BLIND CHANNEL IDENTIFICATION OF MISO SYSTEMS BASED ON THE CP DECOMPOSITION OF CUMULANT TENSORS

Ignat Domanov, Lieven De Lathauwer

Group Science, Engineering and Technology, KULeuven Campus Kortrijk
Etienne Sabbelaan 53, 8500 Kortrijk, Belgium
ignat.domanov, lieven.delathauwer @kuleuven-kortrijk.be
and
Department of Electrical Engineering (ESAT), SCD, Katholieke Universiteit Leuven
Kasteelpark Arenberg 10, postbus 2440, B-3001 Heverlee (Leuven), Belgium
ignat.domanov, lieven.delathauwer @esat.kuleuven.be

ABSTRACT

We consider a Multiple-Input Single-Output Finite Impulse Response system. We are interested in the identifying channel parameters only from output signal. Our approach is based on fitting higher-order cumulants in a least squares sense which yields a polynomial optimization problem. Using the link between the problem and the Parallel Factor decomposition of a structured third-order tensor we derive a cheap algorithm for identifying the parameters of MISO system.

1. INTRODUCTION

We consider the following multichannel convolutive mixing model:

$$y[n] = \sum_{l=0}^{L} \mathbf{H}[l]\mathbf{s}[n-l] \tag{1}$$

where $\mathbf{s}[n] = (s_1[n],\dots,s_P[n])$ are random source vectors, $\mathbf{H}[n] = (h_{1n}\ \dots\ h_{Pn})$ is the $1\times P$ matrix of impulse responses of the mixing Multiple-Input Single-Output (MISO) channel, $y[n]$ is the output signal. Signals and system are assumed to be complex-valued.

We consider the problem of blind channel identification: identify channel parameters $\{h_{p,l}\}_{p,l=1,0}^{P,L}$ based only on the system output $y[n]$.

There are many papers devoted to the overdetermined case (more sensors than sources). The underdetermined case has only recently been treated, and systems with one single output sensor have received considerably less attention [6, 12].

In this paper (as in [6, 12]) we will concentrate on the MISO case an develop a technique that only makes use of fourth-order statistics. We assume the following:

A1: Each nonobservable discrete input sequence $s_p[n]$, $p\in[1,P]$ is complex-valued, ergodic, stationary, independent and identically distributed (i.i.d.) with symmetric distribution, zero mean, and nonzero kurtosis.

A2: All input signals $s_p[n]$ are mutually independent.

Research supported by: (1) Research Council K.U.Leuven: GOA-Ambiorics, GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT 1/08/23, (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.

A3: The additive noise sequence $v[n]$ is normally distributed with zero mean and unknown autocorrelation function. It is assumed to be statistically independent from $s_p[n]$, $p\in[1,P]$.

A4: The FIR filters representing the channels h_p, $p\in[1,P]$ are assumed to be causal with the same length $L+1$, i.e. $h_{p,l}=0, \forall l\notin[0,L]$, and $h_{p,l}\neq 0$ for $l=0$ and $l=L$.

As an example of an application where these conditions are satisfied one can consider several source signals sharing the same carrier frequency at the neighborhood of the receiving antenna. In this context, the radio communication channel can be modeled as a MISO system where the output signal $y[n]$ is the result of the linear combination of nonobserved input signals $s_p[n]$ filtered by unknown FIR filters $\mathbf{h}_p = (h_{p,0},\dots,h_{p,L})^T$, $p\in[1,P]$. Thus, in the presence of additive noise $v[n]$ model (1) can be rewritten as follows:

$$\begin{cases} y[n] = x_1[n]+\dots+x_P[n]+v[n], \\ x_p[n] = (\mathbf{h}_p * s_p)[n] := \sum\limits_{l=0}^{L} h_{p,l}s_p[n-l] \end{cases}.$$

An interesting property of Higher Order Statistics (HOS) based 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.

