

# Structured low-rank approximation as a rational function minimization

Konstantin Usevich and Ivan Markovsky

School of Electronics and Computer Science,  
University of Southampton, Southampton, SO17 1BJ, UK  
(e-mail: {kdu,im}@ecs.soton.ac.uk)

---

**Abstract:** Many problems of system identification, model reduction and signal processing can be posed and solved as a structured low-rank approximation problem. In this paper a reformulation of the structured low-rank approximation problem as minimization of a multivariate rational function is considered. Using two different parametrizations, we show that the problem reduces to optimization over a compact manifold or to a set of optimization problems over bounded domains of Euclidean space. We make a review of methods of polynomial algebra for global optimization of the rational cost function.

---

## 1. INTRODUCTION

An affine matrix structure  $\mathcal{S}(p)$  is an affine map

$$\mathcal{S}(p) = S_0 + \sum_{i=1}^{n_p} S_i p_i \quad (1)$$

from the structure parameter space  $\mathbb{R}^{n_p}$  to the space of matrices  $\mathbb{R}^{m \times n}$ . In this paper we assume that  $m < n$ . The family  $\{\mathcal{S}(p) : p \in \mathbb{R}^{n_p}\}$  is called the  $\mathcal{S}$ -structured matrices. The structured low-rank approximation (SLRA) problem [10, 19] is formulated as follows.

*Problem 1.* (SLRA). For a given  $p \in \mathbb{R}^{n_p}$ , structure  $\mathcal{S}$  and natural number  $r < m$

$$\underset{\Delta p \in \mathbb{R}^{n_p}}{\text{minimize}} \|\Delta p\|_2 \quad \text{subject to} \quad \text{rank } \mathcal{S}(p - \Delta p) \leq r. \quad (2)$$

Many problems in system identification, signal processing and computer algebra can be posed and solved as SLRA problem [10, 19]. A common case of an affine structure is

$$\mathcal{S}(p) = [\mathcal{B}_1(p) | \dots | \mathcal{B}_q(p)]^T, \quad (3)$$

where each block  $\mathcal{B}_1(p), \dots, \mathcal{B}_q(p)$  is block-Toeplitz, block-Hankel, unstructured, or exact. For structure (3) efficient local optimization methods (based on *variable projection* principle) have been designed [8].

The aim of this paper is to use polynomial algebra algorithms to find the global minimum (and local minima) of the SLRA problem. For this reason we show that the variable projection principle leads to minimization of a multivariate rational function. A similar approach has been recently proposed in [14], but it has more indeterminates involved, and therefore higher computational complexity.

### 1.1 Variable projection for SLRA problem

The rank constraint  $\mathcal{S}(p - \Delta p) \leq r$  is equivalent to

$$R\mathcal{S}(p - \Delta p) = 0, \quad \text{for an } R \in \mathbb{R}^{d \times m}, \text{ rank } R = d.$$

where  $d := m - r$ . Hence the SLRA problem can be reformulated as the following double minimization problem.

$$\begin{aligned} & \underset{R: \text{rank } R = d}{\text{minimize}} \min_{\Delta p} \|\Delta p\|_2 \\ & \text{subject to} \quad G(R)\Delta p = s(R), \end{aligned} \quad (4)$$

where  $G(R) \in \mathbb{R}^{nd \times n_p}$  and  $s(R) \in \mathbb{R}^{nd \times 1}$  are given by

$$s(R) := \text{vec}(R\mathcal{S}(p)),$$

$$G(R) := [\text{vec}(RS_1) \ \dots \ \text{vec}(RS_{n_p})],$$

and are linear functions in the elements of  $R$ .

If the assumption  $n_p \geq nd$  holds, the inner minimization problem in (4) is an underdetermined least squares problem. The norm of its *minimum norm solution* [1, Ch. 1] is the following function of  $R$ :

$$\begin{aligned} f(R) &:= \left( \min_{\Delta p} \|\Delta p\|_2 \quad \text{subject to} \quad G(R)\Delta p = s(R) \right) \\ &= \left( \min_{\Delta p} \|\Delta p\|_2 \quad \text{subject to} \quad R\mathcal{S}(p - \Delta p) = 0 \right) \\ &= s(R)^T \Gamma(R)^\dagger s(R), \end{aligned} \quad (5)$$

where

$$\Gamma(R) := G(R)G(R)^T \in \mathbb{R}^{nd \times nd},$$

and  $\Gamma(R)^\dagger$  is the pseudoinverse of  $\Gamma(R)$ . Therefore we can eliminate  $\Delta p$  from (4) and obtain an equivalent problem

$$\underset{R: \text{rank } R = d}{\text{minimize}} f(R), \quad (6)$$

See [8, Ch. 4] or [19, Ch. 3] for more details on equivalence between problems (4) and (6).

