SYMMETRIC TENSOR DECOMPOSITION

Jerome Brachat, Pierre Comon*†*, Bernard Mourrain*, Elias Tsigaridas*†*

†Laboratoire I3S, UMR6070 CNRS, UNS,
2000, route des Lucioles, BP 121
06903 Sophia Antipolis Cedex - France,
pcomon@unice.fr

*projet GALAAD, INRIA Sophia-Antipolis
2004, route des Lucioles, BP 93
06902 Sophia Antipolis Cedex - France,
Firstname.Lastname@sophia.inria.fr

ABSTRACT

We present an algorithm for decomposing a symmetric tensor of dimension n and order d as a sum of of rank-1 symmetric tensors, extending the algorithm of Sylvester devised in 1886 for symmetric tensors of dimension 2. We exploit the known fact that every symmetric tensor is equivalently represented by a homogeneous polynomial in n variables of total degree d. Thus the decomposition corresponds to a sum of powers of linear forms.

The impact of this contribution is two-fold. First it permits an efficient computation of the decomposition of any tensor of sub-generic rank, as opposed to widely used iterative algorithms with unproved convergence (e.g. Alternate Least Squares or gradient descents). Second, it gives tools for understanding uniqueness conditions, and for detecting the tensor rank.

1. INTRODUCTION

Symmetric tensors show up in applications mainly as high-order derivatives of multivariate functions. For instance in Statistics, cumulant tensors are derivatives of the second characteristic function [1]. Tensors have been widely utilized in Electrical Engineering since the nineties, because of the use of High-Order Statistics [2] [3] [4] [5] [6] [7]. Even earlier in the seventies, tensors have been used in Chemometrics [8] or psychometrics [9]. Another important application field is Data Analysis. For instance, Independent Component Analysis was originally introduced for symmetric tensors whose rank did not exceed dimension [10] [11]. Now, it has become possible to estimate more factors than the dimension. Further references may be found in [12] [13], and numerous applications of tensor decompositions may be found in [14] [15].

The goal of this paper is to devise an *algebraic* technique able to decompose a symmetric tensor of arbitrary order and dimension in an essentially unique manner (i.e. up to scale and permutation) into a sum of rank-one terms. Of course, reaching such a goal requires some conditions, in particular related to its rank, which must be sub-generic. Our algorithm could be seen as an extension of the SVD algorithm from matrices to n-way arrays. We exploit the strong connection of symmetric tensors and homogeneous polynomials. This approach allows us to use effective algebraic geometry techniques, and to tackle the problem of decomposition using Veronese varieties, duality of vector spaces, and algorithms for polynomial system solving. To the best of our knowledge this is the first time that a decomposition algorithm for symmetric tensors is presented.

The rest of the paper is structured as follows: In the remaining of the section we present some historical remarks and we shed light to the connection of symmetric tensors and homogeneous polynomials. Sec. 2 presents Sylvester's approach for the binary case. In Sec. 3 we exploit many different, albeit equivalent, algebraic formulations of the decomposition problem, as well as the necessary algebraic tools. In Sec. 4 we present the algorithm and illustrate it with an example.

1.1 Historical remarks

Despite their obvious practical interest as they can deal with inexact data, numerical algorithms presently used in most scientific communities are suboptimal, in the sense that they either do not fully exploit symmetries [16], minimize different successive criteria sequentially [17] [18], or are iterative and lack a guarantee of *global* convergence [19] [20]. In addition, they often request the rank to be much smaller than generic [21].

On the other hand, the algorithm based on Sylvester's theorem [22], recalled in section 2, provides a complete answer to the questions of uniqueness and computation, for any order [23]. However, the latter is devoted to 2-dimensional symmetric tensors, and techniques based on pairwise processing have a very limited range of use when the rank exceeds the dimension. For the decomposition of fourth-order tensors we refer the reader to [18].

The algorithm proposed in this paper is inspired from Sylvester's theorem, and extends its principle to larger dimensions. In addition, it fully exploits symmetry, and when the solution is essentially unique, it provides the decomposition for any sub-generic rank.