For triples of integers $(\tau_1,\tau_2,\tau_3)\in[-L,L]\times[-L,L]\times[-L,L]=:[-L,L]^3$ define

$$c_{\tau_1,\tau_2,\tau_3} := cum[y^*(n), y(n+\tau_1), y^*(n+\tau_2), y(n+\tau_3)], \tag{2}$$

where $cum(y_1,y_2,y_3,y_4)$ denotes the fourth-order cumulant of y_1,y_2,y_3,y_4 [11]:

$$\begin{aligned} cum(y_1,y_2,y_3,y_4) &:= E(y_1^*y_2y_3^*y_4) - E(y_1^*y_2)E(y_3^*y_4) - \\ &E(y_1^*y_3^*)E(y_2y_4) - E(y_1^*y_4)E(y_2y_3^*). \end{aligned} \tag{3}$$

Let us first consider the Single-Input Single-Output (SISO) system. In this case HOS-based blind channel identification methods exploit the following Barlett-Brillinger-Rosenblatt (BBR) formula [2]:

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

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

Most methods for channel identification based on HOS can be divided into three main categories (see [11] for details): closed form solutions, linear algebra solutions and nonlinear optimization solutions. Closed form solutions and linear algebra solutions are based on a given subset of equations of overdetermined system (4). Nonlinear optimization solutions exploit all equations of (4). These solutions yield a polynomial optimization problem and hence are more expensive than the closed form and linear algebra solutions, but are more accurate and robust.

Optimization based methods for the SISO case were considered in [4, 5, 7, 10, 13, 14]. It was shown in [7] that the problem of blind SISO identification can be reformulated as a problem of computing the Parallel Factor (Parafac known also as CP) decomposition of a third-order $(2L+1)\times(2L+1)\times(2L+1)$ tensor composed of fourth-order output cumulant values $\mathscr{C}=(c_{\tau_1,\tau_2,\tau_3})_{\tau_1,\tau_2,\tau_3=-L}^{L}$. Moreover, factors in the Parafac decomposition have a Hankel structure.

This paper is possibly a first attempt to find non-linear optimization solutions for MISO case. Similar to the SISO case [7], we exploit the Parafac interpretation of the BBR formula (see Proposition 2.1 below).

The algorithms used to find the Parafac decomposition are 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 [3, 8].

The ALS method has two main drawbacks. First, it may take a long time to converge. Second, it does not preserve the symmetry properties of the original tensor and therefore cannot be used for solving non-linear optimization problem appearing in the paper.

The contribution of this paper is twofold. First, we present an explicit representation of the complex gradient of the cost function. This result has recently been obtained for SISO case in [4, 5]. Second, we give a fixed point interpretation of the critical points of the cost function and propose the cheap Single-Step Krasnoselskij (SSK) iteration to find the fixed points. The SSK algorithm will depend on a real parameter λ, so we will refer to it as to SSK_λ. Although the the convergence of SSK_λ is not guaranteed for arbitrary λ, simulations indicate that trying different λ the SSK_λ always converges to at least local minima.

Notation:

- $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ 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 Kronecker products of the corresponding columns of A and B;
- $E(\cdot)$ denotes the mathematical expectation.

2. COST FUNCTION AND THE LINK WITH THE PARAFAC DECOMPOSITION.

In this section we will follow the ideas from [7]. Under assumptions A1-A4, the BBR formulae takes the form:

$$c_{\tau_1,\tau_2,\tau_3}=\gamma_{4,s_p}\sum_{p=1}^{P}\sum_{l=0}^{L}h_{p,l}^*h_{p,l+\tau_1}h_{p,l+\tau_2}^*h_{p,l+\tau_3}, \qquad (5)$$

where as before $(\tau_1,\tau_2,\tau_3)\in[-L,L]^3$, γ_{4,s_p} is the kurtosis of $s_p(n)$, and c_{τ_1,τ_2,τ_3} is given by (2)-(3).

