Exact Line and Plane Search for Tensor Optimization by Global Minimization of Bivariate Polynomials and Rational Functions

Laurent Sorber, Ignat Domanov, Marc Van Barel, *Member, IEEE*, and Lieven De Lathauwer, *Senior Member, IEEE*

***Abstract*—todo**

***Index Terms*—global minimization, bivariate polynomial system, polyanalytic polynomial, exact line search, exact plane search, tensor decomposition, tensor optimization**

I. Introduction

FINDING the real solutions of a bivariate polynomial system is a central problem in robotics [1], computer modeling and graphics [2], [3], computational geometry [4] and numerical optimization [5]. In this article, we propose an efficient and numerically robust algorithm to solve bivariate and polyanalytic polynomial systems with a single generalized eigenvalue decomposition. Compared to existing eigenvalue-based solvers [6]–[8], the proposed algorithm does not require computing the eigenvectors, nor does it require solving any additional eigenvalue problems to recover the solutions. Moreover, accuracy is maximized by means of numerical balancing and Newton refinement, and a robust method to filter out spurious solutions is presented.

Secondly, we develop several types of exact line and plane search algorithms for a broad class of optimization problems, dubbed *tensor optimization* problems, in which the residual function is tensor-valued and multilinear in its variables. In fact, these search algorithms may just as well be applied to the more general class of polynomial optimization problems, in which the objective function and constraints are (possibly polyanalytic) polynomials. However, we restrict our discussion to tensor optimization, which is interesting in and of itself and comprises applications ranging from tensor decompositions to least squares support tensor machines [9]. We show that their line and plane search subproblems are tantamount to solving a bivariate or polyanalytic polynomial system. Finally, we apply the proposed polynomial system solver to compute the global minimizer of the line and plane search subproblems in the context of blind separatation of multi-user DS-CDMA signals using the rank-$(L, L, 1)$ block term decomposition.

Manuscript received March 1, 2013.

L. Sorber and Marc Van Barel are with NALAG, Department of Computer Science, KU Leuven, Celestijnenlaan 200A, BE-3001 Leuven, Belgium (Laurent.Sorber@cs.kuleuven.be, Marc.VanBarel@cs.kuleuven.be).

I. Domanov and L. De Lathauwer are with the Group Science, Engineering and Technology, KU Leuven Kulak, E. Sabbelaan 53, BE-8500 Kortrijk, Belgium (Ignat.Domanov@kuleuven-kulak.be, Lieven.DeLathauwer@kuleuven-kulak.be).

L. Sorber, I. Domanov and L. De Lathauwer are also with iMinds, Future Health Department and SCD-SISTA, Department of Electrical Engineering (ESAT), KU Leuven, Kasteelpark Arenberg 10, BE-3001 Leuven, Belgium.

The paper is organized as follows. In the next subsection we review our notation and introduce some basic definitions. Aanvullen

A. Notation and preliminaries

A tensor is an element of a tensor product of vector spaces. In this article, we refer to a tensor represented as a multidimensional array, given a choice of bases for each of these vector spaces. The order, or the number of modes, of a tensor is the number of indices associated with each element of that tensor. Vectors are denoted by boldface letters and are lower case, e.g., $\boldsymbol{a}$. Matrices are usually denoted by capital letters, e.g., A. Higher-order tensors are denoted by Euler script letters, e.g., $\mathcal{A}$. An entry of a vector $\boldsymbol{a}$, matrix A or tensor $\mathcal{A}$ is denoted by a_i, a_{ij} or $a_{ijk\ldots}$, depending on the number of modes. A colon is used to select all entries of a mode. For instance, $\boldsymbol{a}_{:j}$ corresponds to the jth column of a matrix A. When there is no confusion, we also use $\boldsymbol{a}_j$ to denote the jth column of the matrix A. Sequences are denoted by a superscript in parentheses, e.g., $\{A^{(n)}\}_{n=1}^{N}$. The superscripts $\cdot^{\mathrm{T}}$, $\cdot^{\mathrm{H}}$, $\cdot^{-1}$ and $\cdot^{\dagger}$ are used for the transpose, Hermitian conjugate, matrix inverse and Moore–Penrose pseudo-inverse, respectively. The complex conjugate is denoted by an overbar, e.g., $\overline{a}$ is the complex conjugate of the scalar a. The inner product of two tensors $\mathcal{T}, \mathcal{U} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is defined as $\langle \mathcal{T}, \mathcal{U} \rangle = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \overline{t}_{i_1 \cdots i_N} u_{i_1 \cdots i_N}$. The Frobenius norm is then denoted by $\|\mathcal{T}\|_{\mathrm{F}} = \sqrt{\langle \mathcal{T}, \mathcal{T} \rangle}$. We use parentheses to denote the concatenation of two or more vectors, e.g., $(\boldsymbol{a}, \boldsymbol{b})$ is equivalent to $\begin{bmatrix} \boldsymbol{a}^{\mathrm{T}} & \boldsymbol{b}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$. The $\mathrm{vec}(\cdot)$ operator represents the column-wise vectorization of matrices and tensors. The $n \times n$ identity matrix is denoted by $\mathbb{I}_n$, and the all-zero and all-one $m \times n$ matrices by $0_{m \times n}$ and $1_{m \times n}$, respectively. The Kronecker product of two matrices A and B is denoted by $A \otimes B$.