1.2 Tensors and Polynomials

Any symmetric tensor of dimension n, i.e. the range of each index, and order d, i.e. the number of indices, can be associated with a homogeneous polynomial in n variables of degree d. For instance, a third order tensor T_{ijk} can be associated with the polynomial $\sum_{ijk} c(i,j,k)\, T_{ijk}\, x_i x_j x_k$, where $c(\cdot)$ denotes some fixed symmetric function. See e.g. [23] for further details. We consider a homogeneous polynomial $F^h(\mathbf{x})$

$$F^h(\mathbf{x}) = \sum_{j_0+j_1+\cdots+j_n=d} a_{j_0,j_1,\ldots,j_n} x_0^{j_0} x_1^{j_1} \cdots x_n^{j_n}. \quad (1)$$

[0] This work has been partially supported by contract ANR-06-BLAN-0074 "Decotes"

Our goal is to compute a decomposition of F^h as a sum of dth powers of linear forms, $F^h(\mathbf{x}) = \sum_{i=1}^r \lambda_i (k_{i,0}x_0 + \cdots + k_{i,n}x_n)^d = \lambda_1 \mathbf{k}_1(\mathbf{x})^d + \lambda_2 \mathbf{k}_2(\mathbf{x})^d + \cdots + \lambda_r \mathbf{k}_r(\mathbf{x})^d$, where $\lambda_i \in \mathbb{C}$, such that r is the smallest possible. This smallest r is often referred to as the *tensor rank*, or sometimes the polynomial width.

Let's see now how the decomposition is made possible in the case of homogeneous polynomials in two variables only.

2. THE BINARY CASE

Let $p(x_1, x_2) = \sum_{i=0}^d \binom{d}{i} c_i\, x_1^i\, x_2^{d-i}$ be a homogeneous polynomial of degree d in 2 variables. Denote $H[r]$ the Hankel matrix of dimensions $(d-r+1) \times (r+1)$ with entries $H[r]_{ij} = c_{i+j-2}$:

$$H[r] = \begin{bmatrix} c_0 & c_1 & \cdots & c_r \\ \vdots & & & \vdots \\ c_{d-r} & \cdots & c_{d-1} & c_d \end{bmatrix}.$$

Then we have:

Sylvester, 1886 $p(x_1, x_2)$ can be written as a sum of d^{th} powers of r distinct linear forms in $\mathbb{C}$ as:

$$p(x_1, x_2) = \sum_{j=1}^r \lambda_j\, (\alpha_j\, x_1 + \beta_j\, x_2)^d, \tag{2}$$

if and only if **(i)** there exists a vector $\mathbf{q}$ of dimension $r+1$, with components q_ℓ, such that

$$H[r]\, \mathbf{q} = \mathbf{0}. \tag{3}$$

and **(ii)** the polynomial $q(x_1, x_2) = \sum_{\ell=0}^r q_\ell\, x_1^\ell\, x_2^{r-\ell}$ admits r distinct roots, i.e. it can be written as $q(x_1, x_2) = \prod_{j=1}^r (\beta_j^*\, x_1 - \alpha_j^*\, x_2)$.

It turns out that the proof of this theorem is constructive [24] [23] [25] and yields the algorithm below.

1. Initialize $r = 0$
2. Increment $r \leftarrow r + 1$
3. If the row rank of $H[r]$ is full, then go to step 2
4. Else compute a basis $\{\mathbf{k}_1, \ldots, \mathbf{k}_l\}$ of the right kernel of $H[r]$.
5. Specialization (pick a random vector in the kernel):
   - Take a generic vector $\mathbf{q}$ in the kernel, e.g. $\mathbf{q} = \sum_i \mu_i \mathbf{k}_i$
   - Compute the roots of the associated polynomial $q(x_1, x_2) = \sum_{\ell=0}^r q_\ell\, x_1^\ell\, x_2^{d-\ell}$. Denote them $(\beta_j, -\alpha_j)$, where $|\alpha_j|^2 + |\beta_j|^2 = 1$.
   - If the roots are not distinct in $\mathbb{P}^2$, try another specialization. If distinct roots cannot be obtained, go to step 2.
   - Else if $q(x_1, x_2)$ admits r distinct roots then compute the coefficients λ_j, $1 \le j \le r$, by solving the linear system below

$$\begin{bmatrix} \alpha_1^d & \ldots & \alpha_r^d \\ \alpha_1^{d-1}\beta_1 & \ldots & \alpha_r^{d-1}\beta_r \\ \alpha_1^{d-2}\beta_1^2 & \ldots & \alpha_r^{d-1}\beta_r^2 \\ \vdots & \vdots & \vdots \\ \beta_1^d & \ldots & \beta_r^d \end{bmatrix} \lambda = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_d \end{bmatrix}.$$

6. The decomposition is $p(x_1, x_2) = \sum_{j=1}^r \lambda_j\, \mathbf{k}_j(\mathbf{x})^d$, where $\mathbf{k}_j(\mathbf{x}) = (\alpha_j\, x_1 + \beta_j\, x_2)$.

Note that step 5 is a specialization only if the dimension of the right kernel is larger than one, which will not occur for ranks smaller than generic.

The goal is now to extend this kind of numerical algorithm to polynomials in more variables. This problem was open until now.

