



**HAL**  
open science

# A polynomial formulation for joint decomposition of symmetric tensors of different orders

Pierre Comon, Yang Qi, Konstantin Usevich

► **To cite this version:**

Pierre Comon, Yang Qi, Konstantin Usevich. A polynomial formulation for joint decomposition of symmetric tensors of different orders. LVA/ICA 2015 - 12th International Conference on Latent Variable Analysis and Signal Separation, Aug 2015, Liberec, Czech Republic. pp.8. hal-01168992

**HAL Id: hal-01168992**

**<https://hal.science/hal-01168992>**

Submitted on 26 Jun 2015

**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.

# A polynomial formulation for joint decomposition of symmetric tensors of different orders

Pierre Comon, Yang Qi and Konstantin Usevich

CNRS, GIPSA-Lab, F-38000 Grenoble France  
firstname.lastname@gipsa-lab.grenoble-inp.fr

**Abstract.** We consider two models: simultaneous CP decomposition of several symmetric tensors of different orders and decoupled representations of multivariate polynomial maps. We show that the two problems are related and propose a unified framework to study the rank properties of these models.

**Keywords:** coupled CP decomposition, polynomial decoupling, generic rank, X-rank

## 1 Introduction

Tensor decompositions became an important tool in engineering sciences and data analysis. Several models require tensor decompositions with additional constraints (coupled decompositions or structured tensors), but the properties of these constrained decompositions are not so well understood.

In this paper, we consider two models of this kind: i) simultaneous CP decomposition of symmetric tensors of different orders (motivated by blind source separation) and ii) decoupling of multivariate polynomials (motivated by problems of identification of nonlinear dynamical systems). We show that these two models are strongly related, and that the notion of rank in these models enjoys many properties similar to tensor rank.

First we define a source separation model in Section 1.1, and next the polynomial decomposition model in Section 1.2. Finally, the organization and contributions of the paper are described in Section 1.3.

### 1.1 Blind source separation and independent component analysis

Consider a linear mixing model [6] in source separation

$$\mathbf{x} = A\mathbf{s},$$

---

<sup>1</sup> This work is supported by the ERC project “DECODA” no.320594, in the frame of the European program FP7/2007-2013.

where  $A$  is an (unknown) mixing matrix

$$A = (\mathbf{a}_1 \cdots \mathbf{a}_r) \in \mathbb{K}^{n \times r},$$

$\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ , and  $\mathbf{s} = (s_1 \cdots s_r)^\top$  is the vector of independent (real or complex) random variables. Then the cumulants of  $\mathbf{x}$  up to order  $d$  can be expanded as

$$\begin{aligned} \mathcal{C}_{\mathbf{x}}^{(1)} &= c_{1,1}\mathbf{a}_1 + \cdots + c_{1,r}\mathbf{a}_r, \\ \mathcal{C}_{\mathbf{x}}^{(2)} &= c_{2,1}\mathbf{a}_1 \otimes \mathbf{a}_1 + \cdots + c_{2,r}\mathbf{a}_r \otimes \mathbf{a}_r, \\ &\vdots \\ \mathcal{C}_{\mathbf{x}}^{(d)} &= c_{d,1}\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_1 + \cdots + c_{d,r}\mathbf{a}_r \otimes \cdots \otimes \mathbf{a}_r, \end{aligned} \tag{1}$$

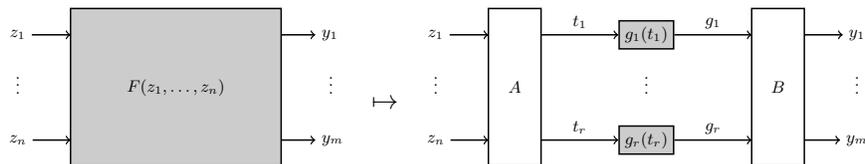
where  $c_{j,k}$  is the  $j$ -th cumulant of the random variable  $s_k$  [9].