### 1.2 SLRA as a rational minimization problem

Since  $G(R)$  is linear in  $R$ , the matrix  $\Gamma(R)$  is a *quadratic polynomial matrix* (having entries which are quadratic functions in elements of  $R$ ). If  $\det(\Gamma(R))$  is not a zero polynomial, then we can define the rational inverse of  $\Gamma(R)$

$$\Gamma(R)^{-1} = \frac{\text{adj}(\Gamma(R))}{\det(\Gamma(R))}, \quad (7)$$

where  $\text{adj}(\Gamma(R))$  is the adjoint polynomial matrix (the matrix of algebraic complements).

If for a fixed  $R$  the matrix  $\Gamma(R)$  is nonsingular, then the pseudoinverse coincides with the polynomial inverse and the cost function (5) is equal to

$$f(R) = \frac{s(R)^T \text{adj}(\Gamma(R)) s(R)}{\det(\Gamma(R))}. \quad (8)$$

*Remark 2.* If  $\det(\Gamma(R)) > 0$  for all matrices  $R$  under consideration, then the SLRA problem is equivalent to minimization of a rational function.

## 2. DIFFERENT PARAMETRIZATIONS OF THE SEARCH SPACE

The cost function (5) is homogeneous in the following sense:  $f(R) = f(UR)$  for any nonsingular matrix  $U \in \mathbb{R}^{d \times d}$ , and therefore  $f(R)$  depends only on subspace of  $\mathbb{R}^m$ , spanned by the rows of  $R$ .

In other words,  $f$  is defined on a *Grassmann manifold*  $\text{Gr}_{\mathbb{R}}(d, m)$  [2, App. C] (the manifold of all  $d$ -dimensional subspaces of  $\mathbb{R}^m$ ). In what follows we consider two parametrizations of the manifold by subsets of the set of  $d \times m$  matrices of full row rank.

### 2.1 Minimization over orthogonal bases

For any subspace of  $\mathbb{R}^m$  there exists an orthonormal basis, hence it is sufficient to consider all  $d \times m$  matrices with orthonormal columns:

$$\mathcal{M}_{\text{ort}} = \{R \in \mathbb{R}^{d \times m} : RR^T = I_d\},$$

and reduce (6) to

$$\underset{R \in \mathcal{M}_{\text{ort}}}{\text{minimize}} f(R). \quad (9)$$

Note that this parametrization is still ambiguous, because for any  $R \in \mathbb{R}^{d \times m}$ ,  $RR^T = I_d$  and  $U \in \mathbb{R}^{d \times d}$ ,  $UU^T = I_d$  holds

$$URR^TU^T = I_d.$$

In particular, optimal  $R$  for problem (9) is not unique.

However,  $\mathcal{M}_{\text{ort}}$  is a submanifold of the  $(dm - 1)$ -sphere in  $\mathbb{R}^{dm}$  defined by

$$\sum_{i,j=1}^{d,m} (R_{ij})^2 = d, \quad (10)$$

and therefore it is compact (*compact Stiefel manifold* [2, App. C]).

If  $\det(\Gamma(R)) \neq 0$  for all  $R \in \mathbb{R}^{d \times m}$  of full row rank, then  $f(R)$  is a continuous and infinitely differentiable rational function on  $\mathcal{M}_{\text{ort}}$ , and therefore it attains a global minimum on  $\mathcal{M}_{\text{ort}}$ . This makes the optimization problem (9) well-defined.

### 2.2 Exploiting input/output partitions

For any  $d \times m$  matrix  $R$  of full row rank one can find a set of  $d$  linearly independent columns. In other words, there exists a permutation matrix  $\Pi$  such that

$$R = [Q \ -P] \Pi,$$

where  $Q \in \mathbb{R}^{d \times r}$  and  $P$  is a  $d \times d$  nonsingular matrix. Therefore the row subspace of  $R$  coincides with the row subspace of the matrix

$$[QP^{-1} \ -I_d] \Pi,$$

and all  $d$ -dimensional subspaces of  $\mathbb{R}^m$  are generated by matrices  $[X \ -I_d] \Pi$ .