The unknown channels $\mathbf{h}_1,\ldots,\mathbf{h}_P$ are defined as the least squares solution of the polynomial system (5). In other words, the goal is to solve the following optimization problem

$$\min_{\mathbf{h}_1,\ldots,\mathbf{h}_P\in\mathbb{C}^{L+1}} f(\mathbf{h}_1,\ldots,\mathbf{h}_P)=\min_{\mathbf{h}\in\mathbb{C}^{(L+1)P}} f(\mathbf{h}), \qquad (6)$$

where

$$\mathbf{h}=(\mathbf{h}_1^T,\ldots,\mathbf{h}_P^T)^T$$

and

$$f(\mathbf{h})=\sum_{|\tau_1|,|\tau_2|,|\tau_3|<L} |c_{\tau_1,\tau_2,\tau_3}-\gamma_{4,s_p}\sum_{p=1}^{P}\sum_{l=0}^{L}h_{p,l}^*h_{p,l+\tau_1}h_{p,l+\tau_2}^*h_{p,l+\tau_3}|^2. \qquad (7)$$

The following Proposition can be easily obtained from the results presented in [7] (as observed in [12]).

Proposition 2.1 *The cost function* (7) *can also be expressed as:*

$$f(\mathbf{h}):=\|\gamma_{4,s_p}\sum_{p=1}^{P}\mathbf{G}(\mathbf{h}_p)\mathbf{h}_p^*-vec(\mathbf{C}_{[1]})\|^2, \qquad (8)$$

where

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

$$\mathbf{H}(\mathbf{h}_p)=\begin{pmatrix} 0 & 0 & \ldots & h_{p,0}\\ \vdots & \vdots & \iddots & \vdots\\ 0 & h_{p,0} & \ldots & h_{p,L_1}\\ h_{p,0} & h_{p,1} & \ldots & h_{p,L}\\ \vdots & \vdots & \iddots & \vdots\\ h_{p,L-1} & h_{p,L} & \ldots & 0\\ h_{p,L} & 0 & \ldots & 0 \end{pmatrix}$$

and $vec(\mathbf{C}_{[1]})\in\mathbb{C}^{(2L+1)^3}$ *denotes the vector whose*

$$(2L+1)^2(\tau_1+L)+(2L+1)(\tau_3+L)+\tau_2+L+1$$

coordinate is equal to c_{τ_1,τ_2,τ_3}.

Formula (7) yields that if $(\mathbf{h}_1,\ldots,\mathbf{h}_P)$ is a global minimizer of f, then for any unit-modulus constants $c_1,\ldots,c_P$, and any permutation $i_1,\ldots,i_P$ of indices $1,\ldots,P$, $(c_1\mathbf{h}_{i_1},\ldots,c_P\mathbf{h}_{i_P})$ is again a global minimizer. We will say that blind MISO identification problem has an essentially unique solution if it is unique up to unit-modulus scaling and permutation. Proposition 2.1 says that formulae (5) can be interpreted as the Parafac decomposition (representation as a sum of $(L+1)P$ rank one terms) of the tensor $\mathscr{C}=(c_{\tau_1,\tau_2,\tau_3})_{\tau_1,\tau_2,\tau_3=-L}^{L}$. This interpretation together with Kruskal's uniqueness condition for Parafac [8] yields that in the MISO case, the blind identification problem has generically an essentially unique solution if $P\le L+1$ (see [12] for details).

3. COMPLEX GRADIENT AND SSK_λ ALGORITHM.

To describe the critical points of f we will use the notion of the complex gradient operator $\frac{\partial f}{\partial \mathbf{h}^*}$, see [9] and references therein. Since f is a polynomial in $\mathbf{h}$ and $\mathbf{h}^*$ it follows that f

is a real-valued function that is analytic with respect to $\mathbf{h}$ and $\mathbf{h}^*$. Hence, $\mathbf{h}$ is a critical point of f if and only if $\frac{\partial f}{\partial \mathbf{h}^*} = 0$ [9]. Now we are ready to present the expression of the complex gradient of the cost function.

The proof of the following results strongly exploits the symmetry properties of the cumulant. It is based on representation (8).

