

Enhanced line search for blind channel identification based on the Parafac decomposition of cumulant tensors

Ignat Domanov and Lieven De Lathauwer

Abstract— In this paper we consider higher-order cumulant based methods for the blind estimation of a single-input single-output finite impulse response system driven by a non-Gaussian signal. This problem can be interpreted as a particular polynomial optimization problem. Using the link between this problem and the parallel factor decomposition of a third-order tensor we present a new representation of the cost function and give an explicit expression for its complex gradient. Then we explore convergence/non convergence of the single-step least-squares algorithm and improve it by enhanced line/plane search procedures.

I. INTRODUCTION

Consider a Single-Input Single-Output (SISO) Finite Impulse Response (FIR) communication channel for which the output signal $y(n)$, after sampling at the symbol rate, is written as follows:

$$y(n) = x(n) + v(n), \quad x(n) = \sum_{l=0}^L h_l s(n-l),$$

where $s(n)$ is the input sequence and $v(n)$ is additive Gaussian noise.

Numerous blind FIR system identification methods have been proposed in the literature. These methods are widely used in signal processing applications such as channel equalization in data communication, time delay estimation, array processing, source separation, etc. An important family of blind equalization algorithms identify a communication channel model based on fitting higher-order cumulants. An interesting property of Higher-Order Statistics (HOS) techniques is that they are insensitive to additive (possibly colored) Gaussian noise. HOS based methods are very useful in dealing with non-Gaussian and/or non-minimum phase linear systems. HOS-based methods pose a nonlinear optimization problem that can be reformulated in multilinear algebra terms as follows [3]:

find the Canonical or Parallel Factor Decomposition (CANDECOMP/PARAFAC) of a third-order tensor (1) composed of fourth-order output cumulant values.

Research supported by: (1) Research Council K.U.Leuven: GOA-Ambiorics, GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, (2) F.W.O.: (a) project G.0427.10N, (b) Research Communities ICCoS, ANMMM and MLDM, (3) the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, “Dynamical systems, control and optimization”, 2007–2011), (4) EU: ERNSI.

I. Domanov and L. De Lathauwer are with the Department of Electrical Engineering - ESAT/SCD, K.U.Leuven, 3001 Leuven, Belgium and with Group Science, Engineering and Technology, K.U.Leuven Campus Kortrijk, 8500 Kortrijk, Belgium
 ignat.domanov@kuleuven-kortrijk.be,
 lieven.delathauwer@kuleuven-kortrijk.be

This third-order tensor has certain symmetry properties, and its factors in the PARAFAC decomposition have a Hankel structure, see Eqs. (polysys–cost function 3) below.

The algorithms used to find the PARAFAC decomposition are most often based on Alternating Least Squares (ALS) initialized by either random values or values calculated by a direct trilinear decomposition based on the generalized eigenvalue problem [2], [4].

ALS has two main drawbacks. First, ALS may take a long time to converge. Second, ALS does not preserve the symmetry properties of the original tensor.

Recently, it has been shown in [7], [8] that ALS-based PARAFAC algorithms can be significantly improved by applying an Enhanced Line Search (ELS) procedure. Namely, new ELS algorithms are less sensitive to local optima and have higher convergence speed. It was also mentioned in [7], [8] that ELS can be combined with any search direction (not necessarily the ALS direction).

On the other hand, a Single-Step Least-Squares (SS-LS) algorithm was proposed to solve (1) [3]. This algorithm preserves the symmetry of the tensor that we need to decompose but it does not necessarily converge monotonically.

We compute the PARAFAC decomposition by means of an ELS algorithm. This method converges monotonically. It preserves the symmetry and the Hankel structure. We derive an explicit solution for the optimal complex step. We compare the computation of the optimal complex step with alternating between updates of the real and imaginary part of the complex step. Moreover, we give a new representation of the cost function and find an explicit expression of its complex gradient. This allows us to design several cheap gradient based optimization algorithms.

Notation:

- $\langle a, b \rangle$ is the scalar product of vectors a and b ;
- $\mathbf{e}_1^{(N)}, \dots, \mathbf{e}_N^{(N)}$ is the canonical basis in \mathbb{C}^N ;
- $V_N \in M^{N \times N}$ is the shift matrix defined by $V_N : \mathbf{e}_N^{(N)} \rightarrow \mathbf{e}_{N-1}^{(N)} \rightarrow \dots \rightarrow \mathbf{e}_1^{(N)} \rightarrow 0$;
- $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^\#$ denote the conjugate, transpose, conjugate transpose and Moore-Penrose pseudoinverse, respectively;
- $A \odot B$ denotes the Khatri-Rao product of matrices A and B : the columns of $A \odot B$ are the tensor products of the corresponding columns of A and B ;
- $E(\cdot)$ denotes the mathematical expectation.