*Proposition 3.* The SLRA problem is equivalent to

$$\underset{\Pi}{\text{minimize}} \underset{X \in \mathbb{R}^{d \times r}}{\text{min}} f_{\Pi}(X), \quad (11)$$

where

$$f_{\Pi}(X) := f([X \ -I_d] \Pi).$$

For a fixed  $\Pi$  different matrices of type  $[X \ -I_d] \Pi$  correspond to different subspaces. In other words, the map  $X \mapsto [X \ -I_d]$  represents a chart of the manifold  $\text{Gr}_{\mathbb{R}}(d, m)$  (a standard coordinate chart [2, App. C]).

Moreover, for a fixed  $\Pi$  the inner minimization problem in (11)

$$\underset{X \in \mathbb{R}^{r \times d}}{\text{minimize}} f_{\Pi}(X). \quad (12)$$

is, in fact, a *structured total least-squares* (STLS) problem [8] for the structure  $\Pi \mathcal{S}(p)$ . Indeed

$$[X \ -I_d] \Pi \mathcal{S}(p) = 0 \iff AX^T = B,$$

where matrices  $A \in \mathbb{R}^{n \times r}$  and  $B \in \mathbb{R}^{n \times d}$  defined by

$$[A \ B]^T := \Pi \mathcal{S}(p)$$

The STLS problem has advantage of being an unconstrained optimization problem with fewer optimization variables than the problem (9). However, the STLS problem may be ill-posed (or ill-conditioned), because  $\mathbb{R}^{d \times r}$  is not compact and all global minima of  $f_{\Pi}(X)$  may not be attained (or attained for very large  $X$ ). The following theorem shows that we can restrict the search space of the STLS problems (11) to compact subsets of  $\mathbb{R}^{d \times r}$ .

*Theorem 4.* The SLRA problem is equivalent to

$$\underset{\Pi}{\text{minimize}} \underset{X \in T_{d,r}}{\text{min}} f_{\Pi}(X), \quad (13)$$

where

$$T_{d,r} := \{X \in \mathbb{R}^{d \times r} : |X_{jk}| \leq 2^{j-1}\}.$$

*Proof.* It suffices to show that for any  $m \times d$  matrix  $R$  of full row rank there exist a permutation matrix  $\Pi$  and a nonsingular matrix  $U$ , such that  $UR = [X \ -I_d] \Pi$  and  $X \in T_{d,r}$ .

- (1)  $d = 1$ . Let  $V = [v_1 \ \dots \ v_m]$  be a rank one  $1 \times m$  matrix. Then there exists  $j$  such that  $|v_k| \leq |v_j|$  for any  $k$ ,  $1 \leq k \leq m$ . If we take  $u = -v_j^{-1}$  and  $\Pi$  the permutation matrix interchanging  $j$ -th and  $m$ -th elements, then  $uV\Pi = [x_1 \ \dots \ x_{m-1} \ -1]$  and  $|x_k| \leq 1$ .
- (2)  $d > 1$ . Define  $R_1 = R$ . Consider the last row  $(R_1)_d$  of the matrix  $R_1$ . By the previous case ( $d = 1$ ) there exist  $u_1 \in \mathbb{R}$  and  $\Pi_1$  such that

$$u_1(R_1)_d = [v_{d,1} \ \dots \ v_{d,m-1} \ -1] \Pi_1$$

and  $|v_{d,k}| \leq 1$ . By subtracting the last row of the matrix  $R$  from others we obtain that

$$U_1 R_1 = \left[ \begin{array}{c|c} R_2 & 0 \\ \hline v_{d,1} \ \dots \ v_{d,n-1} & -1 \end{array} \right] \Pi_1, \quad (14)$$

where  $U_1$  is a nonsingular matrix and  $R_2$  is a  $(d-1) \times (m-1)$  matrix of full row rank. Repeating this procedure for  $R_k$ ,  $k = 2, \dots, d-1$ , we obtain that exist  $U_0$  and  $\Pi_0$  such that  $R_0 := U_0 R \Pi_0$  has the form

$$R_0 = \left[ \begin{array}{cccc|cc} v_{1,1} & \dots & v_{1,r} & -1 & 0 & 0 \\ \vdots & & \vdots & \ddots & \ddots & 0 \\ v_{d,1} & \dots & v_{d,r} & \dots & v_{d,m-1} & -1 \end{array} \right], \quad (15)$$

where  $|v_{j,k}| \leq 1$ .