We overload notation and represent (analytic) univariate polynomials $p(z)$ by matrices $p \in \mathbb{C}^{1 \times d_p + 1}$ as

$$p(z) := p \begin{bmatrix} 1 & z & \cdots & z^{d_p} \end{bmatrix}^{\mathrm{T}}.$$

Analogously, bivariate polynomials $q(x, y)$ and polyanalytic univariate polynomials $r(z, \overline{z})$ are represented as

$$q(x, y) := \begin{bmatrix} 1 & y & \cdots & y^{d_q^{(y)}} \end{bmatrix} q \begin{bmatrix} 1 & x & \cdots & x^{d_q^{(x)}} \end{bmatrix}^{\mathrm{T}} \text{ and}$$
$$r(z, \overline{z}) := \begin{bmatrix} 1 & \overline{z} & \cdots & \overline{z}^{d_r^{(\overline{z})}} \end{bmatrix} r \begin{bmatrix} 1 & z & \cdots & z^{d_r^{(z)}} \end{bmatrix}^{\mathrm{T}},$$

where $q \in \mathbb{C}^{d_q^{(y)}+1 \times d_q^{(x)}+1}$ and $r \in \mathbb{C}^{d_r^{(\overline{z})}+1 \times d_r^{(z)}+1}$, respectively. If $r(z, \overline{z})$ is real-valued for all $z \in \mathbb{C}$, then r is square and $r = r^{\mathrm{H}}$. Polynomial coefficients can also be extracted from these matrix representations. For instance, the polynomial coefficient of the term $\overline{z}^i$ in the polyanalytic polynomial $r(z, \overline{z})$ is $r_{i+1,:}(z)$, which we also denote more compactly as $r_i(z)$.

II. Exact line and plane search for tensor optimization

A. Tensor optimization

Tensor optimization is a special case of the more general class of polynomial optimization, in which the objective function and constraints are polynomials. Let $\mathcal{T}$ be a given tensor in $\mathbb{C}^{I_1 \times \cdots \times I_N}$ and let $\mathcal{M}(\boldsymbol{z}) : \mathbb{C}^n \to \mathbb{C}^{I_1 \times \cdots \times I_N}$ be a multilinear tensor-valued function. A function is said to be multilinear in its argument $\boldsymbol{z} \in \mathbb{C}^n$ if for all i it is linear in z_i when the remaining variables z_j ($j \neq i$) are fixed. We define the degree $d_{\mathcal{M}}$ of a multilinear map $\mathcal{M}$ as the maximum degree of all univariate polynomials $(\mathcal{M}(x))_{i_1 \ldots i_N}$ obtained by setting $z_i = x$ for all i. Unconstrained tensor optimization problems are then of the form

$$\underset{\boldsymbol{z} \in \mathbb{C}^n}{\text{minimize}} \quad \frac{1}{2}\|\mathcal{M}(\boldsymbol{z}) - \mathcal{T}\|_{\mathrm{F}}^2. \tag{1}$$

Among others, this model comprises the canonical polyadic decomposition (CPD) [10], [11], low multilinear rank approximation or Tucker approximation (LMLRA) [12], [13], block term decompositions (BTD) [14]–[16] and least squares support tensor machines (LS-STM) [9]. For instance, in the case of the CPD, $\boldsymbol{z}$ could be defined as as $(\mathrm{vec}(A^{(1)}), \ldots, \mathrm{vec}(A^{(N)})) \in \mathbb{C}^{\sum I_n R}$ and $\mathcal{M}_{\text{CPD}}(\boldsymbol{z})$ by $(\mathcal{M}_{\text{CPD}}(\boldsymbol{z}))_{i_1 \ldots i_N} = \sum_{r=1}^{R} a_{i_1 r}^{(1)} \cdots a_{i_N r}^{(N)}$, which has degree $d_{\mathcal{M}_{\text{CPD}}} = N$. Data fusion as coupled matrix and tensor factorizations (CMTF) [17] corresponds to a weighted sum of tensor optimization objective functions. Another example is the Gauss–Newton two-dimensional subspace minimization trust region subproblem [5]

