

**On generic nonexistence of the  
Schmidt-Eckart-Young decomposition  
for complex tensors**

*Nick Vannieuwenhoven      Johannes Nicaise*  
*Raf Vandebril              Karl Meerbergen*

*Report TW 628, May 2013*



**Katholieke Universiteit Leuven**  
Department of Computer Science  
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

# On generic nonexistence of the Schmidt-Eckart-Young decomposition for complex tensors

*Nick Vannieuwenhoven*      *Johannes Nicaise*  
*Raf Vandebril*      *Karl Meerbergen*

*Report TW 628, May 2013*

Department of Computer Science, K.U.Leuven

## Abstract

The Schmidt-Eckart-Young theorem for matrices states that the optimal rank- $r$  approximation to a matrix is obtained by retaining the first  $r$  terms from the singular value decomposition of that matrix. This work considers a generalization of this optimal truncation property to the CANDECOMP/PARAFAC decomposition of tensors and establishes a necessary orthogonality condition. We prove that this condition is not satisfied at least by an open set of positive Lebesgue measure in complex tensor spaces. It is proved moreover that for complex tensors of small rank this condition can only be satisfied by a set of tensors of Lebesgue measure zero.

**Keywords :** Canonical Polyadic decomposition, CANDECOMP, PARAFAC, orthogonal rank decomposition, Tensor singular value decomposition, Eckart-Young theorem

**MSC :** Primary : 14A10, 14A25, 14Q15, 15A21, 15A69, 15B99.

# ON GENERIC NONEXISTENCE OF THE SCHMIDT-ECKART-YOUNG DECOMPOSITION FOR COMPLEX TENSORS

N. VANNIEUWENHOVEN<sup>†</sup>, J. NICAISE<sup>‡</sup>, R. VANDEBRIL<sup>†</sup>, AND K. MEERBERGEN<sup>†</sup>

**Abstract.** The Schmidt-Eckart-Young theorem for matrices states that the optimal rank- $r$  approximation to a matrix is obtained by retaining the first  $r$  terms from the singular value decomposition of that matrix. This work considers a generalization of this optimal truncation property to the CANDECOMP/PARAFAC decomposition of tensors and establishes a necessary orthogonality condition. We prove that this condition is not satisfied at least by an open set of positive Lebesgue measure in complex tensor spaces. It is proved moreover that for complex tensors of small rank this condition can only be satisfied by a set of tensors of Lebesgue measure zero.

**Key words.** Canonical Polyadic decomposition, CANDECOMP, PARAFAC, orthogonal rank decomposition, Tensor singular value decomposition, Eckart-Young theorem

**AMS subject classifications.** 14A10, 14A25, 14Q15, 15A21, 15A69, 15B99

**1. Introduction.** Data-sparse representations of elements living in the tensor product of (finite dimensional) vector spaces have become an intensely studied subject in recent years, yielding a myriad of representations, such as the higher-order singular value decomposition [23, 55, 56], CANDECOMP/PARAFAC decomposition (CPD) [15, 31], Block-term decompositions [22],  $\mathcal{H}$ -Tucker [28, 30] and Tensor Trains [44], each with varying assumptions and divergent applications. Among these, the CPD is the oldest; according to [11], its roots (for symmetric tensors) can be traced back to algebraic geometry in the middle of 19th century as featured in the work of Sylvester. In contemporary terminology, a tensor  $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times \dots \times n_d}$  is said to admit an  $r$ -term CPD if it can be written as

$$\mathcal{A} = \sum_{i=1}^r \mathbf{a}_i^{(1)} \otimes \mathbf{a}_i^{(2)} \otimes \dots \otimes \mathbf{a}_i^{(d)} \quad (1.1)$$

where  $\mathbf{a}_i^{(k)} \in \mathbb{F}^{n_k}$  and  $\otimes$  denotes the Kronecker product; it is the topic of this paper.

The CPD is employed in a wide variety of scientific disciplines as a tool for data-driven analysis; much of the current interest in the CPD originated in the psychometrics community from the works of Carroll and Chang [15] and Harshman [31]. This body of work became known around 1980 in the field of chemometrics where it is now well-entrenched [49]. Appellof and Davidson [4] proposed the CPD as an analytical technique for fluorescence spectroscopy; they showed that the underlying physical process exactly admits a CPD. Given the sampled tensor, its CPD reveals the excitation and emission spectra for each of the true chemical compounds in the fluorescent mixture. We refer the reader to the review articles [37, 41] for other applications within a data-driven setting.

More recently, the CPD has been used as a data-sparse representation for high-dimensional problems. For instance, Beylkin and Mohlenkamp [7] tackle the solution of linear systems whereby the solution is sought in CPD format; similarly such endeavors have been undertaken for other data-sparse representations, see [5, 38, 43]. Proper generalized decompositions methods [2, 20] seek iteratively extended CPDs for

---

<sup>†</sup>Numerical Approximation and Linear Algebra Group, Department of Computer Science, KU Leuven, Leuven, Belgium. (nick.vannieuwenhoven@cs.kuleuven.be, raf.vandebril@cs.kuleuven.be, karl.meerbergen@cs.kuleuven.be).

<sup>‡</sup>Department of Mathematics, KU Leuven, Leuven, Belgium. (johannes.nicaise@wis.kuleuven.be)

obtaining low-rank approximations to the solution of high-dimensional partial differential equations (PDE), allowing one to reduce the solution of one high-dimensional PDE to the solution of several low-dimensional PDEs.

The subject of this paper concerns a CPD that mirrors the fundamental property of the singular value decomposition of matrices: the approximation theorem of Schmidt [47, 53] and Eckart and Young [26]. Consider the singular value decomposition of  $A \in \mathbb{F}^{m \times n}$ , with  $\mathbb{F} = \mathbb{C}$  or  $\mathbb{R}$ :

$$A = USV^T = \sum_{i=1}^R \sigma_i \mathbf{u}_i \otimes \mathbf{v}_i,$$

where  $R \leq \min\{m, n\}$  is the rank of  $A$ ,  $S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_R) \in \mathbb{R}^{R \times R}$  where the diagonal elements are assumed to be nonzero and sorted by decreasing magnitude, and  $U$  and  $V$  are matrices with orthogonal columns in the Euclidean inner product. Note that we follow the convention from [29] using the transpose rather than the Hermitian conjugate so that the above definition coincides with the definitions in the case of higher-order tensors. According to the Schmidt-Eckart-Young theorem, the best rank- $r$  approximation in the Euclidean topology is given by retaining the first  $r$  terms in the above sum. We are interested in a natural generalization of the Schmidt-Eckart-Young theorem to the CPD of tensors; i.e., we ask ourselves whether there exists a CPD of  $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times \dots \times n_d}$  as in (1.1) such that retaining the first  $r' < r$  terms yields an optimal solution to the approximation problem:

$$\sum_{i=1}^{r'} \mathbf{a}_i^{(1)} \otimes \dots \otimes \mathbf{a}_i^{(d)} \in \arg \min_{\mathbf{u}_j^{(k)} \in \mathbb{F}^{n_k}} \left\| \mathcal{A} - \sum_{i=1}^{r'} \mathbf{u}_i^{(1)} \otimes \mathbf{u}_i^{(2)} \otimes \dots \otimes \mathbf{u}_i^{(d)} \right\|, \quad (1.2)$$

where the norm is the Frobenius norm, also known as the Hilbert-Schmidt and  $\ell_2$  norm. We propose to call such a decomposition a Schmidt-Eckart-Young (SEY) decomposition, and sometimes refer to it as an optimally truncatable decomposition.

The definition of the SEY decomposition is not vacuous; it is shown in this paper that orthogonally diagonalizable tensors [18, 58] satisfy the above conditions. Such tensors appear in several applications and have been extensively studied [3, 6, 14, 18, 46, 58]; nevertheless, it does not appear to be known that this decomposition is optimal in the above sense. In fact, we prove that orthogonal diagonalizability is not a necessary condition for an SEY decomposition; a new class of optimally truncatable tensors is revealed in this paper.

It appears to be a widespread conviction among researchers that an SEY decomposition is not feasible for tensors; however, general results on the (non)existence of the SEY decomposition are scarce over the *real field* and lacking over the *complex field*. In this paper, we settle the case of complex tensors, and prove that a generic tensor indeed does not admit an SEY decomposition. We cover the known results in the next paragraphs.

*Orthogonal tensor decompositions.* Kolda [35, 36] investigated several orthogonal tensor decompositions as possible candidates for a Schmidt-Eckart-Young decomposition. In [36], it was proved that an orthogonal tensor decomposition may not be optimal for a specific set of tensors over the real field, providing the first direct evidence that an SEY decomposition does not always exist. While [35] considers several orthogonal decompositions, it has, to our knowledge, never been proven that orthogonality is a necessary condition for optimal truncatability. We prove that a form of orthogonality not considered before in the literature is necessary, while another unconsidered form of orthogonality, generalizing results from [18, 35], is sufficient.