By subtracting a scaled first row of  $R_0$  from each other row of  $R_0$ , we obtain the matrix

$$W_1 R_0 = \left[ \begin{array}{ccc|ccc} v_{1,1} & \dots & v_{1,r} & -1 & 0 & 0 & \dots & 0 \\ v_{2,1,1} & \dots & v_{2,r,1} & 0 & -1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & v_{3,r+1} & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ v_{d,1,1} & \dots & v_{d,r,1} & 0 & v_{d,r+1} & \dots & v_{d,m-1} & -1 \end{array} \right],$$

such that  $|v_{j,k,1}| \leq 2$ . If we repeat this process for second row, third row, etc., we can eliminate with nonsingular transformations all elements under the diagonal of the right block of  $R_0$

$$W_{d-1} \cdots W_2 W_1 R_0 = [X | -I_d],$$

such that  $|X_{jk}| \leq 2^{j-1}$ .  $\square$

*Remark 5.* (Local optimization methods). From the viewpoint of numerical optimization, the problem (13) may seem more difficult, because it involves a combinatorial problem of choosing the best permutation  $\Pi$ . For local optimization with initial approximation  $R_0$  the exhaustive search over all possible permutations, however, can be avoided. The following strategy may be suggested based on Theorem 4.

- For  $R_0$  choose parametrization  $\Pi$ , such that

$$UR_0 = [X | -I_d] \Pi \quad \text{and} \quad X \in T_{d,r}.$$

- Perform local optimization of  $f_\Pi$  until the optimization variable  $X$  significantly exceeds the domain  $T_{d,r}$ . (This can be checked for example by monitoring the size of  $\|X\|_F$ .) In this case, restart the optimization with initial approximation equal to  $R = [X | -I_d] \Pi$ , where  $X$  is the current value of the parameter  $X$ , by choosing a new permutation matrix  $\Pi$ .

### 3. GLOBAL SOLUTION OF THE SLRA PROBLEM

In this section we consider methods of optimization of a rational functions that exploit the algebraic structure of the problem and allow to find a global minimum.

#### 3.1 Overview of optimization methods

*SDP relaxations* Problem (9) or each subproblem of (13) can be reduced to minimization of a rational function  $\frac{Q(z)}{H(z)}$ ,  $z \in \mathbb{R}^M$  subject to polynomial constraints  $C_k(Z) \geq 0$ ,  $1 \leq k \leq s$ . This problem is equivalent to minimization over the set of Borel measures

$$\begin{aligned} & \underset{\mu}{\text{minimize}} \quad \int_{\mathbb{R}^M} Q(Z) d\mu, \text{ subject to} \\ & \int_{\mathbb{R}^M} H(Z) d\mu = 1, \int_{\mathbb{R}^M} C_k(Z) d\mu \geq 0, \end{aligned}$$

where the minimum corresponds to an atomic measure [11]. This problem can be expressed as a generalized moment problem and solved by semidefinite programming (SDP) relaxations of fixed order that involves a truncated generalized moment matrix.

It is known that the relaxations converge to a global minimum [9], however the speed of convergence is not guaranteed. Optimization of a rational function is implemented in package GloptiPoly [11].

*Systems of polynomial equations* The second, more general, approach consists of finding stationary points of the cost function as solutions of polynomial systems of equations.

Consider first the STLS problem (12) for a fixed  $\Pi$ . Since  $f_\Pi(X)$  is a rational function with no singularities, the local

minima of the function are attained at stationary points, i.e. solutions of the system

$$\frac{\partial}{\partial X_{ij}} f_\Pi(X) = 0, \quad \text{for } i = 1, \dots, d, j = 1, \dots, r. \quad (16)$$

Denote

$$f_\Pi(X) = \frac{Q_\Pi(X)}{H_\Pi(X)},$$

where  $H(X) \neq 0$ . Then the system (16) is equivalent to

$$q_{ij}(X) := \frac{\partial Q_\Pi}{\partial X_{ij}}(X) H_\Pi(X) - Q_\Pi(X) \frac{\partial H_\Pi}{\partial X_{ij}}(X) = 0, \quad (17)$$

for  $i = 1, \dots, d, j = 1, \dots, r$ ,

which is a system of polynomial equations. The number of equation coincides with the number of variables  $dr$ .

The problem (9) can be also reduced to finding stationary points, introducing Lagrange multipliers for the constraint  $RR^T = I_d$ . The number of equations in this case is again equal to the number of indeterminates.

