Identification of Nonminimum Phase Systems Using Nonlinear Optimization Algorithms and a Joint-Diagonalization Based Method

M. Boulouird ⋆, C. E. R. Fernandes∗•, G. Favier ∗, M. M. Hassani ⋆, and J. C. M. Mota•
∗Laboratoire I3S (CNRS/UNSA)
2000 route des Lucioles, B.P 121, 06903 Sophia Antipolis, France
⋆Laboratoire d'Electronique et Instrumentation LEI (FSSM/UCAM)
Faculty of Sciences Semlalia, B.P 2390, 40001 Marrakesh, Morocco
•Laboratoire GTEL (DETI/UFC)
Campus do Pici, CP 6005-60455-760, Fortaleza-CE, Brazil
m.boulouird@ucam.ac.ma

Abstract

In this paper, we compare different approaches of blind identification of nonminimum phase systems. These approaches are both based on third and fourth-order cumulants. Two nonlinear optimization algorithms, namely the Gradient Descent *and the* Gauss-Newton *algorithms are exposed. An algorithm based on the joint-diagonalization of the fourth-order cumulants matrices (FOSI) is also considered, as well as an improved version of the classical $C(q, k)$ algorithm based on the choice of the* Best 1-D Slice *of fourth-order cumulants. Simulation results are shown to illustrate and compare the considered algorithms.*

Keywords : Blind identification, Higher-Order Statistics, Gradient descent algorithm, Gauss-Newton algorithm, Joint-diagonalization.

1. Introduction

Blind identification of linear systems using Higher-Order Statistics (third and fourth-order cumulants) has a wide applicability in many fields ; e.g., sonar, radar, seismic data processing, adaptive filtering, blind equalization, array processing, data communication, time daily estimation, image and speech processing [9]. These statistics are very useful in problems where either non-Gaussianity, nonminimum phase assumptions, and additive Gaussian noise are present [7].

Signal processing techniques using Higher-Order Statistics (HOS) or cumulants have attracted considerable attention in the literature ([3], [8], [11]). There are several motivations behind this interest [2]. First, higher-order cumulants are blind to all kinds of Gaussian noise ; that is, HOS for a Gaussian process are identically zero. Hence, when the processed signal is non-Gaussian and the additive noise is Gaussian, the noise will vanish in the cumulants domain. Thus, a greater degree of noise immunity is possible. Second, cumulants are useful in identifying non-minimum phase systems and in reconstructing non-minimum phase signals when the signals are non-Gaussian. That is because cumulants preserve the phase information of the signal. Third, cumulants are useful in detecting and characterizing the properties of nonlinear systems.

In this paper, we compare blind identification methods using the nonlinear optimization algorithms proposed in [5], with the well known *Fourth-Order System Identification* algorithm proposed in [4]. The first approach has the advantage of estimating a non redundant parameters vector, while the second one exploits all the fourth-order cumulants through a joint-diagonalization procedure. A third approach consists in selecting the best 1-D slice of fourth-order cumulants in order to improve estimation quality using the classical $C(q, k)$ algorithm [10]. These algorithms are used to identify some communication channels and also solar processes.

This paper is organized as follows : The problem statement is given in Section 2. In Section 3, we expose the solutions using Gradient Descent and Gauss-Newton algorithms. FOSI and modified $C(q, k)$ algorithms are briefly introduced in Sections 4 and 5, respectively. In Section 6, simulation results are discussed. Finally, conclusions are drawn in Section 7.