3. PROBLEM FORMULATIONS

Notation. If $\mathbf{a} = (a_1, \ldots, a_n)$ is a vector in $\mathbb{N}^n$, then $|\mathbf{a}|$ is the sum of its elements, i.e. $|\mathbf{a}| = \sum_{i=a}^n a_i$. By $\mathbf{x}^{\mathbf{a}}$ will denote the monomial $x_1^{a_1} \cdots x_n^{a_n}$.

Let R be the ring of polynomials $\mathbb{C}[x_1, \ldots, x_n]$, while R_d will denote the vector space of polynomials of (total) degree at most d. The set $\{\mathbf{x}^{\mathbf{a}}\}_{|\mathbf{a}| \le d} = \{x_1^{a_1} \cdots x_n^{a_n}\}_{a_1 + \cdots + a_n \le d}$ represents the elements of the monomial basis of the vector space R_d. It contains $\binom{n+d-1}{d}$ elements. The corresponding basis of the dual space $\widehat{R}_d$, that is the set of linear forms that compute the coefficients of a polynomial in the primal basis, is the set $\{\mathbf{d}^{\mathbf{a}}\}_{|\mathbf{a}| \le d}$, where $\mathbf{d}^{\mathbf{a}} : R_d \to \mathbb{C}$ and $\mathbf{d}^{\mathbf{a}}(f) = \frac{1}{\prod_{i=1}^n a_i!} \left(\frac{d^{a_1}}{dx_1} \cdots \frac{d^{a_n}}{dx_n} f \right)(0)$.

The superscript h denotes the homogeneous version of the polynmomial. Let S be the set of homogeneous polynomials in $n+1$ variables. S_d represents the homogeneous polynomials of degree d, and $\mathbb{P}(S_d)$, the corresponding projective space. Similarly interpretations hold for the dual spaces $\widehat{S}$ and $\widehat{S}_d$. Analogous to the affine case, we can define primal and dual bases for the homogeneous case.

3.1 Direct approach by polynomial fitting

The first idea is merely to solve, in a given polynomial basis, the polynomial system, induced by the equation

$$F^h(\mathbf{x}) - \sum_i \mathbf{k}_i(\mathbf{x})^d = 0,$$

with respect to the coefficients of the linear forms $\mathbf{k}_i$. We call this the *direct approach*. In the tensor framework, even if the rank is supposed to be known, attempts to solve this problem have not entailed efficient algorithms (cf. section 1). In the polynomial framework, it is easy to see that we end up with an over-determined polynomial system of $\binom{n+d}{d}$ equations in $r(n+1)$ unknowns. This description of the problem is not optimal, since it introduces $r!$ redundant solutions corresponding to permutations of the linear forms. Another drawback is that polynomials involved are of degree d in the coefficients $k_{i,n}$, which are too high from the computational point of view. In fact, our approach does not involve the solution of polynomial systems of degree higher than 2.

3.2 Different views using duality

We consider the following Veronese map of degree d

$$\begin{array}{cccc} \nu & : & S_1 & \to & S_d \\ & & \mathbf{k}(\mathbf{x}) & \mapsto & \mathbf{k}(\mathbf{x})^d. \end{array}$$

which sends a linear (homogeneous) polynomial to its $d-$th power. Recall that the (monomial) basis of S_1 is the set $\{x_0, x_1, \ldots, x_n\}$, while the basis of S_d is the set $\{\mathbf{x}^{\mathbf{a}}\}_{|\mathbf{a}|\leq d} = \{x_0^{a_0} x_1^{a_1} \cdots x_n^{a_n}\}_{a_1+\cdots+a_n=d}$, viz. the set of all the monomials in $x_0, x_1, \ldots, x_n$ of total degree d. The cardinality of the basis is $\binom{n+d}{d} - 1$. Under the action of ν, a linear polynomial $\mathbf{k}(\mathbf{x}) = k_0 x_0 + \cdots + k_n x_n$ corresponds to $\mathbf{k}(\mathbf{x})^d = \sum_{i_0+\cdots+i_n=d} \binom{d}{i_o,\ldots,i_n} k_0^{i_0} \cdots k_n^{i_n} x_0^{i_0} \cdots x_n^{i_n}$. In terms of vectors, $\mathbf{k} = [k_0, \ldots, k_n]^\top$ corresponds to the vector $[\ldots, \binom{d}{i_0,\ldots,i_n} k_0^{i_0} \cdots k_n^{i_n}, \ldots]^\top$.