$$\begin{aligned} \underset{\alpha,\, \beta \in \mathbb{R}}{\text{minimize}} \quad & \frac{1}{2}\left\| \frac{\partial F(\boldsymbol{z}_k)}{\partial \boldsymbol{z}^T} \cdot (\alpha \Delta \boldsymbol{z}_1 + \beta \Delta \boldsymbol{z}_2) - F(\boldsymbol{z}_k) \right\|^2, \\ \text{subject to} \quad & \|\alpha \Delta \boldsymbol{z}_1 + \beta \Delta \boldsymbol{z}_2\|^2 \leq \delta^2 \end{aligned}$$

where F is a vector-valued residual function, $\boldsymbol{z}_k$ is the current iterate, $\Delta \boldsymbol{z}_1$ and $\Delta \boldsymbol{z}_2$ are two descent directions and δ is the trust region radius. In the context of tensor optimization, the residual function F could be $\mathrm{vec}(\mathcal{M}(\boldsymbol{z}) - \mathcal{T})$. However, we shall see that if F is multilinear, one can solve a plane search variant of (1) instead with little additional effort.

Given two descent directions $\Delta \boldsymbol{z}_1$ and $\Delta \boldsymbol{z}_2$, we refer to the following optimization problems

$$\underset{\alpha}{\text{minimize}} \quad \frac{1}{2}\|\mathcal{M}(\boldsymbol{z} + \alpha \Delta \boldsymbol{z}_1) - \mathcal{T}\|_{\mathrm{F}}^2 \tag{LS}$$

$$\underset{\alpha,\, \gamma}{\text{minimize}} \quad \frac{1}{2}\|\mathcal{M}(\gamma \boldsymbol{z} + \alpha \Delta \boldsymbol{z}_1) - \mathcal{T}\|_{\mathrm{F}}^2 \tag{SLS}$$

$$\underset{\alpha,\, \beta}{\text{minimize}} \quad \frac{1}{2}\|\mathcal{M}(\boldsymbol{z} + \alpha \Delta \boldsymbol{z}_1 + \beta \Delta \boldsymbol{z}_2) - \mathcal{T}\|_{\mathrm{F}}^2 \tag{PS}$$

$$\underset{\alpha,\, \beta,\, \gamma}{\text{minimize}} \quad \frac{1}{2}\|\mathcal{M}(\gamma \boldsymbol{z} + \alpha \Delta \boldsymbol{z}_1 + \beta \Delta \boldsymbol{z}_2) - \mathcal{T}\|_{\mathrm{F}}^2 \tag{SPS}$$

as a line search (LS), scaled line search (SLS), plane search (PS) and scaled plane search (SPS), respectively. We append $\mathbb{R}$ or $\mathbb{C}$ to the optimization problem's acronym to distinguish over which field the solution is sought, e.g., a real line search is denoted by LS-$\mathbb{R}$, while a complex scaled line search would be denoted by SLS-$\mathbb{C}$.

Line, plane and space search problems play an important part in algorithms that solve tensor optimization problems. Many optimization-based algorithms rely on either line or plane searches [18] and other popular algorithms for tensor decompositions such as alternating least squares (ALS) have been combined with line searches to mitigate the effect of so-called swamps [19]–[22]. In particular, LS-$\mathbb{R}$ and LS-$\mathbb{C}$ have received some attention recently [19], [23], [24].

Table I gives an overview of the type of objective function the above search problems correspond to for the two most common choices of the solution field. From the table, it is clear that LS-$\mathbb{R}$ and SLS-$\mathbb{R}$ are easy problems to solve since they both amount to computing the roots of a univariate polynomial. Obtaining the global minimizer of the other four search problems is less straightforward because their objective functions are either bivariate or polyanalytic. The following subsections explain how the search problems can be recast in the form given by Table I. Section III then describes how to obtain the global minimizers of (S)LS-$\mathbb{C}$ and (S)PS-$\mathbb{R}$ as the eigenvalues of a generalized eigenvalue problem (GEP).