2. Problem Statement

We consider a discrete time, causal, nonminimum phase linear time-invariant process represented on figure 1 and described by equations (1) and (2),

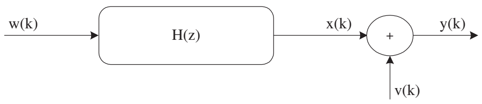

FIG. 1. Single channel system

$$x(k) = \sum_{i=0}^{q} h(i)\, w(k-i); \qquad \{h(0) = 1\} \tag{1}$$

$$y(k) = x(k) + v(k) \tag{2}$$

with the following assumptions :

H.1. The input signal $w(k)$ is a real non measurable sequence, driven from a zero-mean, independent and identically distributed (i.i.d), stationary non-Gaussian process, with unknown distribution, and :

$$C_{m,w}(\tau_1, \tau_2, \ldots, \tau_{m-1}) = \gamma_{m,w}\ \delta(\tau_1, \tau_2, \ldots, \tau_{m-1})$$

where :

- ◇ $C_{m,w}(\tau_1, \tau_2, \ldots, \tau_{m-1})$ is the m**th**-order cumulant of the input signal.
- ◇ $\gamma_{m,w} = C_{m,w}(\underbrace{0, 0, \ldots, 0}_{m-1}) \neq 0, \quad \forall \ \ m \geq 2$
- ◇ $\gamma_{2,w} = \sigma_w^2 = E\left\{w(k)^2\right\}$
- ◇ $\gamma_{3,w} = E\left\{w(k)^3\right\}$ is the skewness of $w(k)$.
- ◇ $\gamma_{4,w} = E\left\{w(k)^4\right\} - 3\left[E\left\{w(k)^2\right\}\right]^2$ is the kurtosis of $w(k)$.

H.2. The additive noise $v(k)$ is assumed to be an i.i.d Gaussian zero-mean sequence with unknown variance, and independent of $w(k)$.

H.3. The order q of the model is assumed to be known.

The objective is to estimate the coefficients $\{h(i)\}_{i=1,\ldots,q}$ (1) from the cumulants of the observations $\{y(k),\ 1 \leq k \leq N\}$ (2).

3. Methods based on nonlinear optimization algorithms

For the nonminimum phase system described by equation (1) with the assumptions **H.1**, **H.2**, and **H.3**, the m**th** and n**th**-order cumulants of the system output (2) are linked by the following relation [1] :

$$\sum_{i=i_{min}}^{i_{max}} h(i) \left[\prod_{k=1}^{m-s-1} h(i+\tau_k)\right] \times$$
$$C_{n,y}(\beta_1, \beta_2, \ldots, \beta_{n-s-1}, i+\alpha_1, i+\alpha_2, \ldots, i+\alpha_s) =$$
$$\frac{\gamma_{n,w}}{\gamma_{m,w}} \sum_{j=j_{min}}^{j_{max}} h(j) \left[\prod_{k=1}^{n-s-1} h(j+\beta_k)\right] \times$$
$$C_{m,y}(\tau_1, \tau_2, \ldots, \tau_{m-s-1}, j+\alpha_1, j+\alpha_2, \ldots, j+\alpha_s) \tag{3}$$

where $m > 2, n > 2$ and s is an arbitrary integer satisfying : $1 \leq s \leq min(m,n) - 2$,

$$and \quad \begin{cases} i_{min} = max(0, -\tau_1, \cdots, -\tau_{m-s-1}) \\ i_{max} = min(q, q-\tau_1, \cdots, q-\tau_{m-s-1}) \\ j_{min} = max(0, -\beta_1, \cdots, -\beta_{n-s-1}) \\ j_{max} = min(q, q-\beta_1, \cdots, q-\beta_{n-s-1}) \end{cases}$$

Setting $n = 3$, $m = 4$, and $s = 1$ in equation (3), yields

$$\sum_{i=i_{min}}^{i_{max}} h(i)h(i+\tau_1)h(i+\tau_2)C_{3,y}(\beta_1, i+\alpha_1) =$$
$$\frac{\gamma_{3,w}}{\gamma_{4,w}} \sum_{j=j_{min}}^{j_{max}} h(j)h(j+\beta_1)C_{4,y}(\tau_1, \tau_2, j+\alpha_1) \tag{4}$$

$$where \quad \begin{cases} i_{min} = max(0, -\tau_1, -\tau_2) \\ i_{max} = min(q, q-\tau_1, q-\tau_2) \\ j_{min} = max(0, -\beta_1) \\ j_{max} = min(q, q-\beta_1) \end{cases}$$