Proposition 3.1 *Let $\mathbf{M}(\mathbf{h})$ be the $P(L+1)\times P(L+1)$ matrix (or $P\times P$ block matrix) defined by*

$$\mathbf{M}(\mathbf{h}) = (\mathbf{G}(\mathbf{h}_i)^H \mathbf{G}(\mathbf{h}_j))_{i,j=1}^P.$$

Then the complex gradient of the cost function $f(\mathbf{h})$ is

$$\frac{\partial f}{\partial \mathbf{h}^*} = 4\gamma_{4,s}^2 \mathbf{M}(\mathbf{h})^* \mathbf{h} - 4\gamma_{4,s} \begin{pmatrix} \mathbf{G}(\mathbf{h}_1)^H vec(\mathbf{C}_{[1]}) \\ \vdots \\ \mathbf{G}(\mathbf{h}_P)^H vec(\mathbf{C}_{[1]}) \end{pmatrix}^* . \quad (9)$$

Let us define the mappings Φ, Φ_λ, and ν as follows

$$\Phi, \Phi_\lambda, \nu : \mathbb{C}^{(L+1)P} \rightarrow \mathbb{C}^{(L+1)P},$$

$$\nu(\mathbf{h}) = \left(\frac{\mathbf{h}_1^T}{\|\mathbf{h}_1\|}, \ldots, \frac{\mathbf{h}_P^T}{\|\mathbf{h}_P\|} \right)^T,$$

$$\Phi(\mathbf{h}) = \frac{1}{\gamma_{4,s}} \left[\mathbf{M}(\nu(\mathbf{h}))^{-1} \begin{pmatrix} \mathbf{G}\left(\frac{\mathbf{h}_1}{\|\mathbf{h}_1\|}\right)^H \mathrm{vec}(\mathbf{C}_{[1]}) \\ \vdots \\ \mathbf{G}\left(\frac{\mathbf{h}_P}{\|\mathbf{h}_P\|}\right)^H \mathrm{vec}(\mathbf{C}_{[1]}) \end{pmatrix} \right]^* ,$$

$$\Phi_\lambda(\mathbf{h}) = \lambda \Phi(\mathbf{h}) + (1-\lambda)\mathbf{h}.$$

For a given $\mathbf{h}_0 \in \mathbb{C}^{(L+1)P}$ consider the sequence of iterates $\{\mathbf{h}^k\}_{k=0}^\infty$ determined by the successive iteration method

$$\mathbf{h}^0,\ \mathbf{h}^1 := \Phi(\mathbf{h}^0), \ldots, \mathbf{h}^k := \Phi(\mathbf{h}^{k-1}), \ldots$$

This sequence is known as Picard iteration (see [1]) and is used for finding the fixed points of Φ (recall that $\mathbf{h}$ is a fixed point for Φ if $\Phi(\mathbf{h}) = \mathbf{h}$). The sequence $\{\mathbf{h}^k\}_{k=0}^\infty$ given by

$$\mathbf{h}^0,\ \mathbf{h}^1 := \Phi_\lambda(\mathbf{h}^0), \ldots, \mathbf{h}^k := \Phi_\lambda(\mathbf{h}^{k-1}), \ldots \quad (10)$$

is called Krasnoselskij iteration (see [1]). It is clear that the mappings Φ and Φ_λ have the same fixed points.

The following Proposition indicates the link between the fixed points of Φ_λ and the critical points of f.

Proposition 3.2 *Let $\lambda \in (0,1]$ and $\mathbf{h}_0 \in \mathbb{C}^{(L+1)P}$. If the Krasnoselskij iteration converges to $\mathbf{h}^\infty := (\mathbf{h}_1^\infty, \ldots, \mathbf{h}_P^\infty)$, then for some positive numbers $d_1, \ldots, d_P$ the vector $(d_1\mathbf{h}_1^\infty, \ldots, d_P\mathbf{h}_P^\infty)$ is a critical point of f.*

