

Laguerre-Volterra Filters Optimization Based on Laguerre Spectra

Alain Y. Kibangou

Laboratoire I3S, CNRS-UNSA, Les algorithmes-Bât. Euclide B, 2000 route des lucioles, BP 121, 06903
Sophia Antipolis Cedex, France
Email: kibangou@i3s.unice.fr

Laboratoire d'Electronique et Instrumentation, FSSM-UCAM, BP 2390, 40000 Marrakesh, Morocco

G rard Favier

Laboratoire I3S, CNRS-UNSA, Les algorithmes-Bât. Euclide B, 2000 route des lucioles, BP 121, 06903
Sophia Antipolis Cedex, France
Email: favier@i3s.unice.fr

Moha M. Hassani

Laboratoire d'Electronique et Instrumentation, FSSM-UCAM, BP 2390, 40000 Marrakesh, Morocco
Email: hassani@ucam.ac.ma

Received 29 July 2004; Revised 1 April 2005; Recommended for Publication by Markus Rupp

New batch and adaptive methods are proposed to optimize the Volterra kernels expansions on a set of Laguerre functions. Each kernel is expanded on an independent Laguerre basis. The expansion coefficients, also called Fourier coefficients, are estimated in the MMSE sense or by applying the gradient technique. An analytical solution to Laguerre poles optimization is provided using the knowledge of the Fourier coefficients associated with an arbitrary Laguerre basis. The proposed methods allow optimization of both the Fourier coefficients and the Laguerre poles.

Keywords and phrases: Volterra filters, Laguerre basis, basis selection, nonlinear system identification.

1. INTRODUCTION

Truncated Volterra filters constitute a class of nonrecursive polynomial models. The main drawback of these models is their over-parameterization. During the last decade, in order to reduce the parametric complexity, that is, the number of parameters, three main approaches have been considered:

- (i) approximations by means of parallel-cascade structures composed of linear filters and memoryless nonlinearities [1];
- (ii) algebraic decompositions of matrices or tensors formed with kernels coefficients [2, 3, 4];
- (iii) expansions on discrete orthonormal bases of functions (OBFs) [5, 6, 7, 8].

The class of OBFs generally used for modeling purposes is that of rational orthonormal bases, such as Laguerre basis [9].

The Laguerre functions have the property of being completely characterized by a single parameter, the Laguerre pole. When expanding a Volterra kernel on a Laguerre basis, the

parsimony of the expansion is strongly linked to the choice of the Laguerre pole. Expansion of Volterra kernels on Laguerre basis was first suggested by Wiener in the 1960s [10]. To the best of our knowledge, Campello et al. [11] were the first to derive an analytical solution to the Laguerre pole optimization for Volterra models. They have generalized the work in [12] and have also shown that using independent bases to expand each kernel gives better results than the use of a single basis. However the obtained analytical solution requires the knowledge of the Volterra kernels. Consequently, a step of Volterra kernels estimation or reconstruction is necessary before computing the Laguerre pole. Note again that this method is applicable if and only if the kernels are strictly causal, that is, the Volterra kernels must satisfy the unit delay condition.

In order to circumvent these limitations, a new approach is proposed in this paper. It is based on the knowledge of the estimated expansion coefficients, also called Fourier coefficients, rather than the Volterra kernels coefficients. The requirement of a unit delay is relaxed. This approach can be viewed as a generalization of both [11, 13].

The organization of the paper is as follows. In the next section, the principle of Volterra model expansion on OBFs is recalled and the analytical expression of the optimal Laguerre pole is given for each kernel. In Section 3, this pole is expressed in terms of the Fourier coefficients relative to the corresponding kernel. Then batch and adaptive identification methods are described in Section 4 and illustrated by means of simulation results in Section 5, before concluding the paper in Section 6.

2. BACKGROUND

A discrete-time P th order Volterra filter is generally described as

$$y(n) = \sum_{p=1}^P \sum_{n_1=0}^{\infty} \cdots \sum_{n_p=0}^{\infty} h_p(n_1, \dots, n_p) \prod_{j=1}^p u(n - n_j), \quad (1)$$

where u , y , and h_p denote respectively the input, the output, and the p th-order Volterra kernel. Boyd and Chua [14] showed that any causal, time invariant, nonlinear system, with fading memory, can be represented to an arbitrary degree of accuracy by a finite expansion in Volterra series.

Considering the expansion of the kernel h_p on an OBF, $\mathfrak{B}_p = \{b_{k,p}\}_{k=0}^{\infty}$, yields

$$h_p(n_1, \dots, n_p) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} g_{k_1, \dots, k_p} b_{k_1, p}(n_1) \cdots b_{k_p, p}(n_p), \quad (2)$$

where the coefficients

$$g_{k_1, \dots, k_p} = \sum_{n_1=0}^{\infty} \cdots \sum_{n_p=0}^{\infty} h_p(n_1, \dots, n_p) b_{k_1, p}(n_1) \cdots b_{k_p, p}(n_p) \quad (3)$$

are called the Fourier coefficients associated with the p th-order kernel. When the used OBFs are Laguerre functions, the set of Fourier coefficients $\{g_{k_1, \dots, k_p}\}$ constitutes the Laguerre spectrum of the p th-order kernel.

Then the input-output equation (1) can be rewritten as

$$y(n) = \sum_{p=1}^P \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} g_{k_1, \dots, k_p} \prod_{j=1}^p s_{k_j, p}(n), \quad (4)$$

where $s_{k_j, p}(n) = \sum_{i=0}^{\infty} b_{k_j, p}(i) u(n - i)$.

The OBFs used in this work are discrete-time Laguerre functions defined by their z -transforms as follows:

$$B_{k,p}(z) = \sqrt{1 - \xi_p^2} \frac{z}{z - \xi_p} \left(\frac{1 - \xi_p z}{z - \xi_p} \right)^k. \quad (5)$$

One can note that only the parameter ξ_p , called the Laguerre pole, characterizes the set of Laguerre functions $\{b_{k,p}(i)\}$.

The derivations presented in this paper are mainly based on the two following properties of Laguerre functions [13, 15, 16]:

$$kb_{k,p}(i) = -\frac{(i+1)\xi_p}{1-\xi_p^2} b_{k,p}(i+1) + \frac{i(1+\xi_p^2) + \xi_p^2}{1-\xi_p^2} b_{k,p}(i) - \frac{i\xi_p}{1-\xi_p^2} b_{k,p}(i-1), \quad (6)$$

$$\frac{\partial b_{k,p}(i)}{\partial \xi_p} = \frac{1}{1-\xi_p^2} ((k+1)b_{k+1,p}(i) - kb_{k-1,p}(i)). \quad (7)$$

In the sequel, the basis associated with each kernel is separately optimized. We define the following cost function:

$$J_p = \frac{1}{p \|h_p\|^2} \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} (k_1 + \cdots + k_p) g_{k_1, \dots, k_p}^2, \quad (8)$$

where $\|h_p\|^2 = \sum_{n_1=0}^{\infty} \cdots \sum_{n_p=0}^{\infty} h_p^2(n_1, \dots, n_p)$. It was shown in [11] that this cost function is an upper bound of the modeling squared error due to the truncation with a finite order of the Laguerre expansion.

By defining

$$Q_{j,p} = \frac{1}{p} \sum_{l=1}^p M_{j,l}, \quad (9)$$

with $j = 1, 2$ and

$$M_{1,l} = \frac{1}{\|h_p\|^2} \sum_{n_1=0}^{\infty} \cdots \sum_{n_p=0}^{\infty} n_l h_p^2(n_1, \dots, n_p),$$

$$M_{2,l} = \frac{1}{\|h_p\|^2} \sum_{n_1=0}^{\infty} \cdots \sum_{n_{l-1}=0}^{\infty} \sum_{n_l=1}^{\infty} \sum_{n_{l+1}=0}^{\infty} \cdots \sum_{n_p=0}^{\infty} n_l h_p(n_1, \dots, n_{l-1}, n_l, n_{l+1}, \dots, n_p) \times h_p(n_1, \dots, n_{l-1}, n_l - 1, n_{l+1}, \dots, n_p), \quad (10)$$

where $l = 1, \dots, p$, the cost function (8) can be explicitly expressed in terms of the Laguerre pole as follows:

$$J_p = \frac{(1 + Q_{1,p})\xi_p^2 - 2Q_{2,p}\xi_p + Q_{1,p}}{1 - \xi_p^2}. \quad (11)$$

This result follows the same steps of calculation as in [11].