TABLE I
THE TYPE OF THE OBJECTIVE FUNCTION EACH OF THE SEARCH PROBLEMS CORRESPOND TO FOR TWO CHOICES OF THE SOLUTION FIELD

Problem \ Field	$\mathbb{R}$	$\mathbb{C}$
LS	degree $2d_{\mathcal{M}}$ analytic univariate polynomial	coordinate degree $d_{\mathcal{M}}$ polyanalytic univariate polynomial
SLS	degree $2d_{\mathcal{M}}$ analytic univariate rational function	coordinate degree $d_{\mathcal{M}}$ polyanalytic univariate rational function
PS	total degree $2d_{\mathcal{M}}$ bivariate polynomial	—
SPS	total degree $2d_{\mathcal{M}}$ bivariate rational function	—

B. (Scaled) Line Search

Since LS-$\mathbb{R}$, LS-$\mathbb{C}$ and SLS-$\mathbb{R}$ are special cases of SLS-$\mathbb{C}$, we first focus on SLS-$\mathbb{C}$. Expand the objective function in (SLS) as

$$\begin{aligned} \frac{1}{2}(\|\mathcal{T}\|_{\mathrm{F}}^2 + \|\mathcal{M}(\gamma \boldsymbol{z} + \alpha \Delta \boldsymbol{z}_1)\|_{\mathrm{F}}^2 - \langle \mathcal{T}, \mathcal{M}(\gamma \boldsymbol{z} + \alpha \Delta \boldsymbol{z}_1)\rangle \\ - \langle \mathcal{M}(\gamma \boldsymbol{z} + \alpha \Delta \boldsymbol{z}_1), \mathcal{T}\rangle) \end{aligned}$$

and use the fact that $\mathcal{M}$ is multilinear with degree $d_{\mathcal{M}}$ to introduce a change of variables $c := \gamma^{d_{\mathcal{M}}}$ and $a := \alpha/\gamma$ so that we have

$$f(a, \overline{a}, c, \overline{c}) := \frac{1}{2}(\|\mathcal{T}\|_{\mathrm{F}}^2 + |c|^2 q(a, \overline{a}) - cp(a) - \overline{cp(a)}),$$

where $p(a) := \langle \mathcal{T}, \mathcal{M}(\boldsymbol{z} + a\Delta\boldsymbol{z}_1)\rangle$ and $q(a,\overline{a}) := \|\mathcal{M}(\boldsymbol{z} + a\Delta\boldsymbol{z}_1)\|_{\mathrm{F}}^2$ are two univariate polynomials. In the case of a line search, $c \equiv \gamma \equiv 1$, $a \equiv \alpha$ and (LS) is equivalent to

$$\underset{\alpha\,\in\,\mathbb{R}}{\text{minimize}} \quad f(\alpha) \tag{LS-$\mathbb{R}$}$$

$$\underset{\alpha\,\in\,\mathbb{C}}{\text{minimize}} \quad f(\alpha,\overline{\alpha}) \tag{LS-$\mathbb{C}$}$$

for LS-$\mathbb{R}$ and LS-$\mathbb{C}$, respectively. In the case of a scaled line search, notice that f is quadratic in c so that a global optimum (a^*, c^*) must satisfy $c^* = \mathrm{Re}\{p(a^*)\}/q(a^*)$ and $c^* = \overline{p(a^*)}/q(a^*,\overline{a}^*)$ for SLS-$\mathbb{R}$ and SLS-$\mathbb{C}$, respectively. Subsituting back in f reveals that a^* is the solution of

$$\underset{a\,\in\,\mathbb{R}}{\text{minimize}} \quad -\frac{\mathrm{Re}\{p(a)\}^2}{q(a)} \tag{SLS-$\mathbb{R}$}$$

$$\underset{a\,\in\,\mathbb{C}}{\text{minimize}} \quad -\frac{|p(a)|^2}{q(a,\overline{a})} \tag{SLS-$\mathbb{C}$}$$

for SLS-$\mathbb{R}$ and SLS-$\mathbb{C}$, respectively.