Remark 3.3 *Since $\mathbf{G}(d\mathbf{h})(d\mathbf{h})^* = |d|^4 \mathbf{G}(\mathbf{h})(\mathbf{h})^*$, the constants $(d_1, \ldots, d_P)$ can be easily found as a positive solution of the optimization problem*

$$\min_{d_1,\ldots,d_P \in \mathbb{R}^+} \|\gamma_{4,s} \sum_{p=1}^P \mathbf{G}(d_p \mathbf{h}_p^\infty) d_p \mathbf{h}_p^{\infty *} - vec(\mathbf{C}_{[1]})\|^2.$$

Proposition 3.2 yields the following Single Step Krasnoselskij (SSK) algorithm

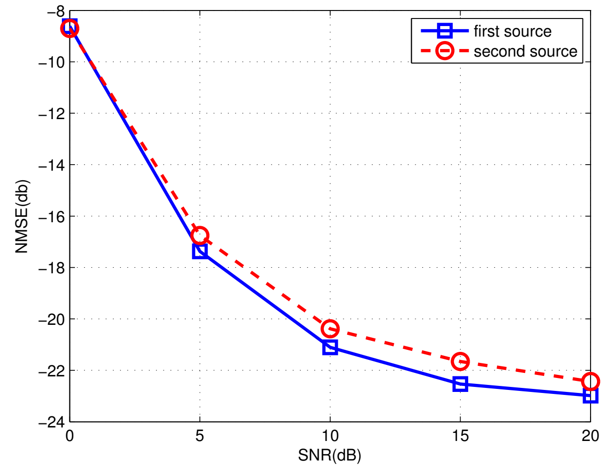

Figure 1: plots of NMSE_1 and NMSE_2

Algorithm 3.4

- *Given :* $\lambda \in (0,1]$, $\mathbf{h}_0 \in \mathbb{C}^{(L+1)P}$, *TOL*, N_{max};
- $k \leftarrow 0$;
- *do until* $\frac{\|h_k - h_{k-1}\|}{\|h_k\|} < TOL$ *or until the maximum number of iterations* N_{max} *is reached*
  - $k \leftarrow k+1$;
  - $\mathbf{h}_k \leftarrow \Phi_\lambda(\mathbf{h}_{k-1})$;
- *end do*

Remark 3.5 *For the SISO case Proposition 3.2 is obtained in [4, 5] where it was used for the analysis of the SS-LS algorithm introduced in [7]. We noted in [5] that for some initializations the SS-LS algorithm does not converge. Moreover, we observed that for some data the SS-LS algorithm does not converge for any initialization: the fixed points of the mapping Φ are unstable. However, the convergence can be significantly improved (see [5]) just by considering the Krasnoselskij version of the SS-LS algorithm with different λ.*

For $\lambda = 1$, the SSK_λ algorithm is a version of the SS-LS algorithm for MISO case. Simulations indicate that the SSK_1 often fails after several iterations because the matrix $\mathbf{M}(\mathbf{h}_k)$ becomes singular. In contrast, the SSK_λ with nonzero λ again has good convergence properties.

To minimize f we propose the following

Algorithm 3.6

1. *Given :* $\mathbf{h}_0 \in \mathbb{C}^{(L+1)P}$, *TOL*, N_{max}, $\lambda = 0.5$;
2. *start Algorithm 3.4*
3. **if** *the maximum number of iterations* N_{max} *is reached in 2.,*

- **then** *repeat step 2. with another random* $\lambda \in (0,1)$
- **else** *stop;*

4. SIMULATIONS.

We considered a scenario with 2 sources. We used channels $\mathbf{h}_1 = (1,\ 0.31 + j0.22,\ 0.57 - j0.29)^T$, $\mathbf{h}_2 = (-0.57 - j0.50,\ -0.26j,\ 0.62 + j0.71)$ with a sample data length $N = 10^4$ and SNR values ranging from 0 to 20 dB. The source