*Ill-posedness.* One popular argument for dismissing the existence of an SEY decomposition involves the ill-posedness of approximation problem (1.2), i.e., the solution may not exist. This problem arises because the set of rank- $r$  tensors is not closed, a result at least implicitly known to classical algebraic geometers, particularized by Bini *et al.* [8–10], and recently scrutinized by de Silva and Lim [24]. The openness of the set of rank- $r$  tensors implies that some rank- $r$  tensors can be approximated arbitrarily well by a tensor of rank  $r_b < r$ . The smallest of these ranks  $r_b$  is called the *border rank* of the tensor [10]. A tensor  $\mathcal{A}$  whose border rank  $r_b = \underline{\text{rank}}_{\otimes}(\mathcal{A})$  differs from its rank  $r = \text{rank}_{\otimes}(\mathcal{A})$  will be referred to as an *open boundary tensor* (OBT). Approximation problem (1.2) is then ill-posed if all solutions to

$$\min_{\text{rank}_{\otimes}(\mathcal{B}) \leq r} \|\mathcal{A} - \mathcal{B}\|$$

are OBTs. As an SEY decomposition can only exist if a solution exists, the following corollary is established:

**COROLLARY 1.1.** *An order- $d$  tensor  $\mathcal{A} \in \mathbb{F}^{n_1 \times \dots \times n_d}$  does not admit an SEY decomposition if its best rank- $r$  approximation does not exist for some  $r$ .*

Several specific examples of such tensors exist. The occurrence of OBTs was already exploited by Bini *et al.* [8–10] in 1979 to derive original fast algorithms for approximate, but arbitrarily accurate, matrix multiplication. Another example is the rank-3 tensor  $\mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u}$ , with  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$  linearly independent, which can be approximated arbitrarily well by a tensor of rank 2 [24]. Other specific examples of OBTs were investigated by Paatero [45]. In [51], Stegeman considered the occurrence of OBTs in some specific tensor spaces with two typical ranks.<sup>1</sup>

A more general result was obtained by Stegeman in [50, 52], in which it is shown that OBTs can occur in *real* tensor spaces of the form  $2 \times p \times q$  with positive volume when approximating a tensor of supergeneric<sup>2</sup> rank by one of the generic rank. Results for arbitrary third-order real tensor spaces were obtained by de Silva and Lim in 2008. They proved that, in such a space, the set of tensors not admitting an optimal rank-2 approximation has positive measure [24, Thm. 8.4]; in other words, by selecting “random” tensors, one has a nonzero probability of obtaining one without a best rank-2 approximation. de Silva and Lim’s theorem cannot be extended straightforwardly to complex tensors, as is clearly stated by its authors [24, §9]. Nevertheless, as an immediate corollary of this theorem, one obtains the following result.

**COROLLARY 1.2.** *The set of third-order tensors in  $\mathbb{R}^{n_1 \times n_2 \times n_3}$  that do not admit an SEY decomposition is nonempty and of positive Lebesgue measure.*

*Beyond ill-posedness.* The ill-posedness of approximation problem (1.2) appears to be cited often as the definite reason why an SEY decomposition cannot exist. For instance, the review article [37] states:

“de Silva and Lim [69] show, moreover, that the set of tensors of a given size that do not have a best rank- $k$  approximation has positive volume (i.e., positive Lebesgue measure) for at least some values of  $k$ , so this problem of a lack of a best approximation is not a “rare” event.”

The above statement is valid only for *third-order real tensors*; Corollary 1.2 cannot be generalized to higher orders or complex tensor spaces by using the techniques

<sup>1</sup>Due to the lack of algebraic closedness of the real field, there may be multiple ranks occurring with a positive measure in the space [24, 39].

<sup>2</sup>The generic rank of a real or complex tensor space is, by definition, the smallest  $R$  such that the Zariski closure of the set of tensors of rank  $\leq R$  considered over  $\mathbb{C}$  is the encompassing space  $\mathbb{C}^{n_1 \times \dots \times n_d}$  [39, §5.2.1]. A tensor of supergeneric, respectively subgeneric, rank is one whose rank is larger, respectively smaller, than the generic rank.

from [24, §8] because they are founded on a classification of the (finite number of) orbits of the general linear group  $\mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$ , which is, as the authors clearly state in [24, p. 1116],

“in general, not possible for tensors of arbitrary size and order simply because the dimension or “degrees of freedom” of  $\mathbb{R}^{d_1 \times \dots \times d_k}$  exceeds that of  $\mathrm{GL}_{d_1, \dots, d_k}(\mathbb{R})$  as soon as  $d_1 \cdots d_k > d_1^2 + \dots + d_k^2$  (which is almost always the case).

However, even by allowing additional parameters, a practical complete classification of the invariants and orbits of finite dimensional  $\mathbb{F}$ -algebras is unlikely to exist due to “wild” algebras, at least in the case of complex tensor spaces. Drozd’s famous Tame and Wild theorem [25] namely states that algebras can be subdivided into two sets: “tame” and “wild”. The irreducible representations of tame algebras consist of a finite set of discrete representations and a finite number of families of representations with one parameter;<sup>3</sup> wild algebras have an infinite number of irreducible representations that cannot be subdivided into one-parameter families. For example, Table 7.1 in [24] presents the 8 distinct irreducible representations of  $\mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$ . A classical example of a tame algebra admitting a finite number of one-parameter families is the Kronecker–Weierstrass–Ja’Ja’ normal form of matrix pencils [27, 34], which describes the irreducible representations of matrix pencils under a simultaneous equivalence transformation; see, e.g., [12]. An important class of wild algebras is  $\mathrm{GL}_{n_1} \times \dots \times \mathrm{GL}_{n_d}$  for  $d \geq 5$  and  $n_i \geq 2$  [42]; i.e., *all tensor spaces of order higher than 5*. A consequence of Drozd’s theorem is that classifying the irreducible representations of wild algebras entails the classification of all other algebras: classifying one wild algebra is as difficult as obtaining the irreducible representations of all algebras [48, §XIX.1]. This is considered to be infeasible by representation theorists. It is thus likely that other proof strategies<sup>4</sup> will have to be pursued to show that approximation problem (1.2) is ill-posed for orders higher than five.<sup>5</sup>

While the ill-posedness argument proves, for third-order real tensors, that a set of nonzero measure wherein no tensor admits an SEY decomposition exists, the general case is still open. All results discussed above explicitly rely on the existence of a set of tensors with positive measure that admit a supergeneric rank. This class of arguments fails for complex tensor spaces, as no such sets of positive volume exist.<sup>6</sup>

The main contribution of this paper is a proof that in every *complex* tensor space there exists a set of positive Lebesgue measure wherein no tensor admits an SEY decomposition; moreover, it is shown that for tensors of small rank this set is dense. The main result we set out to prove in this paper is:

**THEOREM 1.3.** *Let  $d \geq 3$ . The set of order- $d$  tensors in  $\mathbb{C}^{n_1 \times n_2 \times \dots \times n_d}$  that do not admit an SEY decomposition is nonempty and of positive Lebesgue measure.* Compared with Corollary 1.2, our main theorem extends the known results to complex tensor spaces of arbitrary order.

<sup>3</sup>A tame algebra can thus admit an infinite number of irreducible representations if its irreducible representations include at least one one-parameter family.

<sup>4</sup>We do not claim that an approach based on classifying irreducible representations is doomed to fail; for instance, classifying the irreducible representations of the elements of the  $r$ th order secant variety is possible for all tensor spaces, provided that  $r$  is sufficiently small [39, p. 244]; see, e.g., [13]. It is unclear whether such an approach could be employed to prove that the set of tensors for which problem (1.2) is ill-posedness for higher-order tensor spaces over an arbitrary field has positive volume.

<sup>5</sup> $\mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$  is a tame algebra with an infinite number of irreducible representations; in principle, the technique from [24] could be applied.

<sup>6</sup>The validity of this statement can be assessed by replacing the Veronese by the Segre variety in Corollary 6.11 of [21]; alternatively, see [39, p. 69].

The outline of this paper is as follows. In the next section, some terminology and notation is fixed. Then, in section 3, the SEY decomposition is formally proposed. Both a necessary and a sufficient condition for its existence are investigated. In section 4, the main result is proved. Finally, we present conclusions in section 5.

**2. Preliminaries.** Throughout this paper, the symbol  $\mathbb{F}$  denotes either the real field  $\mathbb{R}$  or the complex field  $\mathbb{C}$ . Tensors are typeset in calligraphic upper-case letters ( $\mathcal{A}$ ,  $\mathcal{B}$ ), matrices in upper-case letters ( $A$ ), vectors in boldface lower-case letters ( $\mathbf{a}$ ,  $\mathbf{u}$ ,  $\mathbf{v}$ ) and scalars in lower-case letters ( $a$ ,  $\sigma$ ,  $\lambda$ ). The scalar  $d$  is reserved for the order of a tensor. The Euclidean inner product is denoted by  $\langle \cdot, \cdot \rangle$ , the Euclidean norm by  $\|\cdot\|$ , and the tensor product by  $\otimes$ . The complex unity will be denoted by  $\iota$ .

*Multilinear algebra.* A tensor of order  $d$  is an element of the tensor product of  $d$  vector spaces:  $\mathcal{A} \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \dots \otimes \mathbb{F}^{n_d}$ . Defining the standard tensor basis of order  $d$  as  $\{\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \dots \otimes \mathbf{e}_{i_d}\}_{i_1, i_2, \dots, i_d=1}^{n_1, n_2, \dots, n_d}$ , where  $\mathbf{e}_{i_k}$  is the  $i_k$ th standard basis vector of  $\mathbb{F}^{n_k}$ , we can represent

$$\mathcal{A} = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_d=1}^{n_d} a_{i_1, i_2, \dots, i_d} \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \dots \otimes \mathbf{e}_{i_d},$$

with respect to the standard tensor basis as the  $d$ -array

$$[a_{i_1, i_2, \dots, i_d}]_{i_1, i_2, \dots, i_d=1}^{n_1, n_2, \dots, n_d} \in \mathbb{F}^{n_1 \times n_2 \times \dots \times n_d}.$$

The Euclidean inner product of two tensors  $\mathcal{A}, \mathcal{B} \in \mathbb{F}^{n_1 \times \dots \times n_d}$  can then be defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle := \sum_{i_1=1}^{n_1} \dots \sum_{i_d=1}^{n_d} a_{i_1, \dots, i_d} \bar{b}_{i_1, \dots, i_d}, \quad (2.1)$$

with  $\bar{x}$  the complex conjugate of  $x$ . The corresponding Euclidean norm is  $\|\mathcal{A}\| := \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$ . Note that this definition coincides with the Euclidean norm of  $\mathcal{A}$  when considered as an element over  $\mathbb{F}^{n_1 \times \dots \times n_d}$ . If

$$\mathcal{A} = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)} \otimes \dots \otimes \mathbf{a}^{(d)} \quad \text{and} \quad \mathcal{B} = \mathbf{b}^{(1)} \otimes \mathbf{b}^{(2)} \otimes \dots \otimes \mathbf{b}^{(d)},$$

then, the inner product takes a special form:

$$\langle \mathcal{A}, \mathcal{B} \rangle = \langle \mathbf{a}^{(1)}, \mathbf{b}^{(1)} \rangle \dots \langle \mathbf{a}^{(d)}, \mathbf{b}^{(d)} \rangle,$$

see, e.g., [29, Section 4.5.1]. Two tensors  $\mathcal{A}$  and  $\mathcal{B}$  are orthogonal,  $\mathcal{A} \perp \mathcal{B}$ , iff  $\langle \mathcal{A}, \mathcal{B} \rangle = 0$ . Two simple tensors  $\mathcal{A}$  and  $\mathcal{B}$  are thus orthogonal iff there is at least one mode  $k$  wherein the mode- $k$  vectors are orthogonal,  $\mathbf{a}^{(k)} \perp \mathbf{b}^{(k)}$ .