Note that the numerator of the cost function is a convex, differentiable, and nonnegative function defined on the open convex set $\mathcal{P} = \{\xi_p \in \mathfrak{X} : |\xi_p| < 1\}$ while its denominator is a concave function, differentiable and positive on \mathcal{P} . Consequently, J_p is a pseudoconvex function inside \mathcal{P} . Then any solution to $\partial J_p / \partial \xi_p = 0$ is a global minimum of J_p [17]. It is straightforward to show that this minimum is reached for

$$\xi_{p,\text{opt}} = \begin{cases} \rho_{o,p} - \sqrt{\rho_{o,p}^2 - 1}, & \text{if } \rho_{o,p} \geq 1, \\ \rho_{o,p} + \sqrt{\rho_{o,p}^2 - 1}, & \text{if } \rho_{o,p} \leq -1, \end{cases} \quad (12)$$

where

$$\rho_{o,p} = \frac{2Q_{1,p} + 1}{2Q_{2,p}}. \quad (13)$$

It should be noted that $\rho_{o,p}$ is a characteristic of the kernel h_p since it depends only on $Q_{j,p}$, $j = 1, 2$, and therefore on the Volterra kernel coefficients $h_p(n_1, \dots, n_p)$.

3. OPTIMAL POLES EXPRESSIONS BASED ON THE LAGUERRE SPECTRA

The optimal pole (12) is related to $\rho_{o,p}$ that depends on the Volterra kernel h_p , which means that it is necessary to carry out an estimation of the kernel before determining the optimal pole. In this section, an expression of $\rho_{o,p}$ in terms of the Laguerre spectrum of the p th-order kernel expanded on any Laguerre basis is investigated. Such an expression will enable us to determine the optimal pole directly from the estimated Laguerre spectrum without computing the Volterra kernel coefficients.

Similarly to definitions (9) and (10) associated with the Volterra kernel coefficients, we define the following quantities that depend on the Laguerre spectrum of the p th-order kernel:

$$R_{j,p} = \sum_{l=1}^p T_{j,l}, \quad (14)$$

with $j = 1, 2$, and

$$T_{1,l} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} (2k_l + 1) g_{k_1, \dots, k_p}^2, \quad (15)$$

$$T_{2,l} = 2 \sum_{k_1=0}^{\infty} \cdots \sum_{k_{l-1}=0}^{\infty} \sum_{k_l=1}^{\infty} \sum_{k_{l+1}=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} k_l g_{k_1, \dots, k_{l-1}, k_l, k_{l+1}, \dots, k_p} \times g_{k_1, \dots, k_{l-1}, k_l-1, k_{l+1}, \dots, k_p}, \quad (16)$$

where $l = 1, \dots, p$.

Example 1. For the quadratic case, that is, $p = 2$, we get

$$\begin{aligned} T_{1,l} &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} (2k_l + 1) g_{k_1, k_2}^2, \quad l = 1, 2, \\ T_{2,1} &= 2 \sum_{k_1=1}^{\infty} \sum_{k_2=0}^{\infty} k_1 g_{k_1, k_2} g_{k_1-1, k_2}, \\ T_{2,2} &= 2 \sum_{k_1=0}^{\infty} \sum_{k_2=1}^{\infty} k_2 g_{k_1, k_2} g_{k_1, k_2-1}, \\ R_{1,2} &= T_{1,1} + T_{1,2}, \quad R_{2,2} = T_{2,1} + T_{2,2}. \end{aligned} \quad (17)$$

Now, the objective is to express $Q_{j,p}$, $j = 1, 2$, as a function of $R_{j,p}$. First of all, one can notice that, as described in Appendix A, the property (7) combined with formula (3) yields

$$\begin{aligned} \frac{\partial}{\partial \xi_p} g_{k_1, \dots, k_p} &= \frac{1}{1 - \xi_p^2} \sum_{l=1}^p ((k_l + 1) g_{k_1, \dots, k_{l-1}, k_l+1, k_{l+1}, \dots, k_p} \\ &\quad - k_l g_{k_1, \dots, k_{l-1}, k_l-1, k_{l+1}, \dots, k_p}). \end{aligned} \quad (18)$$

For example,

$$\begin{aligned} \frac{\partial}{\partial \xi_p} g_{k_1, k_2} &= \frac{1}{1 - \xi_p^2} ((k_1 + 1) g_{k_1+1, k_2} - k_1 g_{k_1-1, k_2} \\ &\quad + (k_2 + 1) g_{k_1, k_2+1} - k_2 g_{k_1, k_2-1}). \end{aligned} \quad (19)$$

Thanks to relation (18), the following lemma is derived.

Lemma 1. $R_{1,p}$ and $R_{2,p}$ are linked by means of their derivatives with respect to ξ_p as follows:

$$\frac{\partial R_{1,p}}{\partial \xi_p} = \frac{-2}{1 - \xi_p^2} R_{2,p}, \quad (20)$$

$$\frac{\partial R_{2,p}}{\partial \xi_p} = \frac{-2}{1 - \xi_p^2} R_{1,p}. \quad (21)$$

The proof of this lemma is based on a simple but long calculation summarized in Appendix B.

By remembering that $\|h_p\|^2$ is a characteristic of the nonlinear system to be identified, and consequently independent of the Laguerre basis, the orthonormality of the Laguerre basis allows to get the following expression: $\|h_p\|^2 = \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} g_{k_1, \dots, k_p}^2$. Then, from the definitions of $R_{1,p}$ and of J_p , a simple calculation yields

$$R_{1,p} = 2p \|h_p\|^2 J_p + p \|h_p\|^2, \quad (22)$$

thus

$$\frac{\partial R_{1,p}}{\partial \xi_p} = 2p \|h_p\|^2 \frac{\partial J_p}{\partial \xi_p}. \quad (23)$$

From (22), one can note that $R_{1,p}$ is strictly positive. Consequently $R_{2,p}$ is strictly decreasing inside \mathcal{P} .

We derive a second lemma proved in Appendix C.

Lemma 2. *The terms $Q_{j,p}$ associated with the Volterra kernel h_p and $R_{j,p}$, $j = 1, 2$, associated with its Laguerre spectrum are linked by*

$$R_{1,p} = \frac{p \|h_p\|^2}{1 - \xi_p^2} (2(1 + \xi_p^2)Q_{1,p} - 4\xi_p Q_{2,p} + 1 + \xi_p^2), \quad (24)$$

$$R_{2,p} = \frac{2p \|h_p\|^2}{1 - \xi_p^2} (-2\xi_p Q_{1,p} + (1 + \xi_p^2)Q_{2,p} - \xi_p). \quad (25)$$

Obviously, $Q_{j,p}$ can be expressed as a function of $R_{j,p}$, $j = 1, 2$, by solving the linear system of (24) and (25):

$$\begin{aligned} Q_{1,p} &= \frac{(1 + \xi_p^2)R_{1,p} + 2\xi_p R_{2,p}}{2(1 - \xi_p^2)p \|h_p\|^2} - \frac{1}{2}, \\ Q_{2,p} &= \frac{2\xi_p R_{1,p} + (1 + \xi_p^2)R_{2,p}}{2(1 - \xi_p^2)p \|h_p\|^2}. \end{aligned} \quad (26)$$

Thus $\rho_{o,p}$, given by (13), can be expressed as a function of $R_{1,p}$ and $R_{2,p}$ that depend only on the Laguerre spectrum as follows:

$$\rho_{o,p} = \frac{(1 + \xi_p^2)R_{1,p} + 2\xi_p R_{2,p}}{2\xi_p R_{1,p} + (1 + \xi_p^2)R_{2,p}}. \quad (27)$$

This result is summarized by the following theorem.

Theorem 1. *The pole of the optimal Laguerre basis, associated with the expansion of the Volterra kernel h_p , is obtained from the Laguerre spectrum, associated with the expansion of the same kernel on an arbitrary Laguerre basis characterized by the pole ξ_p , as follows:*

$$\xi_{p,\text{opt}} = \begin{cases} \rho_{o,p} - \sqrt{\rho_{o,p}^2 - 1}, & \text{if } \rho_{o,p} \geq 1, \\ \rho_{o,p} + \sqrt{\rho_{o,p}^2 - 1}, & \text{if } \rho_{o,p} \leq -1, \end{cases} \quad (28)$$

where $\rho_{o,p}$ is given by the formula (27).

This result is particularly useful. When the Volterra kernel expansion on an arbitrary Laguerre basis characterized by the pole ξ_p is infinite, $\rho_{o,p}$ can be recovered in using the associated Laguerre spectrum as stated by formula (27) with (14), (15), and (16) for the calculation of $R_{1,p}$ and $R_{2,p}$.

4. PARAMETERS ESTIMATION METHODS FOR LAGUERRE-VOLTERRA MODELS

In practical cases, the expansion of the Volterra kernels on a Laguerre basis is truncated to a finite-order K . Consequently, for a given Laguerre pole ξ_p , by using truncated expressions of (15) and (16), we get $\hat{\rho}_{o,p}$, which is an approximation of $\rho_{o,p}$, the actual characteristic of the system. This approximation depends on the Laguerre pole and on the Laguerre spectrum, via $R_{1,p}$ and $R_{2,p}$. For a fixed pole ξ_p , the optimization of the Laguerre spectrum improves the approximation $\hat{\rho}_{o,p}$. It is then possible to determine the optimal pole corresponding to the current approximation $\hat{\rho}_{o,p}$. This process is iterated until convergence. For a fixed Laguerre pole, the optimization of Fourier coefficients, therefore of $R_{1,p}$ and $R_{2,p}$, is done either by using the MMSE estimator (batch method) or by using a stochastic-gradient-based algorithm (adaptive method). The Laguerre pole optimization is done by using (28) with the approximated value $\hat{\rho}_{o,p}$.