C. (Scaled) Plane Search

Analogously to the (scaled) line search, expand the objective function in (SPS) as

$$\begin{aligned}\frac{1}{2}(\|\mathcal{T}\|_{\mathrm{F}}^2 + \|\mathcal{M}(\gamma\boldsymbol{z} + \alpha\Delta\boldsymbol{z}_1 + \beta\Delta\boldsymbol{z}_2)\|_{\mathrm{F}}^2 \\ - \langle \mathcal{T}, \mathcal{M}(\gamma\boldsymbol{z} + \alpha\Delta\boldsymbol{z}_1 + \beta\Delta\boldsymbol{z}_2)\rangle \\ - \langle \mathcal{M}(\gamma\boldsymbol{z} + \alpha\Delta\boldsymbol{z}_1 + \beta\Delta\boldsymbol{z}_2), \mathcal{T}\rangle)\end{aligned}$$

and use the fact that $\mathcal{M}$ is multilinear with degree $d_{\mathcal{M}}$ to introduce a change of variables $c := \gamma^{d_{\mathcal{M}}}$, $a := \alpha/\gamma$ and $b := \beta/\gamma$ so that we have

$$\begin{aligned}f(a,\overline{a},b,\overline{b},c,\overline{c}) := \frac{1}{2}(\|\mathcal{T}\|_{\mathrm{F}}^2 + |c|^2 q(a,\overline{a},b,\overline{b}) \\ - cp(a,b) - \overline{c}\overline{p(a,b)}),\end{aligned}$$

where $p(a,b) := \langle \mathcal{T}, \mathcal{M}(\boldsymbol{z} + a\Delta\boldsymbol{z}_1 + \beta\Delta\boldsymbol{z}_2)\rangle$ and $q(a,\overline{a},b,\overline{b}) := \|\mathcal{M}(\boldsymbol{z} + a\Delta\boldsymbol{z}_1 + \beta\Delta\boldsymbol{z}_2)\|_{\mathrm{F}}^2$ are two bivariate polynomials. In the case of a real plane search, $c \equiv \gamma \equiv 1$, $a \equiv \alpha \in \mathbb{R}$, $b \equiv \beta \in \mathbb{R}$ and (PS) is equivalent to

$$\underset{\alpha,\,\beta\,\in\,\mathbb{R}}{\text{minimize}} \quad f(\alpha,\beta). \tag{PS-$\mathbb{R}$}$$

In the case of a real scaled line search, notice that f is quadratic in c so that a global optimum (a^*, b^*, c^*) must satisfy $c^* = \mathrm{Re}\{p(a^*,b^*)\}/q(a^*,b^*)$. Subsituting back in f reveals that (a^*, b^*) is the solution of

$$\underset{a,\,b\,\in\,\mathbb{R}}{\text{minimize}} \quad -\frac{\mathrm{Re}\{p(a,b)\}^2}{q(a,b)}. \tag{SPS-$\mathbb{R}$}$$

III. Solving systems of bivariate and polyanalytic polynomials

A. As a polynomial eigenvalue problem

The stationary points of the search problems (S)PS-$\mathbb{R}$ and (S)LS-$\mathbb{C}$ are the roots of a bivariate polynomial system

$$\frac{\partial p}{\partial x} = \frac{\partial p}{\partial y} = 0,$$

and of a polyanalytic univariate polynomial system

$$\frac{\partial q}{\partial z} = \frac{\partial q}{\partial \overline{z}} = 0$$

for certain polynomials $p(x,y)$ and $q(z,\overline{z})$, respectively. In the latter, $\frac{\partial}{\partial z}$ ($\frac{\partial}{\partial \overline{z}}$) is a Wirtinger derivative [25], and can be interpreted as a partial derivative w.r.t. z ($\overline{z}$) while treating $\overline{z}$ (z) as constant. There exists a wide range of techniques to solve systems of (analytic) polynomials, none of which are ideal in every situation:

- *Numeric:* techniques based on Newton's method [18] are useful for finding solutions locally, interval methods [26] can suffer from slow convergence and require bounds on the solutions, and various convex relaxation methods such as the semidefinite programming relaxation [27]–[29] generally only find a lower bound on a polynomial objective function.
- *Symbolic:* methods based on Gröbner bases [30], [31] and resultants [32] originate from algebraic geometry and are used to eliminate variables. The former are considered numerically unstable and expensive, and the latter introduce spurious solutions and have historically not been numerically robust either.
- *Hybrid:* homotopy continuation methods [33], [34] combine homotopy methods with numerical continuation methods and are quite efficient, but can suffer from some problems such as path crossing [35].

The Sylvester resultant is a well-known resultant formulation for systems of multivariate polynomials. By repeated application of the resultant, variables may be eliminated one by one. Analogously, the resultant may also be used to eliminate z or $\overline{z}$ from two polyanalytic univariate polynomials as a result of the following lemma.

Lemma 1: Let $f(z,\overline{z})$ and $g(z,\overline{z})$ be polyanalytic univariate polynomials. If $f(z,\overline{z})$ and $g(z,\overline{z})$ have a common zero z^*, then there are polynomials $s(\overline{z})$ and $t(\overline{z})$ satisfying

$$s(\overline{z})f(z^*,\overline{z}) + t(\overline{z})g(z^*,\overline{z}) = 0, \tag{2}$$

where $d_s < d_g^{(\overline{z})}$ and $d_t < d_f^{(\overline{z})}$.