*Algebraic geometry.* We say that the tensor  $\mathcal{A}$  is *simple* if it can be written as

$$\mathcal{A} = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)} \otimes \dots \otimes \mathbf{a}^{(d)}, \quad \text{with} \quad \mathbf{a}^{(k)} \in \mathbb{F}^{n_k}.$$

$\mathcal{A}$  is of *rank*  $r$  if it can be written as

$$\mathcal{A} = \sum_{i=1}^r \lambda_i \mathbf{a}_i^{(1)} \otimes \mathbf{a}_i^{(2)} \otimes \dots \otimes \mathbf{a}_i^{(d)}, \quad \text{with} \quad \mathbf{a}_i^{(k)} \in \mathbb{F}^{n_k} \quad \text{and} \quad \lambda_i \in \mathbb{F}, \quad (2.2)$$

but not with fewer than  $r$  terms. This rank will be denoted by  $\text{rank}_{\otimes}(\mathcal{A})$ .

It is often useful to consider tensors up to scalar multiplication, i.e., as  $\mathbb{F}$ -rational points on the projective space

$$P = \mathbb{P}(\mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \dots \otimes \mathbb{F}^{n_d}).$$

Then, the simple tensors correspond precisely to the  $\mathbb{F}$ -rational points of the Segre variety

$$\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{F}} = \mathbb{P}\mathbb{F}^{n_1} \times \mathbb{P}\mathbb{F}^{n_2} \times \dots \times \mathbb{P}\mathbb{F}^{n_d} \subset P.$$

If  $\mathbb{F} = \mathbb{C}$ , the tensors of border rank at most  $r$  are described by points on the  $r$ th secant variety  $\sigma_r(\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}})$  of the Segre variety  $\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}}$  [39]. Recall that the secant variety of a subvariety  $\mathcal{S}$  of some projective space  $\mathbb{P}\mathbb{C}^N$  is defined as the Zariski closure of the union of the linear spans of  $r$  points on the variety  $\mathcal{S}$  [39]:

$$\sigma_r(\mathcal{S}) := \overline{\bigcup_{p_1, \dots, p_r \in \mathcal{S}} \text{span}(p_1, p_2, \dots, p_r)}.$$

A  $\mathbb{C}$ -rational point  $\mathcal{A}$  of  $P$  is of border rank  $r$  if it belongs to  $\sigma_r(\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}})$ , but not to  $\sigma_{r-1}(\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}})$ . By construction, the tensors of rank  $r$  form a Zariski-dense constructible subset of the  $r$ th secant variety; that is, rank- $r$  tensors are generic within the set of tensors of border rank  $r$ .

For more information about the connection between algebraic geometry and multilinear algebra, see Landsberg's book [39].

**3. Schmidt-Eckart-Young decomposition.** We consider a natural generalization of the Schmidt-Eckart-Young theorem to higher-order tensors.

**DEFINITION 3.1.** *An order- $d$  tensor  $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times \dots \times n_d}$  admits a Schmidt-Eckart-Young decomposition iff it can be written as a linear combination of simple tensors:*

$$\mathcal{A} = \sum_{i=1}^R \sigma_i \mathcal{A}_i = \sum_{i=1}^R \sigma_i \mathbf{a}_i^{(1)} \otimes \mathbf{a}_i^{(2)} \otimes \dots \otimes \mathbf{a}_i^{(d)}, \quad (3.1)$$

with  $R = \text{rank}_{\otimes}(\mathcal{A})$ ,  $\sigma_i \in \mathbb{F}$ ,  $\mathbf{a}_i^{(k)} \in \mathbb{F}^{n_k}$  and  $\|\mathbf{a}_i^{(k)}\| = 1$ , and such that truncating the decomposition is optimal for all  $r = 1, 2, \dots, R$ :

$$\sum_{i=1}^r \sigma_i \mathcal{A}_i \in \arg \min_{\text{rank}_{\otimes}(\mathcal{B})=r} \|\mathcal{A} - \mathcal{B}\|. \quad (3.2)$$

Note that there may be multiple optima for any given rank. We say that a point in  $\mathbb{P}\mathbb{F}^{n_1 \times n_2 \times \dots \times n_d}$  has an SEY decomposition if it has a tuple of homogeneous coordinates in  $\mathbb{F}^{n_1 \times n_2 \times \dots \times n_d}$  with an SEY decomposition.

The definition imposes no orthogonality constraints, and the  $\sigma_i$ 's are not sorted by decreasing magnitude. It will nonetheless be shown in Corollary 3.4 that the singular value decomposition is the only SEY decomposition for second-order tensors.

**3.1. A necessary condition.** We begin by establishing a necessary condition for admitting an SEY decomposition, and claim that they are weak two-orthogonal:

**DEFINITION 3.2.** *A weak two-orthogonal decomposition is an orthogonal decomposition of an order- $d$  tensor  $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times \dots \times n_d}$  in simple tensors:*

$$\mathcal{A} = \sum_{i=1}^R \sigma_i \mathbf{u}_i^{(1)} \otimes \mathbf{u}_i^{(2)} \otimes \dots \otimes \mathbf{u}_i^{(d)}, \quad \text{with } \mathbf{u}_i^{(k)} \in \mathbb{F}^{n_k}, \|\mathbf{u}_i^{(k)}\| = 1 \text{ and } \sigma_i \in \mathbb{F},$$

and all terms are pairwise orthogonal in at least two modes:

$$\forall 1 \leq i < j \leq R : \exists 1 \leq k_1 < k_2 \leq d : \mathbf{u}_i^{(k_1)} \perp \mathbf{u}_j^{(k_1)} \quad \text{and} \quad \mathbf{u}_i^{(k_2)} \perp \mathbf{u}_j^{(k_2)}. \quad (3.3)$$



**THEOREM 3.3.** *Let  $\mathcal{A}$  admit an SEY decomposition as in Definition 3.1, then it is a weak two-orthogonal decomposition.*

*Proof.* We prove the assertion by contradiction. Assume that an SEY decomposition is not weak two-orthogonal. Then, there exists a maximal tuple  $(i^*, j^*) \in \{1, \dots, R\} \times \{1, \dots, R\}$  with  $i^* < j^*$  such that  $\mathcal{A}_{i^*}$  and  $\mathcal{A}_{j^*}$  are not two-orthogonal. The maximality of this tuple should be interpreted with respect to the following complete order

$$(i_1, j_1) < (i_2, j_2) \quad \text{iff} \quad (i_1 < i_2) \text{ or } ((i_1 = i_2) \text{ and } (j_1 < j_2)).$$

We distinguish between two cases. Either  $\mathcal{A}_{i^*}$  and  $\mathcal{A}_{j^*}$  are orthogonal in only one mode, or they are not orthogonal.

Consider first the case in which  $\mathcal{A}_{i^*}$  and  $\mathcal{A}_{j^*}$  are not orthogonal. Let  $1 \leq k \leq d$  be any mode, and consider the following simple tensor

$$\epsilon \mathcal{P}_{i^* j^*}^{(k)} := \mathbf{a}_{i^*}^{(1)} \otimes \dots \otimes \mathbf{a}_{i^*}^{(k-1)} \otimes \epsilon \mathbf{a}_{j^*}^{(k)} \otimes \mathbf{a}_{i^*}^{(k+1)} \otimes \dots \otimes \mathbf{a}_{i^*}^{(d)}, \quad \text{with } \epsilon \in \mathbb{R},$$

which has the property that it only perturbs the  $k$ th mode of the  $i^*$ th term in the SEY decomposition of  $\mathcal{A}$  in the direction of  $\mathbf{a}_{j^*}^{(k)}$ . That is,

$$\mathbf{a}_{i^*}^{(1)} \otimes \dots \otimes \mathbf{a}_{i^*}^{(d)} + \epsilon \mathcal{P}_{i^* j^*}^{(k)} = \mathbf{a}_{i^*}^{(1)} \otimes \dots \otimes \mathbf{a}_{i^*}^{(k-1)} \otimes \left( \mathbf{a}_{i^*}^{(k)} + \epsilon \mathbf{a}_{j^*}^{(k)} \right) \otimes \mathbf{a}_{i^*}^{(k+1)} \otimes \dots \otimes \mathbf{a}_{i^*}^{(d)}.$$

Consequently, adding this perturbation to  $\sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j$  does not increase its rank. Because  $\mathcal{A}$  admits an SEY decomposition, this perturbed sum does not improve the approximation error:

$$\left\| \mathcal{A} - \sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j - \epsilon \mathcal{P}_{i^* j^*}^{(k)} \right\|^2 \geq \left\| \mathcal{A} - \sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j \right\|^2 = \eta^2. \quad (3.4)$$