We consider the Laguerre-Volterra filter described by (4):

$$y(n) = \sum_{p=1}^P \sum_{k_1=0}^{K-1} \cdots \sum_{k_p=0}^{K-1} g_{k_1, \dots, k_p} \prod_{j=1}^p s_{k_j, p}(n) = G^T S(n), \quad (29)$$

where $G = (G_1^T \cdots G_P^T \cdots G_P^T)^T$ and $S(n) = (S_1^T(n) \cdots S_P^T(n) \cdots S_P^T(n))^T$, G_p containing the Fourier coefficients g_{k_1, \dots, k_p} and $S_p(n)$ the p th-order cross-products of the $s_{k,p}(n)$ signals. In the sequel $d(n)$ will denote the actual output of the system to be modeled.

4.1. Batch method

With a batch method, for fixed Laguerre poles, the Fourier coefficients can be estimated using the well-known MMSE (minimum mean square error)-based solution such as

$$\hat{G} = \mathbf{\Gamma}^{-1} C, \quad (30)$$

where $\mathbf{\Gamma} = \mathbb{E}[S(n)S^T(n)]$, $C = \mathbb{E}[d(n)S(n)]$, and \mathbb{E} denotes the mathematical expectation. Note that to avoid eventual ill conditioning of the $\mathbf{\Gamma}$ matrix, an orthogonal formulation of the MMSE estimator can be used instead of the standard formula (30) (see, e.g., [10, 18, 19, 20]).

In order to derive an iterative procedure for Laguerre poles estimation, note that the combination of (20) with (23) yields

$$\frac{\partial J_p}{\partial \xi_p} = \frac{-R_{2,p}}{(1 - \xi_p^2)p \|h_p\|^2}. \quad (31)$$

As stated before, since J_p is a pseudoconvex function inside \mathcal{P} , then $R_{2,p}$ is zero when J_p is minimal and the reciprocal is true; that is, when $R_{2,p}$ is nearly zero, then J_p is close to its minimal value and consequently the pole ξ_p is close to its optimal value. In other words, when $R_{2,p}$ is nearly zero, the approximation $\hat{\rho}_{o,p}$ is the best possible one. Thus the iterative procedure can be stopped. This property enables us to derive the following batch estimation method.

TABLE 1: Computational complexity.

	Operations	Proposed method	Campello's method
Pole estimation	Add.	$3pK^p - pK^{p-1} + 3$	$(2p + 1)M^p - pM^{p-1} - 1$
	Mult. + Div.	$6pK^p - 3pK^{p-1} + 8$	$(3p + 1)M^p - pM^{p-1} + 6$
	Square root	1	1
Kernel reconstruction	Add.	—	$(K^p - 1)M^p$
	Mult.	—	$pK^p M^p$
Total		$O(pK^p)$	$O((p + 1)K^p M^p)$

- (1) Select arbitrary poles in the domain \mathcal{P} and build the corresponding Laguerre bases.
- (2) For $n = 0, \dots, N - 1$, filter the input $u(n)$ by the Laguerre filters defined by (5), calculate the cross products of the filtered inputs $s_{k,p}(n)$ for generating the vectors $S_p(n)$ and then $S(n)$.
- (3) Estimate the Laguerre spectrum for each kernel by means of (30) in replacing $\mathbf{\Gamma}$ and C by estimated values obtained as time-averaged values of the products $S(n)S^T(n)$ and $d(n)S(n)$.
- (4) For each basis, that is, $p = 1, \dots, P$, evaluate $T_{j,l}$ and $R_{j,p}$, $l = 1, \dots, p$, $j = 1, 2$, by using truncated versions of (15), (16), and formula (14).
- (5) For $p = 1, \dots, P$, if $R_{2,p}$ is close to zero, stop; else
 - (a) evaluate

$$\hat{\rho}_{o,p} = \frac{(1 + \xi_p^2)R_{1,p} + 2\xi_p R_{2,p}}{2\xi_p R_{1,p} + (1 + \xi_p^2)R_{2,p}}, \quad (32)$$

- (b) determine a new pole ξ_p using (28),
- (c) build a Laguerre basis associated with the pole ξ_p determined in the previous step, and return to 2 until convergence.

By comparing with Campello's method [11], the main advantage of the algorithm proposed herein is its low computational complexity (See Table 1). Without considering the Laguerre spectrum estimation and the input filtering by the OBFs, because they are common to both the proposed method and the method in [11], at each iteration the estimation of the Laguerre pole associated with the p th-order kernel requires $O(pK^p)$ operations. The computational cost reduction is due to the fact that, unlike Campello's method, no Volterra kernel reconstruction is required.

4.2. Adaptive method

This second method can be seen as a block-NLMS (normalized least mean square) method. The principle of this method is to adapt the Fourier coefficients until a given convergence criterion is satisfied, then the Laguerre poles are estimated, and the calculation is iterated with new data. Note that the Fourier coefficients are adapted by means of the NLMS algorithm (33) while the Laguerre poles are estimated

by using (28):

$$G(n+1) = G(n) + \frac{\mu}{a + \|S(n)\|^2} (d(n) - G^T(n)S(n))S(n), \quad (33)$$

where a is a small positive constant and μ the step size. The estimation of a new Laguerre pole occurs when the estimated Fourier coefficients are nearly constant during a given time window of length N_O . The adaptive procedure is summarized as follows.

(a) Parameters

- (i) K : truncation order.
- (ii) μ : step size ($0 < \mu < 2$).
- (iii) a : small positive constant.
- (iv) ϵ_p : convergence threshold.
- (v) N_O : window length.

(b) Initialization

- (i) Set the vector $G(0)$ equal to zero and $n = 0$.
- (ii) Set arbitrary initial values for ξ_p , $p = 1, \dots, P$, and build the associated Laguerre bases.

(c) Computation

- (1) Calculate the filtered inputs $s_{k,p}(n)$ associated with the Laguerre filters $B_{k,p}(z)$ defined by (5), organize the cross products of the filtered inputs into the vectors $S_p(n)$ to generate the vector $S(n)$.
- (2) Estimate the Laguerre spectra

$$G(n+1) = G(n) + \frac{\mu}{a + \|S(n)\|^2} (d(n) - G^T(n)S(n))S(n). \quad (34)$$

- (3) For $p = 1, \dots, P$,
 - (i) if $\|G_p(n+1) - G_p(n)\|^2 < \epsilon_p$ during the interval $[n+2 - N_O, n+1]$,
 - (a) calculate $R_{1,p,n}$ and $R_{2,p,n}$ by using truncated expressions of (15), (16), and (14),
 - (b) evaluate $\hat{\rho}_{o,p,n}$ by using (27),
 - (c) determine new Laguerre poles according to (28),
 - (d) build the associated Laguerre basis, increment n , and return to step 1;
 - (ii) else increment n and return to step 1.

Note that the experimental convergence of this algorithm is similar to that of a standard NLMS algorithm in a non-stationary environment due to the modification of the estimated Laguerre poles.

5. SIMULATION RESULTS

In this section, we present the results of two experiments that illustrate the good properties of the proposed algorithms. In the first example, the optimized Volterra filter is run with the same structure as that of the system to be identified. The second example involves experimenting with the Volterra filter under conditions of model mismatch. The measurement noise is assumed to be a white Gaussian process with zero mean. All the simulation results are obtained as ensemble averages over 50 independent runs. Note that, for the batch method, the iteration number 1 corresponds to the initialized values of the Laguerre poles in Figures 2, 3, 4, 5, and 9.

5.1. Example 1

The problem under consideration is that of identifying a second-order Volterra system described by

(i) first-order kernel

$$H_1(z) = \frac{z(z+0.5)}{(z-0.3)(z-0.2)}, \quad (35)$$

(ii) second-order kernel

$$h_2(i, j) = 0.25h(i)h(j), \quad (36)$$

where

$$h(i) = \mathcal{Z}^{-1} \left\{ \frac{z(z+1)}{(z-0.8)(z+0.8)} \right\}. \quad (37)$$

This system is simulated as a Volterra system with memory $M = 20$. Taking the symmetry of the quadratic kernel into account, this filter has 230 parameters to estimate. The input signal is a centered, white, Gaussian process with a unit variance. The output signal-to-noise ratio (SNR) is equal to 30 dB. $N = 5000$ input/output data are simulated. The identification performances are evaluated in terms of the normalized mean square error (NMSE):

$$\text{NMSE} = \frac{\sum_{n=0}^{N-1} (d(n) - y(n))^2}{\sum_{n=0}^{N-1} d^2(n)}, \quad (38)$$

$$\text{NMSE}_{\text{dB}} = 10 \log_{10} \text{NMSE}.$$

To validate the theoretical analysis presented in the previous sections and to evaluate the quality of the estimated Laguerre poles, the cost functions J_1 and J_2 defined in (8) are plotted in Figure 1. As stated before, these functions have a single minimum respectively located at $\xi_{1,\text{opt}} = 0.525$ and $\xi_{2,\text{opt}} = 0.733$.