Another Veronese map, also of degree d, is

$$\begin{array}{rcccl} \delta & : & \mathbb{C}^{n+1} & \to & \widehat{S}_d \\ & & \mathbf{z} & \mapsto & \mathbb{1}_{\mathbf{z}}, \end{array}$$

which sends a point $\mathbf{z} = (z_0, z_1, \ldots, z_n)$ to $[\ldots, z_0^{i_0} \cdots z_n^{i_n}, \ldots]$. Recall that the linear functionals $\{\mathbf{d}^{\mathbf{a}}\}_{|\mathbf{a}|=d}$ consist a basis of $\widehat{S}_d$. It holds that $S_1 \cong \mathbb{C}^{n+1}$. The map τ:

$$\begin{array}{rcl} \tau : \widehat{S}_d & \to & S_d \\ [\ldots, z_0^{i_0} \ldots z_n^{i_n}, \ldots]^\top & \mapsto & [\ldots, \binom{d}{i_0,\ldots,i_n} z_0^{i_0} \ldots z_n^{i_n}, \ldots]^\top \end{array}$$

is an isomorphism. The map remains an isomorphism even if we restrict it to the images of the maps ν and δ, that is $\nu(S_1)$ and $\delta(\mathbb{C}^n)$, respectively. The inverse of τ is the map $\tau^{-1} : S_d \to \widehat{S}_d$. Consider a polynomial in S_d, that is the $d-$th power of a linear form, say $\mathbf{k}(\mathbf{x})^d$. If we apply the map τ^{-1} to this polynomial, then we have that $\tau^{-1}(\mathbf{k}(\mathbf{x})^d) = \mathbb{1}_{\mathbf{k}}$; that is the linear form that gives the evaluation of a polynomial (homogeneous of degree d) over the point $\mathbf{k} = [k_0, k_1, \ldots, k_n]^\top$.

Let us now revisit the problem of decomposition. Initially we are given a polynomial $F^h(\mathbf{x}) \in S_d$. The decomposition $F^h(\mathbf{x}) = \sum_{i=1}^r \lambda_i \mathbf{k}_i(\mathbf{x})^d$ corresponds to a secant variety in $\nu(S_1) \subset S_d$. Using the properties of the isomorphism τ and its inverse we can gain another view of the problem. If we apply τ^{-1} to F^h we compute its dual, that is $\Phi = \tau^{-1}(F^h)$. The decomposition of the latter, i.e. $\Phi = \sum_{i=1}^r \lambda_i \mathbb{1}_{\mathbf{k}_i}$, is a linear combination of elements in $\delta(\mathbb{C}^n) \subset \widehat{R}_d$.

Overall, it holds that $\tau^{-1}(F^h(\mathbf{x})) = \tau^{-1}\left(\sum_{i=1}^r \lambda_i \mathbf{k}_i(\mathbf{x})^d\right) = \sum_{i=1}^r \lambda_i \tau^{-1}(\mathbf{k}_i(\mathbf{x})^d) = \sum_{i=1}^r \lambda_i \mathbb{1}_{\mathbf{k}_i} = \Phi$. Moreover, $\tau(\Phi) = \tau\left(\sum_{i=1}^r \lambda_i \mathbb{1}_{L_i}\right) = \sum_{i=1}^r \lambda_i \tau(\mathbb{1}_{\mathbf{k}_i}) = \sum_{i=1}^r \lambda_i \mathbf{k}_i(\mathbf{x})^d = F^h(\mathbf{x})$.

All the previous views of the decomposition problem are equivalent. The results obtained for any of them could be translated for the other.

3.3 Quotient algebra and duality

The idea of the algorithm is to exploit the properties of $\Phi \in \widehat{R}$, that we assume that is known up to degree d. More precisely, we consider the symmetric bilinear form $\mathbb{H}_\Phi : (p, q) \mapsto \Phi(p\,q)$, the matrix of which in the monomial basis is $(\Phi(\mathbf{x}^{\mathbf{a}+\mathbf{b}}))_{\mathbf{a},\mathbf{b}\in\mathbb{N}^n}$. Let I_Φ be the kernel of $\mathbb{H}_\Phi$.

Proposition 3.2 If $\Phi = \sum_{i=1}^r \lambda_i \mathbb{1}_{\mathbf{k}_i}$ with $\lambda_i \neq 0$ and $\mathbf{k}_i \in \mathbb{C}^n$, then $p \in I_\Phi$ iff $p(\mathbf{k}_i) = 0$ for $i = 1, \ldots, r$.

In other words, the common roots of all the polynomials in I_Φ define the linear terms in the tensor decomposition of F.

In order to compute the zeros of I_Φ, we may use a well-known theorem (see e.g. [26, 27, 28]), which we apply to the zero-dimensional ideal I_Φ:

Theorem 3.3 The eigenvalues of the matrices $\mathbb{M}_a$ and $\mathbb{M}_a^\top$, of the linear operators that correspond to the multiplication by a in R modulo I_Φ, and its transposed, are $\{a(\mathbf{k}_1), \ldots, a(\mathbf{k}_r)\}$. The common eigenvectors of the matrices $(\mathbb{M}_{x_i}^\top)_{1\leq i\leq n}$ are (up to a scalar) $\mathbb{1}_{\mathbf{k}_i}$, $i = 1, \ldots, r$.