However, by expanding the norm on the left-hand side and observing that the norm of  $\epsilon \mathcal{P}_{i^* j^*}^{(k)}$  is  $\epsilon^2$ , we obtain

$$\eta^2 - 2\Re \left\langle \mathcal{A} - \sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j, \epsilon \mathcal{P}_{i^* j^*}^{(k)} \right\rangle + \epsilon^2 \geq \eta^2,$$

which after exploiting linearity and reordering the terms becomes

$$\epsilon^2 \geq 2\Re \left( \sum_{j=j^*}^R \sigma_j \langle \mathcal{A}_j, \epsilon \mathcal{P}_{i^* j^*}^{(k)} \rangle \right) = 2\Re \left( \epsilon \sum_{j=j^*}^R \sigma_j \langle \mathbf{a}_j^{(k)}, \mathbf{a}_{j^*}^{(k)} \rangle \prod_{\substack{m=1 \\ m \neq k}}^d \langle \mathbf{a}_j^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle \right), \quad (3.5)$$

where  $\Re(\alpha)$  denotes the real part of  $\alpha$ . From the maximality of  $(i^*, j^*)$ , it follows that if  $j > j^*$  then  $\mathcal{A}_j$  is orthogonal, in at least two modes, to  $\mathcal{A}_{i^*}$ . Otherwise, there would be a  $j' > j^*$  such that  $\mathcal{A}_{j'}$  is not two-orthogonal to  $\mathcal{A}_{i^*}$ , yielding an immediate contradiction to the maximality of  $(i^*, j^*)$ . Consequently,  $\prod_{m=1, m \neq k}^d \langle \mathbf{a}_j^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle = 0$  if  $j \neq j^*$ , because only one mode is excluded from this product, so that  $\mathcal{A}_{i^*}$  is still orthogonal in at least one other mode to  $\mathcal{A}_j$ . Therefore, (3.5) reduces to

$$\epsilon^2 \geq 2\epsilon \cdot \Re \left( \sigma_{j^*} \prod_{\substack{m=1 \\ m \neq k}}^d \langle \mathbf{a}_{j^*}^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle \right). \quad (3.6)$$

If  $\mathbb{F} = \mathbb{C}$ , an additional equation is required to construct our contradiction; to this end, consider (3.4) again, but now apply the permutation  $\iota \mathcal{P}_{i^* j^*}^{(k)}$ , with  $\epsilon \in \mathbb{R}$ , and expand the norm to obtain (3.5). However,  $\epsilon$  should now be replaced with  $-\iota\epsilon$  on the right-hand side. Applying the induction hypothesis as before, we obtain finally

$$\epsilon^2 \geq 2\epsilon \cdot \Im \left( \sigma_{j^*} \prod_{\substack{m=1 \\ m \neq k}}^d \langle \mathbf{a}_{j^*}^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle \right), \quad (3.7)$$

where we have used the identity  $\Re(-\iota\alpha) = \Im(\alpha)$ , and  $\Im(\alpha)$  denotes the imaginary part of  $\alpha$ . If  $\mathbb{F} = \mathbb{R}$ , (3.7) is not required. If the real part, respectively imaginary part, on the right-hand side of (3.6), respectively (3.7), is nonzero, a contradiction to the optimality of the SEY decomposition can be constructed by choosing  $\epsilon$  sufficiently small. Therefore,

$$\sigma_{j^*} \prod_{\substack{m=1 \\ m \neq k}}^d \langle \mathbf{a}_{j^*}^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle = 0.$$

As  $\sigma_{j^*} \neq 0$  because that would contradict  $R = \text{rank}_{\otimes}(\mathcal{A})$ , there should be at least one  $m \neq k$  such that  $\mathbf{a}_{j^*}^{(m)} \perp \mathbf{a}_{i^*}^{(m)}$ .

Consider now the second case;  $\mathcal{A}_{i^*}$  and  $\mathcal{A}_{j^*}$  are orthogonal in one mode, say  $k$ . Repeat the argument for the previous case for mode  $k$ , which is now fixed. Then, we find that this case is also contradictory. Consequently, the assumption of the existence of such a maximal tuple  $(i^*, j^*)$  must be false. This concludes the proof.  $\square$

As an immediate consequence, we obtain

**COROLLARY 3.4.** *The singular value decomposition is the unique SEY decomposition of a second order tensor  $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2}$ .*

*Proof.* In the case of two modes, it is clear that Theorem 3.3 reduces to requiring

$$\mathcal{A} = USV^T \quad \text{with } U \in \mathbb{F}^{n_1 \times R}, V \in \mathbb{F}^{n_2 \times R}, \text{ and } S \in \mathbb{F}^{R \times R}. \quad (3.8)$$

Herein,  $U$  and  $V$  have orthonormal columns, and  $S$  is a diagonal matrix. This is the compact singular value decomposition.  $\square$

Weak two-orthogonality is not a sufficient condition, because it does not exclude the OBT  $\mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u}$  with  $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$  and  $\mathbf{u} \perp \mathbf{v}$  [24].

**3.2. A sufficient condition.** The Tensor SVD, or completely orthogonal CPD, [18, 35, 58] appears in various applications related to blind-source separation [6, 14, 46]. We show that such a decomposition is a special case of the strong two-orthogonal decomposition, which is proved to be an SEY decomposition.

**DEFINITION 3.5.** *A rank- $R$  tensor  $\mathcal{A} \in \mathbb{F}^{n_1 \times \dots \times n_d}$  admits a strong two-orthogonal decomposition of rank  $R$  with splitting point  $s \in \{1, 2, \dots, d\}$  if it can be written as*

$$\mathcal{A} = \sum_{i=1}^R \sigma_i \underbrace{\mathbf{u}_i^{(1)} \otimes \dots \otimes \mathbf{u}_i^{(s)}}_{\mathcal{U}_i^s} \otimes \underbrace{\mathbf{v}_i^{(s+1)} \otimes \dots \otimes \mathbf{v}_i^{(d)}}_{\mathcal{V}_i^s} = \sum_{i=1}^R \sigma_i \mathcal{U}_i^s \otimes \mathcal{V}_i^s, \quad (3.9)$$

with  $\mathbf{u}_i^{(k)} \in \mathbb{F}^{n_k}$ ,  $\mathbf{v}_i^{(k)} \in \mathbb{F}^{n_k}$ ,  $\sigma_i \in \mathbb{R}$ ,  $\|\mathbf{u}_i^{(k)}\| = 1$ ,  $\|\mathbf{v}_i^{(k)}\| = 1$ , and

$$\forall 1 \leq i < j \leq R: \quad \mathcal{U}_i^s \perp \mathcal{U}_j^s \quad \text{and} \quad \mathcal{V}_i^s \perp \mathcal{V}_j^s.$$

The coefficients  $\sigma_i$  are assumed to be sorted:  $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_R^2 > 0$ .

The fact that the  $\sigma_i$  are required to be real does not limit generality; the coefficients can always be chosen to be the norm of  $\sigma_i \mathcal{U}_i^s \otimes \mathcal{V}_i^s$ , which is real. The partitioning of the nodes is also not restricted to consecutive modes because we may arbitrarily renumber the modes; however, for simplicity, we assume a partitioning with consecutive mode numbers in the remainder.

From the above definition, it follows that a strong two-orthogonal decomposition is a weak two-orthogonal decomposition with an additional restriction on the choice of  $k_1$  and  $k_2$  in Definition 3.2; it is, in addition, required that  $1 \leq k_1 \leq s < k_2 \leq d$ . This turns out to be sufficient for obtaining optimality.

**THEOREM 3.6.** *A rank- $R$  strong two-orthogonal decomposition of a rank- $R$  tensor is an SEY decomposition.*

*Proof.* A tensor space  $\mathbb{T} := \mathbb{F}^{n_1} \otimes \dots \otimes \mathbb{F}^{n_s} \otimes \mathbb{F}^{n_{s+1}} \otimes \dots \otimes \mathbb{F}^{n_d}$  is isomorphic, as a vector space, thus ignoring the tensor structure [29], to  $\mathbb{T}_s := \mathbb{F}^{n_1 \dots n_s} \otimes \mathbb{F}^{n_{s+1} \dots n_d}$ . By definition, a simple tensor  $\mathbf{a}^{(1)} \otimes \dots \otimes \mathbf{a}^{(s)} \otimes \mathbf{a}^{(s+1)} \otimes \dots \otimes \mathbf{a}^{(d)} \in \mathbb{T}$  becomes the rank-1 matrix

$$(\mathbf{a}^{(1)} \otimes \dots \otimes \mathbf{a}^{(s)}) \otimes (\mathbf{a}^{(s+1)} \otimes \dots \otimes \mathbf{a}^{(d)}) \in \mathbb{T}_s \cong \mathbb{T},$$

where the products inside the brackets should now be interpreted as Kronecker products. Using multilinearity, it follows that a tensor  $\mathcal{A}$  as in Definition 3.5 with splitting point  $s$  admits the decomposition

$$\mathcal{A} = \sum_{i=1}^R \sigma_i \mathcal{U}_i^s \otimes \mathcal{V}_i^s,$$

which upon closer inspection is *the singular value decomposition* of  $\mathcal{A}$  considered as matrix in  $\mathbb{T}_s$ . Observe that any rank- $r$  CPD over  $\mathbb{T}$  is a matrix of rank at most  $r$  over  $\mathbb{T}_s$ . As the singular value decomposition provides an optimal approximation of rank  $r$  for the matrix  $\mathcal{A} \in \mathbb{T}_s$  it follows that no rank- $r$  CPD over  $\mathbb{T}$  can be a strictly better approximation than the provided strong two-orthogonal decomposition; otherwise, the optimality of the matrix singular value decomposition would be contradicted. Finally, considering a limit of a sum of  $r$  simple tensors over  $\mathbb{T}$  cannot improve the singular value decomposition, because  $\mathcal{A} \in \mathbb{T}_s$ , which is an order-2 tensor product for which it is known that the set of rank- $r$  tensors is closed; hence, limits do not extend the set over which the optimization is defined.  $\square$

The Tensor SVD [18, 58] is a CPD where orthogonality is imposed in every mode; hence, it is an SEY decomposition.

It is an open question whether strong two-orthogonality is also necessary.