By setting $\tau_1 = \tau_2 = 0$ in (4), we get the relation used in this paper for estimating the parameters $\{h(i)\}_{i=1,2,\ldots,q}$ of the model.

$$\sum_{i=0}^{q} h^3(i) C_{3,y}(\beta_1, i+\alpha_1) = \frac{\gamma_{3,w}}{\gamma_{4,w}}$$
$$\sum_{j=j_{min}}^{j_{max}} h(j)h(j+\beta_1)C_{4,y}(0, 0, j+\alpha_1) \tag{5}$$

It is important to determine the range of values of α_1 and β_1 so that the cumulants $\{C_{3,y}(\beta_1, i+\alpha_1)\}_{i=0,\cdots,q}$, $\{C_{4,y}(0,0,j+\alpha_1)\}_{j=j_{min},\cdots,j_{max}}$, and the coefficients $\{h(j+\beta_1)\}$ are not all zero for each equation.

By taking into account the property of causality of the model and the domain in which third and fourth-order cumulants of a nonminimum phase (q) process are non-zero [7], we obtain :

$$\begin{cases} -q \leq \beta_1 \leq q \\ -2q \leq \alpha_1 \leq q \\ -2q + \beta_1 \leq \alpha_1 \leq q + \beta_1 \end{cases} \tag{6}$$

Using the symmetry properties of cumulants [9], the set of values for α_1 and β_1 is defined by :

$$\begin{cases} -q \leq \beta_1 \leq 0 \\ -2q \leq \alpha_1 \leq q + \beta_1 \end{cases} \tag{7}$$

Concatenating (5) for all the values of α_1 and β_1 defined by (7), we obtain the following system of equations :

$$M\theta = r \tag{8}$$

where :

$$\begin{aligned} \theta = [h(1) \cdots h(q) \quad h^2(1) \quad h(1)h(2) \\ \cdots h(1)h(q) \quad h^2(2) \cdots h(2)h(q) \quad h^2(3) \\ \cdots h^2(q) \quad \epsilon_{4,3} \quad \epsilon_{4,3}h^3(1) \cdots \epsilon_{4,3}h^3(q)]^T \end{aligned} \tag{9}$$

- ◇ $\epsilon_{4,3} = \gamma_{4,w}/\gamma_{3,w}$.
- ◇ M is a matrix of dimension $\left[\frac{5q^2+7q+2}{2}, \frac{q^2+5q+2}{2}\right]$.
- ◇ θ is a vector of dimension $\left[\frac{q^2+5q+2}{2}, 1\right]$.
- ◇ r is a vector of dimension $\left[\frac{5q^2+7q+2}{2}, 1\right]$.

3.1. Gradient Descent Algorithm (GDA)

The idea of this algorithm is to reduce the dimension of the estimated parameter vector θ which has $\left(\frac{q^2+5q+2}{2}\right)$ components, as seen in (9). The new parameters vector θ_{NL} is a $(q+1)$ length vector :

$$\theta_{NL} = [h(1), \cdots, h(q), \epsilon_{4,3}]^T \tag{10}$$

The criterion to be minimized is :

$$J_{LS} = \|r - \phi(\theta_{NL})\|^2$$

The **GDA** solution has the following form :

$$\hat{\theta}^{i+1}_{NL_{gr}} = \hat{\theta}^{i}_{NL_{gr}} + \lambda J^T(r - \phi(\hat{\theta}^{i}_{NL_{gr}})) \tag{11}$$

where :

- ◇ ϕ is the system of equations obtained by concatenating (5) for all the values of α_1 and β_1 defined by (7) :

$$\phi(\theta_{NL}) = M\theta$$

- ◇ J is the Jacobian matrix of ϕ,

$$J = \left[\frac{\partial \phi_k}{\partial \theta_{NL_l}}\right]_{(k,l)}$$

  where $k = 1, \cdots, \frac{5q^2+7q+2}{2}$, and $l = 1, \cdots, q+1$.
- ◇ λ is the step-size.

