , and  $Q_{i,j}$  denote the quotient bundle

$$0 \rightarrow \gamma^* \cdot \beta_j^* \cdot \alpha_{i,j}^*(T_{i,j}) \rightarrow F_{i,j} \rightarrow Q_{i,j} \rightarrow 0.$$

Let  $H_{i,j} = \text{Hom}(\gamma^* \cdot \beta_j^* (\bigotimes_{k \neq i} \alpha_{k,j}^*(T_{k,j})), F_{i,j})$ ,  $\tilde{H}_{i,j} = \text{Hom}(\gamma^* \cdot \beta_j^* (\bigotimes_{k \neq i} \alpha_{k,j}^*(T_{k,j})), Q_{i,j})$

and  $E = \bigoplus_{j=1}^r (H_{1,j} \oplus \tilde{H}_{2,j} \oplus \dots \oplus \tilde{H}_{n,j})$ . So  $\text{rank} E = \dim M = r \sum_{i=1}^n d_i - r(n-1)$ .

Now we fix  $r$  such that  $\text{Seg}(\mathbb{P}V_1 \times \dots \times \mathbb{P}V_n)$  is not  $r$ -defective. Let  $C_{i,j}$  be the quadric hypersurface in  $V_{i,j}$  defined by  $\{v \in V_{i,j} | v^\top v = 0\}$ . Let  $U$  be the open

subset of  $M$  consisting of  $p = (\otimes_{i=1}^n u_{i,1}, \dots, \otimes_{i=1}^n u_{i,r})$  such that each  $u_{i,j} \notin C_{i,j}$ , and

$$(5.2) \quad \dim \left( \sum_{i,j} \left( \otimes_{k \neq i} u_{k,j} \right) \otimes V_{i,j} \right) = r \sum_{i=1}^n d_i - r(n-1).$$

**Lemma 29.** For  $p \in U$ , let  $\tilde{A} = \sum_{j=1}^r \otimes_{i=1}^n u_{i,j}$ , then for any  $x_{i,j} \in V_{i,j}$  and  $[y_{i,j}] \in V_{i,j}/\langle u_{i,j} \rangle$ ,

- (1) There is some  $A \in V_1 \otimes \dots \otimes V_n$  such that  $\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle = x_{i,j}$  if and only if  $u_{1,j}^\top x_{1,j} = \dots = u_{n,j}^\top x_{n,j}$  for all  $j$ .
- (2) There is some  $A \in V_1 \otimes \dots \otimes V_n$  such that  $\langle A - \tilde{A}, \otimes_{k=2}^n u_{k,j} \rangle = x_{1,j}$  and  $[\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle] = [y_{i,j}]$  for all  $i \geq 2$ .

*Proof.* 1. If  $\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle = x_{i,j}$ , then  $\langle A - \tilde{A}, \otimes_{k=1}^n u_{k,j} \rangle = u_{i,j}^\top x_{i,j}$ .

For the linear system of  $A$ :

$$(5.3) \quad \langle A, \otimes_{k \neq i} u_{k,j} \rangle = \langle \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle + x_{i,j},$$

let  $U$  denote the coefficient matrix of  $A$  formed by  $\otimes_{k \neq i} u_{k,j}$ , and  $z$  be the vector corresponding to  $\langle \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle + x_{i,j}$ .

Conversely, the linear system (5.3) is solvable if and only if the matrices  $U$  and  $[U, z]$  have the same rank. By (5.2),  $\text{rank}(U) = r \sum_{i=1}^n d_i - r(n-1)$ , and  $u_{i,j} \otimes (\otimes_{k \neq i} u_{k,j}) = u_{i,j} \otimes (\otimes_{k \neq i} u_{k,j})$  are generators of the linear relations in  $[\otimes_{k \neq i} u_{k,j}]$ . So if  $u_{1,j}^\top x_{1,j} = \dots = u_{n,j}^\top x_{n,j}$ , then  $\text{rank}([U, z]) \leq r \sum_{i=1}^n d_i - r(n-1) = \text{rank}(U)$ .