**4. Generic nonexistence.** In this section, the prime result is presented, which states that a set of nonzero Lebesgue measure exists in  $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \dots \times n_d}$ ,  $d \geq 3$ , such that its elements do *not* exhibit an SEY decomposition. Note that we restrict ourselves to the complex case  $\mathbb{F} = \mathbb{C}$  in this section; however, recall from the introduction that this is the interesting case as no general results are known. The results in this section apply for  $d \geq 3$ ; matrices admit an SEY decomposition as per Corollary 3.4.

Throughout this section, we consider secants of the Segre variety  $\mathcal{S}_{n_1, \dots, n_d}^{\mathbb{C}}$  embedded in  $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \dots \times n_d}$ . The set of tensors that admit a rank- $r$  weak two-orthogonal decomposition, write  $\sigma_r^{\perp}$ , is a subset of the set of tensors of rank at most  $r$ , write  $\sigma_r'$ , which, in turn, is a subset of the tensors of border rank at most  $r$ , i.e., the  $r$ th order secant variety  $\sigma_r$  of the Segre variety  $\mathcal{S}_{n_1, \dots, n_d}^{\mathbb{C}}$ . Furthermore, a generic tensor of border rank  $r$  also has rank  $r$  over  $\mathbb{C}$ ; that is,  $\sigma_r'$  is dense in  $\sigma_r$  [39].

Before proceeding with our proof strategy, we relate an interesting alternative strategy communicated to us by G. Ottaviani. From [19] we know that generic rank- $r$  tensors in  $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$  are identifiable, at least if  $r$  is sufficiently small. This entails that the points on the Segre variety  $\mathcal{S}_{n_1, \dots, n_d}^{\mathbb{C}}$  are uniquely determined, and in general configuration. However, it is intuitively clear that the necessary condition we derived in the previous section imposes a certain configuration on the points, hereby contradicting generality. This immediately entails that an SEY decomposition does not exist for such generic rank- $r$  tensors. The strategy we follow in this paper can result in stronger statements, as we are able to use the conditions for nondefectivity of Segre varieties, rather than the more restricted conditions for identifiability. In particular, we can demonstrate that a generic tensor in a cubic tensor space  $\mathbb{C}^{n \times \cdots \times n}$  ( $d$  times), with  $n$  sufficiently large, does not admit an SEY decomposition; a result that cannot be derived from the strategy sketched in this paragraph. It is an important observation worth stressing, however: *optimal truncation and identifiability cannot occur simultaneously in generic rank- $r$  tensors*.

As  $\sigma_r$  is an algebraic variety, the natural setup for comparing the “size” of  $\sigma_r$  and  $\sigma_r^\perp$  is by comparing their dimensions as algebraic varieties. For nonalgebraic sets, such as  $\sigma_r^\perp$ , the dimension is defined as the dimension of the smallest enclosing algebraic set, i.e., the closure in the Zariski topology. So, to show that  $\sigma_r^\perp$  is substantially smaller than  $\sigma_r$ , we would need to show that the dimension of the Zariski closure of  $\sigma_r^\perp$  in  $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$  is strictly smaller than  $\dim(\sigma_r)$ . Unfortunately, it is difficult to control the dimension of this Zariski closure because  $\sigma_r^\perp$  is not an algebraic subset of  $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$  due to the complex conjugation that appears in the definition of the inner product. However, we may still investigate the underlying real algebraic structure of  $\sigma_r^\perp$  by applying a Weil restriction of scalars from  $\mathbb{C}$  to  $\mathbb{R}$ ; this simply means that we write each complex coordinate  $x$  as  $x = u + iv$  and consider  $u$  and  $v$  as new coordinates over  $\mathbb{R}$ . Let  $\prod_{\mathbb{C}/\mathbb{R}}$  denote the Weil restriction functor, then  $\sigma_r^\perp$  can be considered as a Zariski-closed subset of the real algebraic manifold

$$\left(\prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}\right)(\mathbb{R}) \cong \mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d};$$

herein,  $(A)(\mathbb{R})$  denotes the set of real points of the complex algebraic variety  $A$ , and  $\cong$  denotes the canonical isomorphism. The (complex) variety obtained by applying the Weil restriction to  $\sigma_r^\perp$  is the complexification of a real algebraic variety isomorphic to the real algebraic structure of  $\sigma_r^\perp$  [33]. Considering then the set of real points of this Weil restriction yields a real algebraic variety over the reals admitting, in addition to the Zariski topology, the familiar Euclidean topology. We will exploit this underlying Euclidean structure to prove our main theorem.

We begin by investigating the dimension of  $\sigma_r^\perp$ . The precise definition of this dimension will become clear from the well-defined construction of the auxiliary variety  $\mathcal{J}$  in the proof of the next proposition.

PROPOSITION 4.1. *We have*

$$\dim(\sigma_r^\perp) \leq 2r \sum_{\ell=1}^d (n_\ell - 1) + 2\delta - 2,$$

with  $\delta = r - 2\lfloor r/2 \rfloor$ .

*Proof.* We will need an adapted version of the join operation. Let  $V_1, \dots, V_s$  be subvarieties of  $\prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N$ , for some positive integer  $N$ . Applying the Weil restriction functor to the projection morphism

$$\mathbb{A}_{\mathbb{C}}^N \setminus \{(0, \dots, 0)\} \rightarrow \mathbb{P}\mathbb{C}^N,$$

where  $\mathbb{A}_{\mathbb{C}}^N$  denotes the  $N$ -dimensional affine space over  $\mathbb{C}$ , yields a morphism of  $\mathbb{R}$ -varieties

$$\pi : \prod_{\mathbb{C}/\mathbb{R}} (\mathbb{A}_{\mathbb{C}}^N \setminus \{(0, \dots, 0)\}) \rightarrow \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N$$

whose source is canonically isomorphic to  $\mathbb{A}_{\mathbb{R}}^{2N} \setminus \{(0, \dots, 0)\}$ . We define the affine cone of  $V_i$  as

$$\tilde{V}_i := \pi^{-1}(V_i) \cup \{(0, \dots, 0)\},$$

for each  $i$  in  $\{1, \dots, s\}$ . This is a closed subvariety of  $\mathbb{A}_{\mathbb{R}}^{2N}$ . Now the Zariski closure of the image of  $\tilde{V}_1 \times_{\mathbb{R}} \dots \times_{\mathbb{R}} \tilde{V}_s$  under the addition morphism

$$(\mathbb{A}_{\mathbb{R}}^{2N})^s \rightarrow \mathbb{A}_{\mathbb{R}}^{2N} : (v_1, \dots, v_s) \mapsto v_1 + \dots + v_s$$

is a union of the origin and fibers<sup>7</sup> of the projection morphism  $\pi$ . Removing the origin and taking the image under  $\pi$ , we obtain a closed subvariety of  $\prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N$  that we call the join of  $V_1, \dots, V_s$  and denote by  $J(V_1, \dots, V_s)$ . The fibers of  $\pi$  are real planes, so that the dimension of  $J(V_1, \dots, V_s)$  is at most

$$\dim(V_1) + \dots + \dim(V_s) + 2(s - 1).$$

The simple tensors with complex homogeneous coordinates

$$\left[ u_1^{(1)} + w_1^{(1)} \quad \dots \quad u_{n_1}^{(1)} + w_{n_1}^{(1)} \right]^T \otimes \dots \otimes \left[ u_1^{(d)} + w_1^{(d)} \quad \dots \quad u_{n_d}^{(d)} + w_{n_d}^{(d)} \right]^T$$

correspond to the real points of the  $\mathbb{R}$ -variety

$$V = \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^{n_1} \times_{\mathbb{R}} \dots \times_{\mathbb{R}} \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^{n_d}.$$

For all  $i, j$  in  $\{1, \dots, d\}$  with  $i \neq j$ , the pairs of simple tensors that are orthogonal in modes  $i$  and  $j$  correspond to the real points of the subvariety  $W_{ij}$  of  $V \times_{\mathbb{R}} V$  defined by the equations

$$\left\{ \begin{array}{l} \sum_{\ell=1}^{n_i} (u_{\ell}^{(i)} \dot{u}_{\ell}^{(i)} + v_{\ell}^{(i)} \dot{v}_{\ell}^{(i)}) = 0 \\ \sum_{\ell=1}^{n_i} (u_{\ell}^{(i)} \dot{v}_{\ell}^{(i)} - \dot{u}_{\ell}^{(i)} v_{\ell}^{(i)}) = 0 \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} \sum_{\ell=1}^{n_j} (u_{\ell}^{(j)} \dot{u}_{\ell}^{(j)} + v_{\ell}^{(j)} \dot{v}_{\ell}^{(j)}) = 0 \\ \sum_{\ell=1}^{n_j} (u_{\ell}^{(j)} \dot{v}_{\ell}^{(j)} - \dot{u}_{\ell}^{(j)} v_{\ell}^{(j)}) = 0 \end{array} \right. ,$$

where we used the coordinates  $\dot{u}$  and  $\dot{v}$  for points on the second factor of the product  $V \times_{\mathbb{R}} V$ . The variety  $V \times_{\mathbb{R}} V$  is irreducible of dimension  $4(n_1 - 1) + \dots + 4(n_d - 1)$  and we claim that  $W_{ij}$  has codimension 4. This is easy, but somewhat tedious, to check by covering the projective spaces  $\mathbb{P}\mathbb{C}^{n_a}$  by their standard affine charts, which gives rise to an open covering of  $V \times_{\mathbb{R}} V$  by affine spaces.

We define  $W$  as the union of all the varieties  $W_{ij}$  with  $i \neq j$  inside  $V \times_{\mathbb{R}} V$ . This is again a variety of dimension  $4(n_1 - 1) + \dots + 4(n_d - 1) - 4$ , whose  $\mathbb{R}$ -rational points correspond to pairs of simple tensors with complex homogeneous coordinates that are orthogonal in at least two modes.