#### 3.2 Solving systems of polynomial equations

In what follows, we consider a polynomial system with the number of equations equal to the number of variables

$$\begin{aligned} q_1(z_1, \dots, z_M) &= 0, \\ &\vdots \\ q_M(z_1, \dots, z_M) &= 0, \end{aligned} \quad (18)$$

also called a *complete intersection system*.

For multivariate system of equations it is possible that the system has infinite number of solutions. Most multivariate solvers deal with the case of finite number of solutions and we consider only this case. The number of complex solutions  $N$  (and therefore possible local minima) is bounded by the Bezout bound [7, Ch. 8]:

$$N \leq \prod_{i=1}^M \deg(q_i). \quad (19)$$

*Resultant-based methods* These methods are based on multivariate resultants. They reduce the solution to univariate polynomial equations, however they are applicable only for small degrees or sparse polynomials (see for example [5]). Our polynomials are usually dense and have large degrees, so we do not consider these methods.

*Stetter-Moller matrix methods* These methods are based on the following fact: the coordinates of the roots of the system of equations coincide with the eigenvalues of *multiplication matrices*  $A_{z_k}$ , corresponding to joint eigenspaces of the matrices. The multiplication matrices are representation of multiplication by variables  $z_k$  in a basis of the residue ring  $\mathbb{R}[X]/\mathcal{I}$  of the ideal of the system. The problem therefore reduces to finding multiplication matrices and computing their joint eigenspaces. See [7, Ch. 2] for more details.

The residue ring  $\mathbb{R}[X]/\mathcal{I}$  has dimension equal to  $N$  (the number of complex solutions), and the matrices are sparse  $N \times N$  matrices. In addition, to obtain the multiplication matrices one needs to calculate a Grobner basis. Having a Grobner basis, the matrices can be calculated using, for example, the ApCoCoA package [15].

*Triangular systems and rational univariate representations*  
One can try to find an equivalent to (18) triangular system

$$\begin{aligned} g_1(z_1) &= 0 \\ g_2(z_1, z_2) &= 0 \\ &\vdots \\ g_M(z_1, \dots, z_M) &= 0 \end{aligned}$$

and sequentially eliminate variables, solving at each step polynomial equations in one variable. However, this procedure is numerically inaccurate and is being replaced by the rational univariate representation approach of [4].

Both approaches require computation of a Grobner basis. Efficient computations with Grobner bases, in their turn, involve linear algebra computations with  $N \times N$  matrices. State of art algorithm for the above approaches are implemented as packages for Maple [17].

*Subdivision methods* Note that we are interested to find only real roots of a polynomial system in a bounded box  $[a_1, b_1] \times \dots \times [a_M, b_M]$ . For this particular case one can exploit properties of Bernstein polynomials to efficiently locate real zeros [12].

The complexity of the algorithm depends on the number of coefficients in the Bernstein representation, which can be bounded by  $(\deg(q_i))^M$  for a single polynomial  $q_i$ . The Bernstein subdivision algorithms are implemented in the package SYNAPS [16].

*Homotopy continuation* This method is based on considering a homotopy from a simpler system to the system which is to be solved, and tracking the solutions. In general, this is a merely heuristic method and there is no guarantee of obtaining all the roots.

All the complex solutions are tracked, therefore the complexity depends on  $N$ . This method is implemented in the package PHCPack [3].

*Perturbed systems* For large degree of polynomials the computation of the Grobner basis may become infeasible. In [6] it is proposed to solve a perturbed system already in a form of a Grobner basis

$$\begin{aligned} \lambda z_1^{D+1} + q_1(z_1, \dots, z_M) &= 0 \\ &\vdots \end{aligned} \quad (20)$$

$$\lambda z_M^{D+1} + q_M(z_1, \dots, z_M) = 0$$

where  $D = \max \deg q_i$ . This method is motivated by the problem of minimization of a polynomial  $H(Z)$  of degree  $2d$ , where the perturbation of the system of equations is equivalent to perturbing a polynomial itself

$$H_\lambda(Z) := \lambda(z_1^{2d+2} + \dots + z_M^{2d+2}) + H(Z). \quad (21)$$

The global minimum of  $H(X)$  can be approximated by global minima of  $H_\lambda(X)$  when  $\lambda \rightarrow 0$  (see [6]).

We were not able to find a perturbation of a rational function that leads to a system of the form (20). However, we may perturb the system (17), hoping that the roots of the perturbed system (20) do not diverge too far.