If we denote by $\mathbb{H}_\Phi^E$ the restriction of $\mathbb{H}_\Phi$ to a vector space E of dimension r on which $\mathbb{H}_\Phi$ is invertible, we have the relation $\mathbb{H}_{a\star\Phi}^E = \mathbb{M}_a^\top \mathbb{H}_\Phi^E$, where $\mathbb{H}_{a\star\Phi} : (p, q) \mapsto \Phi(a\,p\,q)$. Thus the solution of the generalized eigenvalue problem $(\mathbb{H}_{a\star\Lambda}^E - \zeta\, \mathbb{H}_\Lambda^E)\mathbf{v} = \mathbb{O}$ yields the eigenvector $\mathbb{H}_\Lambda^E \mathbf{v}$ of $\mathbb{M}_a^\top$, which are by Th. 3.3, the evaluations $\mathbb{1}_{\mathbf{k}_i}$. From these eigenvectors, we deduce the linear factors in the tensor decomposition. The coefficients λ_i $(i = 1, \ldots, r)$ can then be computed by solving a linear system of size r.

4. ALGORITHM

The algorithm that we will present for decomposing a symmetric tensor as sum of rank 1 symmetric tensors generalizes the algorithm of Sylvester [24], devised for dimension 2 tensors, see also [29].

4.1 Overview

Algorithm 1: SYMMETRIC TENSOR DECOMPOSITION

Input: A homogeneous polynomial $f(x_0, x_1, \ldots, x_n)$ of degree d.

Output: A decomposition of f as $f = \sum_{i=1}^r \lambda_i\, \mathbf{k}_i(\mathbf{x})^d$ with r minimal.

- Compute the coefficients of f^*: $c_\alpha = a_\alpha \binom{d}{\alpha}^{-1}$, for $|\alpha| \leq d$, $\alpha = (\alpha_1, \ldots, \alpha_n)$;
- $r := 1$;
- **Repeat**
  1. Compute a set B of monomials of degree $\leq d$ connected to 1 with $|B| = r$;
  2. Find parameters $\mathbf{h}$ s.t. $\det(\mathbb{H}_\Phi^B) \neq 0$ and the operators $\mathbb{M}_i = \mathbb{H}_{x_i\Phi}^B (\mathbb{H}_\Phi^B)^{-1}$ commute.
  3. If there is no solution, restart the loop with $r := r + 1$.
  4. Else compute the $n \times r$ eigenvalues $\zeta_{i,j}$ and the eigenvectors $\mathbf{v}_j$, s.t. $\mathbb{M}_i \mathbf{v}_j = \zeta_{i,j} \mathbf{v}_j$, $i = 1, \ldots, n$, $j = 1, \ldots, r$.

  until the eigenvalues are simple.
- Solve the linear system in $(\lambda_j)_{j=1,\ldots,k}$: $\Phi = \sum_{j=1}^r \lambda_j \mathbf{1}_{\mathbf{v}_j}$ where $\mathbf{v}_j \in \mathbb{K}^n$ are the eigenvectors found in step 4.

Let us briefly comment on the computation process. A basis connected to 1, is a basis containing 1 where each element different from 1, is the product of a variable by another element of the basis. Consider the homogeneous polynomial $f(\mathbf{x})$ in (1) that we want to decompose. We may assume without loss of generality, that for at least one variable, say x_0, all its coefficients in the decomposition are non zero, i.e. $k_{i,0} \neq 0$, for $1 \leq i \leq r$. We dehomogenize f with respect to this variable and we denote this polynomial by $f^a := f(1, x_1, \ldots, x_n)$. We want to decompose the polynomial $f^a(\mathbf{x}) \in R_d$ as a sum of powers of linear forms, i.e. $f(\mathbf{x}) = \sum_{i=1}^{r} \lambda_i \, (1 + k_{i,1}x_1 + \cdots + k_{i,n}x_n)^d = \sum_{i=1}^{r} \lambda_i \, \mathbf{k}_i(\mathbf{x})^d$. Equivalently, we want to decompose its corresponding dual element $f^* \in \widehat{R}_d$ as a linear combination of evaluations over the distinct points $\mathbf{k}_i := (k_{i,1}, \cdots, k_{i,n})$: $f^* = \sum_{i=1}^{r} \lambda_i \, \mathbb{1}_{\mathbf{k}_i}$ (we refer the reader to the end of Section 3.2).