<sup>7</sup>Informally, the fiber in  $p \in Y$  of a morphism of varieties  $\kappa : X \rightarrow Y$  is the preimage of the map in  $p$ . Consider  $\pi : \mathbb{A}_{\mathbb{C}}^N \rightarrow \mathbb{P}\mathbb{C}^N$ , and let  $p' \in \mathbb{A}_{\mathbb{C}}^N$  and  $p = \pi(p') \in \mathbb{P}\mathbb{C}^N$ , then the fiber in  $p$  consists of all vectors  $v \in \mathbb{A}_{\mathbb{C}}^N$  such that  $p'$  and  $v$  are linearly dependent. For a precise scheme-theoretic definition, we refer to [32].

Set  $N = n_1 \cdots n_d$ . Applying the Weil restriction functor to the Segre embedding, we get a closed immersion of  $\mathbb{R}$ -varieties

$$V \times_{\mathbb{R}} V \rightarrow \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N \times_{\mathbb{R}} \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N.$$

Consider the product of projections

$$(\pi, \pi) : (\mathbb{A}_{\mathbb{R}}^{2N} \setminus \{(0, \dots, 0)\}) \times_{\mathbb{R}} (\mathbb{A}_{\mathbb{R}}^{2N} \setminus \{(0, \dots, 0)\}) \rightarrow \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N \times_{\mathbb{R}} \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N.$$

Similarly to the construction of the join, we consider the Zariski closure of the image of  $(\pi, \pi)^{-1}(W)$  under the addition morphism

$$\mathbb{A}_{\mathbb{R}}^{2N} \times_{\mathbb{R}} \mathbb{A}_{\mathbb{R}}^{2N} \rightarrow \mathbb{A}_{\mathbb{R}}^{2N} : (v_1, v_2) \mapsto v_1 + v_2,$$

remove the origin, and take the image under  $\pi$ . The result of these operations is a closed subvariety  $X$  of  $\prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N$  of dimension at most

$$4(n_1 - 1) + \dots + 4(n_d - 1) - 2.$$

By construction, the set of real points  $(X)(\mathbb{R})$  corresponds to a subset of  $\mathbb{P}\mathbb{C}^N$  that contains all rank-2 tensors admitting a weak two-orthogonal decomposition.

Now we write  $r$  as  $2r_0 + \delta$ , with  $r_0$  a nonnegative integer and  $\delta$  an element in  $\{0, 1\}$ . We set

$$\mathcal{J} = J(\underbrace{X, \dots, X}_{r_0 \text{ copies}}, V^\delta) \subset \prod_{\mathbb{C}/\mathbb{R}} \mathbb{P}\mathbb{C}^N,$$

where the notation  $V^\delta$  means that we include the variety  $V$  only in the case where  $\delta = 1$ . The join variety  $\mathcal{J}$  has dimension at most

$$\begin{aligned} & r_0 \dim(X) + \delta \dim(V) + 2(r_0 + \delta - 1) \\ & \leq r_0(4(n_1 - 1) + \dots + 4(n_d - 1) - 2) + \delta(2(n_1 - 1) + \dots + 2(n_d - 1)) + 2(r_0 + \delta - 1) \\ & = 2r \sum_{\ell=1}^d (n_\ell - 1) + 2\delta - 2, \end{aligned}$$

and the set of real points  $\mathcal{J}(\mathbb{R})$  corresponds to a subset of  $\mathbb{P}\mathbb{C}^N$  that contains the set  $\sigma_r^\perp(\mathcal{S})$  of rank  $r$  tensors admitting a weak two-orthogonal decomposition.  $\square$

The established upper bound may be coarse, because the construction of  $X$  and  $\mathcal{J}$  in the above proof only takes the weak two-orthogonality into account for the successive rank-1 terms  $2k$  and  $2k+1$ , for all  $k$ , in a weak two-orthogonal decomposition. Nevertheless, this bound is sufficient for proving the main theorem.

With the understanding of the dimension of  $\sigma_r^\perp$  in place, we can now understand

LEMMA 4.2. *If*

$$\dim(\sigma_r^\perp) < \dim\left(\prod_{\mathbb{C}/\mathbb{R}} \sigma_r\right) = 2 \dim(\sigma_r),$$

*then the set of rank- $r$  tensors not admitting a rank- $r$  weak two-orthogonal decomposition, and, consequently, not admitting an SEY decomposition, i.e.,*

$$\left(\prod_{\mathbb{C}/\mathbb{R}} \sigma_r \setminus \sigma_r^\perp\right)(\mathbb{R}),$$

is a Zariski-open subset, and, hence, dense open subset in the Euclidean topology, of

$$\left(\prod_{\mathbb{C}/\mathbb{R}} \sigma_r\right)(\mathbb{R}) \cong (\sigma_r)(\mathbb{C}).$$

If the above lemma applies, we will say that a *generic* rank- $r$  tensor has no SEY decomposition. Beware that the term “generic” refers to the algebraic structure on the *real* variety  $\prod_{\mathbb{C}/\mathbb{R}} \sigma_r$ ; this does not necessarily imply the existence of a Zariski-dense open subset of the *complex* variety  $\sigma_r$  whose points do not have an SEY decomposition.

To apply Lemma 4.2, we still need a lower bound on the dimension of  $\sigma_r$ . The dimensions of this variety have been studied for over a century now, but, unfortunately, they still elude the scientific community. The *expected dimension* of  $\sigma_r$  is well-known:

$$\dim_E \sigma_r(\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}}) = \min \left\{ N - 1, (r - 1) + r \sum_{\ell=1}^d (n_\ell - 1) \right\},$$

but in some instances  $\dim \sigma_r(\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}})$  may be strictly smaller than the expected dimension; then,  $\sigma_r$  is called a *defective*  $r$ th order secant variety, and  $\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}}$  a defective Segre variety. Only a limited number of defective secant varieties of Segre varieties are known, see, e.g., [1, §6.1] and [39, §5.5], while several secant varieties have been proven to be nondefective [1, 16, 17, 40, 54].<sup>8</sup> It is important to note that

$$\dim \sigma_r^\perp \leq 2r \sum_{\ell=1}^d (n_\ell - 1) < 2 \dim_E \sigma_r,$$

whenever  $d \geq 3$  and  $r \geq 2$ . That is, *whenever the  $r$ th order secant variety  $\sigma_r$  of a  $d$ -factor Segre variety is nondefective, a generic element of  $\sigma_r$  cannot admit a weak two-orthogonal decomposition of rank  $r$ .* Remark, further, that the discrepancy in dimension is  $2(r - 1)$ , which provides ample leeway in the defectivity of  $\sigma_r$  before the approach outlined in this paper becomes moot. Combining the above observations with Lemma 4.2, and using the known results from the literature, we obtain:<sup>9</sup>

**COROLLARY 4.3.** *Assume without loss of generality that  $2 \leq n_1 \leq n_2 \leq \dots \leq n_d$  with  $d \geq 3$ . Then, a generic rank- $r$  tensor in  $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \dots \times n_d}$  does not admit an SEY decomposition if:*

$$2 \leq r \leq \max \left\{ 2, \left\lfloor \frac{n_1^d}{dn_1 - d + 1} \right\rfloor - n_1 + 1, \min \left\{ n_d, \prod_{\ell=1}^{d-1} n_\ell - \sum_{\ell=1}^{d-1} (n_\ell - 1) \right\} \right\}.$$

*Proof.* From [1, Thm. 5.2] we know that  $\sigma_r(\mathcal{S}_{n_1, \dots, n_1}^{\mathbb{C}})$  is nondefective whenever  $r$  is smaller than the first item in the maximization. Then, using [1, Prop. 3.11], it follows that  $\sigma_r(\mathcal{S}_{n_1, \dots, n_d}^{\mathbb{C}})$  is nondefective because  $n_i \geq n_1$ . Applying Lemma 4.2 concludes this case. The second item in the maximization follows from combining Theorems 4.3 and 4.4 from [1], which summarize [16], with Lemma 4.2.  $\square$

<sup>8</sup>It should be noted that a defective  $r$ th order secant variety also implies that all secants of order  $r' > r$  are defective, unless  $\sigma_{r'}$  fills the ambient space  $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \dots \times n_d}$ , so finding defective secant varieties should be an easy task; yet both theoretical and probabilistic [57] results show that few varieties are defective.

<sup>9</sup>Keep in mind that “generic” should be interpreted with respect to the underlying real topology of  $\sigma_r$ .

Corollary 4.3 provides an easy-to-check condition on the rank  $r$  for which we know that admittance of an SEY decomposition is impossible for generic rank- $r$  tensors in the tensor space of the stated dimensions. From the formula it is clear that for sufficiently small  $r$ , an SEY decomposition cannot be admitted by a generic rank- $r$  tensor. This condition is by no means necessary; it is a weak sufficient condition.