Proof: Write $f(z,\overline{z})$ and $g(z,\overline{z})$ as

$$f(z,\overline{z}) = \sum_{i=0}^{d_f^{(\overline{z})}} f_i(z)\overline{z}^i \quad \text{and} \quad g(z,\overline{z}) = \sum_{i=0}^{d_g^{(\overline{z})}} g_i(z)\overline{z}^i.$$

If $f(z^*,\overline{z}^*) = g(z^*,\overline{z}^*) = 0$ for a certain z^*, then $f(z^*,\overline{z})$ and $g(z^*,\overline{z})$ will have a common factor $h(\overline{z}) := \overline{z} - \overline{z}^*$. We have that $\frac{f(z^*,\overline{z})}{h(\overline{z})}g(z^*,\overline{z}) = \frac{g(z^*,\overline{z})}{h(\overline{z})}f(z^*,\overline{z})$. Set $s(\overline{z}) := g(z^*,\overline{z})/h(\overline{z})$ and $t(\overline{z}) := -f(z^*,\overline{z})/h(\overline{z})$. ∎

The Sylvester matrix of two polyanalytic univariate polynomials $f(z,\overline{z})$ and $g(z,\overline{z})$ is

$$S(f,g) := \left[\begin{array}{ccc|ccc} f_0(z) & & & g_0(z) & & \\ \vdots & \ddots & & \vdots & \ddots & \\ f_{d_f^{(\overline{z})}}(z) & & f_0(z) & g_{d_g^{(\overline{z})}}(z) & & g_0(z) \\ & \ddots & \vdots & & \ddots & \vdots \\ & & f_{d_f^{(\overline{z})}}(z) & & & g_{d_g^{(\overline{z})}}(z) \end{array}\right] \begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix} \cdot \quad (3)$$

$$\underbrace{\qquad\qquad}_{d_g^{(\overline{z})} \text{ columns}} \quad \underbrace{\qquad\qquad}_{d_f^{(\overline{z})} \text{ columns}}$$

By Lemma 1, the Sylvester resultant $r(z) := \det(S(f,g))$ vanishes when z is a common zero of $f(z,\overline{z})$ and $g(z,\overline{z})$. This leaves us with a necessary but not sufficient condition for the common zeros of $f(z,\overline{z})$ and $g(z,\overline{z})$. In practice, many of the resultant's roots may not be solutions of the system $f(z,\overline{z}) = g(z,\overline{z}) = 0$. Assuming we have a method to compute the roots of $r(z)$ and remove those which are not solutions of the corresponding polyanalytic polynomial system, finding the real roots of a bivariate polynomial system $f(x,y) = g(x,y) = 0$ is now also within reach. By encoding real solutions $x, y \in \mathbb{R}$ in the complex field with $z \in \mathbb{C}$ as

$$\begin{aligned} x &\to (z+\overline{z})/2 \\ y &\to i(\overline{z}-z)/2 \end{aligned} \quad (4)$$

and solving the polyanalytic polynomial system $f((z+\overline{z})/2, i(\overline{z}-z)/2) = g((z+\overline{z})/2, i(\overline{z}-z)/2) = 0$, we avoid many of the difficulties and additional complexity with the standard approach of eliminating x or y separately. For example, some other eigenvalue-based methods [6]–[8] project the roots of the system $f(x,y) = g(x,y) = 0$ on the x-coordinates by eliminating y, after which they are lifted to their corresponding y-coordinates. First, the multiplicities of the x-coordinates must be estimated, taking into account that some x-coordinates may have a small imaginary part due to numerical inaccuracies, and then an additional (smaller) eigenproblem must be solved for each of the x-coordinates based on the eigenvectors of the parent eigenproblem. Another approach [36]–[38] is to eliminate both x and y separately and then for each pair of projected coordinates check whether it constitutes a solution of the given system or not.