In algebraic algorithms for blind source separation, typically a relaxed version of the decomposition problem (1) is considered. For example, in some approaches, a single cumulant (e.g., fourth order) is considered; in others the problem is reduced to decomposition of a partially symmetric tensor, see [6,9] for an overview. In most methods the structure of the joint decomposition (1) is lost, which we aim to avoid in this paper.

We should note that there exist few algorithms for blind source separation which use simultaneous diagonalization of symmetric tensors. In [8] a special case of  $d = 4$ ,  $n = 2$  is considered, and fourth- and third-order cumulants are simultaneously diagonalized by finding a common kernel of two matrices. In [7], a similar idea is used for combining cumulants of higher orders. (In [7] the case of  $n > 2$  sensors is also considered, but is treated suboptimally.) A theoretical framework for joint decomposition of cumulant tensors is also addressed in [4], but without proposing numerical algorithms.

## 1.2 Block-structured models of nonlinear systems

A common problem in nonlinear system identification is to decompose a multivariate nonlinear mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  in a block-structured form as a linear map followed by univariate nonlinear transformations, the outputs of which are linearly mixed again, see Fig. 1. This problem appears in identification of nonlinear state-space models [18] and parallel Wiener-Hammerstein systems [16].



**Fig. 1.** Decomposition of a multivariate function in a block-structured form.

If the multivariate function is represented as a polynomial, and the scalar nonlinear functions are also polynomials, then the decomposition in Fig. 1 becomes a polynomial decomposition problem, which we describe formally below.

Let  $\mathbb{K}$  be  $\mathbb{R}$  or  $\mathbb{C}$ . By  $\mathbb{K}_d[\mathbf{z}]$  we denote the space of homogeneous polynomials of degree  $d$ , and by  $\mathbb{K}_{\leq d}[\mathbf{z}]$  the space of polynomials of degree  $\leq d$ . Consider a multivariate polynomial map  $F : \mathbb{K}^n \rightarrow \mathbb{K}^m$ , *i.e.*, a vector  $F(\mathbf{z}) = (f_1(\mathbf{z}) \cdots f_m(\mathbf{z}))^\top$  of multivariate polynomials ( $f_i \in \mathbb{K}_{\leq d}[\mathbf{z}]$ ) in variables  $\mathbf{z} = (z_1 \cdots z_n)^\top$ . We say that  $F$  has a *decoupled representation*, if it can be expressed as

$$F(\mathbf{z}) = B \cdot \mathbf{g}(A^\top \mathbf{z}), \quad (2)$$

where

$$A = (\mathbf{a}_1 \cdots \mathbf{a}_r) \in \mathbb{K}^{n \times r}, \quad B = (\mathbf{b}_1 \cdots \mathbf{b}_r) \in \mathbb{K}^{m \times r},$$

are transformation matrices, and  $\mathbf{g} : \mathbb{K}^r \rightarrow \mathbb{K}^m$  is defined as

$$\mathbf{g}(t_1, \dots, t_r) = (g_1(t_1) \cdots g_r(t_r))^\top$$

where  $g_k$  are nonhomogeneous univariate polynomials of degree  $\leq d$ .

The decomposition (2) is exactly the one depicted in Fig. 1, and can be also equivalently represented as

$$F(\mathbf{z}) = \mathbf{b}_1 g_1(\mathbf{a}_1^\top \mathbf{z}) + \cdots + \mathbf{b}_r g_r(\mathbf{a}_r^\top \mathbf{z}). \quad (3)$$

Recently, two different, but related methods were proposed for solving the decoupling problem in the case  $m > 1$ , see [18] and [11]. Both methods are based on CP decomposition of a non-symmetric tensor constructed from the coefficients of the polynomial mapping. We also should note that there exist other tensor-based methods for identifying block-structured systems [13], which operate with structured tensors.

### 1.3 Contributions of this paper

The first aim of this paper is to show that the joint CP decomposition described in Section 1.1 is a special case of the polynomial decomposition from Section 1.2. Next, we show that both models can be viewed as a special case of  $X$ -rank decomposition: a powerful concept proposed recently in [2]. This concept provides a unified framework for studying properties of rank of the models (minimal  $r$  in (1) or (3)), and reformulate these questions in the language of algebraic geometry. Finally, we prove that underlying algebraic varieties are irreducible. As a consequence, the following results (proved in [2]) hold true.