Assume that we know the value of r. If we know the value of Φ on polynomials of degree high enough, it allows us to compute the tables of multiplication modulo the kernel of $\mathbb{H}_\Phi$. By Theorem 3.3, if we solve the generalized eigenvector problem $(\mathbb{H}_{x_1 \star \Phi} - \zeta \, \mathbb{H}_\Phi)\mathbf{v} = \mathbb{O}$, then we can recover the points of evaluation $\mathbf{k}_i$. By solving a linear system, we will then deduce the value of $\lambda_i, \ldots, \lambda_r$. For certain (big) values of r it can happen that not all the elements of the corresponding matrices are known. In this case, we use the property that the matrices of multiplication commute, and we form a system, the solutions of which are these unknown elements. We refer the reader to [12, 13] for details, and we present an example to illustrate the algorithm.

4.2 Example

(1) Convert the symmetric tensor to the corresponding homogeneous polynomial.

Assume that we are given a tensor of dimension 3 and order 5, and that the corresponding homogeneous polynomial is $f = -1549440\, x_0 x_1 {x_2}^3 + 2417040\, x_0 {x_1}^2 {x_2}^2 + 166320\, {x_0}^2 x_1 {x_2}^2 - 829440\, x_0 {x_1}^3 x_2 - 5760\, {x_0}^3 x_1 x_2 - 222480\, {x_0}^2 {x_1}^2 x_2 + 38\, {x_0}^5 - 497664\, {x_1}^5 - 1107804\, {x_2}^5 - 120\, {x_0}^4 x_1 + 180\, {x_0}^4 x_2 + 12720\, {x_0}^3 {x_1}^2 + 8220\, {x_0}^3 {x_2}^2 - 34560\, {x_0}^2 {x_1}^3 - 59160\, {x_0}^2 {x_2}^3 + 831840\, x_0 {x_1}^4 + 442590\, x_0 {x_2}^4 - 5591520\, {x_1}^4 x_2 + 7983360\, {x_1}^3 {x_2}^2 - 9653040\, {x_1}^2 {x_2}^3 + 5116680\, x_1 {x_2}^4$.

(2) Compute the matrix of the linear form.

We form a $\binom{n+d-1}{d} \times \binom{n+d-1}{d}$ matrix, the rows and the columns of which correspond to the evaluation of the dual of the polynomial over all the monomial $\{\mathbf{x}^{\mathbf{a}}\}_{|\mathbf{a}| \leq d}$, using the map $a_{j_0\, j_1 \ldots j_n} \mapsto a_{j_0\, j_1 \ldots j_n} \frac{j_0!\, j_1! \ldots j_n!}{d!}$, where $a_{j_0\, j_1 \ldots j_n}$ is the coefficient of the monomial $x_1^{j_0} \cdots x_n^{j_n}$.

Part of the corresponding matrix follows.

$$
\begin{array}{c}
\begin{array}{rrrrrrr} 1 & x_1 & x_2 & x_1^2 & x_1 x_2 & x_2^2 & x_1^3 \end{array} \\
\left[\begin{array}{rrrrrrr}
38 & -24 & 36 & 1272 & -288 & 822 & -3456 \\
-24 & 1272 & -288 & -3456 & -7416 & 5544 & 166368 \\
36 & -288 & 822 & -7416 & 5544 & -5916 & -41472 \\
1272 & -3456 & -7416 & 166368 & -41472 & 80568 & -497664 \\
-288 & -7416 & 5544 & -41472 & 80568 & -77472 & -1118304 \\
822 & 5544 & -5916 & 80568 & -77472 & 88518 & 798336 \\
-3456 & 166368 & -41472 & -497664 & -1118304 & 798336 & h_{6,0,0} \\
-7416 & -41472 & 80568 & -1118304 & 798336 & -965304 & h_{5,1,0} \\
5544 & 80568 & -77472 & 798336 & -965304 & 1023336 & h_{4,2,0} \\
-5916 & -77472 & 88518 & -965304 & 1023336 & -1107804 & h_{3,3,0}
\end{array}\right]
\end{array}
$$

The whole matrix is 21×21. For reasons of space we present only the first 7 columns. Notice that we do not know the elements in some positions of the matrix (in the 7th column). In general we do not know the elements that correspond to monomials with (total) degree higher than 5.

(3) Extract a principal minor of full rank.

We should re-arrange the rows and the columns of the matrix so that the first principal minor is of full rank, r. We call this minor $\mathbb{H}_\Phi$. In order to do that we try to put the matrix in row echelon form, using elementary row and column operations. In our example the 4×4 principal minor is of full rank, so there is no need for re-arranging the matrix. The matrix $\mathbb{H}_\Phi$ is

$$
\mathbb{H}_\Phi = \begin{bmatrix} 38 & -24 & 36 & 1272 \\ -24 & 1272 & -288 & -3456 \\ 36 & -288 & 822 & -7416 \\ 1272 & -3456 & -7416 & 166368 \end{bmatrix}
$$

Notice that the columns of the matrix correspond to the monomials $\{1, x_1, x_2, x_1^2\}$.