3.2. Gauss-Newton Algorithm (GNA)

This algorithm can be written as :

$$\hat{\theta}^{i+1}_{NL_{gn}} = \hat{\theta}^{i}_{NL_{gn}} + \mu(J^TJ)^{-1}J^T(r - \phi(\hat{\theta}^{i}_{NL_{gn}})) \tag{12}$$

where :

- ◇ r, ϕ, and J are defined in section 3.1.
- ◇ $\hat{\theta}_{NL_{gn}}$ has the form (10).
- ◇ μ is the step-size.

The parameter $\epsilon_{4,3}$ must be estimated since we suppose we don't know the nature of the distribution of the input signal $w(k)$.

4. A Joint diagonalization-based algorithm

The *Fourth-Order System Identification* (FOSI) algorithm [4] proposes a solution to the blind identification problem based on the joint diagonalization of a set of fourth-order cumulant matrices via a Jacobi technique. The existing relationships between the taps of a nonminimum phase system driven by a non-Gaussian white input, and the (sample) fourth-order cumulant matrices of the output process make possible the recovery of the parameters of the system.

The procedure of joint-diagonalization exploits the fact that any orthonormalized fourth-order cumulant matrix is diagonal in the basis of the columns of a unitary matrix $\mathbf{Q}$, which under certain conditions is unique (up to a permutation matrix and phase factors). Moreover, it is easy to show that the entire set of orthonormalized fourth-order cumulant matrices can be *approximatively* simultaneously diagonalized under the same unitary transformation $\mathbf{Q}$. So, after a preliminary orthonormalization step, a new set of orthonormalized matrices is simultaneously diagonalized, giving rise to the determination of the matrix $\mathbf{Q}$.

The solution of this joint-diagonalization problem is equivalent to the minimization of the following criterion :

$$\phi(\mathbf{Q}, \mathcal{M}) \stackrel{\text{def}}{=} \sum_{k=1}^{K} |\ \text{diag}(\mathbf{Q}^H\bar{\mathbf{M}}(k)\mathbf{Q})\ |^2, \tag{13}$$

where $\mathcal{M} = \{\bar{\mathbf{M}}(k) | k = 1, \ldots, K\}$ is the set of orthonormalized cumulant matrices. The system parameters estimates are obtained from an estimate unitary matrix $\hat{\mathbf{Q}}$ minimizing the criterion (13), plus the orthonormalizing matrix, determined from the eigendecomposition of a positive definite fourth-order cumulant matrix. The amount of fourth-order statistical information required by this method is $(2q+1)^3$.

5. Best 1-D Slice $C(q,k)$ algorithm

The classical $C(q,k)$ algorithm is written as follows [6] :

$$h(k) = \frac{C_{4,y}(q,0,k)}{C_{4,y}(q,0,0)}, \quad k = 1,\ldots,q. \tag{14}$$

This algorithm is very sensitive to cumulants estimation errors and requires exact knowledge about the system order q. Nonetheless, the amount of statistical information required is very small, which makes it a very simple and attractive estimation method. Actually, all the needed information may be arranged into a vector $\mathbf{c}_0$, defined entrywise as $\mathbf{c}_{\tau_2}(k) = C_{4,y}(q,\tau_2,k), k = 1,\ldots,q$, where τ_2 is fixed to zero. Thus, (14) may be rewritten as

$$\mathbf{h} = \mathbf{c}_0/\mathbf{c}_0(0), \tag{15}$$

where $\mathbf{h} = [h(1)\ \ldots\ h(q)]^T$.