1. For  $\mathbb{K} = \mathbb{C}$ , a generic (*i.e.*, drawn with probability 1) collection of tensors (a generic polynomial), has the same rank, called complex generic rank  $r_{gen,\mathbb{C}}$ .
2. For  $\mathbb{K} = \mathbb{R}$ , the rank of a generic collection of tensors (or a generic polynomial) is at least  $r_{gen,\mathbb{C}}$ .
3. For the both real and complex fields, the maximal rank is at most twice the generic complex rank, *i.e.*,  $r_{max,\mathbb{R}}, r_{max,\mathbb{C}} \leq 2r_{gen,\mathbb{C}}$ .

## 2 Polynomial decompositions

### 2.1 Symmetric tensors and polynomials

There is a one-to-one correspondence between symmetric  $\overbrace{n \times \cdots \times n}^s$  tensors and homogeneous polynomials of degree  $s$  [5]:

$$T(\mathbf{z}) = \mathcal{C} \times_1 \mathbf{z} \cdots \times_s \mathbf{z} \in \mathbb{K}_s[\mathbf{z}]. \quad (4)$$

Now assume that the tensor  $\mathcal{C}$  admits a CP decomposition

$$\mathcal{C} = c_1 \mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_1 + \cdots + c_r \mathbf{a}_r \otimes \cdots \otimes \mathbf{a}_r. \quad (5)$$

Then, by (4), decomposition (5) is equivalent to the decomposition

$$T(\mathbf{z}) = c_1 \ell_1^d(\mathbf{z}) + \cdots + c_r \ell_r^d(\mathbf{z}), \quad (6)$$

where  $\ell_k(\mathbf{z}) := \mathbf{a}_k^\top \mathbf{z}$  is a linear form. The decomposition (6) is called Waring decomposition [5].

### 2.2 Decomposition of polynomials

By equivalence between (5) and (6), the system (1) can be rewritten as

$$\begin{aligned} T^{(1)}(\mathbf{z}) &= c_{1,1} \ell_1(\mathbf{z}) + \cdots + c_{1,r} \ell_r(\mathbf{z}), \\ T^{(2)}(\mathbf{z}) &= c_{2,1} \ell_1^2(\mathbf{z}) + \cdots + c_{2,r} \ell_r^2(\mathbf{z}), \\ &\vdots \\ T^{(d)}(\mathbf{z}) &= c_{d,1} \ell_1^d(\mathbf{z}) + \cdots + c_{d,r} \ell_r^d(\mathbf{z}). \end{aligned} \quad (7)$$

Now define the non-homogeneous polynomial  $F \in \mathbb{K}_{\leq d}[\mathbf{z}]$  as

$$F(\mathbf{z}) = T^{(1)}(\mathbf{z}) + \cdots + T^{(d)}(\mathbf{z}), \quad (8)$$

Then from (7) it is easy to see that simultaneous Waring decomposition (7) (hence, the simultaneous symmetric CP decomposition (1)) is equivalent to the following problem: Given a multivariate polynomial  $F \in \mathbb{K}_{\leq d}[\mathbf{z}]$ , find minimal  $r$ ,  $g_k \in \mathbb{K}_{\leq d}[t]$  (univariate polynomials) and  $\mathbf{a}_k \in \mathbb{K}^n$  such that

$$F(\mathbf{z}) = \sum_{k=1}^r g_k(\ell_k(\mathbf{z})), \quad (9)$$

where  $\ell_k = \mathbf{a}_k^\top \mathbf{z}$  and  $g_k(t) = c_{0,k} + c_{1,k}t + \cdots + c_{d,k}t^d$ .

*Note 1.* Evidently, decomposition (9) is a special case of (3) with  $m = 1$ . Vice versa, any decomposition of the form (3) with  $m = 1$  can be reduced to (9). Indeed, we can always assume that the linear transformation  $B$  is equal to  $B = (1 \cdots 1)$ , without loss of generality.