(4) We compute the "shifted" matrix $\mathbb{H}_{x_1 \star \Phi}$.

If the columns of $\mathbb{H}_\Phi$ correspond to set of some monomials, say $\{\mathbf{x}^{\mathbf{a}}\}$, then the columns of $\mathbb{H}_{x_1 \star \Phi}$ correspond to the set of monomials $\{x_1\, \mathbf{x}^{\mathbf{a}}\}$. In our example

$$
\mathbb{H}_{x_1 \star \Phi} \begin{bmatrix} -24 & 1272 & -288 & -3456 \\ 1272 & -3456 & -7416 & 166368 \\ -288 & -7416 & 5544 & -41472 \\ -3456 & 166368 & -41472 & -497664 \end{bmatrix},
$$

the columns of which correspond to the monomials $\{x_1, x_1^2, x_1 x_2, x_1^3\}$, i.e. the monomials of $\mathbb{H}_\Phi$, $\{1, x_1, x_2, x_1^2\}$, multiplied by x_1.

We assume for the moment that all the elements of the matrices $\mathbb{H}_{x_1 \star \Phi}$ and $\mathbb{H}_\Phi$ are known. If this is not the case, then we can compute the unknown entries of the matrix, using necessary and sufficient conditions of the quotient algebra; it holds that $\mathbb{M}_{x_i} \mathbb{M}_{x_j} - \mathbb{M}_{x_j} \mathbb{M}_{x_i} = \mathbb{O}$. We refer the reader to [12] for details.

(5) We solve the equation $(\mathbb{H}_{x_1 \star \Phi} - \zeta \, \mathbb{H}_\Phi)\mathbf{v} = 0$.

We solve the generalized eigenvalue/eigenvector problem using one of the well-known techniques [30]. We multiply the (generalized) eigenvectors by $\mathbb{H}_\Phi$ and we normalize the resulting vectors so that the first element is 1, and we read the solutions from the coordinates of the (normalized) eigenvectors, according to Th. 3.3. In our example the normalized eigenvectors are

$$
\begin{bmatrix} 1 \\ -12 \\ -3 \\ 144 \end{bmatrix}, \begin{bmatrix} 1 \\ 12 \\ -13 \\ 144 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}
$$

The coordinates of the eigenvectors correspond to the elements $\{1, x_1, x_2, x_1^2\}$. Thus, we can recover the coefficients of x_1 and x_2 in the decomposition from coordinates of the eigenvectors. Thus, The polynomial admits a decomposition $f = \lambda_1(x_0 - 12x_1 - 3x_2)^5 + \lambda_2(x_0 - 2x_1 + 3x_2)^5 + \lambda_3(x_0 + 2x_1 + 3x_2)^5 + \lambda_4(x_0 + 12x_1 - 13x_2)^5$.

It remains to compute λ_i's. We can do this easily by solving an over-constrained linear system, which we know that always has a solution, since the decomposition exists. Doing that, we deduce that $\lambda_1 = 5$, $\lambda_2 = 15$, $\lambda_3 = 15$ and $\lambda_4 = 3$. We obtain the following minimum decomposition of the polynomial as a sum of powers of linear forms: $f = 5(x_0 - 12x_1 - 3x_2)^5 + 15(x_0 - 2x_1 + 3x_2)^5 + 15(x_0 + 2x_1 + 3x_2)^5 + 3(x_0 + 12x_1 - 13x_2)^5$ that is the corresponding tensor is of rank 4.

5. CONCLUSIONS AND FUTURE WORK

We proposed an algorithm for symmetric tensor decomposition, extending the algorithm of Sylvester to dimensions higher than 2. The algorithm decomposes symmetric tensors when the rank is sub-generic and when the decomposition is unique. In order for the algorithm to work for any rank, we should be able to extend the quotient matrix defined in Sec. 3.3. We will report on this in the near future. We are currently working on an efficient `C++` implementation of the algorithm.

We thank the anonymous referees for their comments that helped us improve the paper.

REFERENCES

[1] P. McCullagh, *Tensor Methods in Statistics*, Monographs on Statistics and Applied Probability. Chapman and Hall, 1987.

[2] A. Swami, G. Giannakis, and S. Shamsunder, "Multichannel ARMA processes," *IEEE Trans. Sig. Proc.*, vol. 42, no. 4, pp. 898–913, Apr. 1994.