From here on, we assume the substitution (4) has been carried out for bivariate polynomial systems and focus on computing the roots of a polyanalytic polynomials systems $f(z,\overline{z}) = g(z,\overline{z}) = 0$. One possible approach is to explicitly compute the resultant $r(z)$ as the determinant of a polynomial matrix [39], after which the roots can be obtained as the eigenvalues of its companion matrix. However, it is well known that a polynomial's roots can be very sensitive to small changes in the polynomial's coefficients [40]. A better approach is to compute the resultant's roots implicitly as the solution of a polynomial eigenvalue problem (PEP). When $r(z) = \det(S(f,g)) = 0$, there must exist a nonzero vector $\boldsymbol{x}$ in the kernel of $S(f,g)$:

$$S(f,g)\cdot\boldsymbol{x} = (S_0 + S_1 z + \cdots + S_d z^d)\cdot\boldsymbol{x} = \mathbf{0}, \quad (5)$$

where $S_k \in \mathbb{C}^{n\times n}$ for all k, $n := d_f^{(\overline{z})} + d_g^{(\overline{z})}$ and $d := \max(d_f^{(z)}, d_g^{(z)})$. The PEP (5) can now be linearized into a GEP in several ways, one of which is the pencil (A,B) where

$$A := \begin{bmatrix} -S_{d-1} & \mathbb{I}_n & \\ \vdots & & \ddots \\ -S_1 & & & \mathbb{I}_n \\ -S_0 & & & \end{bmatrix} \text{ and } \quad (6a)$$

$$B := \begin{bmatrix} S_d & & & \\ & \mathbb{I}_n & & \\ & & \ddots & \\ & & & \mathbb{I}_n \end{bmatrix}. \quad (6b)$$

Zeg misschien iets van "generalized companion form". Vermeld complex Newton–Raphson [18], [41].

B. Balancing the system and its associated pencil

todo

C. Filtering spurious solutions

todo

IV. Numerical experiments

A. Bivariate polynomial systems

todo

B. Blind separation of multi-user DS-CDMA signals

todo

V. Conclusion

Acknowledgments

References

[1] K. Zhou and S. I. Roumeliotis, "Multirobot Active Target Tracking With Combinations of Relative Observations," *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 678–695, Aug. 2011.

[2] J. T. Kajiya, "Ray tracing parametric patches," *ACM SIGGRAPH Comp. Graph.*, vol. 16, no. 3, pp. 245–254, Jul. 1982.

[3] L. Gonzales-Vega and I. Necula, "Efficient topology determination of implicitly defined algebraic plane curves," *Comput. Aided Geom. Design*, vol. 19, no. 9, pp. 719–743, Dec. 2002.

[4] C. Bajaj, T. Garrity, and J. Warren, "On the applications of multi-equational resultants," Department of Computer Science, Purdue University, Technical Report 88-826, Nov. 1988.

[5] R. H. Byrd, R. B. Schnabel, and G. A. Schultz, "Approximate solution of the trust region problem by minimization over two-dimensional subspaces," *Math. Program.*, vol. 40, no. 1, pp. 247–263, 1988.

[6] L. Busé, H. Khalil, and B. Mourrain, "Resultant-based methods for plane curves intersection problems," in *Proceedings of the 8th international conference on Computer Algebra in Scientific Computing*, ser. Lecture Notes in Computer Science, vol. 3718. Springer-Verlag, 2005, pp. 75–92.

[7] R. M. Corless, P. M. Gianni, and B. M. Trager, "A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots," in *Proceedings of the International Symposium on Symbolic and Algebraic Computation*. ACM Press, 1997, pp. 133–140.

[8] D. Manolakis and J. W. Demmel, "Algorithms for intersecting parametric and algebraic curves II: Multiple intersections," *Graph. Models and Image Proc.*, vol. 57, no. 2, pp. 81–100, Mar. 1995.

[9] D. Tao, X. Li, X. Wu, W. Hu, and J. Maybank, "Supervised tensor learning," *Knowledge and Information Systems*, vol. 13, no. 1, pp. 1–42, 2007.

[10] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart–Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.

[11] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, no. 1, pp. 84–84, 1970.

[12] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[13] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Sep. 2009.

[14] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — Part I: Lemmas for partitioned matrices," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1022–1032, 2008.

[15] ——, "Decompositions of a higher-order tensor in block terms — Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, 2008.

[16] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms — Part III: Alternating least squares algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1067–1083, 2008.

[17] E. Acar, M. A. Rasmussen, F. Savorani, T. Næs, and R. Bro, "Understanding data fusion within the framework of coupled matrix and tensor factorizations," Nov. 2012, submitted.

[18] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., ser. Springer Series in Operations Research. Springer, 2006.

[19] M. Rajih, P. Comon, and R. A. Harshman, "Enhanced line search: a novel method to accelerate PARAFAC," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1128–1147, 2008.

[20] R. Bro, "Multi-way analysis in the food industry: models, algorithms, and applications," Ph.D. dissertation, University of Amsterdam, 1998.