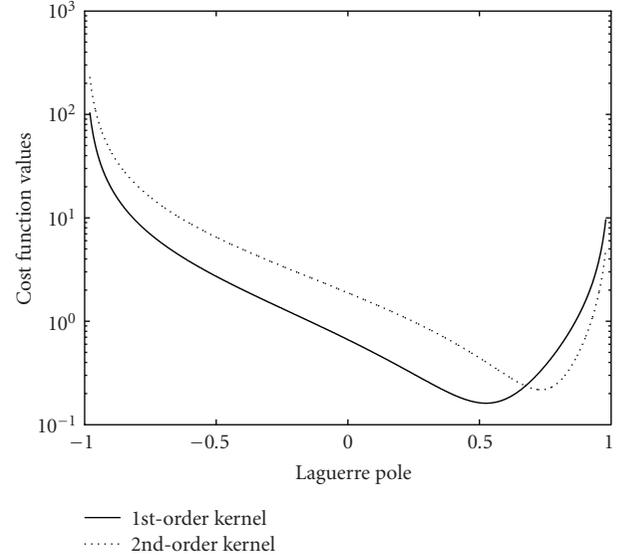


FIGURE 1: Cost functions J_1 (solid line) and J_2 (dotted line).

TABLE 2: Optimal poles and NMSE for different values of the truncation order.

K	ξ_1	ξ_2	NMSE_{dB}
3	0.543 ± 0.013	0.820 ± 0.012	-9.394 ± 0.443
4	0.573 ± 0.018	0.824 ± 0.014	-12.010 ± 0.587
5	0.540 ± 0.013	0.806 ± 0.006	-14.495 ± 0.199
6	0.526 ± 0.011	0.819 ± 0.003	-14.868 ± 0.182
7	0.511 ± 0.011	0.812 ± 0.006	-15.212 ± 0.220
8	0.517 ± 0.010	0.815 ± 0.003	-15.377 ± 0.219
9	0.504 ± 0.010	0.813 ± 0.005	-15.645 ± 0.195
10	0.509 ± 0.008	0.801 ± 0.003	-15.870 ± 0.165

One can note that the described system is not strictly causal, so the method of [11] cannot be applied. In incorporating a unit delay in the transfer functions, the resulting system fills the requirements of [11], and by applying this method we find, respectively, $\xi_{1,\text{cam}} = 0.525$ and $\xi_{2,\text{cam}} = 0.734$. These values are similar to the theoretical optimal values previously determined and are to be compared with those provided by the proposed estimation methods.

5.1.1. Batch estimation

The system identification is considered in a batch mode. The behavior of the proposed algorithm is particularly studied to point out the influence of three factors: the truncation order of the Laguerre expansions of the Volterra kernels, the initial choice of the Laguerre poles, and the stop criterion.

To evaluate the truncation order effects, several runs are done with different truncation orders. The Laguerre poles are initialized with the value 0.001. The obtained results are given in Table 2. Figure 2 illustrates the behavior of the estimated Laguerre poles during the iterations of the proposed

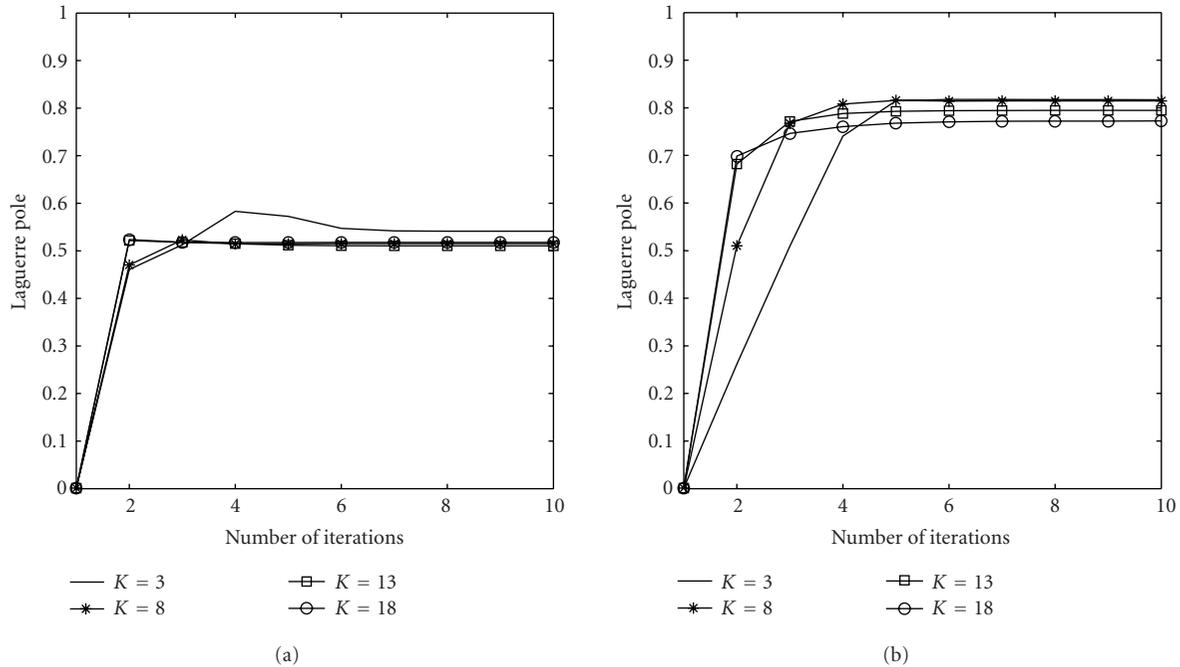


FIGURE 2: Laguerre poles estimation for different truncation orders: (a) first-order kernel and (b) second-order kernel.

algorithm. From the figure we can conclude that the algorithm converges in relatively few iterations. For a high truncation order ($K = 18$, e.g.), one iteration is sufficient to reach the optimal value while more iterations are needed for smaller truncation orders. Table 2 shows that the variations of the estimated poles values are very small when the truncation order varies between 3 and 10. However, for a larger truncation order, one can note an improvement of the value of the pole, more remarkable in the case of the quadratic kernel (Figure 2), and that of the convergence speed of the algorithm, with a higher computation cost. On the other hand, in the context of system identification, obviously the truncation order has a great influence on the precision of the identified model as indicated by the NMSE (Table 2). With a fixed truncation order, one can see that increasing the number of iterations improves not only the estimation of the Laguerre pole but also the overall precision of the identified model (Figure 3).

Figure 4 shows the effects of the Laguerre poles initialization on the convergence speed of the proposed batch algorithm, when the truncation order is fixed at $K = 7$. Although the initial values of the Laguerre poles are different, the algorithm converges toward the same Laguerre poles estimated values in few iterations. However, the convergence is faster when the initial value of the Laguerre pole is closer to the optimal value.

We have shown that when the estimated pole ξ_p reaches its optimal value, $R_{2,p}$ is equal to zero. We have also shown that $R_{2,p}$ is strictly decreasing in \mathcal{P} . Figure 5a plots the variations of $R_{2,1}$ and $R_{2,2}$. One can note that when these terms become almost null, the poles reach their optimal values (Figure 5b).

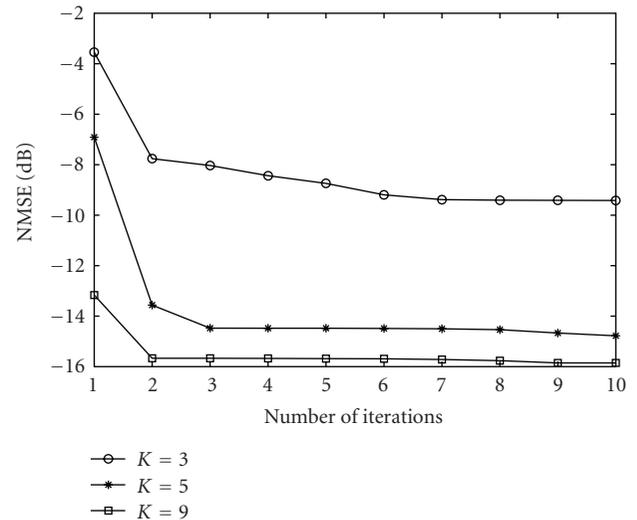


FIGURE 3: NMSE for different truncation orders.

From these simulation results we can conclude that

- (i) generally few iterations are needed to find the optimal poles;
- (ii) the convergence is faster when high truncation orders are considered and when the initial values of the Laguerre poles are close to their optimal values;
- (iii) when the Laguerre poles reach their optimal value, $R_{2,1}$ and $R_{2,2}$ are almost null, so these quantities can be used as a stop criterion for the proposed algorithm;
- (iv) the estimated poles are close to the theoretical values and to those obtained by applying the method in [11].