[3] P. Chevalier, L. Albera, A. Ferreol, and P. Comon, "On the virtual array concept for higher order array processing," *IEEE Trans. Sig. Proc.*, vol. 53, no. 4, pp. 1254–1271, Apr. 2005.

[4] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Sig. Proc.*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[5] P. Chevalier, "Optimal separation of independent narrow-band sources - concept and performance," *Signal Processing*, vol. 73, no. 1, pp. 27–48, Feb. 1999, (special issue on blind separation and deconvolution).

[6] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. on Sig. Proc.*, vol. 48, no. 3, pp. 810–823, 2000.

[7] L. de Lathauwer and J. Castaing, "Tensor-based techniques for the blind separation of ds-cdma signals," *Signal Processing*, vol. 87, no. 2, pp. 322–336, Feb. 2007.

[8] R. Bro, "Parafac, tutorial and applications," *Chemom. Intel. Lab. Syst.*, vol. 38, pp. 149–171, 1997.

[9] H. A. L. Kiers and W. P. Krijnen, "An efficient algorithm for Parafac of three-way data with large numbers of observation units," *Psychometrika*, vol. 56, pp. 147, 1991.

[10] P. Comon, "Independent Component Analysis," in *Higher Order Statistics*, J-L. Lacoume, Ed., pp. 29–38. Elsevier, Amsterdam, London, 1992.

[11] J. F. Cardoso, "Blind signal separation: statistical principles," *Proc. of the IEEE*, vol. 90, pp. 2009–2025, Oct. 1998, special issue, R.W. Liu and L. Tong eds.

[12] P. Comon, B. Mourrain, and E. Tsigaridas, "Symmetric tensor decomposition," Tech. Rep., I3S, Aug 2008.

[13] J. Brachat, P. Comon, B. Mourrain, and E. Tsigaridas, "Symmetric tensor decomposition," Tech. Rep., (submitted), 2009, arXiv:0901.3706.

[14] A. Cichocki and S-I. Amari, *Adaptive Blind Signal and Image Processing*, Wiley, New York, 2002.

[15] A. Smilde, R. Bro, and P. Geladi, *Multi-Way Analysis*, Wiley, 2004.

[16] L. Albera, A. Ferreol, P. Comon, and P. Chevalier, "Blind identification of overcomplete mixtures of sources (BIOME)," *Lin. Algebra Appl.*, vol. 391, pp. 1–30, Nov. 2004.

[17] A. Yeredor, "Non-orthogonal joint diagoinalization in the LS sense with application in blind source separation," *IEEE Trans. Sig. Proc.*, vol. 50, no. 7, pp. 1545–1553, 2002.

[18] L. de Lathauwer, J. Castaing, and J.F. Cardoso, "Fourth-order cumulant-based identification of underdetermined mixtures," *IEEE Trans. Sig. Proc.*, vol. 55, no. 6, pp. 2965–2973, June 2007.

[19] R. A. Harshman and M. Lundy, "PARAFAC: Parallel factor analysis," *Computational Stat. Data Analysis*, vol. 18, pp. 39–72, 1994.

[20] P. Paatero, "The multilinear engine: A table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model," *Journal of Computational and Graphical Statistics*, vol. 8, no. 4, pp. 854–888, Dec. 1999.

[21] T. Jiang and N. Sidiropoulos, "Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models," *IEEE Trans. Sig. Proc.*, vol. 52, no. 9, pp. 2625–2636, Sept. 2004.

[22] J. Kung and G. Rota, "The invariant theory of binary forms," *Bulletin of the American Mathematical Association*, pp. 27–85, 1984.

[23] P. Comon and B. Mourrain, "Decomposition of quantics in sums of powers of linear forms," *Signal Processing*, vol. 53, no. 2-3, pp. 93–107, 1996.

[24] J. J. Sylvester, "Sur une extension d'un théorème de Clebsch relatif aux courbes du quatrième degré," *Comptes Rendus, Math. Acad. Sci. Paris*, vol. 102, pp. 1532–1534, 1886.

[25] P. Comon, G. Golub, L-H. Lim, and B. Mourrain, "Symmetric tensors and symmetric tensor rank," *SIAM Journal on Matrix Analysis Appl.*, vol. 30, no. 3, pp. 1254–1279, 2008.

[26] D. Cox, J. Little, and D. O'Shea, *Using Algebraic Geometry*, Number 185 in Graduate Texts in Mathematics. Springer, New York, 2nd edition, 2005.

[27] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2nd edition, 1997.

[28] M. Elkadi and B. Mourrain, *Introduction á la résolution des systémes polynomiaux*, vol. 59 of *Mathḿatiques et Applications*, Springer, 2007.

[29] G. Comas and M. Seiguer, "On the rank of a binary form," 2001.

[30] G.H. Golub and C.F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 1996.