We note that it should be possible to change $\mathbf{c}_0$ in (15) by any other $\mathbf{c}_j, j = 1,\ldots,q$ in order to find different parameter estimations $\hat{\mathbf{h}}_j$. Indeed, it is known that the smaller estimation error ($\epsilon_j = |\mathbf{h} - \hat{\mathbf{h}}_j|^2$) is obtained by replacing $\mathbf{c}_0$ in (15) by the vector $\mathbf{c}_\lambda$ with the maximum two-norm ($\max[\mathbf{c}_j^H\mathbf{c}_j]$) [10]. This procedure consists in an improved algorithm that uses $\tau_2 = \lambda$ instead of $\tau_2 = 0$ in (15). The new identification formula is then written as

$$h(k) = \frac{\mathbf{c}_\lambda(k)}{\mathbf{c}_\lambda(0)} = \frac{C_{4,y}(q,\lambda,k)}{C_{4,y}(q,\lambda,0)}, \quad k = 1,\ldots,q. \tag{16}$$

This method makes use of only $(q+1)^2$ statistical information, providing a reduction rate bounded by $8q$ regarding the amount of statistics used by FOSI.

6. Simulations

In the simulations presented in this Section, the available data $\{y(k)\}$ was generated by two different models, shown below. In both models the input signal $w(k)$ is a zero-mean exponentially distributed i.i.d sequence with $\gamma_{2,w} = \sigma_w^2 = 1$ and $\gamma_{3,w} = 2$. The additive noise sequence $\{v(k)\}$ is an i.i.d zero-mean Gaussian sequence.

We carried out 200 Monte Carlo simulations with different noise sequences. For each run, we computed the Normalized Mean Square Error (NMSE) defined as

$$NMSE = \frac{\sum_{i=1}^{q}\left(h(i) - \widehat{h}(i)\right)^2}{\sum_{i=1}^{q} h^2(i)}$$

where $h(i)$ and $\widehat{h}(i)$ are respectively the actual and the estimated impulse responses. The results present the fluctuations of the mean $NMSE$, in dB, against the noise level (SNR).

Model 1 :

$$y(k) = w(k) - 2.333w(k-1) + 0.667w(k-2) + v(k)$$

The zeros of the system transfer function $H(z)$ are located at 1.9994 and 0.3336. This model has also been used in [1], [5], and [10]. In this case $N = 10240$ samples for each run. The simulation results are summarized in figure 2.

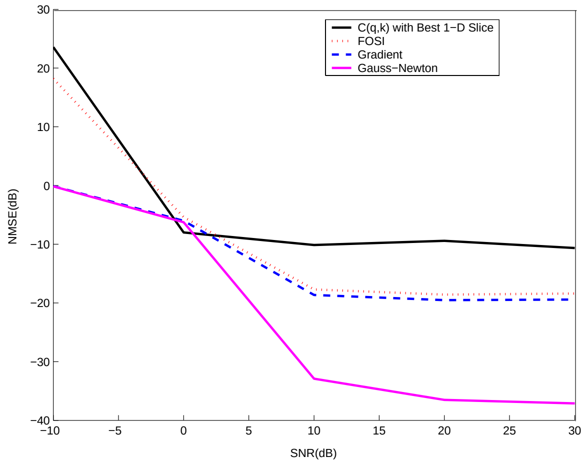

FIG. 2. Performance of the blind identification methods for the model 1

Model 2 :

$$y(k) = w(k) + 0.1w(k-1) - 1.87w(k-2) + 3.02w(k-3) - 1.435w(k-4) + 0.49w(k-5) + v(k)$$

The zeros of the system transfer function $H(z)$ are located at $-2, 0.7 \pm j0.7$ and $0.25 \pm j0.433$. This model has also been used in [5]. In this case $N = 40960$. The simulation results are given in figure 3.

The figures 2 and 3 demonstrate the effectiveness of the first approach, concerning the methods using nonlinear optimization techniques. In figure 3, the Gradient and Gauss-Newton algorithms are much more powerful than FOSI and $C(q,k)$ with Best 1-D Slice. Notice, however, that these latter ones use only fourth-order cumulants, while the nonlinear optimization approaches utilize both third and fourth-order cumulants. We note that increasing model order q severely affects the performance of these algorithms.