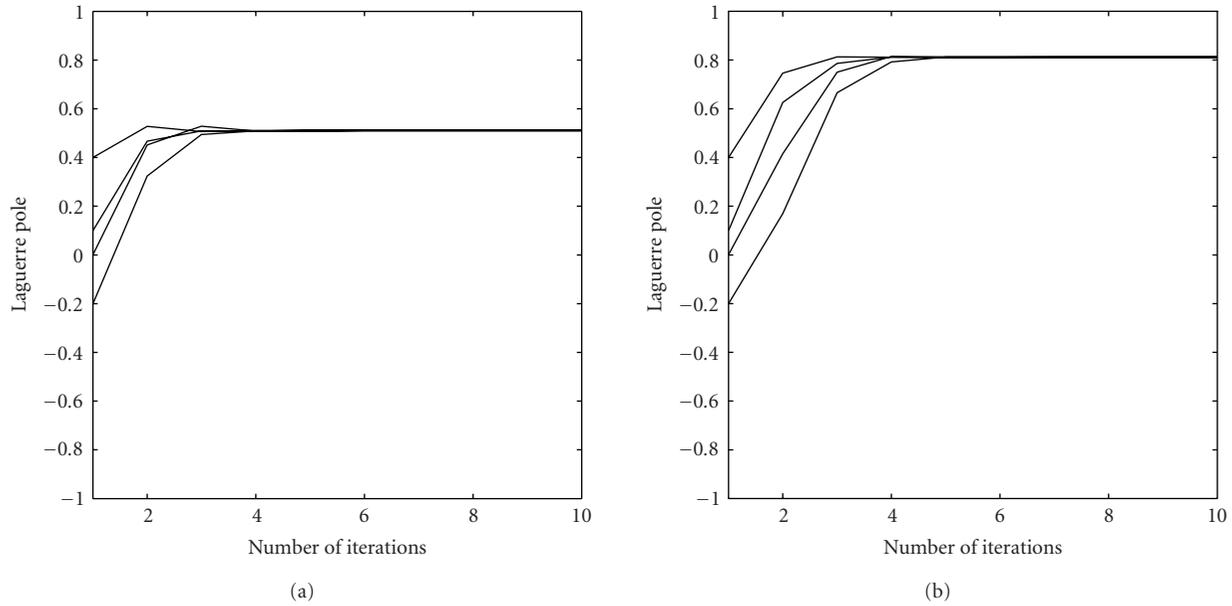


FIGURE 4: Laguerre poles estimation for different initial values. (a) Pole associated with the first-order kernel and (b) pole associated with the second-order kernel.

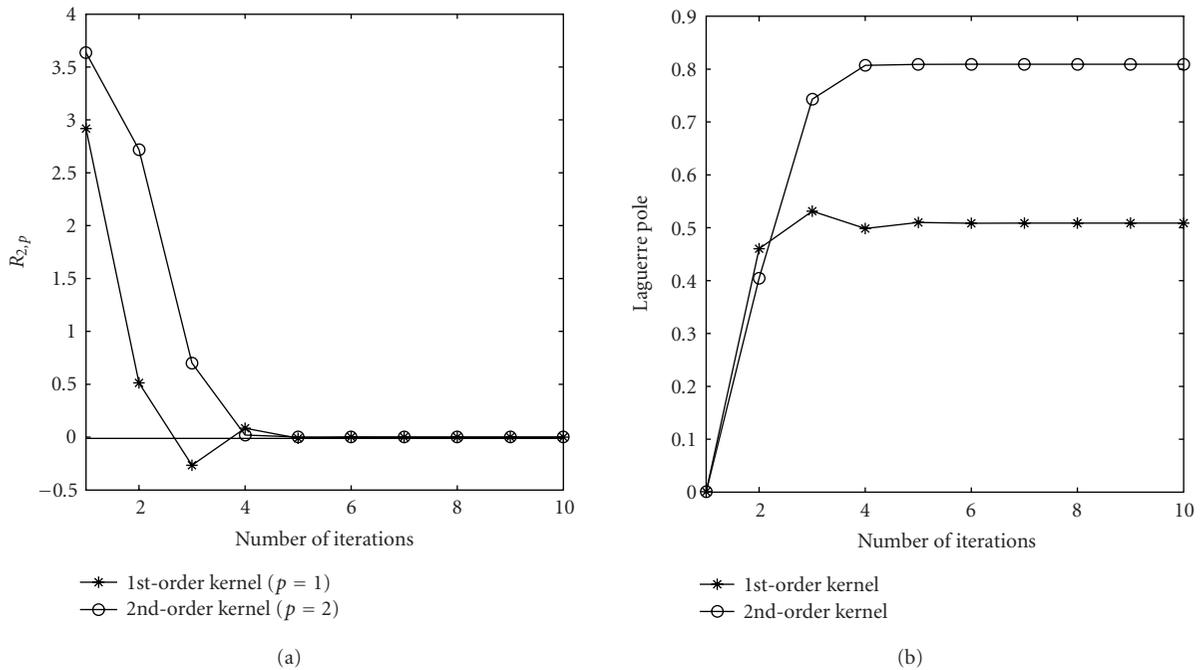


FIGURE 5: Variation of $R_{2,p}$ (a) and estimated poles (b) for $K = 7$.

5.1.2. Adaptive estimation

The same Volterra system is simulated with the same experimental conditions. Laguerre poles are initialized to 0. The Volterra kernels expansions on Laguerre bases are truncated to $K = 7$; thus the resulting Laguerre-Volterra filter has

only 35 Fourier coefficients to estimate while the standard Volterra model has 230 parameters. The step size of the NLMS algorithm is chosen as $\mu = 0.3$. The convergence thresholds are chosen equal to $\epsilon_p = 10^{-3}$, $p = 1, 2$, and the window length is $N_O = 50$.

The variations of the estimated Laguerre poles and of two Fourier coefficients are plotted in Figure 6 for both the linear and the quadratic kernels. One can note that the poles converge towards values close to those obtained with the block estimation method. One important question is that of the choice of the convergence threshold ϵ_p and of the window length N_O . When the value of ϵ_p is chosen too small, a huge number of iterations can be necessary for the convergence of a new estimated Laguerre pole. Consequently, particular care should be taken for choosing these parameters.

To illustrate the convergence of the overall identification procedure, the NMSE is also plotted (Figure 7). The evolution of the NMSE associated with the proposed adaptive algorithm is compared with that obtained when the Laguerre poles are arbitrarily chosen equal to 0.2 and when the Laguerre poles are chosen equal to their optimal values obtained from the plots of the cost functions J_1 and J_2 (Figure 1).

We also compare the proposed algorithm with the gradient method [5, 13] for Laguerre pole estimation. Recall that the gradient method consists in minimizing the output squared error both with respect to the Laguerre pole and the Fourier coefficients. Two different initializations are considered for both methods: ($\xi_1 = \xi_2 = 0$) and ($\xi_1 = 0.3; \xi_2 = 0.5$). The step sizes of the gradient method are chosen equal to 1/300 and 1/1000 for the adaptation of the Laguerre pole respectively associated with the linear and the second-order kernel.

As depicted in Figure 8, the proposed adaptive method performs much better than the gradient method. The gradient method is very sensitive to the choice of the step size, and its convergence to the optimal value is not guaranteed. The convergence can lead to local minima, under the influence of the initialization.

From these simulation results we can note the following.

- (i) In steady state, the performances of the proposed algorithm are similar to those obtained with the Laguerre-Volterra filter the poles of which are optimal. Obviously its convergence is slower.
- (ii) Both the adaptive and batch methods provide estimated poles close to the poles obtained with Campello's method and also close to those deduced from the plots of Figure 1. As previously stated, the advantage of the proposed methods is that the a priori knowledge of the Volterra kernels is not needed, and both Fourier coefficients and Laguerre poles are simultaneously estimated.
- (iii) The adaptive method performs better than the gradient method.

5.2. Example 2

Power amplifiers (PAs) are important elements in radio communication systems and they are inherently nonlinear. Some of PAs are modeled by a static nonlinearity such as a polynomial. However, for wideband applications (e.g., in wideband CDMA) and/or with high-power amplifiers (HPA) (e.g., base stations PAs), memory effects show up in the PA [21]. Having a memory means that the output of the PA is not only a function of the current input but also of past inputs. A relatively simple baseband behavioral model that accommodates memory as well as nonlinear behavior is the Wiener model, that is, a linear filter followed by a memoryless nonlinearity given by the Saleh model [22, 23]:

$$A(r) = \frac{2r}{1+r^2}, \quad (39)$$

where $A(r)$ and r are respectively the amplitude of the signal at the output and at the input of the memoryless HPA. The linear filter is a fourth-order Butterworth filter described in [24]:

$$H(z) = \frac{(0.078 + 0.1559z^{-1} + 0.078z^{-2})(0.0619 + 0.1238z^{-1} + 0.0619z^{-2})}{(1.0 - 1.3209z^{-1} + 0.6327z^{-2})(1.0 - 1.0486z^{-1} + 0.2961z^{-2})}. \quad (40)$$

The input signal is real and Gaussian. We consider its transmission over an AWGN (additive white Gaussian noise) channel with SNR= 30 dB. The overall channel constituted by an HPA with memory followed by an AWGN channel is nonlinear. In this example, we consider the identification of the HPA with memory modeled as a linear-cubic Volterra system by using $N = 15000$ data. We first evaluate the batch estimation method.