*Complexity of the methods* As we discussed, the complexity of the polynomial system solving usually depends on the number of complex solutions.

Let us first calculate the total degree of the involved polynomials. The polynomial matrix  $\Gamma(R)$  consists of quadratic polynomials, and therefore

$$\deg \det(\Gamma(R)) \leq 2nd.$$

The degree of the numerator is also bounded by  $2nd$ , since

$$\deg \text{adj}(\Gamma(R)) \leq 2(nd-1)$$

and  $\deg s(R) = 1$ . The degree bounds are almost always attained.

Hence, the degrees of polynomial equations (18) is bounded by  $4nd-1$ . The total number of solutions is bounded by  $(4nd-1)^{dr}$  for the STLS problem.

## 4. EXAMPLES

### 4.1 Identification of autonomous systems

SLRA of a Hankel matrix with  $m = r + 1$  rows and rank reduction by 1 is equivalent to identification of an autonomous linear-time-invariant system of order less than or equal to  $r$  [10].

A Hankel matrix is defined by

$$\mathcal{S}(p) = \begin{bmatrix} p_1 & p_2 & \dots & p_n \\ p_2 & p_3 & \dots & p_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ p_m & \dots & \dots & p_{n+p} \end{bmatrix},$$

where  $n_p = m + n - 1$ . The matrices  $S_i$  in (1) are given by

$$(S_i)_{jk} = \begin{cases} 1, & \text{if } j+k = 1-i, \\ 0, & \text{if } j+k \neq 1-i. \end{cases} \quad (22)$$

Due to the constraint  $n_p \geq nd$ , only rank reduction by one is possible, and therefore  $d = 1$ , and  $R = [r_1 \dots r_m]^T$ . Then

$$G(R) = \begin{bmatrix} r_1 & r_2 & \dots & r_m & 0 & \dots & 0 \\ 0 & r_1 & r_2 & \dots & r_m & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & 0 \\ 0 & \dots & 0 & r_1 & r_2 & \dots & r_m \end{bmatrix}, \quad (23)$$

is of full row rank for  $R$  with  $\text{rank} R = 1$  (i.e.  $R \neq 0$ ). Hence  $f(R)$  and  $f_{\Pi}(X)$  are rational functions without singularities.

Our running example is with  $m = 3$  rows and varying  $n_p$ . We denote  $X = [x_1 \ x_2]$  and consider three permutation maps

$$[x_1 \ x_2 \ -1] \Pi_1 = [x_1 \ x_2 \ -1],$$

$$[x_1 \ x_2 \ -1] \Pi_2 = [-1 \ x_1 \ x_2],$$

$$[x_1 \ x_2 \ -1] \Pi_3 = [x_1 \ -1 \ x_2],$$

which exhaust all possible row spaces of full row rank  $1 \times 3$  matrices. By Theorem 4 it is sufficient to find minima of functions  $f_{\Pi_k}$  in the box  $[-1, 1] \times [-1, 1]$ .

(1) *Case of multiple local minima* Consider a time series

$$\begin{aligned} p_k &= s_k + \varepsilon_k, \quad \varepsilon_k \text{ i.i.d. } N(0, \sigma) \\ s_k &= \left(\frac{3}{4}\right)^{k-1} \sin\left(\frac{2\pi(k-1)}{6}\right) + \sigma \varepsilon_k, \end{aligned} \quad (24)$$

with  $n_p = 12$ . In this case  $n = 10$  and polynomials have degree 40, which leads to 1600 possible solutions of the polynomial system. This makes the computation of the Grobner bases prohibitive. Therefore, we compare the approaches that are left:

homotopy continuation, Bernstein subdivision, SDP relaxations and perturbed system.

For the choice  $\sigma = 0.35$ , the 2-norm of the noise  $\varepsilon_k$  is more than two times higher than the variance of the signal  $s_k$ , and multiple local minima of the cost function are likely to exist. We consider a particular realisation, and also introduce rounding.

$$p^{(1)} = [-0.14, 1, 0.21, -0.42, 0.255, -0.62, 0.315, -0.1, -0.2, -0.21, 0.835, 0.005]^T.$$

Experiments show that the global minimum is attained for  $f_{\Pi_1}$  in  $[-1, 1] \times [-1, 1]$ . Table 1 demonstrates the performance of the methods for  $p^{(1)}$ .

Table 1. Performance of the methods in Example 1.