2. By 1. the system

$$(5.4) \quad \begin{cases} \langle A, \otimes_{k \geq 2} u_{k,j} \rangle = \langle \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle + x_{1,j} \\ \langle A, \otimes_{k \neq i} u_{k,j} \rangle = \langle \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle + y_{i,j} + t_{i,j} u_{i,j} \end{cases}$$

for  $i \geq 2$  is solvable for some  $A$  and  $t_{i,j}$  if and only if

$$(5.5) \quad u_{1,j}^\top x_{1,j} = u_{2,j}^\top (y_{2,j} + t_{2,j} u_{2,j}) = \dots = u_{n,j}^\top (y_{n,j} + t_{n,j} u_{n,j}).$$

Since  $u_{i,j} \notin C_{i,j}$ , when  $t_{i,j} = \frac{u_{1,j}^\top x_{1,j} - u_{i,j}^\top y_{i,j}}{u_{i,j}^\top u_{i,j}}$ , (5.4) is solvable.  $\square$

Let  $S = \{s \in H^0(M, E) \mid s = \left( \bigoplus_{j=1}^r \langle A - \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle, \bigoplus_{j=1}^r \bigoplus_{i=2}^n [\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle] \right)\}$ .

**Lemma 30.** For  $\alpha = \{(i_1, j_1), \dots, (i_l, j_l) \mid 1 \leq i_1, \dots, i_l \leq n, 1 \leq j_1, \dots, j_l \leq r\}$ , let  $F_\alpha = X_{1,1} \times \dots \times X_{n,r}$ , where  $X_{i,j} = \mathbb{P}(C_{i,j})$  if  $(i, j) \in \alpha$  and  $X_{i,j} = \mathbb{P}V_{i,j}$  otherwise. For  $p \in M|_{F_\alpha}$ ,  $\langle A - \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle = x_{1,j}$  and  $[\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle] = [y_{i,j}]$  for  $i \geq 2$  if and only if

$$(5.6) \quad u_{i,j}^\top y_{i,j} = u_{1,j}^\top x_{1,j}$$

for  $(i, j) \in \alpha$  and  $i \geq 2$ .

*Proof.* By Lemma 29, the system (5.4) for  $i \geq 2$  is solvable for some  $A$  and  $t_{i,j}$  if and only if

$$(5.7) \quad u_{1,j}^\top x_{1,j} = u_{2,j}^\top (y_{2,j} + t_{2,j} u_{2,j}) = \cdots = u_{n,j}^\top (y_{n,j} + t_{n,j} u_{n,j}).$$

When  $(i,j) \notin \alpha$ , let  $t_{i,j} = \frac{u_{1,j}^\top x_{1,j} - u_{i,j}^\top y_{i,j}}{u_{i,j}^\top u_{i,j}}$ , when  $(i,j) \in \alpha$ .  $u_{i,j}^\top y_{i,j} = u_{1,j}^\top x_{1,j}$ .  $\square$

Let  $R = r \sum_{i=1}^n d_i - r(n-1)$  and  $D_q = \{p \in M \mid \dim \left( \sum_{i,j} (\otimes_{k \neq i} u_{k,j}) \otimes V_{i,j} \right) \leq R - q\}$ , choose a basis  $a_{i,j}^l$  for  $V_{i,j}$ , consider the matrix  $N$  formed by  $\otimes_{k \neq i} u_{k,j} \otimes a_{i,j}^l$ , then  $D_q$  is defined by  $(R-q) \times (R-q)$  minors of  $N$ . If  $D_q \setminus D_{q+1} \neq \emptyset$ , for each  $p \in D_q \setminus D_{q+1}$ , some  $(R-q-1) \times (R-q-1)$  minor does not vanish, which gives an open covering of  $D_q \setminus D_{q+1}$ . If  $D_q = \cdots = D_{q+l}$ , then for each  $p \in D_q \setminus D_{q+l+1}$ , some  $(R-q-l-1) \times (R-q-l-1)$  minor does not vanish, which forms an open covering of  $D_q \setminus D_{q+l+1}$ . Without loss of generality we assume  $D_q \setminus D_{q+1} \neq \emptyset$ , then for each  $p \in D_q \setminus D_{q+1}$ , there are  $R-q$  independent vectors  $\sum_{i=1}^r (\otimes_{k \neq i} u_{k,j}) \otimes w_{i,j}$  which span  $\sum_{i,j} (\otimes_{k \neq i} u_{k,j}) \otimes V_{i,j}$ , and the linear space of linear equations that  $p$  satisfies are spanned by  $r(n-1) + q$  linear equations  $\sum (\otimes_{k \neq i} u_{k,j}) \otimes v_{i,j} = 0$ , where  $v_{i,j} \in V_{i,j}$  is algebraic in  $u_{k,l}$ . Among these linear equations we can choose  $r(n-1)$  generators to be  $\otimes_{k \neq i} u_{k,j} \otimes u_{i,j} = \otimes_{h \neq l} u_{h,j} \otimes u_{l,j}$  for  $i \neq l$ , and we choose and denote the other  $q$  linear equations by  $\sum (\otimes_{k \neq i} u_{k,j}) \otimes v_{i,j}^l = 0$  for  $1 \leq l \leq q$ .