[21] J. B. Kruskal, R. A. Harshman, and M. E. Lundy, "How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships," in *Multiway data analysis*, R. Coppi and S. Bolasco, Eds. North-Holland Publishing Co., 1989, pp. 115–121.

[22] B. C. Mitchell and D. S. Burdick, "Slowly converging PARAFAC sequences: swamps and two-factor degeneracies," *J. Chem.*, vol. 8, no. 2, pp. 155–168, Mar. 1994. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/cem.1180080207/abstract

[23] D. Nion and L. De Lathauwer, "An enhanced line search scheme for complex-valued tensor decompositions. Application in DS-CDMA," *Signal Processing*, vol. 88, no. 3, pp. 749–755, 2008.

[24] Y. Chen, D. Han, and L. Qi, "New ALS methods with extrapolating search directions and optimal step size for complex-valued tensor decompositions," *IEEE Trans. Sig. Proc.*, vol. 59, no. 12, pp. 5888–5898, Dec. 2011.

[25] R. Remmert, *Theory of Complex Functions*. Springer-Verlag, 1991.

[26] R. E. Moore, *Methods and applications of interval analysis*. SIAM, 1987.

[27] J. B. Lasserre, "Global optimization with polynomials and the problems of moments," *SIAM J. Optim.*, vol. 11, no. 3, pp. 796–817, 2001.

[28] N. Z. Shor, "Class of global minimum bounds of polynomial functions," *Cybern. and Sys. Anal.*, vol. 23, no. 6, pp. 731–734, 1987.

[29] P. A. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," Ph.D. dissertation, California Institute of Technology, 2000.

[30] J.-C. Faugère, "A new efficient algorithm for computing grbner bases without reduction to zero (F5)," in *Proceedings of the 2002 international symposium on Symbolic and algebraic computation*, ser. ISSAC '02. ACM, 2002, pp. 75–83.

[31] B. Buchberger, "Gröbner bases and systems theory," *Multidim. Systems and Signal Proc.*, vol. 12, pp. 223–251, 2001.

[32] I. Z. Emiris and B. Mourrain, "Matrices in elimination theory," *J. Symb. Comput.*, vol. 28, no. 1-2, pp. 3–43, 1999.

[33] T.-Y. Li, "Numerical solution of multivariate polynomial systems by homotopy continuation methods," *Act. Num.*, vol. 6, pp. 399–436, 1997.

[34] J. Verschelde, J. Verlinden, and R. Cools, "Homotopies exploiting newton polytopes for solving sparse polynomial systems," *SIAM J. Numer. Anal.*, vol. 31, pp. 915–930, 1994.

[35] A. P. Morgan, *Polynomial continuation and its relationship to the symbolic reduction of polynomial systems*, ser. Symbolic and Numerical Computation for Artificial Intelligence. Academic Press, New York, 1992, pp. 23–45.

[36] E. Berberich, P. Emeliyanenko, and M. Sagraloff, "An elimination method for solving bivariate polynomial systems: Eliminating the usual drawbacks," in *Proceedings of the Thirteenth Workshop on Algorithm Engineering and Experiments (ALENEX)*, M. Müller-Hannemann and R. Werneck, Eds. San Francisco, CA, USA: SIAM, Jan. 2011, pp. 35–47.

[37] R. Seidel and N. Wolpert, "On the exact computation of the topology of real algebraic curves," in *Proceedings of the twenty-first annual symposium on Computational geometry*, ser. SCG 2005. New York, NY, USA: ACM, 2005, pp. 107–115.

[38] D. I. Diochnos, I. Z. Emiris, and E. P. Tsigaridas, "On the asymptotic and practical complexity of solving bivariate systems over the reals," *J. Symb. Comput.*, vol. 44, no. 7, pp. 818–835, Jul. 2009.

[39] M. Šebek, S. Pejchová, D. Henrion, and H. Kwakernaak, "Numerical methods for zeros and determinant of polynomial matrix," in *Proceedings of the IEEE Mediterranean Symposium on New Directions in Control and Automation*, Chania, Crete, Greece, Jun. 1996, pp. 488–491.

[40] J. H. Wilkinson, "The evaluation of the zeros of ill-conditioned polynomials. part I," *Numer. Math.*, vol. 1, no. 1, pp. 150–166, 1959.

[41] L. Sorber, M. Van Barel, and L. De Lathauwer, "Unconstrained optimization of real functions in complex variables," *SIAM J. Optim.*, vol. 22, no. 3, pp. 879–898, 2012.