Method	Time (s)	Solution	$f(R)$
SDP, ort	55.6	failed	1.45585
SDP, u, 10	1.9	failed	1.45293
SDP, u, 20	19.3	failed	1.45672
SDP, b, 10	4.2	(-0.83663, -0.96014)	1.45290
PHCPack	373.9	(-0.83661, -0.96015)	1.45290
Perturbed	$> 4 \cdot 10^3$	(-0.20134, 0.13092)	1.45536
SBDV	$< 1$	(-0.83661, -0.96015)	1.45290
STLS	$< 0.01$	(-0.83661, -0.96015)	1.45290

- $SDP_n$  denotes the GloptiPoly method with the order of relaxation  $n$ . The tolerance of the underlying SDP solver (SeDuMi) is set to  $10^{-10}$ . “ort” corresponds to problem (9), “u” — to subproblem (12) for fixed  $\Pi$ , “b” — to subproblem of (13) with box constraints (domain bounded by polynomial inequalities  $x^2 \leq 1$  and  $y^2 \leq 1$ ). If the method fails to extract the solution, an upper bound on the minimum value is displayed.
- Perturbed stands for the solution of (20) by Rational Univariate Representation implementation in Maple.  $\lambda$  is taken to be 4000. (For comparison, the maximal absolute value of the polynomials’ coefficients in the unperturbed system is  $9.9828 \cdot 10^6$  and is attained at the monomial of degree 30.)

The multiplication matrices even for  $\lambda = 4000$  are ill-conditioned. (The maximal absolute value of the coefficients of multiplication matrices is  $8.1851 \cdot 10^{16}$ .)

- For PHCPack and SBDV (Bernstein subdivision solver) the global minimum is computed by selecting the smallest value of the cost function in the box  $[-1, 1] \times [-1, 1]$ .
- STLS denotes the solution of STLS problem by a local optimization method [18] started from unstructured low-rank approximation.

GloptiPoly with default order of relaxation (SDP, 10) and with increased order (SDP, 20) fails to extract the solution. Bounding of domain helps in this case, whereas solving problem (9) fails as well. SYNAPS and PHCPack find the same solution, however the perturbed system approach fails to find the true minimum because it is close to the boundary of the box.

On Fig. 1 on can see that PHCPack and SYNAPS identify correctly all stationary points. The roots of the perturbed system remain unchanged in the vicinity of 0, but are more likely to move or disappear if they are close to the boundary of the box.

(2) *Case of single local minimum,  $n_p = 19$*  Consider the time series (24)

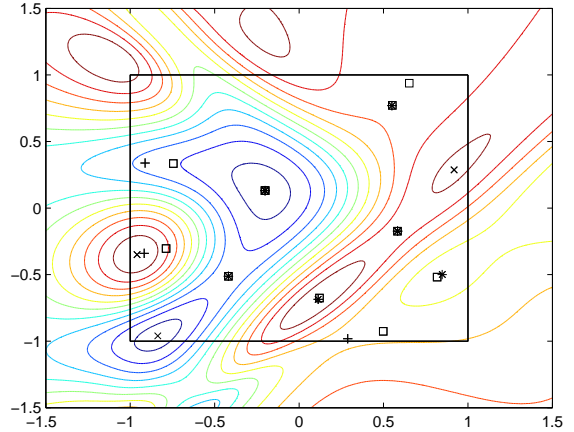


Fig. 1. Cost function for  $p^{(1)}$ , stationary points found by SYNAPS and PHCPack for the original system (x), and stationary points found by SYNAPS for the perturbed system with  $\lambda = 4000$  (+) and  $\lambda = 10^6$  (box)

$$p^{(2)} = [-0.051, 0.570, 0.478, -0.075, -0.348, -0.166, 0.040, 0.068, 0.052, 0.049, -0.071, 0.171, 0.074, -0.115, -0.001, -0.021, -0.012, -0.014, 0.063]^T$$

with  $n_p = 19$ . In this case  $n = 17$  and polynomials have degree at most 68, which leads to 4624 possible solutions. But in this case we add noise with  $\sigma = 0.1$  to ensure that the cost function has exactly one local minimum. The results are shown in Table 2.

Table 2. Performance of the methods in Example 2.

Method	Sec.	Solution	$f(R)$
SDP, 17	8.4	failed	failed
SDP, 24	46.5	failed	failed
SDP, b, 17	26.1	failed	failed
PHCPack	2883.9	(0.74802, 0.87727)	0.74087
SBDV	$< 3$	(-0.55547, 0.63950)	0.07822
STLS	$< 0.01$	(-0.55548, 0.63951)	0.07822

GloptiPoly fails to find the global minimum, reporting “numerical problems”. PHCPack also fails to find the global minimum, possibly due to large number of complex solutions. The Bernstein subdivision solver happened to find the solution with a good accuracy, as well as all stationary points in the box (see Fig. 2).