The authors are aware of only one work [1] which studies the theoretical properties of (9), and more precisely the maximal rank. Also, a practical algorithm for computation of the decomposition (9) was proposed recently in [17].

### 3 X-rank decompositions

Here we recall a general definition of  $X$ -rank [2]. We will try to show how the decompositions in Sections 1.1 and 1.2 may fit in the  $X$ -rank framework.

Let  $W$  be a vector space over  $\mathbb{K}$ , and  $\mathbb{P}W$  be the corresponding projective space. Let  $X \subset \mathbb{P}W$  be a nondegenerate projective variety and  $\widehat{X}$  be an affine cone over  $X$ . Then for any  $v$  in  $W \setminus \{0\}$  we can define the  $X$ -rank

$$\text{rank}_X(v) = \min r : v = \widehat{x}_1 + \cdots + \widehat{x}_r, \quad \widehat{x}_k \in \widehat{X}. \quad (10)$$

The variety  $X$  (and its affine cone  $\widehat{X}$ ) represents the set of rank-one terms.

Let us fix the variety  $X$ . The maximal  $X$ -rank is defined as

$$r_{max} := \max_{v \in W} \text{rank}_X(v).$$

The typical ranks  $r_{typ,k}$ ,

$$r_{typ,0} < \cdots < r_{typ,n_{typ}} \leq r_{max},$$

are all the numbers such that the sets  $\{v \in W \mid \text{rank}_X(v) = r_{typ,k}\}$  have non-empty interior in Euclidean topology (see also [2]). Informally speaking, the typical ranks are the  $X$ -ranks that appear with non-zero probability if we draw randomly the vector  $v$  from a continuous probability distribution on  $W$ .

For  $X$ -ranks, the following basic results are known [2].

**Theorem 1 ([2, Theorems 1,3]).** *If  $r_{typ,0}$  is the smallest typical (real or complex) rank, then  $r_{max} \leq 2r_{typ,0}$ .*

**Theorem 2 ([2, page 1]).** *If  $\mathbb{K} = \mathbb{C}$  and  $X$  is an irreducible variety, then there exists a unique typical rank (called generic rank, and denoted by  $r_{gen}^{\mathbb{C}}$ ).*

**Theorem 3 ([2, Theorem 2]).** *If  $\mathbb{K} = \mathbb{R}$  and  $X$  is an irreducible variety, and  $X_{\mathbb{C}}$  is its complexification, then the smallest typical real rank equals the generic rank, i.e.,  $r_{typ,0}^{\mathbb{R}} = r_{gen}^{\mathbb{C}}$ .*

It is easy to show that decompositions (3) and (9) can be viewed as special cases of (10), as pointed out below.

1. **Rank-one polynomials** (9): take  $W = \mathbb{K}_{\leq d}[\mathbf{z}]$  and

$$\widehat{X} := \{f(\mathbf{z}) \in W \mid f(\mathbf{z}) = g(\mathbf{a}^\top \mathbf{z}), g(t) = \sum_{j=0}^d c_j t^j, \mathbf{a} \in \mathbb{K}^n\}. \quad (11)$$

2. **Rank-one polynomial maps** (3): take  $W = (\mathbb{K}_{\leq d}[\mathbf{z}])^m$  and

$$\widehat{X} := \{F(\mathbf{z}) \in W \mid F(\mathbf{z}) = \mathbf{b}g(\mathbf{a}^\top \mathbf{z}), g(t) = \sum_{j=0}^d c_j t^j, \mathbf{a} \in \mathbb{K}^n, \mathbf{b} \in \mathbb{K}^m\}. \quad (12)$$

Although we expressed the rank-one sets in (11) and (12), it is not immediate that we can use Theorems 1–3. We still need to prove that these sets are algebraic varieties and are irreducible. This is exactly the goal of the following section.