**Lemma 31.** For  $p \in M|_{D_q \setminus (\cup_\alpha F_\alpha \cup D_{q+1})}$ ,

- (1)  $\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle = x_{i,j}$  if and only if  $u_{1,j}^\top x_{1,j} = \cdots = u_{n,j}^\top x_{n,j}$  and  $\sum_{i,j} x_{i,j}^\top v_{i,j}^l = 0$  for  $1 \leq l \leq q$ .
- (2)  $\langle A - \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle = x_{1,j}$  and  $[\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle] = [y_{i,j}]$  for  $i \geq 2$  if and only if

$$(5.8) \quad \sum_j \left( x_{1,j}^\top v_{1,j}^l + \sum_{i \geq 2} (y_{i,j} + t_{i,j} u_{i,j})^\top v_{i,j}^l \right) = 0,$$

where  $t_{i,j} = \frac{u_{1,j}^\top x_{1,j} - u_{i,j}^\top y_{i,j}}{u_{i,j}^\top u_{i,j}}$  and  $1 \leq l \leq q$ . For each  $p$ , the linear subspace formed by the tuples  $(x_{1,j}, [y_{i,j}])$  satisfying (5.8) is independent of the choice of  $v_{i,j}^l$ .

*Proof.* 1. If  $\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle = x_{i,j}$ , then  $\langle A - \tilde{A}, \otimes_{k=1}^n u_{k,j} \rangle = u_{i,j}^\top x_{i,j}$ , and  $\sum_{i,j} x_{i,j}^\top v_{i,j}^l = \langle A - \tilde{A}, \sum_{i,j} \otimes_{k \neq i} u_{k,j} \otimes v_{i,j}^l \rangle = 0$ .

We use the same notation as Lemma 29. The linear system of  $A$

$$(5.9) \quad \langle A, \otimes_{k \neq i} u_{k,j} \rangle = \langle \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle + x_{i,j}$$

is solvable if and only if the matrices  $U$  and  $[U, z]$  have the same rank  $R - q$ . Since  $u_{i,j} \otimes (\otimes_{k \neq i} u_{k,j}) = u_{l,j} \otimes (\otimes_{k \neq l} u_{k,j})$  and  $\sum (\otimes_{k \neq i} u_{k,j}) \otimes v_{i,j}^l = 0$  generate the linear relations in  $U$ . So if  $u_{1,j}^\top x_{1,j} = \cdots = u_{n,j}^\top x_{n,j}$ , and  $\sum_{i,j} x_{i,j}^\top v_{i,j}^l = 0$ , then  $\text{rank}([U, z]) \leq R - q = \text{rank}(U)$ .

2. By 1. the system (5.4) is solvable for some  $A$  if and only if

$$(5.10) \quad \begin{cases} u_{1,j}^\top x_{1,j} = u_{2,j}^\top (y_{2,j} + t_{2,j} u_{2,j}) = \cdots = u_{n,j}^\top (y_{n,j} + t_{n,j} u_{n,j}), \\ \sum_j \left( x_{1,j}^\top v_{1,j}^l + \sum_{i \geq 2} (y_{i,j} + t_{i,j} u_{i,j})^\top v_{i,j}^l \right) = 0. \end{cases}$$

Since  $u_{i,j} \notin C_{i,j}$ , let  $t_{i,j} = \frac{u_{1,j}^\top x_{1,j} - u_{i,j}^\top y_{i,j}}{u_{i,j}^\top u_{i,j}}$ , then (5.4) is solvable for  $i \geq 2$  if and only if and (5.8) holds.

The subset formed by  $\{x_{1,j}, [y_{i,j}]\}$  such that (5.10) holds is a linear space, and does not depend on the choice of  $v_{i,j}^l$ .  $\square$

**Lemma 32.** *For  $p \in M|_{(D_q \cap F_\alpha) \setminus D_{q+1}}$ ,  $\langle A - \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle = x_{1,j}$  and  $\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle = [y_{i,j}]$  for  $i \geq 2$  if and only if (5.6) holds for  $(i,j) \in \alpha$ , and (5.8) holds for  $t_{i,j} = \frac{u_{1,j}^\top x_{1,j} - u_{i,j}^\top y_{i,j}}{u_{i,j}^\top u_{i,j}}$  when  $(i,j) \notin \alpha$ , and for some  $t_{i,j}$  when  $(i,j) \in \alpha$ . For each  $p$ , the linear subspace consisting of  $\{x_{1,j}, [y_{i,j}]\}$  satisfying these conditions is independent of the choice of  $v_{i,j}^l$ .*