## 5. CONCLUSIONS

Reformulating the structured low-rank approximation problem as a rational function minimization problem and deriving its optimality conditions, we used polynomial algebra packages to find all local and global minima. Unfortunately, using this procedure, the current state-of-the-art packages one can solve only structured low-rank approximation problem of small dimension, say less than 20 structure parameters. Our experiments also showed that for rational function optimization problem methods that consider only real solutions of the polynomial system are more efficient.

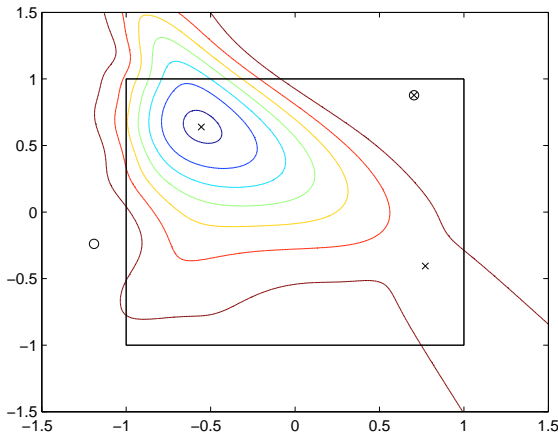


Fig. 2. Cost function for  $p^{(2)}$  and stationary points found by SYNAPS (x) and PHCPack (o)

#### ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement no. 258581 "Structured low-rank approximation: Theory, algorithms, and applications". We also would like to thank Bernard Hanzon for pointing us in direction of this research.

#### REFERENCES

- [1] A. Björk *Numerical Methods for Least Squares Problems*. SIAM, 1996.
- [2] U. Helmke and J. B. Moore *Optimization and Dynamical Systems*, 2nd ed., Springer-Verlag, 1996.
- [3] J. Verschelde Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Trans. on Math. Software* 25, 1999.
- [4] F. Rouillier Solving Zero-Dimensional Systems Through the Rational Univariate Representation. *Appl. Algebra Eng. Commun. Comput.* 9, 1999
- [5] B. Sturmfels *Solving Systems of Polynomial Equations (CBMS Regional Conference Series in Mathematics)*. American Mathematical Society, 2002.
- [6] B. Hanzon, D. Jibeteau Global Minimization of a Multivariate Polynomial using Matrix Methods. *Journal Name: Journal of Global Optimization*, 27:1 1–23, 2003.
- [7] H. J. Stetter *Numerical Polynomial Algebra*. SIAM, 2004.
- [8] I. Markovsky, J. C. Willems, S. Van Huffel, and B. De Moor *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. SIAM, 2006.
- [9] D. Jibeteau, E. De Klerk Global optimization of rational functions: a semidefinite programming approach. *Math. Program.* 106:93–109, 2006.
- [10] I. Markovsky Structured low-rank approximation and its applications. *Automatica*, 44:891–909, 2008.
- [11] D. Henrion, J. B. Lasserre, J. Loeferberg. GloptiPoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software*, 24:–5 761–779, 2009.
- [12] B. Mourrain, J. P. Pavone, Subdivision methods for solving polynomial equations. *Journal of Symbolic Computation*, 44:292–306, 2009,

- [13] I. Bleylvens *Algebraic polynomial system solving and applications*. PhD thesis, Maastricht University, 2010.
- [14] P. Dreesen, B. De Moor, "Back to the Roots - From Polynomial Equations to Linear Algebra". Internal Report 10-191, ESAT-SISTA, Katholieke Universiteit Leuven, 2010. Accepted for publication in 16th IFAC Symposium on System Identification (SYSID 2012).
- [15] *CoCoA: Applied Computations in Commutative Algebra*. Available at <http://www.apcocoa.org/>
- [16] *SYNAPS (SYmbolic Numeric ApplicationS)*. <http://www-sop.inria.fr/galaad/software/synaps/>
- [17] *Maple™* <http://www.maplesoft.com/products/maple/>
- [18] *SLRA package for structured low-rank approximation* <http://github.com/kdu/slra>
- [19] I. Markovsky *Low Rank Approximation: Algorithms, Implementation, Applications*. Springer, 2012.