*Proof.* Let  $\mathbb{P}V$  be a complex projective space of dimension  $n - 1$ . Define  $\phi : \mathbb{P}V \rightarrow \mathbb{P}V \times \cdots \times \mathbb{P}S^dV$  by  $\phi([v]) = ([v], \dots, [v^d])$ , then  $\phi$  is an embedding (*i.e.*,  $\mathbb{P}V$  is isomorphic to  $\phi(\mathbb{P}V)$ ). Denote the image of  $\phi$  by  $M$ .

Let  $T_W$  be the tautological line bundle (called canonical line bundle in [15, §2]) on a projective space  $\mathbb{P}W$ , *i.e.*,  $T_W = \{([w], w) \in \mathbb{P}W \times W \mid w \in W\}$ , where  $\mathbb{P}W \times W$  is a trivial vector bundle over  $\mathbb{P}W$ .

Next, let  $p_i : \mathbb{P}V \times \cdots \times \mathbb{P}S^dV \rightarrow \mathbb{P}S^iV$  be the  $i$ -th natural projection, then  $D = \bigoplus_{i=1}^d p_i^* T_{S^iV}$  is a vector bundle of rank  $d$  over  $\mathbb{P}V \times \cdots \times \mathbb{P}S^dV$ , where  $p_i^*$  is the pull-back map induced by  $p_i$  [15]. Let  $E$  denote the restriction of  $D$  on  $M$ , thus  $E$  is a closed sub-variety of  $D$ .

Finally, define  $\tilde{\psi}_1 : E \rightarrow V \oplus \cdots \oplus S^dV$  by

$$\tilde{\psi}_1([v], \dots, [v^d], c_1v, \dots, c_dv^d) = (c_1v, \dots, c_dv^d).$$

It is easy to see that  $\widehat{X}_1 = \tilde{\psi}_1(E)$ .

Since each  $\mathbb{P}S^iV$  is complete ([14, Def. 7.1]) by [14, Thm. 7.22],  $\mathbb{P}V \times \cdots \times \mathbb{P}S^dV$  is complete by [14, §7.5], then

$$\begin{aligned} \bigoplus_{i=1}^d p_i^*(\mathbb{P}S^iV \times S^iV) &\rightarrow V \oplus \cdots \oplus S^dV \\ (([\alpha_1], \dots, [\alpha_d]), \beta_1, \dots, \beta_d) &\mapsto (\beta_1, \dots, \beta_d) \end{aligned} \quad (16)$$

is proper by [14, §7.16a], where  $\alpha_i, \beta_i \in S^iV$ . Thus by [14, §7.17], the restriction to  $D \rightarrow V \oplus \cdots \oplus S^dV$  is proper, and then  $\tilde{\psi}_1 : E \rightarrow V \oplus \cdots \oplus S^dV$  is proper. By definition of properness,  $\tilde{\psi}_1$  is universally closed, so  $\widehat{X}_1 = \tilde{\psi}_1(E)$  is closed, *i.e.*,  $\widehat{X}_1$  is an affine variety. Because  $E$  is irreducible,  $\widehat{X}_1$  is also irreducible.  $\square$

By Note 2 from Section 4.1 we have the following corollary.

**Corollary 1.**  $\widehat{X}_m$  defined in (15) is also irreducible, and is an affine cone over a projective variety.

Finally, from Proposition 1 and Corollary 1, we have that Theorems 1–3 can be applied for decompositions (9) and (3). In particular, for these decompositions there exists a complex generic rank (equal to the minimal real typical rank).

### 4.3 Generic rank for bivariate polynomials

Finally, in the case  $n = 2$  and  $m = 1$ , the variety  $X_1$  (corresponding to the affine cone  $\widehat{X}_1$  defined in (14)) is a special case of the rational normal scroll [12,10]. Using the results [3] on dimension of the  $r$ -th secant variety  $\sigma_r(X)$  of a rational normal scroll  $X$ , we can explicitly find the complex generic rank for this case.

**Proposition 2.** *The generic rank for bivariate polynomials is equal to*

$$r_{gen} := \left\lceil \frac{2d+7}{2} - \frac{\sqrt{8d+17}}{2} \right\rceil - 1.$$

*Proof.* By [3, p. 359], the dimension of  $\sigma_r(X_1) \subset \mathbb{P}^N$  is  $\min\{N, N - \frac{(d-r+1)(d-r+2)}{2} + r\}$  (note that there is an incorrect sign in the original paper [3]).