The truncation order is chosen equal to $K = 9$. This choice corresponds to a trade-off between the model

complexity and the model precision. From Figure 9 we can conclude that the Laguerre poles respectively associated with the linear and the cubic kernels converge to the same value. This behavior was predictable since the dominant dynamics associated with the linear and the cubic kernels are the same for a Wiener model.

Moreover, we evaluate the proposed adaptive algorithm for two kinds of input: a white Gaussian signal and a uniformly distributed 8-PAM signal, which is a signal often used in digital communications. One can see that the estimated

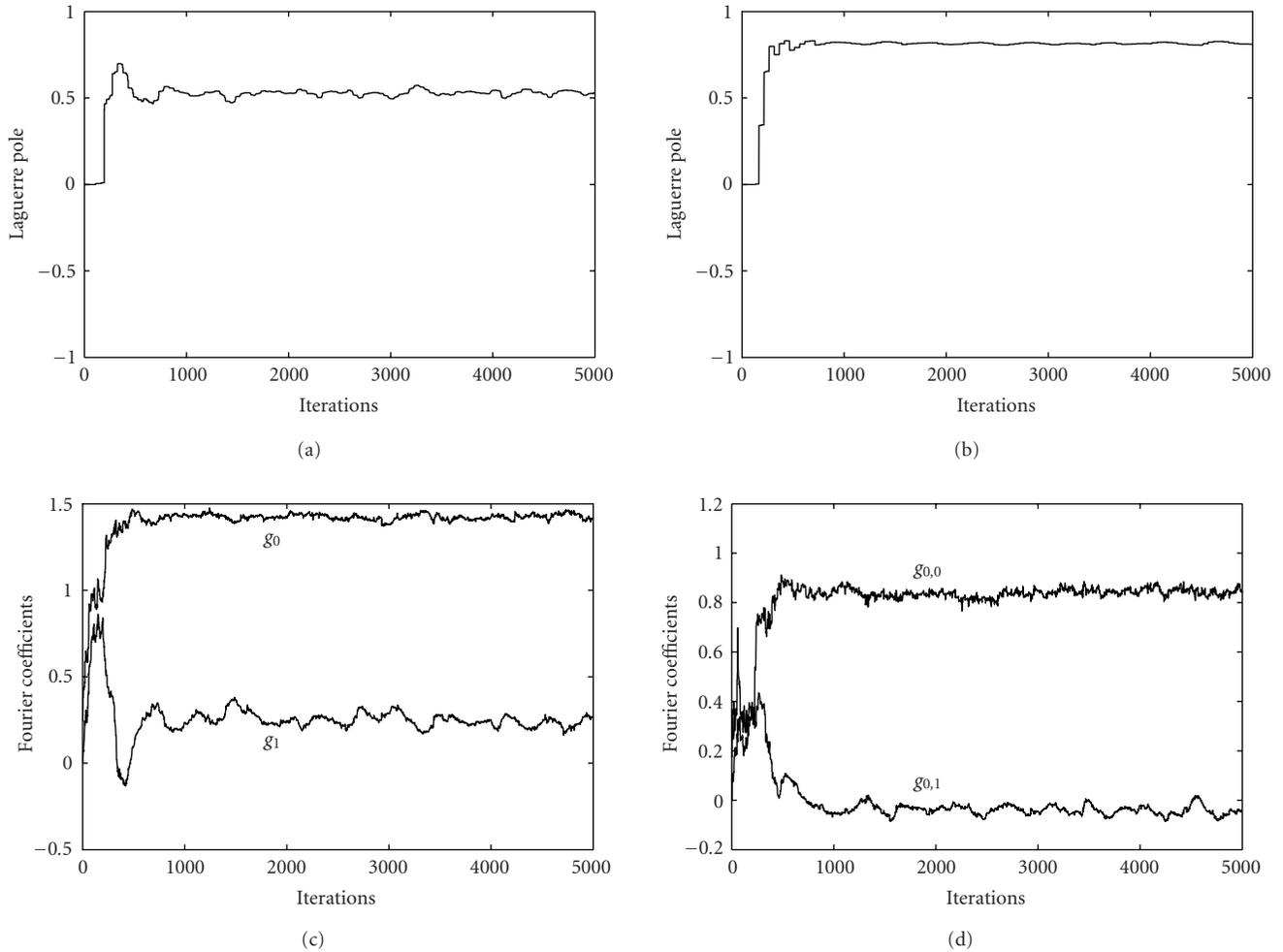


FIGURE 6: (a), (b) Estimated Laguerre poles and (c), (d) two Fourier coefficients corresponding to the (c) linear and (d) quadratic kernels.

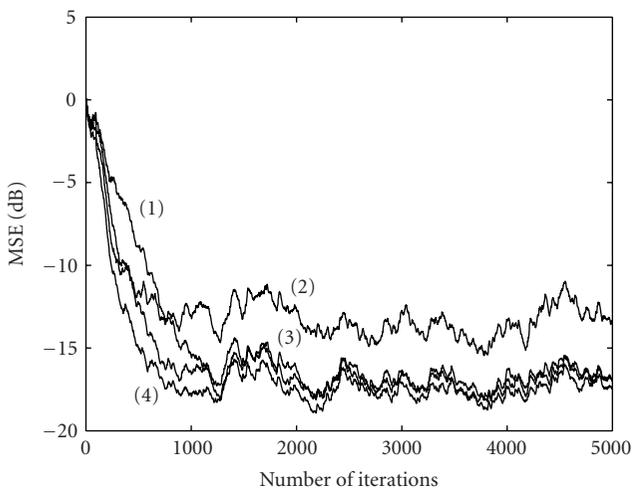


FIGURE 7: Comparison of four configurations of Laguerre-Volterra filters: (1) Laguerre poles optimized with the proposed adaptive algorithm, (2) arbitrarily fixed Laguerre poles for $K = 7$, (3) arbitrarily fixed Laguerre poles for $K = 9$, and (4) Laguerre poles fixed at their optimal values.

poles converge to the same values as those obtained with the block estimation method (Figure 10). The performances, in terms of NMSE, of the Laguerre-Volterra filter with optimized poles are plotted in Figure 11. Obviously the Gaussian input allows to get better performance since it is more exciting than the 8-PAM input. However, in terms of Laguerre poles estimation, the proposed algorithm delivers quite similar estimated poles with a Gaussian signal and a PAM signal, the algorithm convergence speed being faster with the Gaussian signal.

6. CONCLUSION

In this paper, the optimization of Volterra kernels expansions on Laguerre bases has been addressed. An analytical solution to this optimization problem has been obtained using Volterra kernels expansions on independent Laguerre bases. This solution has been expressed in terms of the estimated Laguerre spectrum associated with each Volterra kernel. Then batch and adaptive methods have been proposed to optimize both Laguerre spectra and Laguerre

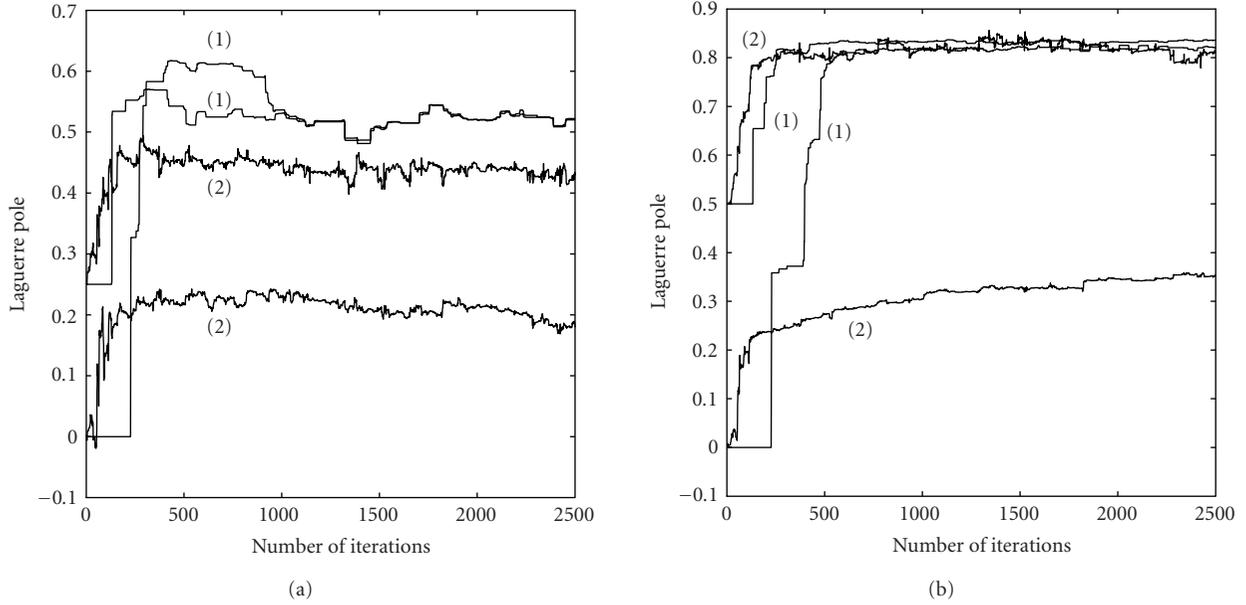


FIGURE 8: Adaptive estimation of Laguerre poles: (1) proposed algorithm and (2) gradient method. (a) First-order kernel and (b) second-order kernel.