II. PROBLEM FORMULATION

We assume that the output signal $y(n)$ is known. We assume for simplicity that $y(n)$ is zero-mean. For triples

of integers $(\tau_1, \tau_2, \tau_3) \in [-L, L] \times [-L, L] \times [-L, L] =: [-L, L]^3$ define $c_{\tau_1, \tau_2, \tau_3} := \text{cum}[y^*(n), y(n + \tau_1), y^*(n + \tau_2), y(n + \tau_3)]$, where $\text{cum}(y_1, y_2, y_3, y_4)$ denotes the fourth-order cumulant of y_1, y_2, y_3, y_4 [6]:

$$\text{cum}(y_1, y_2, y_3, y_4) := E(y_1^* y_2 y_3^* y_4) - E(y_1^* y_2) E(y_3^* y_4) - E(y_1^* y_3^*) E(y_2 y_4) - E(y_1^* y_4) E(y_2 y_3^*).$$

HOS-based blind channel identification methods are based on the Barlett-Brillinger-Rosenblatt formula [1]:

$$c_{\tau_1, \tau_2, \tau_3} = \gamma_{4,s} \sum_{l=0}^L h_l^* h_{l+\tau_1} h_{l+\tau_2}^* h_{l+\tau_3}, \quad (2)$$

where $(\tau_1, \tau_2, \tau_3) \in [-L, L]^3$ and $\gamma_{4,s}$ is the kurtosis of $s(n)$.

The unknown channel h is defined as the least squares solution of the polynomial system (2). In other words, the goal is to solve the following optimization problem

$$\min_{h \in \mathbb{C}^{L+1}} f(\mathbf{h}), \quad (3)$$

where

$$f(\mathbf{h}) = \sum_{|\tau_1|, |\tau_2|, |\tau_3| < L} |c_{\tau_1, \tau_2, \tau_3} - \gamma_{4,s} \sum_{l=0}^L h_l^* h_{l+\tau_1} h_{l+\tau_2}^* h_{l+\tau_3}|^2. \quad (4)$$

III. ALGORITHMS AND ANALYSIS

The PARAFAC interpretation of (4) was obtained in [3]. We recall the vectorized version of this interpretation:

$$f(\mathbf{h}) := \|\gamma_{4,s} \mathbf{G}(\mathbf{h}) \mathbf{h}^* - \text{vec}(\mathbf{C}_{[1]})\|^2, \quad (5)$$

where

$$\mathbf{G} = \mathbf{G}(\mathbf{h}) = \mathbf{H} \odot \mathbf{H} \odot \mathbf{H}^*,$$

$$\mathbf{H} = \mathbf{H}(\mathbf{h}) = \begin{pmatrix} 0 & 0 & \dots & h_0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & h_0 & \dots & h_{L-1} \\ h_0 & h_1 & \dots & h_L \\ \vdots & \vdots & \ddots & \vdots \\ h_{L-1} & h_L & \dots & 0 \\ h_L & 0 & \dots & 0 \end{pmatrix}$$

and $\text{vec}(\mathbf{C}_{[1]}) \in \mathbb{C}^{(2L+1)^3}$ denotes the vector whose $(2L+1)^2(\tau_1 + L) + (2L+1)(\tau_3 + L) + \tau_2 + L + 1$ coordinate is equal to $c_{\tau_1, \tau_2, \tau_3}$.

Based on representation (5) Fernandes et al. [3] proposed the following algorithm for the minimization of f :

Algorithm 1. (SS-LS algorithm)

1. build $\hat{\mathbf{H}}^{(r-1)} = \mathbf{H}(\hat{\mathbf{h}}^{(r-1)})/\hat{h}_0^{(r-1)}$
2. Compute $\hat{\mathbf{G}}^{(r-1)}$ using

$$\hat{\mathbf{G}}^{(r-1)} = \hat{\mathbf{H}}^{(r-1)} \odot \hat{\mathbf{H}}^{(r-1)} \odot \hat{\mathbf{H}}^{(r-1)*}.$$

3. Minimize the cost function

$$\psi(\mathbf{h}^*, \mathbf{h}^{(r-1)}) = \|\text{vec}(\mathbf{C}_{[1]}) - \gamma_{4,s} \hat{\mathbf{G}}^{(r-1)} \mathbf{h}^*\|^2$$

so that

$$\hat{\mathbf{h}}^{(r)} = (\gamma_{4,s}^{-1} \hat{\mathbf{G}}^{(r-1)})^\# \text{vec}(\mathbf{C}_{[1]})^*.$$

4. Iterate until $\|\hat{\mathbf{h}}^{(r)} - \hat{\mathbf{h}}^{(r-1)}\|/\|\hat{\mathbf{h}}^{(r)}\| \leq \epsilon$.

The SS-LS algorithm is very cheap, but its convergence is not guaranteed. We found that there exist values of $\mathbf{C}_{[1]}$ so that for some initial guesses the algorithm does not converge.