Thus the generic rank  $r_{gen}$  is the maximal  $r \in \{1, \dots, d\}$  such that

$$N - \frac{(d-r+2)(d-r+3)}{2} + r - 1 < N$$

which is equivalent to  $r < \frac{2d+7}{2} - \frac{\sqrt{8d+17}}{2}$ , i.e.,  $r_{gen} = \left\lceil \frac{2d+7}{2} - \frac{\sqrt{8d+17}}{2} \right\rceil - 1$ .  $\square$

## References

1. Białynicki-Birula, A., Schinzel, A.: Representations of multivariate polynomials as sums of polynomials in linear forms. *Colloq. Mathematicum* 112(2), 201–233 (2008)
2. Blekherman, G., Teitler, Z.: On maximum, typical and generic ranks. *Mathematische Annalen* pp. 1–11 (2014), in press.
3. Catalano-Johnson, M.L.: The possible dimensions of the higher secant varieties. *American Journal of Mathematics* pp. 355–361 (1996)
4. Comon, P.: Contrasts, independent component analysis, and blind deconvolution. *Int. Journal Adapt. Control Sig. Proc.* 18(3), 225–243 (2004)
5. Comon, P., Golub, G., Lim, L.H., Mourrain, B.: Symmetric tensors and symmetric tensor rank. *SIAM. J. Matrix Anal. Appl.* 30(3), 1254–1279 (2008)
6. Comon, P., Jutten, C.: *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press (Elsevier) (2010)
7. Comon, P., Rajih, M.: Blind Identification of Under-Determined Mixtures based on the Characteristic Function. *Signal Processing* 86(9), 3334–3338 (2006)
8. De Lathauwer, L., Comon, P., De Moor, B., Vandewalle, J.: ICA algorithms for 3 sources and 2 sensors. In: *Higher-Order Statistics, 1999. Proceedings of the IEEE Signal Processing Workshop on*. pp. 116–120 (1999)
9. De Lathauwer, L.: Algebraic methods after prewhitening. In: Comon, P., Jutten, C. (eds.) *Handbook of Blind Source Separation, Independent Component Analysis and Applications*, p. 831. Academic Press (Elsevier) (2010)
10. De Poi, P.: On higher secant varieties of rational normal scrolls. *Le Matematiche* 51(1) (1997)
11. Dreesen, P., Ishteva, M., Schoukens, J.: Decoupling multivariate polynomials using first-order information. Tech. rep. (2014), <http://arxiv.org/abs/1410.4060>.
12. Harris, J.: *Algebraic geometry. A first course*. Berlin etc.: Springer-Verlag (1992)
13. Kibangou, A.Y., Favier, G.: Tensor analysis-based model structure determination and parameter estimation for block-oriented nonlinear systems. *IEEE J. Sel. Topics Sig. Proc.* 4(3), 514–525 (2010)
14. Milne, J.: *Algebraic Geometry*. <http://www.jmilne.org/math/CourseNotes/AG.pdf>, version 6.00, August 24, 2014
15. Milnor, J., Stasheff, J.: *Characteristic Classes*. Annals of mathematics studies, Princeton University Press (1974)
16. Tiels, K., Schoukens, J.: From coupled to decoupled polynomial representations in parallel Wiener-Hammerstein models. In: *52nd IEEE Conference on Decision and Control, Florence, Italy, December 10-13, 2013*, pp. 4937–4942 (2013)
17. Usevich, K.: Decomposing multivariate polynomials with structured low-rank matrix completion. In: *21st Int. Symposium on Mathematical Theory of Networks and Systems, July 7-11, 2014. Groningen, The Netherlands*. pp. 1826–1833 (2014)
18. Van Mulders, A., Vanbeylen, L., Usevich, K.: Identification of a block-structured model with several sources of nonlinearity. In: *Proceedings of the 14th European Control Conference (ECC 2014)*. pp. 1717–1722 (2014)