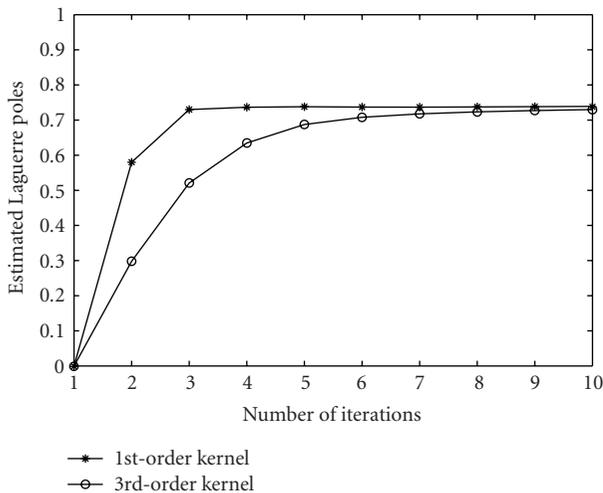


FIGURE 9: Laguerre poles estimated with the block estimation method.

poles. This approach generalizes previous works for linear [13] and nonlinear systems [11]. The performance of the proposed identification methods has been illustrated by means of simulation results that show the usefulness of the proposed methods. Theoretical stability and convergence analysis of the adaptive method will be the subject of further study. Among the OBFs of rational type, the Laguerre basis is the simplest one to optimize. However,

they generally present worse performances for system approximation compared with generalized orthonormal bases (GOB) [25]. The authors have recently proposed some GOB selection procedures for Volterra kernels expansions [7, 26]. These methods based on the minimization of a least squares criterion are exhaustive in nature, that is, they search the optimal poles in a finite set resulting from the discretization of the interval $]-1, 1[$. In future works the authors intend to investigate analytical solution to the problem of GOB optimization for Volterra kernels expansions.

APPENDICES

A. DERIVATION OF FORMULA (18)

By assuming that $\sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} |h_p(i_1, \dots, i_p)| < \infty$, and using formula (3), the derivative of the Fourier coefficient with respect to the pole ξ_p is given by

$$\begin{aligned} & \frac{\partial g_{k_1, \dots, k_p}}{\partial \xi_p} \\ &= \frac{\partial}{\partial \xi_p} \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) \prod_{j=1}^p b_{k_j, p}(i_j) \\ &= \sum_{l=1}^p \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) \frac{\partial b_{k_l, p}(i_l)}{\partial \xi_p} \prod_{j=1, j \neq l}^p b_{k_j, p}(i_j). \end{aligned} \quad (\text{A.1})$$

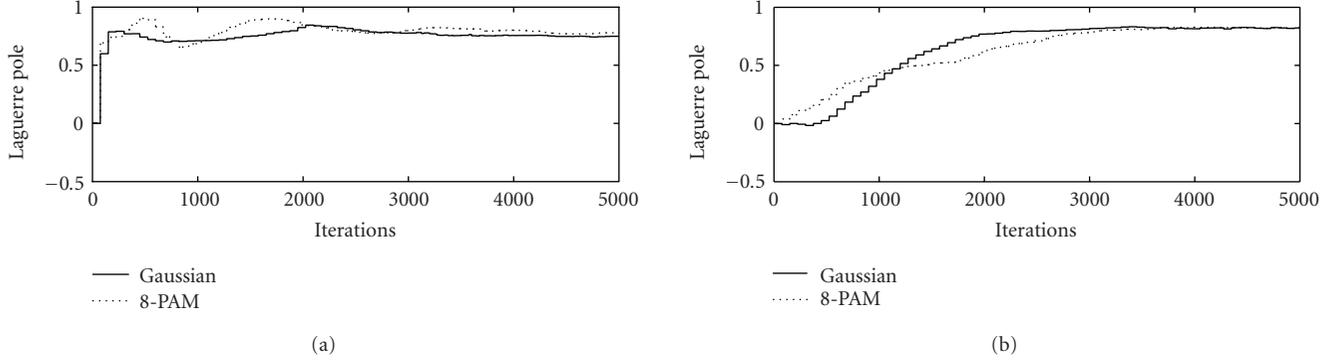


FIGURE 10: Laguerre poles estimated with the adaptive algorithm: (a) first-order kernel and (b) third-order kernel.

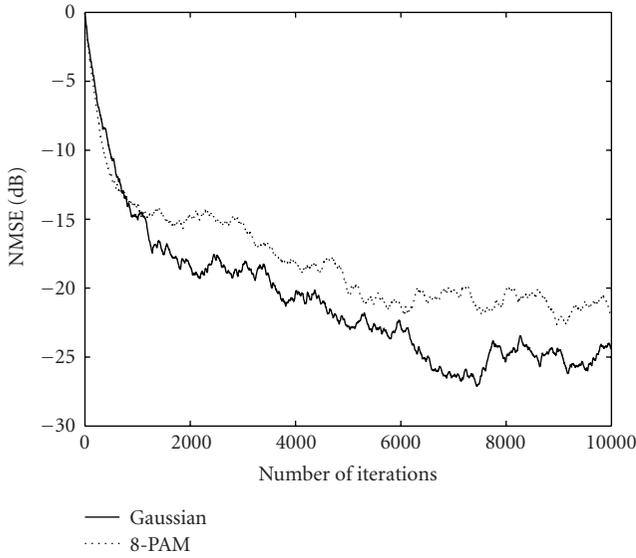


FIGURE 11: NMSE for two different input signals.

Then the relation (7) yields

$$\begin{aligned}
 & \frac{\partial g_{k_1, \dots, k_p}}{\partial \xi_p} \\
 &= \frac{1}{1 - \xi_p^2} \sum_{l=1}^p \left[(k_l + 1) \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) b_{k_l+1, p}(i_l) \right. \\
 & \quad \left. \times \prod_{j=1, j \neq l}^p b_{k_j, p}(i_j) \right] \\
 & - \frac{1}{1 - \xi_p^2} \sum_{l=1}^p k_l \sum_{i_1=0}^{\infty} \cdots \sum_{i_p=0}^{\infty} h_p(i_1, \dots, i_p) b_{k_l-1, p}(i_l) \\
 & \times \prod_{j=1, j \neq l}^p b_{k_j, p}(i_j) \\
 &= \frac{1}{1 - \xi_p^2} \sum_{l=1}^p [(k_l + 1) g_{k_1, \dots, k_l+1, \dots, k_p} - k_l g_{k_1, \dots, k_l-1, \dots, k_p}]. \tag{A.2}
 \end{aligned}$$

B. PROOF OF LEMMA 1

We assume that all series involved in this proof are uniformly convergent. The derivative of $T_{1,l}$, $l = 1, \dots, p$, is given by

$$\frac{\partial T_{1,l}}{\partial \xi_p} = 2 \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} (2k_l + 1) g_{k_1, \dots, k_p} \frac{\partial g_{k_1, \dots, k_p}}{\partial \xi_p}. \tag{B.1}$$

By using (18), we get

$$\begin{aligned}
 \frac{\partial T_{1,l}}{\partial \xi_p} &= \frac{2}{1 - \xi_p^2} \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} (2k_l + 1) g_{k_1, \dots, k_p} \\
 & \times \sum_{j=1}^p ((k_j + 1) g_{k_1, \dots, k_j+1, \dots, k_p} - k_j g_{k_1, \dots, k_j-1, \dots, k_p}). \tag{B.2}
 \end{aligned}$$

Thus, we have

$$\begin{aligned}
 & \frac{1 - \xi_p^2}{2} \frac{\partial T_{1,l}}{\partial \xi_p} \\
 &= \sum_{j=1}^p \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} (2k_l + 1) (k_j + 1) g_{k_1, \dots, k_j+1, \dots, k_p} g_{k_1, \dots, k_p} \\
 & - \sum_{j=1}^p \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} (2k_l + 1) k_j g_{k_1, \dots, k_j-1, \dots, k_p} g_{k_1, \dots, k_p}. \tag{B.3}
 \end{aligned}$$

Expanding the above relation gives

$$\frac{1 - \xi_p^2}{2} \frac{\partial T_{1,l}}{\partial \xi_p} = -2 \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} k_l g_{k_1, \dots, k_l-1, \dots, k_p} g_{k_1, \dots, k_p}. \tag{B.4}$$

From this relation and definition (16) we get

$$\frac{\partial T_{1,l}}{\partial \xi_p} = \frac{-2}{1 - \xi_p^2} T_{2,l}. \tag{B.5}$$

Similarly it can be shown that

$$\frac{\partial T_{2,l}}{\partial \xi_p} = \frac{-2}{1 - \xi_p^2} T_{1,l}. \quad (\text{B.6})$$