7. Conclusion

In this paper, we have compared four different solutions for the problem of blind identification of nonminimum phase systems using third and fourth-order cumulants. In terms of quality of parameter estimation, the

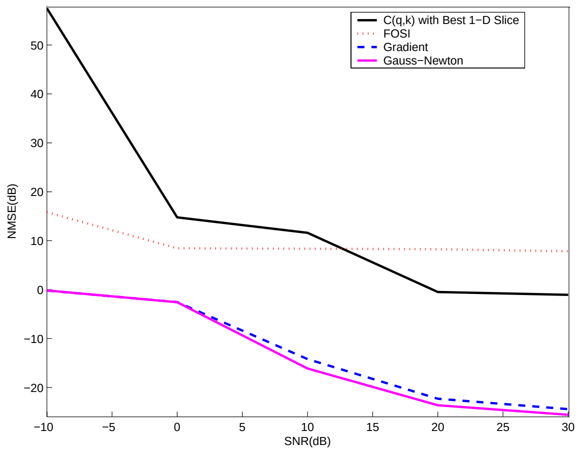

FIG. 3. Performance of the blind identification methods for the model 2

nonlinear optimization-based methods over performed the joint-diagonalization approach as well as the $C(q,k)$ algorithm incorporating Best 1-D Slice, especially for higher order models. Nevertheless, complexity of nonlinear optimization-based algorithms remain much higher comparatively to the other two methods.

References

[1] K. Abderrahim, R. Ben Abdennour, G. Favier, M. Ksouri, and F. Msahli, "New Results on FIR Identification using Cumulants", *Journal Européen des Systèmes Automatisés-JESA*, Vol. 35, No. 5, pp. 601-622, 2001.

[2] A. Al-Smadi and A. Alshamali, "Fitting ARMA Models to Linear non-Gaussian Processes using Higher Order Statistics", *Signal Processing*, Vol. 82, No. 11, pp. 1789-1793, November 2002.

[3] S. A. Alshebeili, A. N. Venetsanopoulos, and A. E. Çetin, "Cumulant Based Identification Approaches for Nonminimum Phase FIR Systems", *IEEE Transactions on Signal Processing*, Vol. 41, No. 4, pp. 1576-1588, April 1993.

[4] A. Belouchrani and B. Derras, "An Efficient Fourth-Order System Identification FOSI Algorithm Utilizing the Joint Diagonalization Procedure", *In Proceedings of the 10-th IEEE Workshop on Statistical Signal and Array Processing*, Pennsylvania, USA, pp. 621–625, August 2000.

[5] M. Boulouird, G. Favier, and M. M. Hassani, "Parameter Estimation of Moving Average Processes using Cumulants and Non-linear Optimization Algorithms", *In Proceedings of the 2nd IFAC International Conference on Informatics in Control, Automation and Robotics-ICINCO'05*, Barcelona, Spain, pp. 11-15, September 14-17, 2005.

[6] G.B. Giannakis, "Cumulants : a Powerful Tool in Signal Processing", *Proceedings of the IEEE*, Vol. 75, No. 9, pp. 1333-1334, September 1987.

[7] 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, March 1991.

[8] Y. J. Na, K. S. Kim, I. Song, and T. Kim, "Identification of Nonminimum Phase FIR Systems Using the Third- and Fourth-Order Cumulants", *IEEE Transactions on Signal Processing*, Vol. 43, No. 8, pp. 2018-2022, December 1995.

[9] C. L. Nikias and J. M. Mendel, "Signal Processing With Higher Order Spectra", *IEEE Signal Processing Magazine*, pp. 10-37, July 1993.

[10] C. E. R. Fernandes, G. Favier, and J. C. M. Mota, "Un critère de sélection des cumulants pour l'identification aveugle de modèles MA à déphasage non-minimal", Submitted to *20th GRETSI Symposium on Signal and Image Processing*, Louvain-la-Neuve, Belgium, September 6-9, 2005.

[11] A. G. Stogioglou and S. McLaughlin, "MA Parameter Estimation and Cumulant Enhancement", *IEEE Transactions on Signal Processing*, Vol. 44, No. 7, pp. 1704-1718, July 1996.