*Proof.* The system (5.4) is solvable for some  $A$  if and only if

$$(5.11) \quad \begin{cases} u_{1,j}^\top x_{1,j} = u_{2,j}^\top (y_{2,j} + t_{2,j} u_{2,j}) = \cdots = u_{n,j}^\top (y_{n,j} + t_{n,j} u_{n,j}), \\ \sum_j \left( x_{1,j}^\top v_{1,j}^l + \sum_{i \geq 2} (y_{i,j} + t_{i,j} u_{i,j})^\top v_{i,j}^l \right) = 0. \end{cases}$$

When  $(i,j) \notin \alpha$ , let  $t_{i,j} = \frac{u_{1,j}^\top x_{1,j} - u_{i,j}^\top y_{i,j}}{u_{i,j}^\top u_{i,j}}$ . Then (5.4) is solvable if and only if (5.6) and (5.8) holds for some  $t_{i,j}$  when  $(i,j) \in \alpha$ .

The subset formed by  $\{x_{1,j}, [y_{i,j}]\}$  satisfying these conditions is a linear space, and does not depend on the choice of  $v_{i,j}^l$ .  $\square$

*Proof of Theorem 21.* We prove the theorem by induction on  $r$ .

When  $r = 1$ , it is showed by Friedland and Ottaviani [9] that a generic  $T$  has a finite number of singular vector tuples. Assume the theorem holds for  $r - 1$ .

Let  $U \subset M$  be the open subset such that each  $\otimes_{k=1}^n u_{k,j} \neq 0$ , then by Lemma 29,  $S$  generates  $E|_{U'}$ , where  $U' = U \setminus (\cup_q D_q \cup \cup_\alpha F_\alpha)$ , and  $\text{rank} = \dim M$ . The dimension of the solutions of (5.4) for any  $(x_{1,j}, [y_{i,j}])$  is the same.

Let  $E_\alpha$  be the subbundle of  $E$  over  $M|_{F_\alpha}$  such that  $(x_{1,j}, [y_{i,j}])$  satisfies (5.6) for  $(i,j) \in \alpha$ , and  $S_\alpha = \{(\oplus_j \langle A - \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle, \oplus_{i,j} [\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle])\}$ , then by Lemma 30,  $S_\alpha$  generates  $E_\alpha|_{U_\alpha}$ , where  $U_\alpha = F_\alpha \setminus (\cup_q D_q \cup \cup_{\beta \subset \alpha} F_\beta)$ , and  $\text{rank} E_\alpha \geq \dim F_\alpha$ . The dimension of the solutions of (5.6) for any  $(x_{1,j}, [y_{i,j}])$  is the same.

Let  $E_q$  be the subbundle of  $E$  over  $M|_{D_q}$  such that  $(x_{1,j}, [y_{i,j}])$  satisfies (5.8), and  $S_q = \{(\oplus_j \langle A - \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle, \oplus_{i,j} [\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle])\}$ , then by Lemma 31,  $S_q$  generates  $E_q|_{U_q}$ , where  $U_q = D_q \setminus (D_{q+1} \cup \cup_\alpha F_\alpha)$ , and  $\text{rank} E_q \geq \dim D_q$ . The dimension of the solutions of (5.8) for any  $(x_{1,j}, [y_{i,j}])$  is the same.

Let  $E_{\alpha,q}$  be the subbundle of  $E$  over  $M|_{F_\alpha \cap D_q}$  such that  $(x_{1,j}, [y_{i,j}])$  satisfies (5.6) and (5.8), and  $S_{\alpha,q} = \{(\oplus_j \langle A - \tilde{A}, \otimes_{k \geq 2} u_{k,j} \rangle, \oplus_{i,j} [\langle A - \tilde{A}, \otimes_{k \neq i} u_{k,j} \rangle])\}$ , then

by Lemma 32,  $S_{\alpha,q}$  generates  $E_{\alpha,q}|_{U_{\alpha,q}}$ , where  $U_{\alpha,q} = (F_{\alpha} \cap D_q) \setminus (D_{q+1} \cup \cup_{\beta \subset \alpha} F_{\beta})$ , and  $\text{rank} E_{\alpha,q} \geq \dim F_{\alpha,q}$ . The dimension of the solutions of (5.6) and (5.8) for any  $(x_{1,j}, [y_{i,j}])$  is the same.