With these relations the derivative of $R_{1,p}$ is given by

$$\frac{\partial R_{1,p}}{\partial \xi_p} = \sum_{l=1}^p \frac{\partial T_{1,l}}{\partial \xi_p} = \frac{-2}{1 - \xi_p^2} \sum_{l=1}^p T_{2,l} = \frac{-2}{1 - \xi_p^2} R_{2,p}. \quad (\text{B.7})$$

Similarly, we have

$$\frac{\partial R_{2,p}}{\partial \xi_p} = \frac{-2}{1 - \xi_p^2} R_{1,p}. \quad (\text{B.8})$$

C. PROOF OF LEMMA 2

Recall that from (23) we have

$$\frac{\partial R_{1,p}}{\partial \xi_p} = 2p \|h_p\|^2 \frac{\partial J_p}{\partial \xi_p}. \quad (\text{C.1})$$

Knowing that

$$\frac{\partial J_p}{\partial \xi_p} = \frac{-2}{(1 - \xi_p^2)^2} ((1 + \xi_p^2) Q_{2,p} - 2\xi_p Q_{1,p} - \xi_p), \quad (\text{C.2})$$

we get

$$\frac{\partial R_{1,p}}{\partial \xi_p} = \frac{-4p \|h_p\|^2}{(1 - \xi_p^2)^2} ((1 + \xi_p^2) Q_{2,p} - 2\xi_p Q_{1,p} - \xi_p). \quad (\text{C.3})$$

Using (20) leads to

$$R_{2,p} = \frac{2p \|h_p\|^2}{1 - \xi_p^2} ((1 + \xi_p^2) Q_{2,p} - 2\xi_p Q_{1,p} - \xi_p), \quad (\text{C.4})$$

which is one of the desired results, that is, formula (25) of Lemma 2.

Moreover, by taking the derivative of $R_{2,p}$, that is,

$$\frac{\partial R_{2,p}}{\partial \xi_p} = 2p \|h_p\|^2 \frac{4\xi_p Q_{2,p} - 2(1 + \xi_p^2) Q_{1,p} - (1 + \xi_p^2)}{(1 - \xi_p^2)^2}, \quad (\text{C.5})$$

and by substituting it in (21), the result (24) follows.

REFERENCES

- [1] M. J. Korenberg, "Parallel cascade identification and kernel estimation for nonlinear systems," *Annals of Biomedical Engineering*, vol. 19, no. 4, pp. 429–455, 1991.
- [2] T. M. Panicker and V. J. Mathews, "Parallel-cascade realizations and approximations of truncated Volterra systems," *IEEE Trans. Signal Processing*, vol. 46, no. 10, pp. 2829–2832, 1998.
- [3] A. Khouaja, A. Kibangou, and G. Favier, "Third-order Volterra kernels complexity reduction using PARAFAC," in *Proc. 1st IEEE International Symposium on Control, Communications and Signal Processing (ISCCSP '04)*, pp. 857–860, Hammamet, Tunisia, March 2004.
- [4] A. Khouaja and G. Favier, "Identification of PARAFAC-Volterra cubic models using an alternating recursive least-squares algorithm," in *Proc. 12th European Signal Processing Conference (EUSIPCO '04)*, pp. 1903–1906, Vienna, Austria, September 2004.
- [5] R. Hacıoglu and G. A. Williamson, "Reduced complexity Volterra models for nonlinear system identification," *EURASIP Journal on Applied Signal Processing*, vol. 2001, no. 4, pp. 257–265, 2001.
- [6] R. J. G. B. Campello, W. C. do Amaral, and G. Favier, "Optimal Laguerre series expansion of discrete Volterra models," in *Proc. European Control Conference (ECC '01)*, pp. 372–377, Porto, Portugal, September 2001.
- [7] A. Y. Kibangou, G. Favier, and M. M. Hassani, "Generalized orthonormal basis selection for expanding quadratic Volterra filters," in *Proc. 13th IFAC Symposium on System Identification (SYSID '03)*, pp. 1119–1124, Rotterdam, The Netherlands, August 2003.
- [8] A. Y. Kibangou, G. Favier, and M. M. Hassani, "Selection of generalized orthonormal bases for second-order Volterra filters," to appear in *Signal Processing*.
- [9] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans. Automat. Contr.*, vol. 36, no. 5, pp. 551–562, 1991.
- [10] N. Wiener, *Nonlinear Problems in Random Theory*, John Wiley & Sons, New York, NY, USA, 1958.
- [11] R. J. G. B. Campello, G. Favier, and W. C. do Amaral, "Optimal expansions of discrete-time Volterra models using Laguerre functions," *Automatica*, vol. 40, no. 5, pp. 815–822, 2004.
- [12] Y. Fu and G. A. Dumont, "An optimum time scale for discrete Laguerre network," *IEEE Trans. Automat. Contr.*, vol. 38, no. 6, pp. 934–938, 1993.
- [13] N. Tanguy, R. Morvan, P. Vilbé, and L. C. Calvez, "Online optimization of the time scale in adaptive Laguerre-based filters," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1184–1187, 2000.
- [14] S. Boyd and L. O. Chua, "Fading memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Trans. Circuits Syst.*, vol. 32, no. 11, pp. 1150–1161, 1985.
- [15] T. Oliveira e Silva, "On the determination of the optimal pole position of Laguerre filters," *IEEE Trans. Signal Processing*, vol. 43, no. 9, pp. 2079–2087, 1995.
- [16] N. Tanguy, P. Vilbé, and L. C. Calvez, "Optimum choice of free parameter in orthonormal approximations," *IEEE Trans. Automat. Contr.*, vol. 40, no. 10, pp. 1811–1813, 1995.
- [17] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1993.
- [18] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, John Wiley & Sons, New York, NY, USA, 1980.
- [19] M. J. Korenberg, S. B. Bruder, and P. J. McIlroy, "Exact orthogonal kernel estimation from finite data records:

extending Wiener's identification of nonlinear systems," *Annals of Biomedical Engineering*, vol. 16, no. 2, pp. 201–214, 1988.

- [20] V. J. Mathews, "Orthogonalization of correlated Gaussian signals for Volterra system identification," *IEEE Signal Processing Lett.*, vol. 2, no. 10, pp. 188–190, 1995.
- [21] J. Kim and K. Konstantinou, "Digital predistortion of wide-band signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417–1418, 2001.
- [22] A. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, 1981.
- [23] M. Ibnkahla, N. J. Bershad, J. Sombrin, and F. Castanié, "Neural network modeling and identification of nonlinear channels with memory: algorithms, applications, and analytic models," *IEEE Trans. Signal Processing*, vol. 46, no. 5, pp. 1208–1220, 1998.
- [24] J. Lee and V. J. Mathews, "A fast recursive least squares adaptive second-order Volterra filter and its performance analysis," *IEEE Trans. Signal Processing*, vol. 41, no. 3, pp. 1087–1102, 1993.
- [25] B. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Trans. Automat. Contr.*, vol. 42, no. 4, pp. 515–521, 1997.
- [26] A. Y. Kibangou, G. Favier, and M. M. Hassani, "A growing approach for selecting generalized orthogonal basis functions in the context of system modelling," in *Proc. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP '03)*, Grado, Italy, June 2003.

Alain Y. Kibangou received the B.S. and M.S. degrees from the Cadi Ayyad University of Marrakesh (Morocco) respectively in physics (1998) and electrical engineering (2000). He received the Ph.D. degree in 2005 jointly from the University of Nice Sophia Antipolis (France) and from the University of Marrakesh in automatic control and signal and image processing. He is currently a postdoctoral researcher at the I3S Laboratory (Sophia Antipolis, France). His research interests include signal processing for communications and nonlinear systems identification and equalization.



Gérard Favier was born in Avignon, France, in 1949. He received the Engineering Diplomas from ENSCM (École Nationale Supérieure de Chronométrie et de Micromécanique), Besançon, and ENSAE (École Nationale Supérieure de l'Aéronautique et de l'Espace), Toulouse, and the Engineering Doctorate and State Doctorate degrees from the University of Nice Sophia Antipolis, in 1973, 1974, 1977 and 1981, respectively. In 1976, he joined the CNRS (Centre National de la Recherche Scientifique), and now he works as a Research Director of CNRS at the I3S Laboratory, Sophia Antipolis. From 1995 to 1999, he was the Director of the I3S Laboratory. His present research interests include nonlinear process modeling and identification, tensor modeling of wireless communication systems, nonlinear and blind equalization, and predictive control.



Moha M. Hassani was born in Morocco, in 1954. He received the Third Cycle Doctorate in automatic and signal processing from The University of Nice, France, in 1982, and the Ph.D. degree in electrical engineering from the University of Sherbrooke, Canada, in 1992. He joined Cadi Ayyad University, Marrakesh, Morocco, in 1982 as an Assistant Professor. Since 1993, he has been a Professor at the Faculty of Sciences Semlalia, Marrakesh, Morocco. His research interests are mainly nonlinear process modeling and identification applied to telecommunication and solar energy fields.