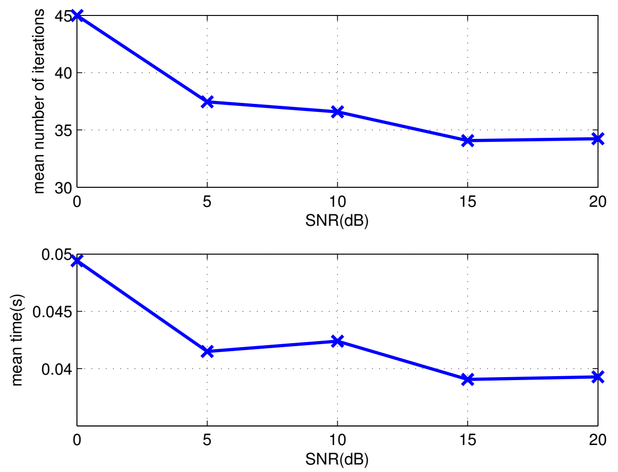

Figure 2: Mean of the number of iterations and mean of the computational time of Algorithm 3.6

signals have been QPSK modulated and R=300 input data blocks have been independently generated. The normalized mean squared error (NMSE) of the results was computed for each channel $p = 1,2$ as follows:

$$NMSE_p = \frac{1}{R}\sum_{r=1}^{R}\frac{\|\hat{\mathbf{h}}_p^r - \mathbf{h}_p\|^2}{\|\mathbf{h}_p\|^2},$$

where $\hat{\mathbf{h}}_p^r$ is the estimated channel vector associated with source p.

The channels have been estimated by Algorithm 3.6. We used $N_{max} = 50$. Fig. 2 demonstrates that Algorithm 3.4 was called only one time in average. We do not compare with the performance of an other algorithm since there does not exist an algorithm that is able to solve the problem in a satisfactory manner.

5. CONCLUSION.

In this paper, we found an optimization solution of the problem of Blind Channel Identification of MISO System. First, we used the link with the Parafac decomposition to find an explicit expression of the complex gradient of the cost function and then, we designed cheap iterative algorithm for identifying the parameters of the MISO system. Although we considered channels with the same length our approach remains valid in general case.

REFERENCES

[1] V. Berinde, "Iterative Approximation of Fixed Points," Springer Verlag, Lectures Notes in Mathematics, 1912, 2007.

[2] D.R. Brillinger, 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.

[3] L. De Lathauwer, "A Survey of Tensor Methods, " in *Proc. of the 2009 IEEE International Symposium on Circuits and Systems (ISCAS 2009)*, Taipei, Taiwan May 24-27. 2009, pp 2773–2776.

[4] I. Domanov, L. De Lathauwer, "Enhanced Line Search for Blind Channel Identification Based on the Parafac Decomposition of Cumulant Tensors, " in *Proc. of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010)*, Budapest, Hungary, Jul. 2010, pp. 1001–1002.

[5] I. Domanov, L. De Lathauwer, "On FIR system identification of SISO systems based only on fourth-order cumulants: cheap optimization algorithm"(in preparation).

[6] C. E. R. Fernandes, P. Comon, and G. Favier, "Blind Identification of MISO-FIR Channels," *Signal Proc.*, 90, pp. 490–503, 2010.

[7] C. E. R. Fernandes, G. Favier, and 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, pp. 1382-1401, 2008.

[8] T. G. Kolda, B. W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, 3, pp. 455-500, 2009.

[9] K. Kreutz-Delgado, "The complex gradient operator and the CR calculus," arXiv Report No. arXiv:0906.4835v1 , available at http://arxiv.org/abs/0906.4835.

[10] Gregory D. Lazear, "Mixed-phase wavelet estimation using fourth-order cumulants," *Geophysics*, v. 58, no. 7, pp. 1042-1051, July 1993.

[11] 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, pp. 278-305, 1991.

[12] Mikael Sørensen, Pierre Comon, and Lieven De Lathauwer, "Cumulant Tensor based Blind Identification of MISO FIR Channels" (in preparation).

[13] J.K. Tugnait, "Identification of linear stochastic systems via second- and fourth- order cumulant matching," *IEEE Trans. Information Theory*, vol. IT-33, pp. 393-407, May 1987.

[14] J.K. Tugnait, "Blind estimation and equalization of digital communication FIR channels using cumulant matching," *IEEE Trans. on Communications*, vol. COM-43, pp. 1240-1245, Part III, Feb/March/April 1995.