Therefore  $S$  almost generates  $E|_U$ , and since for a generic  $\sigma \in S$ ,  $Z_{\sigma}$  is a complex affine algebraic variety, by Lemma 27,  $Z_{\sigma}|_U$  consists of a finite number of points. With the induction assumption, we can conclude that  $Z_{\sigma}$  consists of a finite number of simple points.  $\square$

## REFERENCES

- [1] H. Abo, G. Ottaviani, and C. Peterson, "Induction for secant varieties of Segre varieties," *Trans. Amer. Math. Soc.*, pp. 767–792, 2009, arXiv:math/0607191. [Online]. Available: <http://web.math.unifi.it/users/ottavian/public.html>
- [2] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 1, pp. 6–20, 2009.
- [3] D. Cartwright and B. Sturmfels, "The number of eigenvalues of a tensor," *Linear Algebra and its Applications*, vol. 438, pp. 942–952, 2013.
- [4] K. C. Chang, K. Pearson, and T. Zhang, "Perron-Frobenius theorem for nonnegative tensors," *Commun. Math. Sci.*, vol. 6, no. 2, pp. 507–520, 2008.
- [5] L. Chiantini and C. Ciliberto, "Weakly defective varieties," *Trans. Amer. Math. Soc.*, vol. 354, no. 1, pp. 151–178, 2002.
- [6] L. Chiantini, G. Ottaviani, and N. Vannieuwenhoven, "An algorithm for generic and low-rank specific identifiability of complex tensors," *SIAM J. matrix Ana. Appl.*, 2014, to appear.
- [7] P. Comon, "Tensors: a brief introduction," *IEEE Sig. Proc. Magazine*, vol. 31, no. 3, pp. 44–53, May 2014, special issue on BSS. hal-00923279.
- [8] S. Friedland, S. Gaubert, and L. Han, "Perron-Frobenius theorem for nonnegative multilinear forms and extensions," *Linear Algebra Appl.*, vol. 438, pp. 738–749, 2013.
- [9] S. Friedland and G. Ottaviani, "The number of singular vector tuples and uniqueness of best rank-one approximation of tensors", Mar. 2014, arXiv:1210.8316, to appear in Foundations of Computational Mathematics.
- [10] S. Friedland and M. Stawiska, *Best approximation on semi-algebraic sets and k-border rank approximation of symmetric tensors*, Nov. 2013, arXiv:1311.1561.
- [11] W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, ser. Series in Computational Mathematics. Berlin, Heidelberg: Springer, 2012.
- [12] R. Hartshorne, *Algebraic Geometry*, ser. Graduate Texts in Mathematics. New York: Springer, 1977, vol. 52.
- [13] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, Sep. 2009.
- [14] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions," *Linear Algebra and Applications*, vol. 18, pp. 95–138, 1977.
- [15] J. M. Landsberg, *Tensors: Geometry and Applications*, ser. Graduate Studies in Mathematics. AMS publ., 2012, vol. 128.
- [16] L.-H. Lim and P. Comon, "Nonnegative approximations of nonnegative tensors," *Jour. Chemometrics*, vol. 23, pp. 432–441, Aug. 2009, hal-00410056. [Online]. Available: <http://dx.doi.org/10.1002/cem.1244>
- [17] A. Shashua and T. Hazan, "Non-negative tensor factorization with applications to statistics and computer vision," in *22nd International Conference on Machine Learning*, Bonn, 2005, pp. 792–799.
- [18] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *Jour. Chemo.*, vol. 14, pp. 229–239, 2000.
- [19] V. D. Silva and L.-H. Lim, "Tensor rank and the ill-posedness of the best low-rank approximation problem," *SIAM Journal on Matrix Analysis Appl.*, vol. 30, no. 3, pp. 1084–1127, 2008.
- [20] A. Smilde, R. Bro, and P. Geladi, *Multi-Way Analysis*. Chichester UK: Wiley, 2004.
- [21] A. Stegeman and P. Comon, "Subtracting a best rank-1 approximation does not necessarily decrease tensor rank," *Linear Algebra Appl.*, vol. 433, no. 7, pp. 1276–1300, Dec. 2010, hal-00512275.

- [22] F. L. Zak, *Tangents and secants of algebraic varieties*, ser. Translations of Mathematical Monographs. Providence, RI: American Mathematical Society, 1993, vol. 127, translated from the Russian manuscript by the author. MR 1234494 (94i:14053).
- [23] P. Zhang, H. Wang, R. Plemmons, and P. Pauca, "Tensor methods for hyperspectral data analysis: A space object material identification study," *J. Optical Soc. Amer.*, vol. 25, no. 12, pp. 3001–3012, Dec. 2008.
- [24] J. Zhou, A. Bhattacharya, A. Herring, and D. Dunson, *Bayesian factorizations of big sparse tensors*, Jun. 2013, arXiv:1306.1598.