The upper bound on the rank in Corollary 4.3 can often be improved by observing that the dimension of  $\sigma_r^\perp$  is substantially smaller than  $\sigma_r$ , and that this discrepancy increases proportionally to  $r$ . We will illustrate this with one well-studied case, the cubic tensor spaces  $\mathbb{P}\mathbb{C}^{n \times n \times \dots \times n}$ . Here, it can even be shown that a *generic tensor in this space* (with respect to the Euclidean topology) does not admit an SEY decomposition:

**THEOREM 4.4.** *Let  $d \geq 3$  and  $n \geq 2$ . Then, a generic tensor in  $\mathbb{P}\mathbb{C}^{n \times n \times \dots \times n}$  ( $d$  times) does not admit an SEY decomposition whenever*

$$d \geq 6 \quad \text{or} \quad \begin{cases} n \geq 4 & \text{if } d = 3 \\ n \geq 15 & \text{if } d = 4. \\ n \geq 5 & \text{if } d = 5 \end{cases}$$

*Proof.* Let  $\mathcal{S} = \mathcal{S}_{n,n,\dots,n}^{\mathbb{C}}$ , and let  $\underline{r}$  denote the smallest  $\underline{r}$  such that  $\sigma_{\underline{r}} = \mathbb{P}\mathbb{C}^{n \times n \times \dots \times n}$ ; this is called the *generic rank* [39]. A generic rank- $\underline{r}$  tensor is called a generic tensor. We additionally recall the definition of the *expected generic rank* for cubic tensor spaces:

$$\underline{r}_E = \left\lceil \frac{n^d}{dn - d + 1} \right\rceil \leq \underline{r};$$

the generic rank equals the expected generic rank whenever the Segre variety  $\mathcal{S}$  has no defective secant varieties. We define also the cumulative secant defect

$$\delta_r := \sum_{i=2}^r (\dim \sigma_{i-1} + \dim \mathcal{S} + 1 - \dim \sigma_i);$$

note that this quantity can only increase with  $r$ . Then, one verifies that

$$\dim \sigma_r = r \sum_{\ell=1}^d (n_\ell - 1) + (r - 1) - \delta_r + 1.$$

It follows that Lemma 4.2 may be invoked whenever the middle inequality in

$$\frac{1}{2} \dim \sigma_r^\perp \leq r \sum_{\ell=1}^d (n_\ell - 1) + \delta - 1 < \dim \sigma_r = r \sum_{\ell=1}^d (n_\ell - 1) + (r - 1) - \delta_r + 1$$

holds; note that this is not a *necessary* condition, however, because the upper bound in Property 4.1 is likely to be coarse. The middle inequality is then equivalent with

$$-\delta + 1 + (r - 1) > \delta_r - 1 \quad \text{which is satisfied if } r > \delta_r + \delta - 1.$$

Recalling that  $\delta \in \{0, 1\}$ , we find that  $r > \delta_r$  implies the latter equation regardless of  $\delta$ , and is thus a sufficient condition for  $\dim \sigma_r^\perp < 2 \dim \sigma_r$ . The foregoing discussion is only valid when  $r$  is small enough so that  $\sigma_r^\perp$  does not fill the ambient space, but this will not arise in the remainder.



If  $d = 3$ , Lickteig's classic result [40, p. 97] on the nondefectivity of  $\mathcal{S}$ ,  $n \neq 3$ , can be used. It states that all secant varieties of  $\mathcal{S}$  are nondefective, and thus  $\underline{r} = \underline{r}_E$ . Therefore,  $\delta_r = 0$  for  $r < \underline{r}_E$ , and  $\delta_{\underline{r}_E} < 3n - 3$ . Thus if the last inequality in

$$\underline{r}_E = \left\lceil \frac{n^3}{3n-2} \right\rceil \geq \frac{n^3}{3n-2} > 3n-3$$

is satisfied, then  $\dim \sigma_r^\perp < 2 \dim \sigma_r$  so that Lemma 4.2 applies. It is straightforward to verify using a computer algebra system that the last inequality in the above is satisfied whenever  $n \geq 7$ . The remaining cases can be verified by substituting  $n = 4, 5, 6$  in the more refined inequality  $\underline{r}_E > n^3 \bmod (3n - 2)$ .

If  $d \geq 4$ , we know that  $\delta_r = 0$  for  $r \leq \frac{n^d}{dn-d+1} - n$  from [1, Theorem 5.2]. From the same theorem, it can be deduced that there are at most  $n$  defective secants of  $\mathcal{S}$  before the space is filled. Consequently, the maximum cumulative secant defect in the generic rank is  $n(dn - d + 1)$ . If we compare this with the last rank for which we know that the cumulative defect is zero, then  $r > \delta_r$  for all  $r$  in said range. Therefore, if

$$\frac{n^d}{dn-d+1} - n > n(dn-d+1),$$

it follows that  $\dim \sigma_r^\perp < 2 \dim \sigma_r$ , and Lemma 4.2 applies. Using a computer algebra system it is easy to verify the cases  $d = 4, 5$ . For  $d = 6, 7$ , one would find that  $n \geq 3$ , which can be improved by Catalisano, Geramita, and Gimigliano's recent result on the nondefectivity of  $n = 2$  if  $d \geq 5$  [17]. Finally, if  $d = 8$ , one immediately obtains  $n \geq 2$ , providing the base case for the following inductive proof. Assume that

$$n^k > (kn - k + 1)(n(kn - k + 1) + n),$$

for all  $n \geq 2$ , for some  $k \geq 8$ . Consider then the fraction

$$\begin{aligned} \alpha(k, n) &= \frac{((k+1)n - (k+1) + 1)(n((k+1)n - (k+1) + 1) + n)}{(kn - k + 1)(n(kn - k + 1) + n)} \\ &= \left(\frac{k+1}{k}\right)^2 - \frac{k-1}{k^2(kn - k + 1)} - \frac{2(2+k)}{k^2(kn - k + 2)}. \end{aligned}$$

It is easy to verify that  $\alpha(k, n) < (9/8)^2 < 2$  for  $n \geq 2$  and  $k \geq 8$ . Clearly, if  $\alpha(k, n) < 2$ , we have

$$\begin{aligned} n^{k+1} &> \alpha(k, n)n^k > \alpha(k, n)(kn - k + 1)(n(kn - k + 1) + n) \\ &= ((k+1)n - (k+1) + 1)(n((k+1)n - (k+1) + 1) + n), \end{aligned}$$

proving the inductive case. This concludes the proof.  $\square$

Finally, we prove the main result, already presented in the introduction as Theorem 1.3. It states that in every complex tensor space of order at least three, one can always find a set of positive Lebesgue measure wherein its elements do not admit an SEY decomposition. The previous theorem actually specializes the main theorem for cubic tensor spaces, proving that the set is then dense.

**THEOREM 4.5.** *Let  $d \geq 3$ . Then there exists a non-empty open subset  $V$  of  $\mathbb{C}^{n_1 \cdots n_d} \cong \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$  with respect to the Euclidean topology such that the points in  $V$  do not admit an SEY decomposition;  $V$  has positive Lebesgue measure.*

*Proof.* Set  $N = n_1 \cdots n_d$ . It is known that  $\sigma_2(\mathcal{S}_{n_1, n_2, \dots, n_d}^{\mathbb{C}})$  is never defective for  $d \geq 3$  [1, p. 781]. Let  $\mathcal{A}_0$  be a rank-2 tensor in  $\mathbb{C}^N$  that does not lie in the affine cone

over  $\sigma_2^\perp \subset \mathbb{P}\mathbb{C}^N$ . Since  $\sigma_2^\perp$  is closed in  $\mathbb{P}\mathbb{C}^N$  with respect to the Euclidean topology, there exists a real value  $\epsilon > 0$  such that the open  $\ell_2$ -ball around  $\mathcal{A}_0$  in  $\mathbb{C}^N$  with radius  $\epsilon$  is disjoint from the cone over  $\sigma_2^\perp$ .<sup>10</sup> Now let  $V$  be the open  $\ell_2$ -ball around  $\mathcal{A}_0$  in  $\mathbb{C}^N$  with radius  $\epsilon/2$ . Then for every point  $\mathcal{A}$  in  $\mathbb{C}^N$ , any best rank-2 approximation  $\mathcal{A}^*$  of  $\mathcal{A}$  satisfies

$$\|\mathcal{A} - \mathcal{A}^*\| \leq \|\mathcal{A} - \mathcal{A}_0\| < \frac{\epsilon}{2} \quad \text{so that} \quad \|\mathcal{A}^* - \mathcal{A}_0\| < \epsilon,$$

and, hence,  $\mathcal{A}^*$  does not lie in the affine cone over  $\sigma_2^\perp$ . This means that  $\mathcal{A}^*$  does not admit an SEY decomposition. Alternatively, if  $\mathcal{A}$  does not have a best rank-2 approximation, we again find that it does not admit an SEY decomposition. In both cases,  $\mathcal{A}$  does not admit a best rank-2 approximation that itself admits an SEY decomposition. However, from the definition it follows that a tensor  $\mathcal{A}$  can only admit an SEY decomposition if for *every rank* there exists a best approximation of that rank that itself admits an SEY decomposition; this concludes the proof.  $\square$

Note that in the proof we may substitute  $\sigma_2^\perp$  for any secant variety that would satisfy Lemma 4.2, but such an exercise would only be useful if this would somehow provide information about the value of  $\epsilon$ , and possibly increase it.

**5. Conclusions.** We argued that current approaches for investigating the existence of an SEY decomposition rely explicitly on the existence of a set of tensors admitting supergeneric ranks with positive measure. Such an approach fails in a complex setting, leading us to propose an alternative strategy based on algebraic geometry and a comparison of dimensions of the varieties involved. We showed for every complex tensor space that an SEY decomposition is not admitted at least by a set of positive Lebesgue measure. Notwithstanding these results, we also provided a non-trivial class of tensors, i.e., those admitting a strong two-orthogonal decomposition, which are optimally truncatable.

**Acknowledgements.** J. Landsberg is heartily thanked for introducing us to algebraic geometry by means of his nice talk at the 17th International Linear Algebra Society Conference in Braunschweig, Germany, and for valuable insights and suggestions he provided there and in subsequent e-mail communication. G. Ottaviani is kindly thanked for the feedback he provided while the first author visited the Università di Firenze. We thank F. Cools for identifying an error in the proof of Theorem 3.3. Finally, C. Mertens is kindly thanked for various enlightening discussions we had.

**Funding.** The first author's research is supported by a Ph. D. fellowship of the Research Foundation – Flanders (FWO). The second and third author acknowledge the support by the Research Foundation – Flanders (FWO) project G034212N (Reestablishing smoothness for matrix manifold optimization via resolution of singularities). The third and fourth author acknowledge support by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization). The third author additionally acknowledges the support of the KU Leuven Research Council, projects OT/11/055 (Spectral Properties of Perturbed Normal Matrices and their Applications) and CoE EF/05/006 Optimization in Engineering (OPTEC). The fourth author acknowledges the support of KU Leuven Research Council grants PFV/10/002 (Optimization in Engineering) and OT/10/038 (Multi-parameter model order reduction and its applications).

<sup>10</sup>That is, the set of points at distance at most  $\epsilon$  from  $\mathcal{A}_0$ , as measured in the Euclidean norm.