The proofs of the following results strongly exploit the symmetry properties of the cumulant. They are based on representation (5).

Proposition 1. Another representation of the cost function is

$$f(\mathbf{h}) = \gamma_{4,s}^2 \left(\|\mathbf{h}\|^8 + 2 \sum_{k=1}^L |\langle \mathbf{h}, V_{L+1}^k \mathbf{h} \rangle|^4 \right) - 2\gamma_{4,s} \text{vec}(\mathbf{C}_{[1]})^H \mathbf{G}(\mathbf{h}) \mathbf{h}^* + \|\text{vec}(\mathbf{C}_{[1]})\|^2. \quad (6)$$

To describe the critical points of \mathbf{f} we will use the notion of the complex gradient operator $\frac{\partial \mathbf{f}}{\partial \mathbf{h}^*}$, see [5] and references therein. Since \mathbf{f} is a polynomial in \mathbf{h} and \mathbf{h}^* it follows that \mathbf{f} is a real-valued function that is analytic with respect to \mathbf{h} and \mathbf{h}^* . Hence, h is a critical point of \mathbf{f} iff $\frac{\partial \mathbf{f}}{\partial \mathbf{h}^*} = 0$ [5]. Now we are ready to present the expression of the complex gradient of the cost function.

Proposition 2. The complex gradient of cost function \mathbf{f} is

$$\frac{\partial \mathbf{f}}{\partial \mathbf{h}^*} = 4\gamma_{4,s}^2 [\mathbf{G}(\mathbf{h})^H \mathbf{G}(\mathbf{h})]^* \mathbf{h} - 4\gamma_{4,s} [\mathbf{G}(\mathbf{h})^H \text{vec}(\mathbf{C}_{[1]})]^*. \quad (7)$$

Applying Proposition 2 to step 3 of Algorithm 1 we obtain the following result.

Corollary 1. Let Algorithm 1 converge to \mathbf{h}^∞ . Then \mathbf{h}^∞ is proportional to some critical point of \mathbf{f} .

Based on Propositions 1-2 we designed ELS and Enhanced Plane Search (EPS) versions of Algorithm 1 and several gradient optimization algorithms. Our algorithms are cheap and always converge monotonically. We also show that the ELS and EPS algorithms have the same computational cost.

REFERENCES

- [1] D.R. Brillinger and M. Rosenblatt, Computation and interpretation of k-th order spectra, in: B. Harris(Ed), Spectral analysis of time Series, Wiley, New York, USA, 1967, pp. 189-232.
- [2] L. De Lathauwer, A Survey of Tensor Methods, *Proc. of the 2009 IEEE International Symposium on Circuits and Systems (ISCAS 2009)*, 2009, pp 2773-2776.
- [3] C. E. R. Fernandes, G. Favier, J. C. M. Mota, Blind channel identification algorithms based on the Parafac decomposition of cumulant tensors: The single and multiuser cases, *Signal Processing*, vol. 88, 6, 2008, pp 1382-1401.
- [4] T. G. Kolda and B. W. Bader, Tensor Decompositions and Applications, *SIAM Review*, vol. 51, 3, 2009, pp 455-500.
- [5] K. Kreutz-Delgado, The complex gradient operator and the CR calculus, Lecture Supplement ECE275A, 2006, pp 1-74.
- [6] J.M. Mendel, Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications, *Proceedings of the IEEE*, vol.79, No. 3, 1991, pp 278-305.
- [7] D. Nion and L. De Lathauwer, An enhanced line search scheme for complex-valued tensor decompositions. Application in DS-CDMA, *Signal Processing*, vol. 88, 3, 2008, pp 749-755.
- [8] M. Rajih, P. Comon, R.A. Harshman, Enhanced line search: a novel method to accelerate PARAFAC, *SIAM J. Matrix Anal. Appl.*, vol. 30, 3, 2008, pp 1128-1147.