## REFERENCES

- [1] H. ABO, G. OTTAVIANI, AND C. PETERSON, *Induction for secant varieties of Segre varieties*, Trans. Amer. Math. Soc., 361 (2009), pp. 767–792.
- [2] A. AMMAR, F. CHINESTA, AND A. FALCÓ, *On the convergence of a greedy rank-one update algorithm for a class of linear systems*, Arch. Comput. Methods Eng., 17 (2010), pp. 473–486.
- [3] A. ANANDKUMAR, R. GE, D. HSU, S.M. KAKADE, AND M. TELGARSKY, *Tensor decompositions for learning latent variable models*, 2012. arXiv:1210.7559.
- [4] C.J. APPELLOF AND E.R. DAVIDSON, *Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents*, Anal. Chem., 53 (1981), pp. 2053–2056.
- [5] J. BALLANI AND L. GRASEDYCK, *A projection method to solve linear systems in tensor format*, Numer. Linear Algebra Appl., 20 (2012), pp. 27–43.
- [6] M. BENNANI DOSSE, J.M.F. TEN BERGE, AND J.N. TENDEIRO, *Some new results on orthogonally constrained candecomp*, J. Classification, 28 (2011), pp. 144–155.
- [7] G. BEYLKIN AND M. J. MOHLENKAMP, *Algorithms for numerical analysis in high dimensions*, SIAM J. Sci. Comput., 26 (2005), pp. 2133–2159.
- [8] D. BINI, *Border rank of a  $p \times q \times 2$  tensor and the optimal approximation of a pair of bilinear forms*, in Automata, Languages and Programming, J.W. de Bakker and J. van Leeuwen, eds., Lecture notes in Comput. Sci. 85, Springer, New York, 1980, pp. 98–108.
- [9] ———, *Relations between exact and approximate bilinear algorithms*, Calcolo, 17 (1980), pp. 87–97.
- [10] D. BINI, G. LOTTI, AND F. ROMANI, *Approximate solutions for the bilinear form computational problem*, SIAM J. Comput., 9 (1980), pp. 692–697.
- [11] MC. BRAMBILLA AND G. OTTAVIANI, *On the Alexander–Hirschowitz theorem*, Journal of Pure and Applied Algebra, 212 (2008), pp. 1229–1251.
- [12] T. BRÜSTLE, *Typical examples of tame algebras*, in Representations of finite dimensional algebras and related topics in Lie theory and geometry, V. Dlab and C. M. Ringel, eds., vol. 40 of Fields Institute communications, AMS, 2004, pp. 27–44.
- [13] J. BUCZYŃSKI AND J.M. LANDSBERG, *On the third secant variety*. arXiv:1111.7005.
- [14] J.F. CARDOSO AND A. SOULOUMIAC, *Blind beamforming for non-Gaussian signals*, IEEE Proceedings-F, 140 (1993), pp. 362–370.
- [15] J. CARROLL AND J.-J. CHANG, *Analysis of individual differences in multidimensional scaling via an  $n$ -way generalization of “Eckart–Young” decomposition*, Psychometrika, 35 (1970), pp. 283–319.
- [16] MV. CATALISANO, A.V. GERAMITA, AND A. GIMIGLIANO, *Ranks of tensors, secant varieties of Segre varieties and fat points*, Linear Algebra Appl., 355 (2002), pp. 263–285.
- [17] ———, *Secant varieties of  $\mathbb{P}^1 \times \cdots \times \mathbb{P}^1$  ( $n$ -times) are not defective for  $n \geq 5$* , J. Algebraic Geom., 20 (2011), pp. 295–327.
- [18] J. CHEN AND Y. SAAD, *On the tensor SVD and the optimal low rank orthogonal approximation of tensors*, SIAM J. Matrix Anal. Appl., 30 (2009), pp. 1709–1734.
- [19] L. CHIANTINI AND G. OTTAVIANI, *On generic identifiability of 3-tensors of small rank*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1018–1037.
- [20] F. CHINESTA, A. AMMAR, A. LEYGUE, AND R. KEUNINGS, *An overview of the proper generalized decomposition with applications in computational rheology*, J. Non-Newtonian Fluid Mech., 166 (2011), pp. 578–592.
- [21] P. COMON, G. H. GOLUB, L-H. LIM, AND B. MOURRAIN, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.
- [22] L. DE LATHAUWER, *Decompositions of a higher-order tensor in block terms—part II: definitions and uniqueness*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1033–1066.
- [23] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.
- [24] V. DE SILVA AND L.-H. LIM, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1084–1127.
- [25] J. DROZD, *Tame and wild matrix problems*, in Representation Theory II, Vlastimil Dlab and Peter Gabriel, eds., vol. 832 of Lecture Notes in Mathematics, Springer Berlin / Heidelberg, 1980, pp. 242–258.
- [26] C. ECKART AND G. YOUNG, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.
- [27] F.R. GANTMACHER, *The Theory of Matrices, Volume 2*, Chelsea Publishing Company, New York, USA, 1959.
- [28] L. GRASEDYCK, *Hierarchical singular value decomposition of tensors*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2029–2054.
- [29] W. HACKBUSCH, *Tensor Spaces and Numerical Tensor Calculus*, no. 42 in Springer Series in Computational Mathematics, Springer Berlin / Heidelberg, 2012.

- [30] W. HACKBUSCH AND S. KÜHN, *A new scheme for the tensor representation*, J. Fourier Anal. Appl., 15 (2009), pp. 706–722.
- [31] R. A. HARSHMAN, *Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis*, UCLA Working Papers in Phonetics, 16 (1970), pp. 1–84.
- [32] R. HARTSHORNE, *Algebraic Geometry*, no. 52 in Graduate texts in mathematics, Springer-Verlag New York Inc., New York, NY, USA, 1997.
- [33] J. HUISMAN, *Real Abelian varieties with complex multiplication*, PhD thesis, Vrije Universiteit te Amsterdam, 1992.
- [34] J. JA’ JA’, *An addendum to Kronecker’s theory of pencils*, SIAM J. Appl. Math., 37 (1979), pp. 700–712.
- [35] T. G. KOLDA, *Orthogonal tensor decompositions*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 243–255.
- [36] ———, *A counterexample to the possibility of an extension of the Eckart–Young low-rank approximation theorem for the orthogonal rank tensor decomposition*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 762–767.
- [37] T. G. KOLDA AND B. W. BADER, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.
- [38] D. KRESSNER AND C. TOBLER, *Krylov subspace methods for linear systems with tensor product structure*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1688–1714.
- [39] J. M. LANDSBERG, *Tensors: Geometry and Applications*, vol. 128 of Graduate Studies in Mathematics, AMS, Providence, Rhode Island, 2012.
- [40] T. LICKTEIG, *Typical tensorial rank*, Linear Algebra Appl., 69 (1985), pp. 95–120.
- [41] M. MØRUP, *Applications of tensor (multiway array) factorizations and decompositions in data mining*, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1 (2011), pp. 24–40.
- [42] L.A. NAZAROVA, *Representations of quivers of infinite type*, Math. USSR Izv. 7 (1973), pp. 749–792.
- [43] A. NOUY, *Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems*, Arch. Comput. Methods Eng., 17 (2010), pp. 403–434.
- [44] I.V. OSELEDETS, *Tensor-train decomposition*, SIAM J. Sci. Comput., 33 (2011), pp. 2295–2317.
- [45] P. PAATERO, *Construction and analysis of degenerate PARAFAC models*, J. Chemometrics, 14 (2000), pp. 285–299.
- [46] B. PESQUET-POPESCU, J.-C. PESQUET, AND A.P. PETROPULU, *Joint singular value decomposition - a new tool for separable representation of images*, Proc. ICIP, (2001), pp. 569–572.
- [47] E. SCHMIDT, *Zur Theorie der linearen und nichtlinearen Integralgleichungen*, Math. Annal., 63 (1907), pp. 433–476.
- [48] D. SIMSON AND A. SKOWROŃSKI, *Elements of the Representation Theory of Associative Algebras. Volume 3 Representation-Infinite Tilted Algebras*, Cambridge University Press, Cambridge, United Kingdom, 2007.
- [49] A. SMILDE, R. BRO, AND P. GELADI, *Multiway Analysis: Applications in the Chemical Sciences*, John Wiley & Sons Ltd, Chichester, West Sussex PO19 8SQ, England, 2004.
- [50] A. STEGEMAN, *Degeneracy in CANDECOMP/PARAFAC explained for  $p \times p \times 2$  arrays of rank  $p + 1$  or higher*, Psychometrika, 71 (2006), pp. 483–501.
- [51] ———, *Degeneracy in CANDECOMP/PARAFAC and INDSCAL explained for several three-sliced arrays with a two-valued typical rank*, Psychometrika, 72 (2007), pp. 601–619.
- [52] ———, *Low-rank approximation of generic  $p \times q \times 2$  arrays and diverging components in the CANDECOMP/PARAFAC model*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 988–1007.
- [53] G. W. STEWART, *On the early history of the singular value decomposition*, SIAM Rev., 35 (1993), pp. 551–566.
- [54] V. STRASSEN, *Rank and optimal computation of generic tensors*, Linear Algebra Appl., 52–53 (1983), pp. 645–685.
- [55] L. R. TUCKER, *Some mathematical notes on three-mode factor analysis*, Psychometrika, 31 (1966), pp. 279–311.
- [56] N. VANNIEUWENHOVEN, R. VANDEBRIL, AND K. MEERBERGEN, *A new truncation strategy for the higher-order singular value decomposition*, SIAM J. Sci. Comput., 34 (2012), pp. A1027–A1052.
- [57] ———, *A probabilistic numerical algorithm based on Terracini’s lemma for proving nondefectivity of Segre varieties*, Tech. Rep. TW622, KU Leuven, Leuven, Belgium, 2012.
- [58] T. ZHANG AND G. H. GOLUB, *Rank-one approximation to high order tensors*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 534–550.